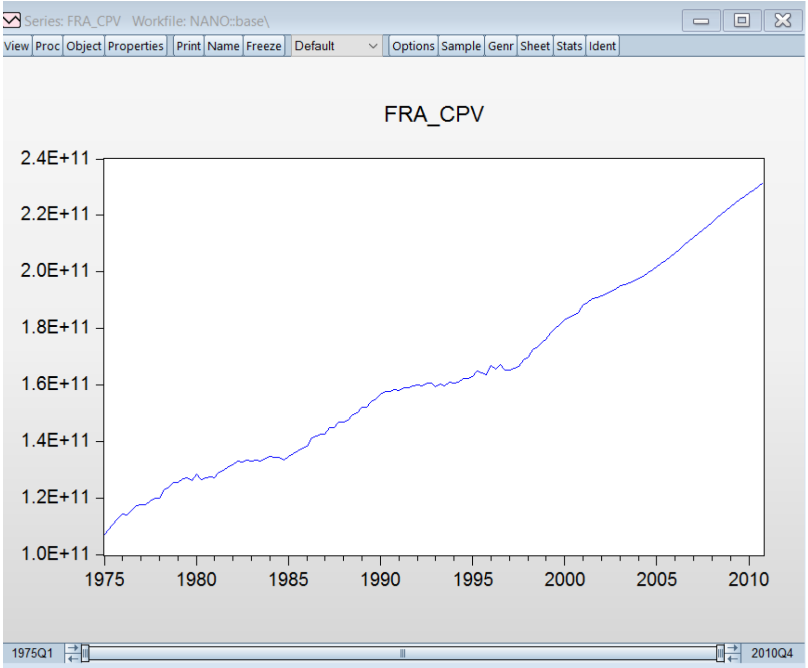
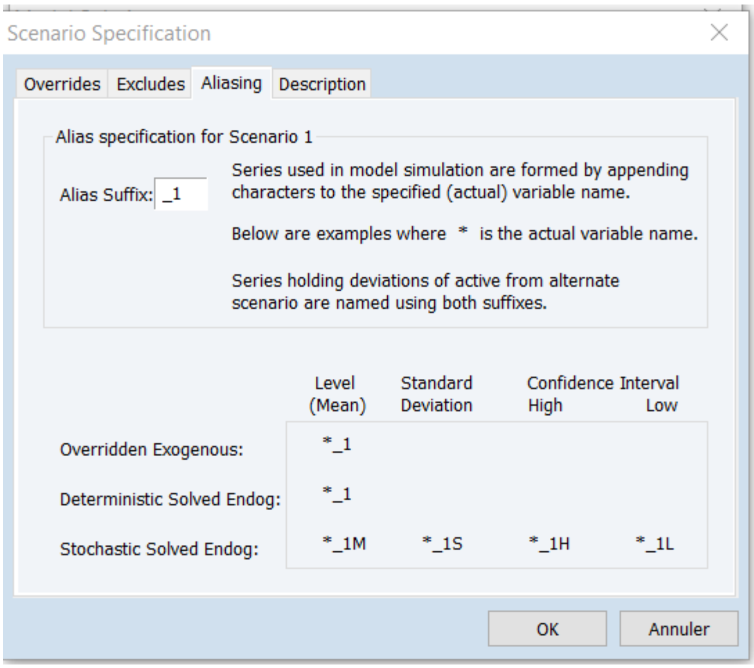
STRUCTURAL ECONOMETRIC MODELLING METHODOLOGY AND TOOLS WITH APPLICATIONS UNDER EVIEWS

JEAN-LOUIS BRILLET.

CONTENTS


Series: FRA_CPV Workfile: NANO::base\
View Proc Object Properties Print Name Freeze Default Options Sample Genr Sheet Stats Ident
FRA_CPV
2.4E+11
2.2E+11
2.0E+11
1.8E+11
1.6E+11
1.4E+11
1.2E+11
1.0E+11
1975 1980 1985 1990 1995 2000 2005 2010
1975Q1
2010Q4


Scenario Specification
Overrides
Excludes
Aliasing
Description
Alias specification for Scenario 1
Alias Suffix: _1
Series used in model simulation are formed by appending characters to the specified (actual) variable name.
Below are examples where * is the actual variable name.
Series holding deviations of active from alternate scenario are named using both suffixes.
Level (Mean)
Standard Deviation
Confidence Interval
High
Low
Overridden Exogenous:
*_1
Deterministic Solved Endog:
*_1
Stochastic Solved Endog:
*_1M
*_1S
*_1H
*_1L
OK
Annuler

FOREWORD

The purpose of this book is a little special.

First, of course, by its subject: structural econometric modelling no longer looks so fashionable, having lost ground to Computable General Equilibrium models and in particular their Dynamic Stochastic versions.

We will contend that while this might be true in the academic field (you just have to look at the program of congresses and symposiums) there is still a lot of room for structural models. Indeed, many institutions are still using them and even building new ones, both in developed and developing countries. We shall try to show that this position is quite justified, and that for a large part of the modelling applications, in particular the analysis and interpretation of macroeconomic interactions, the call for structural models remains a good strategy, arguably the best one.

But we shall not stop at proving the usefulness of these models. For the people we have convinced, or which were so already, we will provide a set of tools facilitating all the tasks in the modelling process. Starting from elementary elements, it will lead by stages the user to a level at which he should be able to build, manage and use his professional, operational model.

This means this book will, as its title says, focus essentially on applied and even technical features, which does not mean it will be so simplistic.

After a necessary description of the field, we shall use the largest part of the book to show the reader how to build his own model, from general strategies to technical details. For this we shall rely on a specific example, presented at the beginning, and which we will follow through all the steps of model development. When the situation becomes more complex (with the addition of product and international dimensions), we shall still find this model at the core of the cases.

This model will also be present in the background, when we address new directions, which we think are quite compatible with our approach: Quasi-Accounting, and Stock-Flow Consistent models.

Our examples will be based on EViews package, the most popular modeling product presently available This will allow us to be more helpful to EViews users, concentrating on its practice (including some tricks).

Finally, just as important if not more so, we shall provide a set of files allowing readers to practice modelling (either alone or as part of a course). And for more advanced users, we shall give access to files allowing to produce operational (if small) models, which they can adapt to their own ideas, with the tedious tasks: producing the data, defining the accounting framework and organizing simulations over the future, being already prepared.

All these elements are provided for free, and downloadable on the EViews site, at the address:

https://www.eviews.com/StructModel/structmodel.html

This version of the book takes into account the features of the last version of EViews; EViews 13. However, most of the text is valid for earlier versions. The main differences come from improvements in the user-friendliness.

INTRODUCTION

Since an early date in the twentieth century, economists have tried to produce mathematical tools which, applied to a given practical problem, formalized a given economic theory to produce a reliable numerical picture. The most natural application is of course to forecast the future, and indeed this goal was present from the first. But one can also consider learning the consequences of an unforeseen event, or measuring the efficiency of a change in the present policy, or even improving the understanding of a set of mechanisms too complex to be grasped by the human mind.

In the beginning (let us say since the 1930s) the field was occupied by the "structural" models. They start from a given economic framework, defining the behaviors of the individual agents according to some globally consistent economic theory. They use the available data to associate to these behaviors reliable formulas, which are linked by identities guaranteeing the consistency of the whole set. These models can be placed halfway between the two above categories: they do rely on statistics, and also on theory. To accept a formula, it must respect both types of criteria.

The use of this kind of models, which occupied the whole field at the beginning, is now mostly restricted to policy analysis and medium-term forecasting. For the latter, they show huge advantages: the full theoretical formulations provide a clear and understandable picture, including the measurement of individual influences. They allow also to introduce stability constraints leading to identified long-term equilibriums, and to separate this equilibrium from the dynamic fluctuations which lead to it.

In the last decades, new kinds of models have emerged, which share the present market.

- The "VAR" models. They try to give the most reliable image of the near future, using a complex estimated structure of lagged elements, based essentially on the statistical quality, although economic theory can be introduced, mostly through constraints on the specifications. The main use of this tool is to produce short-term assessments.
- The Quasi-Accounting models, which rely on very basic behaviors, most of the time calibrated. This allows to treat cases where data is available for extremely limited sample periods, or where the fine detail (generally in products) forbids to apply econometrics with a good chance of global success.
- Stock-Flow Consistent models, which answer two criticisms addressed to structural models: producing incomplete and formally unbalanced models, and not taking enough into account the stocks, in particular of financial assets. By detailing these assets by agent and category, SFCs allow to consider sophisticated financial behaviors, sometimes at the expense of the "real" properties.
- And last (but not the least) Computable General Equilibrium models. They use a detailed structure with a priori formulations and calibrated coefficients to solve a generally local problem, through the application of one or several optimizing behaviors. The issues typically addressed are optimizing resource allocations, or describing the consequences of trade agreements. The mechanisms described contain generally little dynamics.

This is no longer true for the Dynamic Stochastic General Equilibrium models, which dominate the current field. They include dynamic behaviors and take into account the uncertainty in economic evolutions. Compared to the traditional models (see later) they formalize explicitly the optimizing equilibria, based on the aggregated behavior of individual agents. This means that they allow agents to adapt their behavior to changes is the rules governing the behaviors of others, including the State, in principle escaping the Lucas critique. As the model does not rely on traditional estimated equations, calibration is required for most parameters.

Compared to CGEs and DSGEs, optimization behaviors are present (as we shall see later) and introduced in the estimated equations. But they are frozen there, in a state associated with a period, and the behavior of other agents at the time. If these conditions do not change, the statistical validation is an important advantage. But sensitivity to shocks is flawed, in a way which is difficult to measure.

A very good (and objective) description of the issue can be found in:

http://en.wikipedia.org/wiki/Dynamic_stochastic_general_equilibrium

http://en.wikipedia.org/wiki/Macroeconomic_model#Empirical_forecasting_models

It seems to us that the main criterion in the choice between DSGEs and traditional structural models lie in the tradeoff between statistical validation and adaptability of behaviors.

In the last years, popularity of structural econometric modelling seems to have stabilized. A personal hint for this (if not an actual proof) is the sustained demand for participation in structural modelling projects, observed on the sites of companies devoted to international cooperation.

Another issue is that being the first tool produced (in the thirties of the last century) it was applied immediately to the ambitious task of producing reliable forecasts. The complexity of the economy, and the presence of many random shocks makes this completely unrealistic (and this is even more true today). During the golden years of structural modelling, when economy was growing at a regular (and high) rate, forecasting was as easy as riding a tame horse on a straight path: anybody could do it. But when the horse turned into a wild one, the quality of the rider showed, and it did not stay in the saddle too long. Failing to succeed in a task too difficult for any tool (including VAR and CGE models, which do not have to forecast the medium-term), gave discredit to structural models and all their uses, including policy analysis and even the understanding and interpretation of complex economic mechanisms, applications for which neither VAR nor CGE can compete in our opinion.

Also, the role of financial issues has grown, which the initial structural models were not well equipped to address. But Stock-Flow Consistent versions can be an answer to this problem.

Anyway, even with limited ambitions, producing a sound econometric structural model is not a simple task. Even a professional economist, having an excellent knowledge of both economic theory (but not necessarily a complete and consistent picture) and econometric techniques (but not necessarily of their practical application) will find it quite difficult producing a reliable and operational econometric model.

The purpose of this book is to shorten the learning process, in several ways.

After a global presentation of economic models:

- Notations, definitions, mathematical characteristics (dynamics, linearity, continuity, identifiability…).
- Applications: economic theory, forecast, education.
- Classification of existing models.

We shall describe how to organize the sequence of model building tasks, from data production and framework specification to actual operational studies.

For each task, we shall give all the necessary elements of methodology.

We shall present the main economic options available, with some theoretical explanations.

All these explanations will be based on a practical example, the production of a very small model of the French economy. The size will not forbid us to address most of the problems encountered in the process.

The methods, techniques and solutions proposed will be based on the EViews software. This will allow us to present some useful features and tricks, and to provide a sequence of complete programs, which the user can modify at will, but not necessarily too heavily, as all the models of this type share a number of common elements. The main issue is of course the estimation process, each case leading generally to an original version of each behavioral equation.

A set of documented programs is available on demand, following the above principles

- For the small example,
- For a more detailed product, a model for a single country, not far from an operational version.

These programs will allow to:

- Import the original data
- Build the model framework
- Transform the data to conform to the elements in the model.
- Estimate a set of equations, starting with standard behaviors, possibly updated.
- Check the technical and theoretical consistency of the resulting model.
- Produce forecasts and policy studies.

In each case, we shall present programs which actually work. An econometric solution will be found, reliable both in statistical and economic terms. And the properties of the models will be rather satisfying, with a long-term solution and reasonable dynamics leading to it.
Finally, we shall address the more complex problems: multi-sector and multi-country models (and both options combined). The specific issues will be described, and a framework for a three-product model will be provided, following the same lines as the previous example.

The goal of this book is therefore both limited and ambitious. Without getting into theoretically complex features, it should give readers all the elements required to construct their own model. Being relieved of the more technical (and tedious) tasks, they will be allowed to concentrate on the more intelligent (and interesting) ones.

Readers must be aware they will find here neither a full description of econometric and statistical methods, nor a course in economic theory. We shall give basic elements on these fields, and rather focus on their links with the modelling process itself. For more detailed information, one can refer to the list of references provided at the end of the volume.

Concerning Quasi-Accounting and even more Stock-Flow Consistent models, for which our experience is much more limited, we will be even less directive.

THE EXAMPLE: A VERY BASIC MODEL

To present the elements and the framework of a structural econometric model, we shall use a specific example, which we shall address permanently during our presentation. In spite of its limited size, we think it remains quite representative of the class of models we are considering in this manual.

At the start of any model building process, one has to specify in a broad manner the logic of his model, and the behaviors he wants his model to describe. No equation needs to be established at this time. We shall place ourselves in this situation.

In our example, an economist has decided to build a very simple model of the French economy. As our tests will be based on actual data, a country had to be chosen, but the principles apply to any medium sized industrialized country.

Our model includes the following elements.

- Based on their production expectations and the productivity of factors, firms invest and hire workers to adapt their productive capacity. However, they exert some caution in this process, as they do not want to be stuck with unused elements.
- The levels reached in practice define potential production.
- Firms also build up inventories.
- Households obtain wages, based on total employment (including civil servants) but also a share of Gross Domestic Product. They consume part of this revenue.
- Final demand is defined as the sum of its components: consumption, productive investment, housing investment, the change in inventories, and government demand.
- Imports are a share of local demand («domestic demand»). But the less capacities remain available, the more an increase in demand will call for imports.
- Exports follow world demand, but producers are limited by available capacities, and their priority is satisfying local demand.
- Supply is equal to demand.
- Productive capital grows with investment, but is subject to depreciation.

The above framework looks rather straightforward, and certainly simplistic. Obviously, it lacks many elements, such as prices, financial concepts, and taxes. This will be addressed as later developments.

Let us no go further for the time being. One can observe that if we have not built a single equation yet, a few are already implicit from the above text.

CHAPTER 1: NOTATIONS AND DEFINITIONS

Before we start presenting the process of model building, we must define the concepts we shall use. They will be based on individual examples taken from our (future) model.

1.1 THE MODEL AS A SET OF EQUATIONS

In a general way, a model will be defined as a set of fully defined formulas describing the links between a set of concepts.

Formally, a model can be written as the vector function of variables.

$$f(\ldots.) = 0$$

We shall address in turn:

- The nature of elements appearing in the function.
- The nature of the functions themselves.

1.2 THE ELEMENTS IN A MODEL

1.2.1 VARIABLES: ENDOGENOUS AND EXOGENOUS

Obviously, a model will be used to measure economic concepts, depending on other elements.

Two variable types will appear in a model:

- Endogenous variables, or results, whose value will be obtained by solving the system of equations,
- Exogenous variables, or assumptions, whose value is known from outside considerations, and which obviously condition the solution.

If the model is solved over past periods, this value should be known. But in forecasting operations, it will have to be chosen by the model builder (or user).

For the system to be solved, the number of endogenous variables must correspond to the number of equations.

Our formulation becomes:

$$\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$$

with

- x vector of exogenous variables
- y vector of endogenous variable (with the same dimension as f).

For instance, in our model:

- Imports will be endogenous, as they depend on local demand. Exports too, depending on world demand.
- World demand will be exogenous, as we are building a model for a single country, and we are going to neglect the impact of local variables on the world economy. Of course, this impact exists, as France is (still) an important country, and its growth has some influence on the world economy. But the relatively limited improvement can only be obtained at the very high cost of building a world model. This simplification would be less acceptable for a model of the USA, or China, or the European Union as a whole (we shall address this issue in detail later).

Technically, one can dispute the fact that exports are endogenous. As we make them depend only on an exogenous world demand, they are de facto predetermined, apart from an unforecastable error. But they have to be considered endogenous. Our model describes the local economy, and one of its features is the exports of local firms, allowed by the external assumption on foreign demand, but following a local behavior.

As to Government demand, models of the present type will keep it also exogenous, but for different reasons:

- The main goal of this model is to show its user (which can be the State, or a State advising agency, an independent economist playing the role of the State, or even a student answering a test on applied economics) the consequences of its decisions. So these decisions must be left free, and not forced on him.
- The behavior of the State is almost impossible to formalize, as it has few targets (mostly growth, inflation, unemployment, budget and trade balances) and a much larger number of available instruments[1]. If their base values are more or less fixed, it can deviate from them arbitrarily, without too much delay. To achieve the same goal, past French governments have used different global approaches, calling for different panels of individual instruments.2
- The State alone has enough individual power to influence significantly the national economy.

Each of the two exogenous elements is characteristic of a broader category:

- Variables considered as external to the modeled area, on which economic agents considered by the model have no or little influence. In addition to the world environment for a national model, this can mean population3, or meteorological conditions4, or the area available for farming. The theoretical framework of the model can also suppose exogenous structural elements, such as the real interest rate, the evolution of factor productivity, or the depreciation rate of capital.

---

[1] Not here, but in the general case.

[2] For instance, to decrease unemployment, a government can increase demand or reduce firms' taxes, and the tax instrument can be social security contributions, or subsidies

[3] In long-term models growth might affect the death and birthed rates thus population.

[4] Which can depend on growth (glasshouse effect).

- Variables controlled by an agent, but whose decision process the model does not describe. Even if it was formally possible, the model builder wants to master their value, to measure their consequences on the economic balance. These will be referred to as decision variables or «instruments".

Changing the assumptions on these two types of variables, therefore, will relate to questions of very different spirit:

- What happens if perhaps...? (the price of oil increases abruptly).
- What happens if I (the State), decide...? (to decrease VAT rate[5]).

The second type of question can be inverted: what decision do I have to take to obtain this particular result? (By how much should I decrease the rate of employers' social contributions to create 1000 jobs?). This means that the status (exogenous/endogenous) of some variables is changed: the answer calls for specific techniques or solving a transformed model. We will deal with this later.

Sometimes the two approaches can also be combined; by considering first the consequences of an evolution of uncontrolled elements, and then supposing a reaction of the State, for instance a change in policy that would return the situation to normal. For instance, the State could use its own tools to compensate losses in external trade due to a drop in world demand.

From a model to another, the field described can change, but also the separation between endogenous and exogenous. The real interest rate can change its nature depending on the endogeneity of the financial sector, technical progress can be assumed as a trend or depend on growth, and the level of population can depend on revenue.

1.2.2 EQUATIONS: BEHAVIORAL AND IDENTITIES

1.2.2.1 Behaviors

The first role of the model is to describe "behaviors": the model builder, following most of the time an existing economic theory, will establish a functional form describing the behavior of a given agent, and will use econometrics to choose a precise formulation, with estimated parameters.

In describing consumption, one might suppose that its share in household income is determined by

- The level of income (a higher income will make consumption less attractive or necessary, compared to savings6).
- Recent variations of income (consumers take time in adapting their habits to their new status).
- The evolution of unemployment: if it grows, the prospect of losing a job will lead households to increase reserves.
- Inflation: it defines the contribution required to maintain the purchasing power of financial savings.

Once identified, all these elements will be united in a formula, or rather a set of possible formulas (households can consider present inflation, or the average over the last year; the increase in unemployment can use its level or percentage change). These formulas will be confronted with the available data, to find a specification statistically acceptable on the whole, each element participating significantly in the explanation, and presenting coefficient values

[5] Provided the EU commission will allow it.

[6] Let us recall that investment in housing is considered as savings.

consistent with economic theory. Once parameters are estimated, each element of the resulting formulation will contribute to the logical behavior of the associated agent.

But the process is not always so straightforward. Two other cases can be considered.

- The behavior can be formalized, but not directly as estimation-ready formulas. A framework has first to be formalized, then processed through analytical transformations possibly including derivations and maximizations, leading finally to the equation (or set of equations) to estimate. This will be the case for our Cobb-Douglas production function (page 105) for which we compute the combination of labor and capital which maximize profits for a given production level, according to a set of formulas obtained outside the model. Or for the definition of the wage rate as the result of negotiations between workers unions and firm managers, based on their respective negotiating power.
- Often the model builder will not be able to formulate precisely the equation, but will consider a set of potential explanatory elements, waiting for econometric diagnoses to make a final choice between formulations (generally linear). For instance, the exchange rate might depend on the comparison of local and foreign inflation, and on the trade balance.

In any case, even if the exact intensity of influences is unknown to the model builder[7], economic theory generally defines an interval of validity**,** and especially a sign. Whatever the significance of the statistical explanation, it will be rejected if its sign does not correspond to theory. In the example above, the increase of labor demand must generate gains in the purchasing power of the wage rate.

Anyway, the less the theoretical value of the estimated coefficient is known, the more care must be applied to the numerical properties of the model, at simulation time.

The formulation of these theoretical equations often makes use of specific operators, allowing alternative calculations: Boolean variables, operators for maximum and minimum. For instance, in disequilibrium models, the theoretical equation can include a constraint. We can consider also the case of a function of production with complementary factors, where the level of each factor determines an individual constraint:

$$CAP = min(pl.L, pk.K)$$

with CAP production capacity, L employment, K capital, and pl, pk the associated productivities

1.2.2.2 Identities

A model composed only of behavioral equations cannot generally be used as such. Additional equations will be needed, this time describing undisputable and exact relationships.

Several cases can be identified, which can apply simultaneously:

- Some concepts are linked by an accounting formula, and we need to ensure their numerical coherence. For example, once the model has defined household revenue, it cannot estimate savings and consumption

[7] Otherwise he would not have estimated it.

separately as the sum of the two is known[8]. A single element will be estimated: it can be savings, consumption, the savings ratio or the consumption ratio, and the other elements will follow, using an identity.

- Some concepts are linked by a causal sequence of elements, and some elements in the chain are not defined by behaviors. For example, if we estimate firms' employment and household consumption, we must formalize household revenue (as a sum including wages) to make job creation improve consumption. And in our example, defining final demand (as a sum of its components) ensures that imports will follow consumption.

Of course, one can consider eliminating these identities by replacing each element they compute by the corresponding formula. This is not always technically possible, but in any case it would:

- Lead to overly complex formulations, difficult to interpret and slower to compute.
- Discard potentially interesting information.

In addition, one will be led to introduce:

- Intermediate variables simplifying formulations (and speeding up computations). Even if the growth rate of the real wage rate, which uses a slightly complex expression, was not considered interesting as an economic quantity, it will be useful to define it, if it appears as an explanatory element in many equations.
- Purely descriptive elements: the ratio of Government Balance to GDP is a crucial element in evaluating the financial health of the state (and one of the « Maastricht » criteria for entering the European Monetary Union).
- Finally, economic theory is not always absent from this type of equation: the supply – demand equilibrium has to be enforced:

Q (supply from local producers) + M (foreign supply to the country) = FD (demand from local agents) + X (foreign demand to the country).

And the choice of the variable which balances it has a strong theoretical impact on model properties.

- If exports and imports come from behaviors, and demand from the sum of its components, we need to compute Q as:

Q (local output) = (FD-M) (local demand supplied by local producers) +X (foreign demand supplied by local producers)

This means that production will adapt to demand (which itself can depend on the availability of products).

- But we could also suppose that:

The producers chose to limit their output at a level actually lower than demand, because additional production would bring negative marginal profits. In this case Q will be fixed, and we could have:

Q = fixed, X = f(WD), FD = f(economy), M = FD - Q + X

[8] This would also be absurd in terms of household behavior.

- Or the country can only import in foreign currency, which it obtains through exports.

X = f(WD), M = f(X), Q = fixed, FD = Q+(M-X)

1.2.3 PARAMETERS

Parameters can be defined as scalars with a varying value. The only formal difference with exogenous variables is that they lack a time dimension[9].

Two types of parameters can be considered, according to the way their value is established:

- Those estimated by reference to the past: starting from a theoretical but fully defined formula including unknown parameters, the model builder will seek the values which provide the formulation closest to observed reality, according to a certain distance. This means using "econometrics".
- Those decided by the model builder: economic theory or technical considerations can supply a priori assumptions concerning a particular behavior. For instance, if a Central Bank uses a standard Taylor rule to decide its interest rate, its sensitivity to the inflation level should be 0.5. A special case will be represented by a control variable, giving (without changing the formulation) a choice between several types of independent behaviors.

The distinction is not as clear as it may seem: in particular, if estimation fails to provide an economically coherent result, the model builder can be driven to accept the value of a parameter, if it is consistent with economic theory. And even if not, to choose the nearest value within the theoretical range. For instance, an indexation of wages on inflation by 1.1 can lead the modeler to apply 1, if the difference is not significant.

With **a** as a vector of parameters (â estimated) the system becomes:

$$\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y},\hat{\boldsymbol{a}},\boldsymbol{a}) = \boldsymbol{0}$$

And in our example, one could estimate the influence of world demand on exports, for example by supposing that relative variations are proportional (or equivalently that the elasticity of exports to world demand is constant).

$$\boldsymbol{\Delta X/X = a \cdot \Delta WD/WD}$$

[9] In EViews, modifying a parameter value applies to the current model, and taking it into account calls for a new compilation, making the new version operational. This is both tedious and error-prone. One might consider replacing parameters by series with a constant value, which gives access to the much more manageable “scenario” feature.

where a should be close to unity, if the share of the country on the world market is stable[10].

But if the estimated coefficient not significant, we can get back to:

$$\Delta X/X = \Delta WD/WD$$

This choice could also have been made from the start for theoretical reasons, or to ensure the long-term stability of the model.

Clearly, to estimate a parameter it is necessary to define entirely the associated formula.

1.2.4 THE RANDOM TERM

In practice, the behavior of agents does not answer exactly to formalized functions, and the formulation obtained by estimation will not reproduce the reality. It will only approximate this behavior, using elements which conform to some economic theory, each of them providing a significant contribution to the changes in the explained variable. The number of estimated parameters will then generally be much lower than the size of the sample, or the number of observed values. In practice, adding elements to the explanation can:

- In the favorable cases, improve the quality of the explanation given by the elements already present, which can now concentrate on their natural role, instead of trying to participate in the explanation of other mechanisms in which their efficiency is limited11.
- But the new element can compete with the others in explaining a mechanism in which they all have some competence, limiting the improvement and leaving the sharing of the explanation rather undetermined (and therefore limiting the significance of the coefficients).1213

In practice, these correlation problems will always appear, sometimes very early, and generally before the fifth or sixth element. Beyond that figure, the precision of individual coefficients will decrease, and global quality will improve less and less.

This means that a typical econometric equation will contain a maximum of four parameters, while variables will be known on fifty to one hundred quarters.

---

[10] In our model WD stands for world trade (including its expansion), not the aggregate demand of countries.

[11] Just like a worker which has to use his time on two tasks, and is really qualified for one. For example, if an excellent musician but average lyricist is teamed with a good lyricist, the quality of songs (both music and lyrics) will improve.

[12] This can be a problem for the model if the two competing elements have a different sensitivity to a particular variable. For instance, if one is sensitive to a tax rate, the other not: then the role of the tax rate will be undetermined.

[13] If two workers with the same profile complete a task together, it is difficult to evaluate their individual contribution. One might have rested the whole period.

It will be therefore necessary, to formulate an exact model, to accept the presence of non-zero additional terms (residuals). If one believes in the model, this residual should be interpreted as a random perturbation without economic meaning[14]. But if the equation is badly specified, it will also come from other sources: omitting a relevant variable, replacing it by another less relevant, or choosing the wrong form for the equation[15].

The fault will not always lie with the model builder, who might not have been able to apply his original ideas. The variables he needs may not be precisely measured, or only with a slightly different definition, or they may not be available at all, as in, for example, the goals or expectations of a given agent.

Practically speaking, one will often suppose that this residual follows a random distribution, with a null average, a constant standard error, and residuals independent across periods.

Our formulation becomes therefore, in the general case, noting u the vector of residuals:

$$\boldsymbol{f(x, y, \hat{a}, a, u) = 0}$$

In the example, if we want to represent changes in household consumption as a constant share of total production variations, we will write:

$$\boldsymbol{CO = a \cdot Q + b + u}$$

or rather, if we want u to have a constant relative influence:

$$\boldsymbol{CO/Q = a + u}$$

As we shall see later, the second equation avoids heteroscedasticity problems.

1.2.5 RESIDUALS VERSUS ERRORS

[14] One can also take into account that the relationship is not exact. For instance, that an value of an elasticity is only very close to constant.

[15] Of course, as we have said before, one is never able to estimate the « true » equation. This remark should apply to a large conceptual error, leading to behaviors distinctly different from an acceptable approximation of reality.

It is probably the time to bring an important issue about the nature of econometrics.

When he considers a behavioral equation, the economist can have two extreme positions.

- He believes the behavior can be exactly specified according to a formula, which is affected by an error term with a given distribution (maybe a white noise, or a normal law). With an infinite number of observations, we would get an exact measurement of the parameters, and therefore of the error (which remains of course) and its distribution.
- He thinks that the concept he wants to describe is linked to some other economic elements, but the relation is only an application, of which any formula represents only an approximation. To this application a random term can also be added, if one believes that the replication of the same explanatory elements will bring a different result. Additional observations will only get a better mapping.

The debate is made more complex by several facts:

- The data on which he wants to base his estimation is not measured correctly. One cannot expect the statisticians to produce error free information, for many reasons: measurement errors, inappropriate sample, mistaken concepts...
- Even if measured correctly, the concepts he is going to use are not necessarily the right ones. For instance, a given behavior should be applied to the only firms which do make profits, a separation which is not available at the macroeconomic level.
- The discrete lags which he will apply to these concepts are not the right ones either. For instance, it might be known that an agent considers the price index of the last month, but only quarterly data is available.
- The estimation period is not homogenous, and this cannot be explained by the data. For instance, the mood of consumers (and their consumption behavior) can evolve without any link to measurable economic elements.

From the above issues, the logical conclusion should be:

- The first position is illusory, and to a point which is impossible to measure (of course).
- But we have to take it if we want to apply econometric methods.

This means that in the following text we shall put ourselves in the first position, but we will always keep in mind the true situation and give to the difference between the concept and its estimation the less ambitious name of “residual”.

1.2.6 FORMULATIONS

We shall now consider the form of the equations. Let us first approach the time dimension.

1.2.7 THE TIME DIMENSION

Variables in economic models have generally a time dimension, which means they are known through discrete values, almost always with a constant periodicity: generally annual, quarterly or monthly series. This means we will consider models in discrete time.

There are exceptions, however. The most frequent applies to micro-economic models, describing the behavior of a panel of individual firms or households, and the dimension will correspond to items in a set. Sometimes they will be ordered, using the level of one variable, such as the income level for a set of households. Time can be introduced as an additional dimension, but possibly with a varying interval, either predetermined (phases of the moon) or unpredictable (periods of intense cold).

1.2.7.1 Consequences of the discretization

The time discretization of variables will be introduced in several ways, leading to:

- really instantaneous variables, measured at a given point in time: the capital on the 31st of December at midnight, in an annual model (defined as a stock variable).
- averages: the average level of employment observed during a period.
- flows: the goods produced during a period.

The same economic concept might appear under several forms: inflation and price level, stock of debt and balance for the period, average and end-of-period employment levels, plus net job creations. For one household, we can consider the revenue, its yearly change, and the global revenue accumulated during its existence.

1.2.7.2 The seasonality

When models have a less than yearly periodicity, some series can present a specific distortion depending on the sub-period inside the year. This can come from changes in the climate: in winter the consumption of electricity will increase due to heating and lighting, but construction will be mostly stopped. It can be due to social issues: the concentration of holidays in the summer months can reduce production, and the coming of Christmas will increase consumption (in Christian countries). We are going here to provide a basic sketch of the problem, leaving a more serious description to specialized books like Ladiray and Quenneville (2001).

Using unprocessed data can lead to problems: for instance, the level of production in the second quarter will be lower than what we could expect from labor and capital levels. This will disturb estimations and make model solutions more difficult to interpret.

Two solutions can be considered:

- Introducing in the estimated equations “dummy variables associated to each sub-period.
- Extracting from the series their seasonal component and producing a completely new set of values.

Of course, one should not mix the two types of techniques in the same equation (or model).

The second method will be favored, as it also solves the interpretation problem.

Several techniques are available, the most well-known being Census-X13 ARIMA, developed by the US Census Bureau and Statistics Canada[16]. But TRAMO-SEATS[17] is also a common choice. Both are available under EViews.

One must be aware that this process often reduces the statistical quality of estimations. For instance, if demand is particularly high in the last quarter of each year, and imports follow, seasonally adjusting both series will make the link less clear, bringing less precise results. Even more obviously, the relation between demand for heating and temperature will lose power from seasonal adjustment.

These examples show the main issue: in a one-equation model, three situations are possible:

The dependent variable contains a seasonal component, in addition to truly economic features. For instance, agricultural production will be lower in winter, even if the same level of labor, land, fertilizer, machinery is available....

---

[16] https://www.census.gov/srd/www/x13as/

[17] http://www.bde.es/webbde/es/secciones/servicio/software/econom.html

Truly, at the same time, the use of fertilizer will decrease, and probably of labor too, but in a lower way. One either adjust this variable or introduce dummy elements. The internal quality of the relationship should be the same, but the statistical one will improve in appearance, through the correlation of the unadjusted dependent variable with the explanatory dummies.

On the contrary, if all the seasonal explanation comes from the seasonality of explanatory elements, seasonally adjusting is not necessary, and even reduces the quality of estimations (with the variability of elements). One could use raw series to estimate an imports equation, using demand, rate of use of capacities and price competitiveness as explanatory elements.

But what is true for one equation does not apply to the whole model. One cannot mix the two types of series, and this means seasonally adjusting will prevail in practice.

1.2.7.3 Static and dynamic models

To determine the equilibrium for a given period, some models will use only variables from this period: we shall call them static models. They correspond to the formulation

$$f_t(x_t, y_t, a, u_t) = 0$$

The most frequent case is that of input-output models, which use a matrix of "technical coefficients" to compute the detailed production associated to a given decomposition of demand into categories of goods, which itself depends only on instantaneous elements.

$$Q = A \cdot FD$$

(A representing an n-by-n square matrix)

$$FD = f(Q)$$

On the contrary, dynamic models use variables from other periods.

- The reasons are quite numerous. They can be:
- theoretical: some agents will be supposed to base their behavior on the observation of the past. Firms will increase their prices if the profits of the previous quarter have been too low. Or they will build their expectations of demand growth on the previous evolutions of the same variable. These two examples illustrate the main issues: using the past to create an image of the future (backward looking expectations), or to measure a previous gap between actual and target values, which the agent will try to close in the present period.

- institutional: the income tax paid by households can be based on their income of the previous period (this was the case in France, until 2019).
- technical: if a model considers a variable and its growth rate, computing one from the other considers the previous level.

One observes that each of these justifications supposes that influences come only from previous periods: one will speak of (negatively) lagged influences.

The formulation becomes therefore:

$$f_t(y_t, y_{t-1}, \ldots., y_{t-k}, x_t, x_{t-1}, \ldots., x_{t-l}, a, u_t) = 0$$

Let us go back to our model. We can observe already an undisputable lagged influence: most of present capital will come from the remaining part of its previous level. Any other case is still undecided. However, without going too deep into economic theory, one can think of several lagged influences:

- For household consumption, we have already considered that adapting to a new level of revenue takes some time. This means it will depend on previous levels. If we detailed it into products, the previous level can have a positive influence (some consumptions are habit - forming) or a negative one (generally, people do not buy a new car every quarter):

$$CO_t = f(CO_{t-1}, CO_{t-2}, \ldots., HRI_t)$$

- Firms invest to adapt their productive capacities to the level of production needed in the future. We can suppose that they build their expectations on past values.

$$\boldsymbol{I_t = f(Q_t, Q_{t-l}, Q_{t-2}, \ldots.)}$$

It is interesting to note that the previous formulation could be simplified, eliminating any lag larger than one by the addition of intermediate variables:

$$\mathbf{f_t(y_{i,t}, y_{j,t-k}) = 0}$$

(where y_i and y_j represent variables, indexed by time t and t-k)

is equivalent to

$$f_t(y_{i,t}, z_{j,t}) = 0$$

$$z_{1,t} = y_{1,t-1}$$

$$z_{2,t} = z_{1,t-1} \quad (= y_{j,t-2})$$

....

$$z_{k-1,t} = z_{k-2,t-1} \quad (= y_{j,t-k+1})$$

$$z_{k,t} = z_{k-1,t-1} \quad (= y_{j,t-k})$$

in which a lag of k periods on a single variable has been replaced by k one period lags on as many variables (including new ones).

The same method clearly allows eliminating lagged exogenous variables.

On the investment equation of the example, this would give:

$$I_t = f(Q_t, Q1_t, Q2_t, Q3_t)$$

$$Q1_t = Q_{t-1}$$

$$Q2_t = Q1_{t-1}$$

$$Q3_t = Q2_{t-1}$$

But if this method simplifies the theoretical formulation, it has the obvious disadvantage of artificially increasing the size of the model and reducing its readability, without producing additional information. Its interest is reserved to specific studies. For instance, assessing model dynamics can call for the linearization of the model according to present and lagged variables. The above transformation will limit the matrix system to two elements (with lags 0 and 1), which will make further formal computations easier, and independent from the number of lags.

It also allows us to use a simplified formulation in subsequent presentations:

$$f_t(y_t, y_{t-1}, x_t, a, u_t) = 0$$

1.2.7.4 A particular case: rational expectations

It has appeared natural, in previous examples, to consider only negative lags. This will happen if we suppose that the anticipation of agents relies only on the observation of the past (and the present)[18].

To justify positive lag formulations, it is necessary to suppose:

- That agents have the possibility, by their present decisions, to determine the future values of some variables (and the associated behavior can be formalized).
- That agents perfectly anticipate the future (perfect expectations).
- That the expectation by agents of specific evolutions has for consequence the realization of these evolutions (self-fulfilling expectations).
- That agents build their expectations on the behaviors of the other agents19, for which they know the properties (rational expectations). Basically, this means that they are able to apply the model controlling the economy (but not necessarily know its formulas), and the decision process defining its assumptions. For instance, they can forecast the investment program of the Government (depending on economic conditions), they know how firms and households will react, and they know the links between these elements (they are able to consider the supply-demand equilibrium). Actually, this is rather called "model consistent expectations".
- However, they do not necessarily know the unexplained part of the behaviors (which can be associated with the random term). If know only their distribution, we shall speak of stochastic rational expectations. EViews does not provide this feature at present (only one or the other), although this should appear in a future version. They also do not have to know the actual formulas, just be able to compute them.

You do not have to believe in rational expectations to apply them. Producing alternate simulations with different assumptions on expectations will improve greatly the insight in one particular model or on economic mechanisms in general. We shall present this later on a specific case.

1.2.7.5 Other case: continuous time models

This also is a very specific area: some theoretical models will be formulated as a system of equations where variables appear as a function of continuous time, and variations (or growth rates) become exact derivatives. One ends up then with a system of differential equations, which one can be led to integrate.

These models seldom evolve beyond a theoretical stage, if only for lack of statistical information.

But some operational models, describing for instance the stock exchange, can reduce their periodicity to a daily or even shorter value.

1.2.8 LINEARITY

We will consider here the linearity relative to variables. The linearity relative to coefficients will appear in the chapter on estimation.

[18] This use of proxies is made necessary by the absence of direct measurement of anticipations. Exceptionally, they can be obtained by surveys, leading to a specific estimation.

[19] Including the State.

The potential linearity of a model represents a very important property for its analysis as well as its solution. But first we must define the notion of linearity, which can be more or less strict.

The most restrictive will be:

$$\boldsymbol{A} \cdot \boldsymbol{y_t} + \boldsymbol{B} \cdot \boldsymbol{y_{t-1}} + \boldsymbol{C} \cdot \boldsymbol{x_t} + \boldsymbol{b} + \boldsymbol{u_t} = \boldsymbol{0}$$

but one can let matrix elements change with time:

$$A_t \cdot y_t + B_t \cdot y_{t-1} + C_t \cdot x_t + b_t + u_t = 0$$

a definition again less restrictive will suppose linearity relative to the sole endogenous variables:

$$G(x_t, a)y_t + H(x_t, a)y_{t-1} + J(x_t, a) + u_t = 0$$

or even relative to the endogenous of the period:

$$G(y_{t-1}, x_t, a) \cdot y_t + H(y_{t-1}, x_t, a) + u_t = 0$$

Using the multiplier as an example, we can already show that these properties affect the computation of derivatives of model solutions. We will detail later the consequences on convergence properties.

The first property tells that it does not depend on the initial equilibrium, or the period considered. Multiplying the shock by a given factor will have a proportional effect. It is enough to compute it once to know it once and for all.

In the second case, the multiplier will depend only on the period. Starting from different base assumptions will not change the consequences of a given change.

In the third case, the multiplier will depend also on the exogenous values (and the coefficients). It has to be re-computed each time these elements change (or have changed in the past except for one period ahead solutions), but can be stored until the next time they do.

The last case is similar to the third one. But convergence will be affected (see later).

1.2.8.1 Practical cases of non-linearity

It is obvious enough that a single non-linear equation makes the model non-linear, according to one of the previous definitions. Reasons for non - linearity are multiple; one will find in particular:

- Expressions measured in growth rates (therefore possibly linear relative to the endogenous of the period). For example, the growth rate of wages can depend on inflation.
- Expressions formulated as elasticities (generally integrated into logarithms). One will suppose for example that imports and domestic demand show proportional relative variations.
- Ratios entering in behavioral equations.
- Equations using elements at current prices, computed as the product of a quantity by a deflator (which shows the evolution of the price compared to a base year). For example, the trade balance will be obtained as the difference between the products of exports and imports at constant prices by their respective deflators.

Sometimes this distinction is purely formal, and an adequate variable change will allow the return to a linear formulation. However, if we consider the whole model, replacing by its logarithm a variable computed in elasticities will only transfer the problem if the level appears also in the model.

Thus, in our general example, if one uses for the exports equation the formulation:

$$Log(X) = a \cdot Log(WD) + b$$

one can very well introduce variables $LX = Log(X)$ and$LWD = Log(WD)$, which will make the equation linear:

$$LX = a \cdot LWD + b$$

But it will be necessary, to introduce exports in the supply - demand equilibrium:

$$Q + M = FD + X$$

to add the non - linear equation

$$X = Exp(LX)$$

Therefore, most economic models presenting a minimum of realism will not be linear. But numerical computations will generally show that even for models including many formal non - linearities, the approximation by a linearized form around a model solution (denoted by an asterisk):

$$(\partial f_t/\partial y_t)(y_t - y_t^*) + (\partial f_t/\partial y_{t-1})(y_{t-1} - y_{t-1}^*) + (\partial f_t/\partial x_t)(x_t - x_t^*) = 0$$

is acceptable for general purposes.

On the other hand, the stability of the derivatives with time is much more questionable.

Let us suppose the formulation for imports is:

$$Log(M_t) = a \cdot Log(FD_t) + b$$

Linearizing it around a particular solution (noted *), we get

$$(M_t - M_t^*)/M_t^* = a \cdot (FD_t - FD_t^*)/FD_t^*$$

Or

$$(M_t - M_t^*) = a \cdot M_t^*/FD_t^* \cdot (FD_t - FD_t^*)$$

which will represent an adequate linear approximation of the connection between M and FD, provided that M and FD do not move too far away from their base value[20]. This base value might represent a reference path, from which actual values differ due to a change in assumptions.

But, if we restrict further the expression to a constant influence (linearity to constant coefficients),

$$(M_t - M_t^*) = a \cdot (FD_t - FD_t^*)$$

[20] In other words, if the terms of the derivative are negligible beyond the first order.

the approximation can be accepted only if the ratio M / FD does not change too much with time. This is not generally true: the expansion of international trade has led, and still leads, to a sustained growth of the share of imports in domestic demand, for most countries. The ratio M * / FD * will grow strongly with time, and the last formulation will be quite inadequate for forecasts.

1.2.9 OTHER PROPERTIES

1.2.9.1 Continuity

We consider here the continuity of the whole set of endogenous variables relative to assumptions (exogenous variables, parameters). It is almost never verified formally, but should only be considered within the set of acceptable solutions (and assumptions).

For instance, most models use ratios, which is acceptable if the denominator can never become null (like the productivity of labor measured as the ratio of production to employment). Or using logarithms to link imports to demand requires (logically) that those elements are strictly positive. In other words, a fully linear model can produce a negative GDP, but this does not make it less operational if this value is associated with absurd assumptions or coefficients.

So even if all models show non-continuity potential, it should never occur in practice. We can think of only three cases:

- The model framework is correct, but something is wrong with its elements: the numerical assumptions, the estimated coefficients.
- The algorithm used for solving the model leads to absurd values (more on this later).
- The behavioral equations are wrongly specified. As we also shall see later, it can be dangerous to put together elements without a previous assessment of the associated mechanisms (for instance using logarithms as a natural solution).

It is necessary, however, to distinguish these absurd cases from those where the discontinuity applies to the derivative of a variable differentiable by pieces, as we are going to see in the following paragraph.

1.2.9.2 Differentiability

It is less necessary, but its absence can lead to problems in the system solving phase, as well as in the interpretation of results.

Separating from the previous criteria is not always straightforward, as the non-derivability of one variable can correspond to the discontinuity of another: a discontinuous marginal productivity can make the associated production non-differentiable at points of discontinuity.

Returning to the example, we could formalize household consumption in the following manner:

- They receive a constant share - a - of production Q.
- Under an income threshold - R - they consume a share c0.
- On the supplement they consume a share c1.

The consumption equation will become:

$$CO_t = c0 \cdot a \cdot Q_t + (c1 - c0) \cdot max(0, (a \cdot Q_t - R_t))$$

At the point Q = R / a, CO is not differentiable (the derivative to the left is c0.a, to the right c1.a). And the sensitivity of consumption to income is not continuous.

This derivative is not purely formal: it defines the marginal propensity to consume (consumption associated to a unitary income increase), which can appear itself in the model, at least as a descriptive element.

At the household level, the evolution of income tax as a function of revenue (with rates associated to brackets) would represent another example, determining disposable household income.

1.2.9.3 Existence of a solution

It is obviously necessary for the model to have a solution, at least when it is provided with acceptable assumptions[21]. But the potential absence of a solution is present in many formal systems, including linear models. This absence of solution is generally logically equivalent to the existence of an absurd solution, as one can illustrate on the following case.

Let use consider a model with n+1 endogenous variables: X (dimension n) and x (a single variable). We shall describe it as f, a vector of formulas (dimension n+1), in which x appears as an argument of a logarithm,

$$f(x, X, log(x)) = 0$$

If none of the positive values of x ensures the solution of the complete model, it has no solution.

In other words, taking the argument of the logarithm as a parameter α

$$f(x, X, log(\alpha)) = 0$$

and making it vary in R+, solving the associated model on x and X will never provide a value of x equal to this parameter.

The model has obviously no solution.

---

[21] Refusing to provide a solution for absurd assumptions should rather be considered as a quality.

But if the model builder has used a formulation in logarithms, he has probably not considered letting the argument take negative values. By replacing the logarithm by some other expression giving similar values, we would probably have obtained a solution. But if the variable remains negative, this solution would have been unacceptable.

To illustrate this case, we are going to reduce the usual model to a three equations version.

Production adapts to demand corrected by imports and exports, the last being exogenous:

[1] $Q + M = FD + X$

as for demand, one supposes that its relative variations are proportional to those of production:

[2] $Log(FD) = a \cdot Log(Q) + b$

And imports are a share of demand

[3] $M = c \cdot FD$

Let us suppose that one has obtained by estimation in the past: a = 1.05 and b > 0, justified by a level and growth of demand generally superior to production, obviously associated to imports greater (and growing faster) than exports.

Now, let us produce a forecast.

The model can be reduced into:

[1'] $FD = Q/(1-c) - X$

(from (1) and (3))

[2'] $FD = Q^a \cdot exp(\ b)$

from (2)

and

$$[3'] \quad X/Q = (1/(1-c) - Q^{a-1} \cdot exp(b))$$

Obviously, if Q grows (as a-1 = 0.05), the negative element will become eventually higher that the positive one, which means that Q can only be negative, which is impossible as it enters in a logarithm in equation (2). The model has no solution.

Of course, these mathematical observations have an economic counterpart. In the long run, final demand cannot grow continuously faster than production, if imports are a share of demand and exports are fixed. Assumptions, therefore, are not consistent with the estimated formula.

One will notice that the absence of solution is due here to the implicit adoption of a condition verified numerically on the past, but not guaranteed in general. This will be in practice the most frequent case.

1.2.9.4 Uniqueness of the solution

The uniqueness of the solution, for given (and reasonable) values of parameters and assumptions, is also very important. Indeed, we do not see how one could use a model which leaves the choice between several solutions, except maybe if this freedom has a precise economic meaning.

In practice, most models are highly nonlinear if you look at the equations, but the linear approximation is rather accurate within the domain of economically acceptable solutions. This limits the possibility of multiple equilibria: if the system was fully linear, and the associated matrixes regular, there would be indeed a single solution. However, as we move away from this domain, the quasi linearity disappears, and we cannot eliminate the possibility of alternate solutions, probably far enough from the reasonable solution to appear quite absurd. Fortunately, if we start computations inside the domain, an efficient algorithm will converge to the acceptable equilibrium, and we will never even know about any other.

The most significant exception will be that of optimization models, which look for values of variables giving the best result for a given objective (for example the set of tax decreases which will produce the highest decrease in unemployment, given a certain cost): if several combinations of values give a result equal in quality[22], this lack of determination will not undermine the significance of the solution. The existence of several (or an infinity of) solutions will represent an economic diagnosis, which will have to be interpreted in economic terms[23].

Another case appears when the formula represents the inversion of another formula giving a constant value, at least on a c

ertain interval. For example, if over a certain threshold of income households save all of it:

[22] For instance, if the model is too simple to differentiate the role of two taxes.

[23] provided the algorithm used for solving the model is able to manage this indetermination.

$$CO = min(f(Q), CO^{*})$$

Then the income level associated with CO * will represent the total set of values higher than the threshold.

In the general case, the main danger appears in sensitivity studies: if one wants to measure and interpret the economic effects of a modification of assumptions, the existence of a unique reference simulation is an absolute necessity.

Finally, finding several solutions very close to each other might come from purely numerical problems, due to the imprecision of the algorithm: any element of the set can then be accepted, if the difference is sufficiently low.

1.2.9.5 Convexity (or concavity)

The convexity of the system, which is the convexity of the evolution of each endogenous variable with each exogenous variable and parameter taken individually (or of a linear combination of them), can be requested by some algorithms, especially in optimization. In practice it is very difficult to establish, and even rarely verified. At any rate, this characteristic is linked to the definition of variables, and a single change of variables might make it disappear.

1.2.10 CONSTRAINTS THE MODEL MUST MEET

In addition to its theoretical validity, the model will have to meet a set of more technical constraints.

1.2.10.1 Global compatibility

Constraints of compatibility will bear in practice:

a - **on the endogenous between themselves**: one cannot let the model compute variables independently if they are linked by a logical relationship, accounting or theoretical. For example, if the consumer price enters in the determination of the wage rate, it also will have to be influenced directly by the (estimated) price of local production. Or the employment level has to affect household revenue and consumption through a sequence of links.

Accounting balances must be verified: once household revenue has been computed as a sum of elements, an increase in consumption must produce the associated decrease in savings.

Maybe the most important issue lies with the « supply = demand » identity, which will have to enforced both in at constant and current prices. This can lead either to use one of its elements to balance the equation, or to distribute the residual over the global set of elements on one side. By formulating total supply and demand as:

$$O = \sum_{i=1}^{n} O_i$$

And

$$D = \sum_{j=1}^{m} D_j$$

One will use for instance, either

$$D_m = O - \sum_{j=1}^{m-1} D_j$$

Or one will correct the set of demand variables by multiplying each of them by the uncorrected ratio O / D.

In most cases the equilibrium at constant prices will be enforced automatically. It can be written as:

Local production + Imports = Local demand + Exports

Or identifying intermediate consumption:

Local GDP + Intermediate consumption + Imports = Local final demand + Intermediate consumption + Exports

- With only one product, intermediate consumption can be discarded, and one will generally use the equation to compute GDP, checking that it does not get higher that productive capacity24 .
- With several products, we must consider as many equilibrium equations, in which the supply of intermediate consumption goods sums inputs needed for production of the good, and the demand for intermediate consumption goods sums the intermediate uses of the good itself.

$$Q_i + \sum_j IC_{j,i} + M_i = FD_i + \sum_j IC_{i,j} + X_i$$

[24] This can be obtained by a share of imports growing with constraints on local productive capacity.

If we suppose that returns to scale are constant, the vector of value added by good will come from a matrix transformation. The constraint on capacity will be achieved in the same way as above (provided a capacity equation can be obtained).

Defining $c_{i,j}$ as the quantity of good i needed to produce one unit of good j, we get:

$$Q_i + \sum_j c_{j,i}\, Q_i + M_i = FD_i + \sum_j c_{i,j}\, Q_j + X_i$$

Or in matrix terms

$$Q + C \cdot Q + M = FD + {}^{t}C \cdot Q + X$$

or

$$Q = (I - C + {}^{t}C)^{-1}(FD + X - M)$$

Using this framework will automatically enforce the supply-demand equilibrium for all goods.

In practice, most of the problem comes from the equilibrium at current prices. If demand prices are computed individually using behavioral equations, there is no chance the equilibrium will be met. The process described earlier will in practice correct the prices. With S and D as supply and demand elements at constant prices, ps and pd as the associated deflators, we can compute the global values as:

$$\mathrm{SV} = \sum_{\mathrm{i}=1}^{n} \mathrm{ps}_i S_i$$

$$\mathrm{DV} = \sum_{\mathrm{j}=1}^{m} \mathrm{pd}_j D_j$$

The first option will compute a specific price

$$\mathrm{pd}_m = (\mathrm{SV}\text{-}\sum_{\mathrm{j=1}}^{\mathrm{m\text{-}1}} \mathrm{DV}_j)/D_m$$

and the second

$$\mathrm{pd}_j' = (SV/\sum_{\mathrm{j=1}}^{m} \mathrm{pd}_j D_j)\mathrm{pd}_j$$

where the “pd” elements are the independently computed demand prices, and the “pd'” elements the corrected values.

The correcting factor:

$$\mathrm{r} = \mathrm{SV}/\textstyle\sum_{\mathrm{j=1}}^{m} \mathrm{pd}_j D_j$$

can also be written as

$$\mathrm{r} = \mathrm{r}\ (\mathrm{SV}/\textstyle\sum_{\mathrm{j=1}}^{m} \mathrm{pd'}_j D_j)$$

which with

$$\mathrm{pd}_{\mathrm{j}}^{'} = \mathrm{r}\,\mathrm{pd}_{\mathrm{j}}$$

gives a set of equations ensuring the equilibrium. As “r” measures the potential discrepancy between supply and demand, one must check that it is not too different from one.

The following issues appear:

- With the first method, which element should be used to balance the system? The choice is between
  - A small and unimportant variable, to reduce the consequences for model properties; perhaps even a variable which has absolutely no influence on the rest of the model.
  - A variable with large value, to reduce the correcting factor
- The second method represents an extreme application of the first one, where all variables on one side are affected in the same proportional way.

Actually, none of the solutions dominates clearly, the worst being in our sense the very first, which is the same as accepting de facto an imbalance, hidden but with potentially damaging consequences. Also, the second could be associated with a converging economic process, while the first can have no economic interpretation whatsoever.

In fact, one should concentrate on limiting the size of the correction itself. One could represent the problem as eliminating dust: instead of storing it in a specific location (under a carpet), or spreading it all over the place, the best solution is clearly to reduce its production as much as possible. This means that the initially computed prices should be designed to give naturally close values to global supply and demand.

b - **on** exogenous**-> endogenous connections**: Connections must be formulated with care. For example, if the social contributions rate is defined as an exogenous variable in the model, it has to enter in all computations of contribution levels. In particular, it cannot coexist with an exogenous representation of contributions, or one using an estimated coefficient.

To avoid this type of error, a systematic study of model properties must be undertaken before any operational application: in our example, this would mean checking that an increase of the social contribution rate has all expected effects on State revenues as well as on the accounts and behaviors of concerned agents.

Also, the true exogenous concept should be decided. Concerning contributions, the decision variable is clearly its rate, while the associated revenue is influenced by endogenous prices and employment.

c - **on the** exogenous **between themselves**: one should avoid defining two variables as exogenous if they are linked (in any direction) by a logical relationship. If possible, one should endogenize one of them by formalizing this connection.

Let us suppose for example that a model for France uses two exogenous measures of the prices established by its foreign competitors: in foreign currency and in Euros (with a fixed exchange rate). To take into account an increase of foreign inflation, these two variables will have to be modified simultaneously. This is as best more complex, and can lead to errors if one is not careful enough, while it can be avoided simply by endogenizing the price in Euros as the product of the price in foreign currency by the (exogenous) exchange rate.

However, establishing such links is not always possible. For instance, in a national model, foreign prices and foreign production are exogenous, but also clearly influenced by each other. But the nature and importance of the link are highly variable. For instance, a decrease in foreign production can produce world deflation,[25] while inflation can reduce exports and production. To describe them completely one should have to resort to a foreign or world model. An intermediary solution could be to establish a set of linear multipliers linking these elements, but generally the model builder himself will take care of the problem by producing a set of consistent assumptions (with perhaps some help from specialists of the world economy, or from a separate model).

[25] This is the case for the MacSim world model we shall present later.

d **- on endogenous->** exogenous **connections**: they are obviously proscribed, because contrary to the preceding links the model builder cannot master them. They will be found in some models, however, through the presence of the following exogenous:

- Elements measured in constant terms, while they should change with economic activity.
- Deflators, which should depend on other deflators.
- Elements measured in current terms, for both reasons.

If the associated model can possibly produce correct estimates and even forecasts, it runs the risk of showing abnormal sensitivity properties. Let us take an example:

Let us suppose household income HI is composed

- Of the wage revenue, computed as the product of employment by the wage rate: LT . W.
- of other exogenous revenues

Salaries will be indexed perfectly on prices:

$$W = WR \cdot CPI$$

One will have therefore:

$$RHI = WR \cdot CPI \cdot LT + HIQ$$

This equation might perform well in forecasts. But if a change in the assumptions makes prices increase, the purchasing power of total wages will remain unchanged, but for the complement HIQ it will be reduced in the same proportion as the price rise:

$$\Delta(HIQ/CPI) = -(\Delta CPI/CPI) \cdot (HIQ/CPI)$$

One can question this assumption. Some elements in non-wage revenue (social benefits, rents, firm owner's profits, independent workers revenue) are more or less indexed, and can even be over indexed in the case of interest payments (the interest rate should increase with inflation). Others, associated to differed payments (dividends, income tax) will not change immediately. The global sensitivity to prices is not clear, but a null value is obviously not correct.

We will face the same problem with a change in GDP:

$$\Delta(HIQ/Q) = -(\Delta Q/Q) \cdot (HIQ/Q)$$

where we cannot suppose that revenue does not change (grow) with economic activity. Some elements do not or show a limited sensitivity (pensions) but dividends and the revenue of owners of small firms certainly do.

In conclusion, even when a variable measured at current prices has no theoretical content, it should not be kept exogenous, especially if it can be supposed to grow at constant prices. It is general better to consider as exogenous its ratio to another variable, supposed to follow the same trend (in the absence of idea, one can use plain GDP). The model equation will compute the variable by applying the exogenous ratio. This is also can be valid for variables at constant prices (which generally increase with production), to the exception of decision variables identified as such.

In the case above, one could write:

$$HI = WR \cdot CPI \cdot LT + CPI \cdot Q \cdot r\_hiq$$

in which the introduction of Q links additional revenue to the global growth of the economy.

1.2.10.2 Homogeneity

If some equations in a model do not meet homogeneity constraints, this endangers its properties, particularly its sensitivity to shocks. Let us quote some cases:

- Linear relationships between values and quantities. The equation:

CO (consumption at constant prices) = a HRI (current income) + b is not only absurd from a theoretical viewpoint, but will lead in the long-term to a level of savings

$$S \;\; = HRI - CPI \cdot (a \cdot HRI + b)$$

that will become clearly negative over a certain price level.

- Mixing logarithms and levels. Similarly, the equation:

$$CO = a \cdot Log(HRI) + b$$

(this time the two elements will be measured in quantities) makes the ratio CO / HRI decrease to 0, and therefore the savings rate to 1, when HRI grows indefinitely.

This last example shows however a limit to the argument: on short periods the equation can present a satisfactory adjustment, as the consumption to income ratio (propensity to consume, complement to 1 of the savings rate) decreases effectively with income. It is the speed of the decrease, and its long-term evolution, which is questioned here.

1.2.10.3 Constants with dimension

The problem is identical to that of the exogenous with dimension. It invites a careful study of the theoretical content of the constant. Furthermore, as most variables grow with time, the influence of the constant will generally decrease or even disappear in practice. We shall address this issue later, on a practical case.

1.2.11 NORMALIZATION AND IDENTIFICATION

Once equations have been estimated, the problem of normalization remains. We have seen that very often the estimated formula will not explain a variable, but an expression (logarithm, growth rate, ratio, or a more complex expression). But some simulation algorithms will request a model a specific form, called “identified”, in which a single untransformed variable appears on the left-hand side:

$$y_t = f_t(y_t, y_{t-1}, x_t, a, u_t)$$

This means the model builder might have, after estimation, to transform the formulation: this operation is called the normalization of the model.

The advantage is double:

- The application of some solution algorithms is made easier. In some cases (Gauss-Seidel), this form is actually requested.
- This type of formulation allows a better interpretation of the process determining the equilibrium, provided each equation can be interpreted as a causal relation. If the equation describes a behavior, the economist should have placed to the left the element it is supposed to determine, conditional on the elements on the right. This is what we can (and will) do naturally in our example. For instance, the equation describing the choice by households of their consumption level will place naturally the variable "consumption" to the left.

The vast majority of equations will take naturally an identified form. Sometimes, a simple transformation will be necessary, however. Perhaps the most frequent nonlinear operator is the logarithm, associated with the integration of a formula in elasticities.

$$dx/x = f(\dots)$$

represents

$$Log(x) = \int f(\dots.) \cdot dx$$

In this case, one just needs to replace:

$$Log(x) = f(\dots.)$$

by

$$x = exp(f(\dots.))$$

If you use EViews[26], the software will do it for you. You can write the equation using the first form, and the software will normalize the equation itself, computing x. This is also true if the left-hand element contains several variables, but allows straightforward normalization. The most frequent cases are:

A change in logarithm: $Log(x_t/x_{t-1}) = f(\dots.)$

A growth rate: $(x_t - x_{t-1})/x_{t-1} = f(\dots.)$

A ratio: $x_t/y_t = f(\dots.)$

To choose which variable to compute, EViews will take the first variable in the specification of the equation. This simple method will be applied even if the variable has been identified as computed by a previous equation. For instance, in our model, if we introduce the estimation of imports M, then state:

$$M + Q = FD + X$$

EViews will give an error message, as M appears to be computed twice.

[26] Or most packages of the same type.

Moreover, when an equation is forecasted individually, one can chose between the computation of the left-hand term and the element which determines it, for instance M or Δlog (M) for our imports equation.

However, EViews does not solve analytically any equation for the variable. For instance:

$$M/(Q+M) = f(\ldots.)$$

will be translated into:

$$M = (Q+M) \cdot f(\ldots.)$$

introducing a non-recursive process over M.

In any event, normalizing the general equation

$$f(y,\ldots.) = 0$$

is possible by adding on both sides the same variable, which gives:

$$y = y + f(y,\ldots.)$$

However the convergence of a model defined in this manner is often difficult to obtain (for instance if “f” is positively linked to y). In that case, one can use (the value for “a” can be negative):

$$y = y + a \cdot f(y,\ldots.)$$

Stronger simplifications are sometimes possible and will be approached with the numerical solution process.

Identification is not always economically straightforward: in our example, when balancing demand and supply, we can observe that three last variables (Final demand, Exports and Imports) are going to be determined by their own equation (the sum of its elements for the first, estimated equations for the others). This means that balancing must be done through GDP, and we must write the equation as:

$$Q + M = FD + X$$

or

$$Q = (FD - M) + X$$

which makes its theoretical content clearer as: production must (and can) satisfy both exports and the non-imported part of domestic demand.

1.2.12 CONCLUSION

It must be clear by now that the formal definition of the whole set of equations represents with the estimation of behavioral equations an iterative and simultaneous process:

- Behavioral equations start from an initial theoretical formulation to evolve gradually to their final form by reconciling this theory with data and estimation results.
- Accounting equations have been defined as precisely as possible in the preliminary phase, to establish a coherent framework, but they often will have to adapt to the evolution of behavioral equations. Let us suppose for example that the first estimation results suggest excluding from the econometric explanation of exports their agricultural component, setting it as exogenous: a new equation and variable will appear, and the equation for total exports will become an identity.

CHAPTER 2: MODEL APPLICATIONS

We shall now give a panorama of applications using models. Comments will be centered on the example of economic models, and more particularly on the macro-economic ones. But most of the observations can be transposed to the general case.
For each of these applications, technical details shall be left to the "implementation" part (chapter 7). To understand these practical aspects of the use of models, one must first know about the way they are built, described later in chapters 4 to 8.

2.1 OPERATIONAL DIAGNOSES

The most natural use of a model seems to be the evaluation of the economic future, whether as its most probable evolution or as the consequences of some decisions. Assumptions concerning the future will be injected into the model, and its solution will produce the requested diagnosis. Thus, one will seek to anticipate the evolution of the main aggregates of the French economy until the year 2020, taking into account assumptions on the evolution of international economy.

2.1.1 DIFFERENT TYPES OF DIAGNOSES: SCENARIOS AND SHOCKS

Two types of forecasts can be considered: **scenarios** and **shocks**.

- In a **scenario**, one is interested in absolute results, and associating to a full set of assumptions a future evolution of the economic equilibrium. One might seek to obtain

  - forecasts based on the most probable assumptions
  - forecasts associated to a given set (like a party's program)
  - an evaluation of the scope of potential evolutions
  - assumptions allowing to reach specific economic targets.

- On the contrary, with a **shock**, one starts from a base simulation (often called "reference forecast" or “baseline”), or a simulation on the historical period, and measures the sensitivity of the economic equilibrium to a change of assumptions. Two economic paths will then be compared (on the past, one of them can be the historical one).

These shocks can be more or less complex, from the modification of a single assumption to the testing of a new economic policy[27].

These two techniques, scenarios and shocks, before the production of any operational policy diagnosis, will play an important role in the model validation process.

2.1.2 ADVANTAGES OF MODELS

Now that we have described the characteristics of models and their basic use, we shall discuss the advantages they bring (and their failings too).

Relative to the diagnosis provided by a human expert, advantages common to all models will:

[27] However, this new policy should stay within the economic framework of the original model.

- Guarantee the accounting coherence of the resulting equilibrium.
- Consider a practically unlimited number of interdependent influences.
- Provide an explicit formalization of behaviors, allowing an external user to interpret them.
- Produce an exact and instantaneous computation of associated formulas.
- Adapt immediately the full system to a local change of theoretical formulation.

but also

- Allow the stability of reasoning, for human users of an unchanged model.
- Provide the possibility of formal comparisons with other models.

This forecasting ambition was already the basis for the construction of the first models. But this type of use has benefited (since the 1970s) from some evolutions:

- The progress of economic theory, allowing the formalization of more sophisticated mechanisms, better adapted to the observed reality.
- The progress of econometrics, giving access to the statistical method that will produce the most reliable formulation associated with a given problem, and to test more complex assumptions.
- The improvement of numerical algorithms, both for computation speed, and solving more complex systems.
- The simultaneous improvement of computation hardware allowing to process problems of growing size, by increasingly complex methods.
- The progress of modelling science, in producing models better adapted to the original problem, facilitating the production of assumptions, and reducing the cost of reaching acceptable solutions.
- The production of computer software specialized in model building, increasingly efficient, user-friendly, and connected with other packages.
- The improvement of the reliability of data, and the growth of the available sample, regarding both the scope of series and the number of observations (years and periodicity)28.
- The easier communication between modelers, through direct contact and forums, allowing to communicate ideas, programs and methods, and to get the solution to small and large problems already addressed by others.

2.1.3 A CERTAIN REASSESSING

However, the use of models has engendered criticism from the start, using often the term « black box », describing the difficulty in controlling and understanding a set of mechanisms often individually simple but globally very complex.

In recent decades criticism has mounted to the point of calling for a global rejection of traditional (“structural”) models. Surprisingly, critics often find their arguments in the above improvements. One can find:

A utilitarian critique: models have proven unable to correctly anticipate the future. If this observation has appeared (in the beginning of the eighties), it is obviously not because the quality of models has declined. But information on model performance is more accessible (some systematic studies have been produced), and the fluctuations following

[28] However, the size of samples does not necessarily grow with time. In a system of national accounts, the base year has to be changed from time to time, and the old data is not necessarily converted.

the first oil shock have made forecasting more difficult. In periods of sustained and regular growth, extrapolating a tendency is very easy for experts as well as for models. Paradoxically, the emergence of this criticism has followed, rather than preceded, the increasingly direct intervention of model builders and their partners in forecasting results.

An econometric critique: modern techniques require a quantity and a quality of observations that available samples have not followed. A gap has opened between estimation methods judged by econometrics theoreticians as the only ones acceptable, and methods really applicable to a model[29].

A theoretical critique: the development of economic theory often leads to sophisticated formulations that available information has difficulty to validate. And in any event many areas present several alternate theories, between which any choice runs the risk of being criticized by a majority of economists. Thus, in the monetary area, going beyond a basic framework leads to rely on information unavailable in practice, or on formulations too complex to be estimated.

A mixed critique: users of models are no longer passive clients. They criticize formulations, as to their estimated specification, or their numerical properties. This evolution is paradoxically favored by the improvement of the logical interpretation of economic mechanisms, itself fathered essentially by economic knowledge (even the economic magazine articles use implicit macroeconomic relations) and modelling practice (the population of clients includes more and more previous model builders or at least followers of courses on modelling). One could say that model users ask the tool to go beyond their own spontaneous diagnosis, and they want this additional information to be justified.

It is clear that these criticisms grow in relevance as the goal grows in ambition. Forecasts are more vulnerable than simple indicative projections, which seek to cover the field of the possible evolutions. As for policy shock studies, they are not prone to errors on the baseline assumptions, if we discount non-linearities[30].

This relevance will also depend on credit granted to results. One can use figures as such, or be content with orders of magnitude, or even simply seek to better understand global economic mechanisms by locating the most influential interactions (possibly involving complex causal chains). In our sense, it is in this last aspect that the use of models is the most fruitful and the least disputable[31].

2.2 THEORETICAL MODELS

Contrary to previous models, theoretical models may be built for the single purpose of formalizing an economic theory. It may be sufficient to write their equations, associating to a theoretical behavior a coherent and complete system. Reproducing the observed reality is not the main goal of these models, and it is not mandatory to estimate parameters: one can choose an arbitrary value, often dictated by the theory itself. In fact, this estimation will often be technically impossible, when some of the variables used are not observed statistically (the goals or expectations of agents for example).

---

[29] Actually, the sample size required by present techniques (50 or better 100 observations) limits the possibility of estimating equations using deflators or variables at constant prices. Even using quarterly data, separating values into prices and volumes is quite questionable 15 years from the base period.

[30] With a linear model, the consequence of a shock depends only on its size, not on the simulation it starts from.

[31] One example is the impact of a decrease in local tariffs. Ex-ante it increases imports (a negative demand shock). Ex-post it decreases local factor costs (with cheaper investment and cheaper labor, indexed on a lower consumption price). This leads to more local capacity and competitiveness, both on the local scene (limiting the imports increase), and the foreign one. In most models, GDP decreases then grows.

The full interpretation of such a shock provides a lot of information, even if one remains at the non-quantitative level.

However, even based on an artificial series and arbitrary parameters, the numerical simulation of these models can be interesting. Actually, the formulas are often so complex that solving the model numerically will be necessary to observe its solutions as well as properties (such as the sensitivity of solutions to assumptions and to coefficients).

2.3 QUANTIFIED SMALL MODELS

2.3.1 WITH SCIENTIFIC PURPOSE

These models represent an intermediate case. One seeks a realistic representation of the economy, adapted to observed reality, but sufficiently simple to accept the application of complex analysis methods (and the interpretation of their results). In addition to scientific research, this study can be done to measure and to analyze properties of an operational model on a simplified representation (in the eighties Minims, then MicroDMS have been used to characterize the Dynamic Multi Sectorial model of INSEE).

There are two categories of methods:

- "External" methods will use model simulations to observe its quantitative properties, and infer a descriptive comment, both statistical and economic.
- "Internal" methods seek to explain properties of the model by its structural characteristics, using mathematical tools. This does not necessarily call for actual simulations.

2.3.2 WITH AN EDUCATIONAL PURPOSE

Although often of the same type as the ones above, these models try to present economic mechanisms as complete as possible, based on real data, under an interpretable and concise form. If necessary, one will favor the message contained in the presentation over the respect of statistical criteria.

This is the case of the MacSim package, allowing students to interpret international mechanisms and interactions.

CHAPTER 3: MODEL TYPES

We shall now try to establish a classification of models, focusing on the link between the model's characteristics and the goal for which it has been built.

3.1 THE FIELD

The field described by a model is characterized by the variables it computes, but also by assumptions it takes into account.

In the economic model subset, we can consider:

- A geographical field: national models, multinational models, world models. These last can be built in two ways: by putting together preexisting national models, with potentially quite different structures, or by building simultaneously country models of identical structure, possibly with a single team of modelers. We shall deal with this later.
- A theoretical field: the theory used for the formalization of the model may or may not approach specific economic aspects. A Keynesian model might limit the treatment of monetary aspects. A short-term model will not formalize demographic evolutions.
- A field of units: a model might present only variables at constant prices, or physical quantities like barrels of oil or number of pigs.
- A field of agents: a model will describe the behavior of a single agent: households, the State, firms.
- A field of goods: a model might consider only the production and the consumption of one good, for example energy. An energy model can use physical units.

There are other types of fields. However, the distinction is not always easy: some models will describe summarily a global field, except for a certain aspect on which it will concentrate. An energy model, to consider interactions with the rest of the economy, will have to model it also, but not in the same detail. And it can mix physical units (barrels of oil or gigawatts) with national accounts elements, with obvious conversion problems.

On the other hand, it will always be possible, and made easier by some modelling packages, to change (actually to restrict) at the time of simulation the scope of the model. The distinction is then no longer permanent: a multi-national model can be used to simulate a complete evolution of the world economy, but its user can also restrict calculations to the evolution of a group of countries or even a single one, the other elements being fixed. One can simulate a model of the real economy with or without additional monetary features. Or a model using normally rational expectation elements can drop them to become purely backward looking.

3.2 THE SIZE

The history of modelling shows that for a long period new models generally have seen their size grow, for the reasons cited earlier: the progress of model-building techniques, the increased availability of data, the faster computer computations. Additionally, for any given model, size increases regularly in the course of its existence, as new team members want to add their contribution.

However, the last decades have seen a trend favoring a return to models of limited size. Productivity improvements, requested from teams of model builders, are less and less compatible with the management of a large model. Despite the progress of model-building techniques, the desire to reduce costs and delays conflicts with the size, especially (but not only) regarding human operations: elaboration of assumptions and interpretation of results.

Also, the use of a very detailed model can make individual estimations and specifications look too expensive. The attractiveness of a calibrated CGE model will increase.

Finally, the desire to reply to critics comparing models to "black boxes" leads model builders to look for more explicit and manageable instruments.

3.2.1 QUASI-ACCOUNTING MODELS

However, paradoxically, the need for detailed explanations, the availability of more detailed data, and the increased power of computers (both in speed and size of manageable problems) has led to the development of more detailed (often extremely detailed) tools: the Quasi-Accounting versions, considering generally a large number of products (possibly hundreds), with a limit depending only on the availability of data. The framework is generally a full input-output table.

Of course, econometrics are no longer applicable, and formulations are most often rather crude, with behaviors established as exogenous ratios. But this also makes specifications clearer and more manageable, and the properties easier to control. Also, the request for large samples is less strong.

This issue will be treated in a specific part. Among the cases we will present, we will consider a model of more than 15 000 equations, which can be summarized by collapsing the dimensions into a 50 equations presentation.

3.2.2 DETERMINANTS OF THE SIZE

Determinants of the size of the model will be:

- The size of the field covered (see above).
- The degree of aggregation, which can be
  - vertical: number of operations considered (for example one can distinguish several types of subsidies, or social benefits),
  - or horizontal: number of agents listed; one can distinguish more or less sectors of firms, or categories of households.

The degree of aggregation will not be inevitably uniform: an energy model will use a particularly fine detail for energy products.

In fact the same model can appear under several versions of different size, depending especially on the degree of aggregation. Each version has then its proper area of utilization: detailed forecasts, quick simulations, mathematical analysis, and educational uses.

Thus at the end of the 1980s, the 3000 equation D.M.S model (Dynamics Multi Sectorial) used by INSEE for its medium-term forecasts had two companion versions of reduced size: Mini - DMS (200 equations), used for some operational projections and analysis which did not require detailed products, and Micro - DMS (45 equations), with an essentially educational purpose.

This distinction has lost most of its validity, however, following the reduction of the size of operational models.

3.3 THE HORIZON

3.3.1 FOR FORECASTING[32]

If a model is designed for forecasting, its horizon will be defined at the construction of the model. It will be strongly linked to its general philosophy and to the set of mechanisms it implements. A long-term model will be little interested in circumstantial phenomena (such as the lags in the adjustment of wages to prices), while a short-term one will not consider the longest trends (such as the influence of the economic situation on demography).

These differences seem to discard elaborating a model that can be used for both short - and long-term projections. But we shall see that strong reasons, in particular econometric, have made this option appear as the most natural in the present context. We will develop them when we address periodicity, in paragraph 3.4.

In any case, one can find a certain asymmetry in the relevance of this observation. If long-term models can neglect intermediate periods if they do not show significant fluctuations, simulation of the periods beyond the operational horizon can evidence future problems, already present but not visible in the short-term.

It is clear:

- That treating medium or even short-term problems calls for a model with stabilizing properties, which can only be controlled through long-term simulations. This includes in particular controlling the existence and speed of numerical convergence and evidencing cyclical properties.
- That observing long-term properties has to be complemented by the intermediary profile. Again, a long-term stabilization can be obtained through a monotonous or cyclical process.

3.3.2 FOR MODEL ANALYSIS

Here, the horizon depends on the type of analysis one wants to produce. Often, to analyze a model built with a given forecasting horizon, simulation over a longer period must be obtained. Even more than for forecasts, analytic shocks will show and explain anomalies that were not apparent in the normal projection period, but had already a significantly harmful influence. We shall stress these issues later.

3.3.3 A TENTATIVE CLASSIFICATION

One could use the following classification:

- Short-term models: 1 quarter to 2 years.
- Medium-term models: 4 to 7 years.
- Long-term models: 10 years and more.

Obviously, for a dynamic simulation, the full path, including intermediate values, is of interest.

3.4 THE PERIODICITY

[32] One shall notice that we can use several words to characterize these exercises: forecasts, projections, scenarios, simulations. It all depends on the purpose for which the test was made, and perhaps the trust allowed to the results. We favor the last term, which unfortunately has to be completed into: « simulation over future periods ».

The periodicity of a model is linked to the mechanisms it seeks to study and therefore to its horizon.

Short-term models require a short periodicity to consider circumstantial phenomena: delays in the wage indexation on prices, progressive adjustment of the consumption level to an increase of income.

Long-term models can use a sparser periodicity, less for theoretical reasons (long-term behavior can be described by a short-periodicity model), than for technical ones: this choice will reduce constraints on the availability of series, facilitate the production of assumptions, and limit simulation costs.

However, we shall see that the use of "modern" econometrics methods calls for a short periodicity, for all kinds of models, as soon as estimations are considered.

This means that the main determinant of model periodicity comes from the data. Countries which produce quarterly national accounts use quarterly models, which allow them to apply modern techniques with some comfort, and produce both short and long-term studies. Of course, results can be summarized later in yearly tables.

When only yearly accounts are available, the techniques become more simplistic, and true short-term applications are not possible. Unfortunately, this applies most often to countries with a short history of statistics, making the problem the harder.

3.5 OTHER MODELS

We have essentially concentrated on the macro-economic model case. One can also find:

- Micro-economic models: describing the behavior of firms, of households.

These models will sometimes be more theoretical, calling for optimization computations (such as cost minimization) or to elements of strategy (game theory). They will often be estimated on survey data, with very large samples

- Non-economic models: they can apply to biology, physics, chemistry, astronomy, meteorology, ecology, process control, and so on.... and can be used to evaluate the consequences of building a dam, controlling a manufacturing process, looking for the best organization of a project, describing a biological process. These models will often be conceived not as a formalized equation system, but as the maximization of a criterion under some constraints, or as a system of propositions connected by logical operators.

CHAPTER 4: GENERAL ELEMENTS

This part of the book describes the process of development, use and management of a model, taking special interest in technical aspects and particularly computer-oriented features. Application to EViews will be presented in detail, but most of the teachings can be applied to other packages, including those which are not dedicated to econometric structural modelling.

But let us give first a quick description of the organization of the model building process.

4.1 THE STAGES IN THE PROCESS

4.1.1 PREPARING THE MODEL

The first step in the building of any model is producing a draft which ensures some compatibility between available data (wherever it might come from) and the type of model its builder has in mind (goal, field, nature of the variables, underlying theory).

Knowing the scope of available data, the economist will define a model framework for which values can be attributed to all variables, either by using available elements, by computation or as a last resort to expert advice (including the modeler itself). This means that a first decision has to be made as to the field described by the model, the variables used as assumptions, and the variables it shall compute. Moreover, he must divide the equations into identities, which set indisputable links between variables, and equations describing the behavior of agents, for which the final formulation will be based on past evolutions of the associated elements. In the course of model building, this status can change.

The first task will be to gather, by reading from files and transforming the data, the full set of variables needed by the model, to define the form of the identities, and give a first assessment of the behaviors he intends to describe. He shall check for which periods the necessary data is known, and that on these periods identities hold true. If some elements are not available, he will use the best proxies he can get. And if this also fails, he will use his imagination.

He can also make a first economic analysis of the framework implied by model specifications. This is greatly helped by EViews which can give essential information on the model’s logic, even in the absence of any data.

4.1.2 ESTIMATION

The second phase will look for a satisfying description of the behavior of agents, by checking economic theory against available data. The modeler shall define a set of formulations with unknown parameters, compute for each formulation the values which give the best explanation of past evolutions, and make his selection, using as criteria both statistical tests and compliance to economic theory. This process can call for the introduction of new variables, or changes in some definitions, which will mean reformulating some identities.

Of course, both individual and global consistencies must be applied. For instance, using a Cobb-Douglas production function implies considering the global cost in the equation for the output deflator.

4.1.3 SOLVING AND TESTING OVER THE PAST.

Once the full model is defined, one can try to solve it.

- One shall first check for consistency the set of equations, data and parameters, by applying each formula separately on the sample period. If the estimation residuals have been introduced as additional elements, the process should give the historical values in all cases.
- One shall then simulate dynamically the full model on the same period, setting (temporarily) the residuals to zero. This will show if considering current and lagged interactions does not amplify too much the estimation errors, both on the current period and with time. Using an error correction framework should limit the risk of divergence.
- Finally, the reactions of the equilibrium to a change in assumptions, for instance the exogenous component of demand, will be measured. The results will be compared with the teachings of economic theory, and what is known of values given by other models, moderated by the characteristics of the country. However, one should not spend too time here, as simulations over the future will provide a much better context.

Discovering discrepancies can lead to changes in some elements of the model, including the set of its variables. This means going back to step 1 or 2.

4.1.4 SOLVING AND TESTING OVER THE FUTURE

Once the model has passed all tests on the past, further tests will be conducted, under conditions more representative of its actual use: on the future. For this, values will have to be established for future assumptions. Again, the sensitivity of the model to shocks will be studied, this time with a longer and smoother base, better associated with future use. As to the reliability of baseline results, one can rely this time on stochastic simulations.

The results of this step can of course show the necessity to revert to a previous stage, including the introduction of new data, changing causalities, or re-estimation. To limit the number of backward steps, one should introduce in the original data set all potential variables, and decide on behavioral equations considering the global properties.

4.1.5 USING THE MODEL FOR FORECASTS AND POLICY STUDIES

Finally, the model will be considered as fit for economic studies: forecasts and economic policy analysis.

From now on, we shall suppose we are using a dedicated package like EViews. But for people who till model through a spreadsheet), most of our observations will stim apply.

4.2 HOW TO ORGANIZE THE DEVELOPMENT OF THE MODEL

Let us now consider the organization of the model production task.

To create a model, two extreme types of organization can be considered:

- Methodical option:

The model builder

- Specifies completely a coherent model (including accounting equations), precisely separating assumptions from results.
- Looks for the necessary series.
- Estimates behavioral equations.
- Uses the consequent model.

Applying such a framework is obviously illusory, as many backtrackings will be necessary in practice:

- Some series will show up as unavailable, and it will be necessary to replace them or to eliminate them from formulations. Thus, in the absence of series for interests paid by firms one will have to be content with profits before interests.
- Some estimations will give unsatisfactory results: it will be necessary to change formulations, to use additional or alternate series. Thus, a formulation in levels might have to be replaced by a formulation in logarithms (constant elasticities), or in growth rates. Or one will be led to explain the average monthly wage instead of the hourly wage, and to introduce in this last explanation the evolution of the minimal wage. For an oil producing country, it will appear necessary to identify oil (and non-oil products) in both production and exports.[33]
- New ideas will appear during estimation. For example, a recent article on the role of Foreign Direct Investment might lead to test an original formulation.
- Formal errors are going to be identified. Thus, an element (a type of pension) might have been forgotten from households' income.
- Some variables defined as assumptions are going to appear sufficiently influenced by results to see their status modified.
- Some causalities will be questioned when observing numerical properties.
- Simultaneities will have to be replaced by lagged influences.
- The size (or even the sign) of the response to changes in the assumptions will be inconsistent with theory.

- The model will not converge, and the specifications will be considered as the cause.

- Improvisation

To the contrary, a model builder can

- establish general options for the model structure and theoretical framework,
- produce some formulations independent from each other,
- estimate them by accessing to separate series,
- And gradually connect selected elements by completing the model with linking identities and the data set with the necessary exogenous variables.

This framework will be even less effective, if only because the number of single operations on equations and series will present a prohibitive cost. Furthermore, enforcing the accounting and theoretical coherence of the model could prove difficult, and the modelling process might never converge at all to a satisfying version.

- The optimal solution is of course intermediate:

- Define as precisely as possible the field and the classification of the model.
- Define its general theoretical options and its goal.
- Obtain, create and store the total set of presumably useful series, with no limitations.
- Establish domains to estimate, specify associated variables and set formal connections, especially accounting relations.
- Undertake estimations
- And go through changes (hopefully limited) until an acceptable form is obtained.

It is clear that this type of organization is all the easier to implement if:

- The size of the model is small: it is possible to memorize the total set of variable names for a thirty equations model, but for a large model a formal documentation will be necessary, produced from the start and updated

[33] Actually, this should have been evident from the start.

regularly. This framework should be discussed in detail by the modelling team and as many outsiders as possible.

- The number of concerned persons is small (the distinction comes essentially between one and several): for a team project, the role of each participant and his area of responsibility have to be clearly defined. Especially, physical changes (on both data and model specifications) should be the responsibility of one individual, who will centralize requests and apply them. And modifications must be clearly stated and documented, through a formal process.

Individual modifications of the model can be allowed, however, provided a base version is preserved. Thus several members of a team of model builders can test, one a new production function, another an extended description of the financial sector. But even in this case updates will often interfere, at the time modifications generated in separate test versions are applied to the base one. For instance, a new definition of the costs of wages and investment, which define the optimal shares of labor and capital in the production function, will influence the target in the price equation.

CHAPTER 5: PREPARING THE MODEL

We shall now start with the first (and probably most important task): preparing the production of the model.

One might be tempted to start actual model production as soon as possible. But it is extremely important to spend enough time at the start evaluating the options and choosing a strategy. Realizing much later that he has chosen the wrong options, the modeler is faced by two bad solutions: continuing a process leading to a subpar model, or backtracking to the point where the choice was made.

This can concern

- The organization of tasks, like producing at first single country models, for a world modelling project.
- Economic issues, like choosing the complexity of the production function.
- Accounting issues, like deciding the decomposition of products, or the distinction into agents.
- Technical ones, like the number of letters identifying the country in a world model series name.

5.1 PREPARING THE MODEL: THE FRAMEWORK

At the start of the model building process, the modeler (or the team) has at least:

- General ideas about the logic of the model he wants to build.
- Information about the set of available data.

Actually, things can be more advanced:

- The data can be directly available, as a computer file, but not necessarily in the format needed by the modelling package. Many databases (like the World Bank's World Development Indicators) are stored on the producer's website in Excel or CSV format. In more and more cases, access can be provided from inside EViews, but this is not necessarily the best option.
- Equations may have already been established, either as formulas or even estimated items, if the modeling is the continuation of an econometric study, produced by the modeler or another economist.

In any case, the first stage in the process should lead to:

- A fully defined set of equations, except for the actual estimated formulas.
- The corresponding set of potentially relevant data.

Obviously, these two tasks are linked, as equations are established on the basis of available data, and the data is produced to fit the model equations. This means that they are normally processed in parallel. However, it is quite possible:

- To produce most of the data before the equations are defined. Some concepts (the supply - demand equilibrium at constant and current prices, employment, the interest rates) will certainly appear in the model. But some model-specific variables might have to wait.
- To produce the model specification before any data is available. Of course, writing an identity, or stating the equation to be estimated, does not require data. It is only the application (checking the identity is consistent or estimating the equation) which does. But one must be reasonably sure that the data will be available, or that there will be a reasonable technique totargettarget estimate it.

One can even produce a first version of the program transforming into model concepts the original data, once these concepts are completely defined, but before any data is technically available (one just needs their definition).

One can compare the situation with the building of a house: one can draw the plans before the equipment is purchased, but its eventual availability (at the right time) must be certain. And the goods can be obtained before the plans are completely drawn (but the chance of having to use them must be reasonably high)[34].

One can even imagine the data using a random process, and apply an estimation program, without considering the results but checking for the presence of technical mistakes.

These options are not optimal in the general case, but they can help to gain time. Most modelling projects have a deadline, and once the work force is available, the tasks should be processed as soon as possible, if one wants to have the best chance of meeting it.

One can question the feasibility of producing a full set of equations before any estimation. What we propose is to replace the future formulations by a “declaration of intent” which states only the dependent variable on the left, and the explanatory elements on the right. For each equation, the format should be as close as possible to:

Variable=f*(sum of the explanatory variables)

For instance, for exports X depending on world demand WD, the rate of use of capacities UR and price competitiveness COMPX, one will use:

```
scalar f

X = f*(WD+UR+COMPX)
```

The first statement avoids having f considered as an exogenous element.

The advantages of defining a full model are numerous:

- The modeler will be able to check by sight the logic of his model
- The text can be provided to other economists for advice
- The full list of requested variables can be established, allowing to produce a complete transfer program
- Processing the equations through EViews will give interesting advice on several elements. Double clicking on the “model” item one will get:
  - The equations (from ‘Equations” or “Print View”).

[34] As there is a cost to the goods. For free or quasi-free data, the chance can be lowered.

Equations: 14		
"q = fd + x - m "	Eq1:	q = F(fd, m, x)
eq_i	Eq2:	i = F(ec_i, i, k, q, ur)
"log(prle_t) = c_prle "	Eq3:	prle_t = F(t)
"led = q / prle_t "	Eq4:	led = F(prle_t, q)
eq_le	Eq5:	le = F(ec_le, le, led, t)
"lt = le + lg "	Eq6:	lt = F(le, lg)
"rhi = wr * lt + r_ "	Eq7:	rhi = F(lt, q, r_rhiq, wr)
"ih = r_ih * rhi "	Eq8:	ih = F(r_ih, rhi)
"co = rhi * (1 - sr "	Eq9:	co = F(rhi, sr)
eq_ic	Eq10:	ic = F(ec_ic, ic, q, t)
"fd = co + i + gd "	Eq11:	fd = F(co, gd, i, ic, ih)
eq_m	Eq12:	m = F(compm, ct, ec_m, fd, m, q, res_m, ur)
eq_x	Eq13:	x = F(ec_x, res_x, wd, x)
"k = k(-1) * (1 - d "	Eq14:	k = F(dr, i, k)

* The grammatical acceptability of equations will be checked: for instance, if the number of left and right parenthesizes is indeed the same. Erroneous equations will appear in red in “Equations”.

Equations: 341 Errors: 1

Syntax error in "gdp_p + ICB_P + M_P = (FD_P + ICP_P + X_P)) "

"gdp_f + icb_f + m_f = " Eq1: gdp_f = F(fd_f, icb_f, icp_f, m_f, x_f)

* Also, the fact that each endogenous variable is computed only once. The second occurrence will also appear in red.

Equations: 342 Errors: 1 Baselin

"gdp_p + icb_p + m..." Eq1: gdp_p = F(fd_p, icb_p, icp_p, m_p, x_p)

GDP_P is specified as endogenous in another equation. in "gdp_p = (FD_P + ICP_P + X_P) - (+ICB_P + M_P) '

"gdp_f + icb_f + m_f = " Eq2: gdp_f = F(fd_f, icb_f, icp_f, m_f, x_f)

- The variables.

Filter/Sort / Dependencies	All Model Variables Variables: 33 (Endog = 14 , Exog = 19 , Adds = 0)	
ec_m	Exog	EC_M
ec_x	Exog	EC_X
fd	Eq11	FD
gd	Exog	GD
i	Eq2	I
ic	Eq10	IC
ih	Eq8	IH
k	Eq14	K
le	Eq5	LE
led	Eq4	LED
lg	Exog	LG
lt	Eq6	LT
m	Eq12	M
prle_t	Eq3	PRLE_T
q	Eq1	Q
r_ih	Exog	R_IH
r_rhiq	Exog	R_RHIQ
res_m	Exog	RES_M
res_x	Exog	RES_X
rhi	Eq7	RHI
sr	Exog	SR
t	Exog	T
ur	Exog	UR
wd	Exog	WD
wr	Exog	WR
x	Eq13	X

Note: if a variable is currently overridden, its name will appear in red.

* The most important information will come from the list of exogenous: one might find elements which should have been determined by the model, according to its logic. In general, this will mean one has forgotten to state the associated equation. Also, some elements might appear, which should not belong to the model. Normally this corresponds to typing errors.

- The block structure:

```
Number of equations: 14
Number of independent blocks: 3
Number of simultaneous blocks: 1
Number of recursive blocks: 2

Block 1 : 2 Recursive Equations

       prêt(3)                      x(13)

Block 2: 11 Simultaneous Equations (1 feedback var)

       ic(10)                       led(4)                       le(5)
       lt(6)                        rhi(7)                       co(9)
       ih(8)                        i(2)                         fd(11)
       m(12)                        q(1)

Block 3: 1 Recursive Equations

       k(14)
```

It decomposes the set of equations into a sequence of blocks, either recursive (each variable depends only on preceding elements) or simultaneous (variables are used before they are computed). If one is going to succeed in estimating equations which follow the same logic as intended in the preliminary version, the block structure described at this stage will be already fully representative of the future one. One can detect:

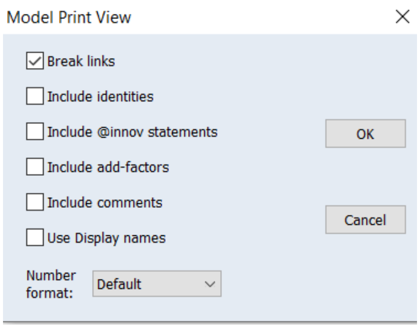
- Abnormal simultaneities: a causal loop might appear, which is not supported by economic theory behind the model.
- Abnormal recursive links: a block of equations containing a theoretical loop (the wage price loop, the Keynesian cross) can appear as recursive. This can come from a forgotten equation, a typing error…

Practical operational examples will be given later.

In any case, observing the causal structure of the model will give some preliminary information about its general logic, and its potential properties.

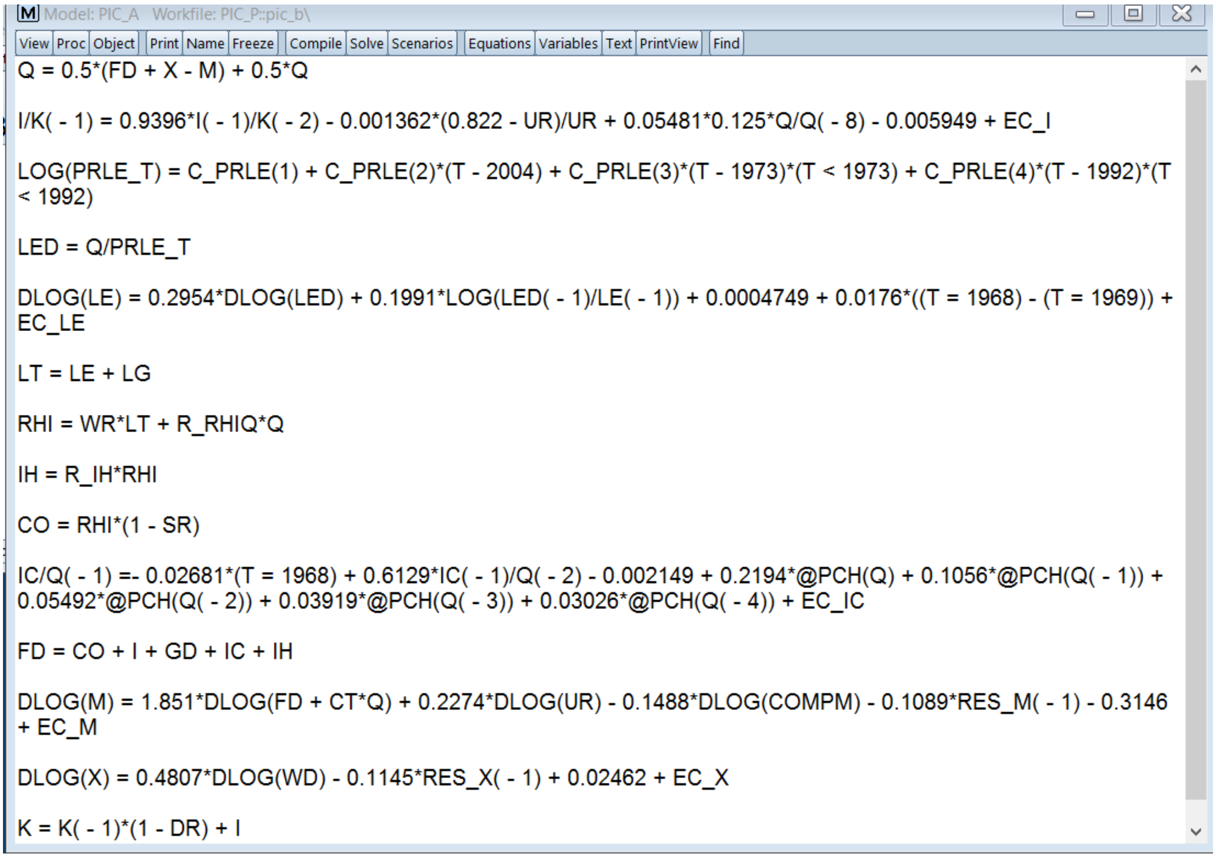
5.1.1 PRINT VIEW

To check on the model specifications, you can use the "View > Print View" will display the following window:



With "OK" we get a presentation more versatile than "Text".

One will note that it is possible to decide on the number of significant digits, which produces clearer displays in a document (the default is 8).



```
Model: PIC_A   Workfile: PIC_P::pic_b\
View Proc Object Print Name Freeze Compile Solve Scenarios Equations Variables Text PrintView Find

Q = 0.5*(FD + X - M) + 0.5*Q

I/K( - 1) = 0.9396*I( - 1)/K( - 2) - 0.001362*(0.822 - UR)/UR + 0.05481*0.125*Q/Q( - 8) - 0.005949 + EC_I

LOG(PRLE_T) = C_PRLE(1) + C_PRLE(2)*(T - 2004) + C_PRLE(3)*(T - 1973)*(T < 1973) + C_PRLE(4)*(T - 1992)*(T < 1992)

LED = Q/PRLE_T

DLOG(LE) = 0.2954*DLOG(LED) + 0.1991*LOG(LED( - 1)/LE( - 1)) + 0.0004749 + 0.0176*((T = 1968) - (T = 1969)) + EC_LE

LT = LE + LG

RHI = WR*LT + R_RHIQ*Q

IH = R_IH*RHI

CO = RHI*(1 - SR)

IC/Q( - 1) =- 0.02681*(T = 1968) + 0.6129*IC( - 1)/Q( - 2) - 0.002149 + 0.2194*@PCH(Q) + 0.1056*@PCH(Q( - 1)) + 0.05492*@PCH(Q( - 2)) + 0.03919*@PCH(Q( - 3)) + 0.03026*@PCH(Q( - 4)) + EC_IC

FD = CO + I + GD + IC + IH

DLOG(M) = 1.851*DLOG(FD + CT*Q) + 0.2274*DLOG(UR) - 0.1488*DLOG(COMPM) - 0.1089*RES_M( - 1) - 0.3146 + EC_M

DLOG(X) = 0.4807*DLOG(WD) - 0.1145*RES_X( - 1) + 0.02462 + EC_X

K = K( - 1)*(1 - DR) + I
```

5.2 PREPARING THE MODEL: SPECIFIC DATA ISSUES

Let us detail the process.

5.2.1 TYPES OF DATA

In the case of a national macroeconomic model, the needed data can be:

- National Accounts elements: operations on goods and services, transfers between agents, measured in value, at constant prices, or at the prices of the previous year. The producer will generally be the national statistical office. For France it would be INSEE (the National Institute for Statistics and Economic Studies).
- The corresponding deflators.
- Their foreign equivalents, using the accounting system and the corresponding base year of the particular country, or rather a synthesis produced by an international organism (OECD, International Monetary Fund, EuroStat....).
- Variables in a greater detail, possibly measured in physical quantities (oil barrels, tons of rice). They can come from a public or private agency, or from the producers themselves. In France energy elements would come from the Observatory of Energy.
- Monetary and financial data, coming mostly from the local National Bank (in France the Bank of France or the European Central Bank. ...), or from an international bank (OECD, World Bank, International Monetary Fund, EBRD, ADB…).

- Data on employment or unemployment. One can get detailed labor statistics (by age, qualification, sex...) from the US Bureau of Labor or the French "Ministère du Travail".
- Demographic data: population, population in age of working, age classes (INSEE in France).
- Survey data: growth and investment prospects according to firm managers, productive capacity, living conditions of households (coming from public or private institutes).
- Qualitative elements: the fact of belonging to a specific set, meeting a specific constraint.
- Micro economic models will generally use survey data (households, firms) with sometimes a time dimension (panels, cohorts) and possibly include some of the above elements as global indicators.

As the area of application of models is unlimited, the field of potentially relevant data is also. A model on the economy of transportation would include technical data on the railway system and on distances between cities, an agricultural model meteorological data and information on varieties of species.

5.2.2 THE ACCESS TO DATA

The medium through which data can be obtained will play an important role. Accessing the necessary data takes into account several features:

- the mode of transmission
- the format used
- the institutional aspects

We shall treat them in turn, then present the most usual cases.

5.2.3 THE MODE OF TRANSMISSION

Several options are available for transferring the data to the model.

5.2.3.1 Physical transmission

Data can be obtained from a physical support, either commercially produced or created for the purpose. This can be either a CD or DVD-ROM, or another rewritable media such as an USB key, or a memory card. For instance, INSEE provides CD-ROMs containing the full National Accounts.

5.2.3.2 E-mail transmission

Files can be transferred from a user to another by e-mail, as an attachment to a message.

5.2.3.3 Sharing files

One can share files using Google Drive, Dropbox, or other means.

The advantage is for participants to a project to share elements in real time. Of course, one must be careful with their status, between read only and read+write. This requires some organization between team members.

In any case, one can transfer the shared file to his computer, allowing any changes.

5.2.3.4 Downloads from a site

Files can be downloaded from a website, commercial or not

An extensive survey of the data available online for free, compiled by John Sloman at the Economics, will be found at the address:

https://www.economicsnetwork.ac.uk/data_sets

with direct access to the relevant sites.

In our experience of building single or multi-country macro econometric models, we are mainly using:

- The OECD Main Economic Indicators.
- The World Bank Development Indicators.

Both institutions provide a large set of series, covering most of the fields associated with our models. There are differences, however:

- The OECD set covers a limited number of countries:

OECD Economies

Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New, Zealand, Norway, Poland, Portugal, Slovak, Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United, Kingdom, United, States, Euro area (17 countries), OECD Total.

World

Non-OECD Economies

Argentina, Brazil, Bulgaria, China (People's Republic of), Colombia, Costa Rica, India, Indonesia, Romania, Russia, South Africa

Dynamic Asian Economies

Rest of the World

Other oil producers

Each set contains 122 quarterly series, and 221 yearly. They begin in 1962 and end at present (April 2020) in 2017.

The data is available in CVS format, easily converted to Excel and then transferred to EViews.

- The World Bank covers a much larger set of countries (267 including groups), and series (1429), but:
  - The periodicity is annual, starting in 1960 at the earliest.
  - Some important elements are lacking, in particular employment and capital (available in principle in the OECD set).

The list of countries and groups can be found at:

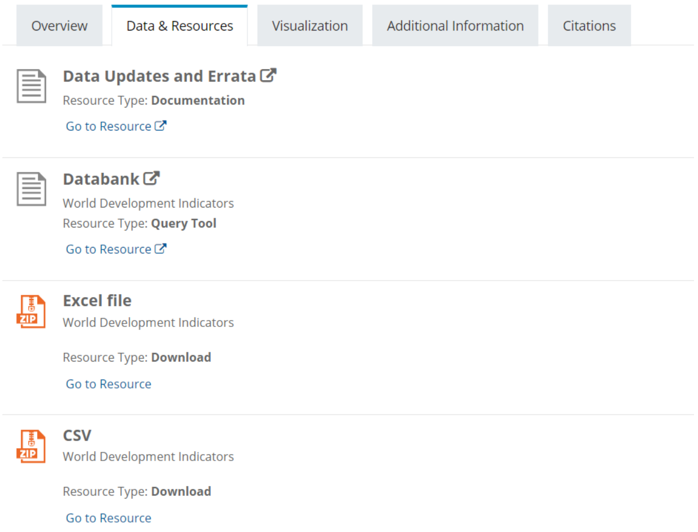
The reason for the large number is that the fields covered are much wider, with a focus on sociological issues, including for instance (potentially):

- People using safely managed drinking water services, rural (% of rural population)
- Women who believe a husband is justified in beating his wife when she burns the food (%)

The set is accessed at:

https://datacatalog.worldbank.org/dataset/world-development-indicators

which displays a page containing:



Clicking on “Download” creates a Zip file for a complete Excel sheet (110 Mb).

Creating an EVIews file containing pages associated with countries is rather straightforward, but involves a little programming (50 lines). The text is available from the author.

Of course, the number of actual values can vary and can be zero.

Other interesting sites are:

- **The IMF site**

http://data.imf.org/

Very interesting for financial data, of course, such as the countries’ Balance of Payments. But most of this information is duplicated on the World Bank site (in annual terms, however).

- **The International Labor Organization**

https://ilostat.ilo.org/data/

for detailed labor series, and employment series missing for the World Bank.

- and for French data, one can access the INSEE site www.insee.fr.

5.2.3.5 Attribute import and export

If documentation is attached to series, it can be imported along with the values and the series names.

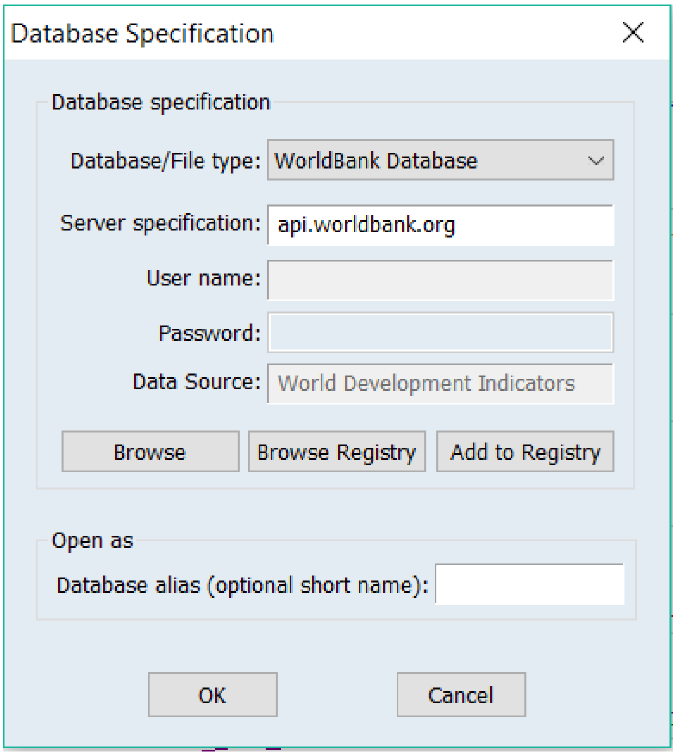
5.2.3.6 Using World Bank data

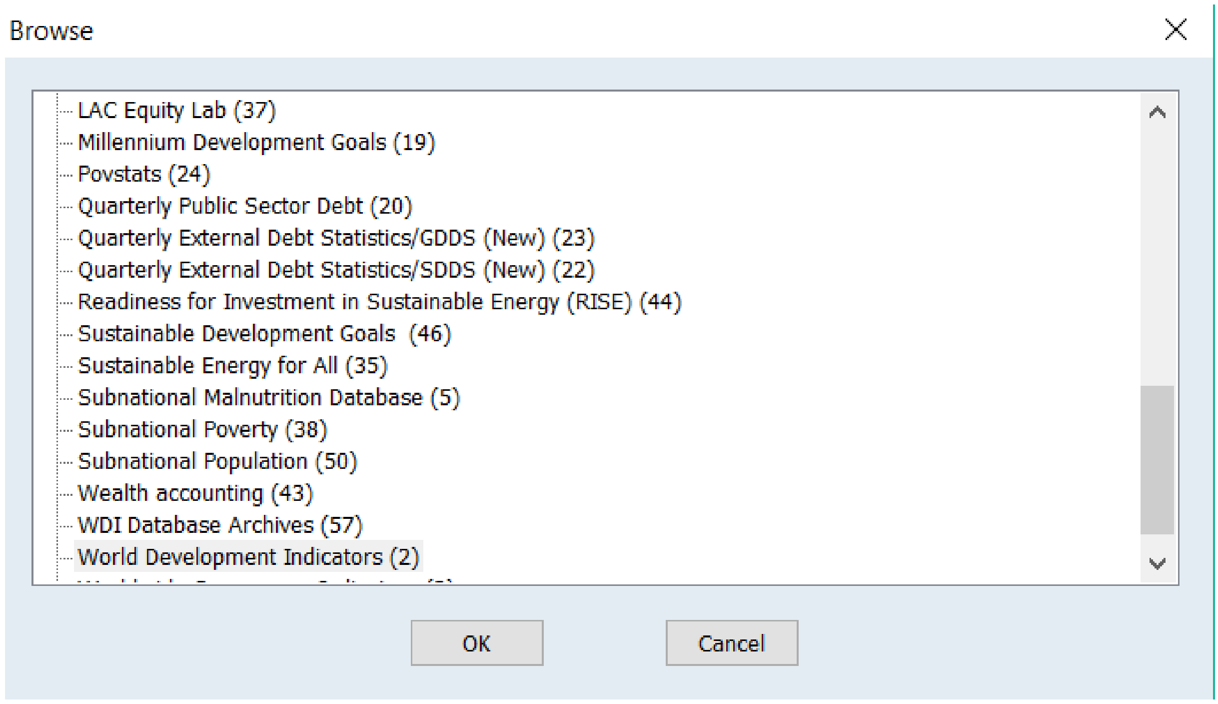
Now you can access the World Bank data from inside a session, through the menu options.

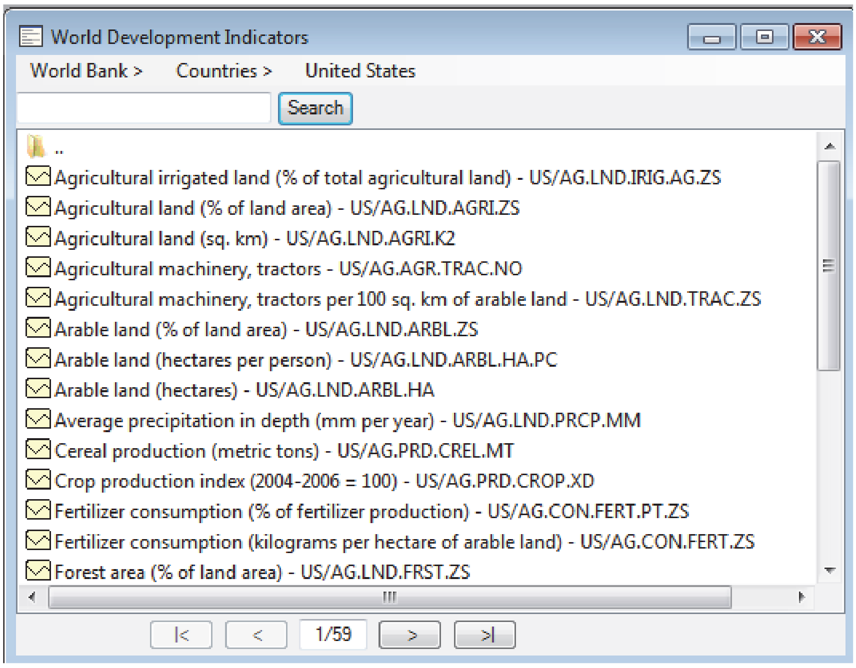
First use:

File > Open> Database

EViews displays this menu in which you select: World Bank Database.





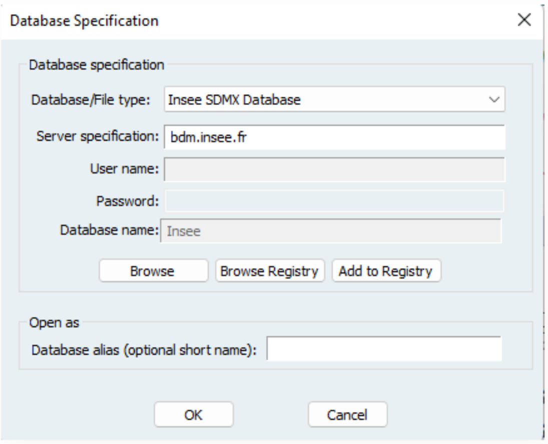


5.2.3.7 Accessing INSEE series

For the French users (or those interested in the French economy), the official INSEE series can be downloaded.

From EViews, one should access:

File>Open database>INSEE SDMX database, which opens:



The information can be obtained at:

https://www.insee.fr/en/information/2868055

5.2.3.8 Other media

In less and less frequent cases, some data will not be available in magnetic form: series will be found in printed or faxed documents, or obtained directly from other experts, or fixed by the user (who then plays the role of expert). This data will have generally to be entered by hand, although a direct interpretation by the computer through optical character recognition (OCR) is quite operational (but this technique calls for documents of good quality).

In this case it is essential not to enter figures directly into the model file, but to create a preliminary file (such as an Excel sheet or even and ASCII file) from which the information will be read. This separates the modelling process from the production of “official” information.

5.2.3.9 Sharing files

It is now quite easy to share files through Internet. This is useful in two cases:

You participate in a project with other modelers.

You follow a project and receive information from the manager.

5.2.3.10 Using the cloud

5.2.3.10.1 Cloud drive support

Although this feature can look unrelated to modelling, it can be quite useful for researchers communicating together on some project. In particular the files can be shared on a Google Drive.

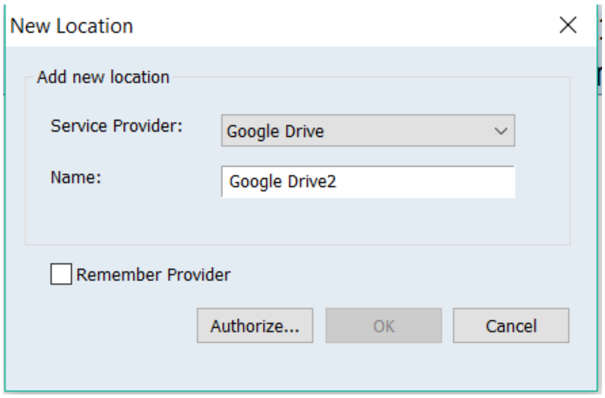
For instance, for a Google Drive, you can work through a Cloud directory, using:

Open>Foreign Data as Workfile>New Location>

The location can be:

- Dropbox.
- Google Drive.
- One Box
- Box.

Authorization can be (has to be) managed.



Then authorize and get access to the whole directory.

5.2.3.11 Change of format

As indicated above, the original data format is generally different from the one used by the model-building software.

In the worst cases, transfer from one software program to another will call for the creation of an intermediate file in a given format, which the model-building software can interpret. The Excel format is the most natural intermediary, as it is read and produced by all packages. In that case, it is not necessary to own a copy of the package to use its format.

In the very worst cases, it is always possible to ask the first program to produce a character file (in the ASCII standard) which can, with minimal editing, be interpreted by the second program as the sequence of statements allowing the creation of the transferred series, including data and definitions[35].

However, the situation has improved in the last years, as more and more packages provide a direct access to the formats

Access
Armes-TSD
Binard
dBase
Excel (through 2003)
Excel 2007 (xml)
EViews Workfile
Gauss Dataset
GiveWin/PcGive
HTML
Lotus 1-2-3
ODBC Dsn File
ODBC Query File

[35] For instance, the sequence:

```
use 1970 to 2007
read x

-----  values  ----
end
can be translated easily by a word processor  into
smpl 1970 2007
series x
----- values ----- ;
```

ODBC Data Source
MicroTSP Workfile
MicroTSP Mac Workfile
RATS 4.x
RATS Portable / TROLL
SAS Program
SAS Transport
SPSS
SPSS Portable
Stata
Text / ASCII
TSP Portable

5.2.3.12 Institutional issues

Of course, one must also consider the relationship between the data producing and modelling institutions. The most technically complex transfers do not necessarily occur between separate institutions. A commercial contract might give the modelling institution direct access (through one of the above means) to information managed by a data producing firm, under the same software format, while a large institution might still use CD-ROMs as a medium between separate units.

However, one must also consider the cost of establishing contracts, including perhaps some bartering between data producing and study producing institutions.

5.2.3.13 How to manage the access to several sources

As a general principle, one should favor using a single source. But this is not always possible. In that case, one should define a primary source, and take from the alternate ones the only additional series. The main problems might come from:

- Deflators and values at constant prices using a different base year.
- Financial and labor data bases sharing elements with national accounts.
- Variables measured in physical units (tons, square meters) having their counterparts in values.

In all these cases, the priority is the consistency of model equations, based on the data from the primary source. Additional elements must be adjusted to provide this consistency. This applies in particular to the balance of equilibrium equations (supply = demand), or sums (total demand= sum of its components).

In addition, for operational models designed to produce official studies and in particular forecasts, it is essential that the results concur with the official local statistics. As forecasts are presented mostly as growth rates (GDP and inflation for example), but provide also the last statistical (official) level, the first value in the forecast must be consistent with both. If the model is built on an outside source, the forecast must be corrected accordingly. This issue will be developed when we present the forecasting task.

For instance, let us suppose that the limits on the availability of local statistics availability forces the modeler to use on external source (like the WDI from the World Bank) to produce a full model, but the local statistical office provides basic elements like GDP. If the model forecast starts in 2020

5.2.4 PREPARING THE DATA FOR THE TRANSFER

Let us now define the best organization for transferring data from the original source to the software (we shall use EViews as an example).

We must guarantee several things:

- The original data must remain available
- It must be updated easily.
- Transfer must be as easy as possible.

To achieve these goals, the best organization should be:

- Storing the original file under another name.
- In this file, creating a new page.
- Copying the original series into this page, using "copy with link" for Excel.

We shall suppose the original data is organized as a matrix (or a set of matrixes) with series either in lines or columns. If not, an additional intermediary phase can be needed[36].

- Insert if necessary a line of series names above the first period data (or a column left of the first column).

It does not matter if the matrix does not start in cell B2. Just insert as asked.

- Read it in EViews using import or copy.

This guarantees that:

- The original data is not modified.
- Updates are easy: just copy the new page into the original one (and drag cells in the second page if new observations have appeared).

The only change in the EViews transfer programs concerns the sample period.

In recent versions, the import statement memorizes the reference to the original (Excel) file. EViews will detect if a change is made and propose updating the page in the workfile accordingly.

5.2.5 THE PRELIMINARY PROCESSING OF SERIES

Very often the nature of available series is not really adapted to the needs of the model. A preliminary processing is then necessary. This can apply to several features.

5.2.5.1 Time transformations

Most of the time the series the model builder will access have the right periodicity. Individual exceptions can occur. New series will have to be computed (inside the modelling package).

The change can be undertaken in two directions: aggregating and disaggregating.

[36] For instance, quarterly data can appear in yearly lines of four columns.

5.2.5.1.1 Aggregation

The easiest case happens if the available periodicity is too short. The nature of the variable will lead naturally to a method of aggregation, giving the exact value of the series:

If we call X_t the aggregated variable in t, and $x_{t,i}$ the variable of the sub - period i in t, we can consider the following techniques:

- Sum, for a flow (such as the production of a branch).

$$X_t = \sum_{i=1}^{n} x_{t,i}$$

- Average, for a level (such as unemployment for a given period).

$$X_t = 1/n \sum_{i=1}^{n} x_{t,i}$$

- First or last value, for a level at a given date (for example the capital on the first day of a year will come from the first day of the first quarter). This will apply to stock variables.

$$X_t = x_{t,1}$$

or

$$X_t = x_{t,n}$$

5.2.5.1.2 Disaggregation

When moving to a shorter periodicity, EViews provides a large list of options, depending on the nature (flow, level or stock) of the variable.

The following table is copied from the EViews Help and applies to the “c=” modifier in the copy statement.

For instance:

```
copy(c=q) quart\x
```

can copy the yearly series X into the quarterly page quart using a quadratic smoothing, in such a way that the average of the quarterly values matches the yearly one.

arg	Low to high conversion methods: "r" or "repeata" (constant match average), "d" or "repeats" (constant match sum), "q" or "quada" (quadratic match average), "t" or "quads" (quadratic match sum), "linearf" (linear match first), "i" or "linearl" (linear match last), "cubicf" (cubic match first), "c" or "cubicl" (cubic match last), "pointf" (point first), "pointl" (point last), "dentonf" (Denton first), "dentonl" (Denton last), "dentona" (Denton average), "dentons" (Denton sum), "chowlinf" (Chow-Lin first), "chowlinl" (Chow-Lin last), "chowlina" (Chow-Lin average),"chowlins" (Chow-Lin sum), "litmanf" (Litterman first), "litmanl" (Litterman last), "litmana" (Litterman average), "litmans" (Litterman sum).
rho=arg	Autocorrelation coefficient (for Chow-Lin and Litterman conversions). Must be between 0 and 1, inclusive.

5.2.5.1.3 Smoothing

Smoothing represents a particular case: preserving the same periodicity as the original series, but with the constraint of a regular evolution, for example a constant growth rate. Instead of *n* free values, the choice is reduced to the value of one (or maybe two) parameters.

EViews provides a large set of methods, some very sophisticated, the most popular being the Hodrick-Prescott and Holt-Winters methods. The methodology and syntax are explained in detail in the User's Manual.

5.2.5.1.4 Seasonal adjustment

As we explained before, one method for dealing with variables presenting a seasonality is to eliminate it, and work with seasonally adjusted series.

Several algorithms can be considered, the best known being probably Census X-13-ARIMA and TRAMO-SEATS, both available in EViews.

Obviously, one should not mix original and adjusted series in the same set of model variables.

5.2.5.2 Change of classification

We have already considered this problem when we addressed the fields of models.

Changing categories will usually correspond to aggregation. In the case of economic models, this will apply essentially to:

- Economic agents: one might separate more or less precisely households' categories (following their income, their occupation, their size...), types of firms (according to their size, the nature of their production...), Government institutions (central and local, social security, agencies...).
- Products (production can be described in more or less detail).
- Operations (one can separate social benefits by risk, by controlling agency, or consider only the global value).
- Geographical units (a world model can aggregate countries into zones).

5.2.5.3 Formal transformations

Some variables needed by the model will not be available as such, but will have to be computed from existing series by a mathematical formula. For example, the rate of use of production capacity will be defined as the ratio between effective production and capacity, coming possibly from different sources. Or the relative cost of wages and capital (used for defining the optimal production process) will take into account the price of the two factors, but also the interest rate, the depreciation rate, the expected evolution of wages, and some tax rates.

5.2.6 UPDATES

Once adapted to needs of the model builder, series often will have to be modified.

Changing the values of existing series can have several purposes:

- Correcting a formal error, made by the model builder or the producer of series: typing errors, or errors of concept.
- Lengthening the available sample: new observations will keep showing up for the most recent periods.
- Improving information: for the last known years, series in the French National Accounts appear in succession as provisional, semi-final and final.
- Changing the definition of some variables. For instance, the work of private doctors in State hospitals can move from the market to the non-market sector or vice-versa.

One can also add to the data set a completely new series

- Which has appeared recently as useful to the model.
- Which has been made available by access to a new source of information, or the creation by data builders of a new, more interesting, concept.

This multiplicity of possible changes prohibits the global set of series used by the model to remain the same even for a short period. Adapting constantly model specifications (in particular the estimated equations) to this evolution would ask a lot from the model builder to the detriment of more productive tasks. This means one should limit the frequency of reconstitutions, for the operational data set (for example once or twice per year for an annual model, or every quarter for a quarterly one), with few exceptions: correcting serious mistakes or introducing really important information.

Without doubt, the best solution is actually to manage two sets of data, one updated frequently enough with the last values, the other built at longer intervals (the periodicity of the model for example). This solution allows to study in advance, by estimations based on the first set, the consequences of the integration of new values on the specifications and properties of the next model version.

5.2.7 SUPPRESSIONS

It is beneficial to delete in the bank those series which have become useless:

- This allows to gain space.
- Searches will be faster.
- The bank will be more coherent with the model.
- The model builder will have less information to memorize, and the architecture of the bank will be easier to master (one will have guessed that this is the most important feature, in our sense).

Useless series that are preserved too long lead to forgetting what they represent, and their later destruction will require a tedious identification process.

For EViews, this presents an additional interest: the elements in the workfile will be display in a single window, and it is essential for this window to concentrate as many interesting elements as possible.

5.2.8 THE DOCUMENTATION

Similarly, investment in the documentation of series produces quick returns. It can concern:

- The definition, possibly on two levels: a short one to display titles in tables or graphs, and a long one to fully describe the concept used.
- The source: original file (and sheet), producing institution and maybe how to contact the producer.
- The units in which the series is measured
- Additional remarks, such as the quality and status (final, provisory, estimated) of each observation.
- The date of production and last update (hours and even minutes also can be useful to determine exactly which set of values an application has used). This information is often recorded automatically by the software.
- If pertinent, the formula used to compute it.

Example: Wage rate = Wages paid / (employment x Number of weeks of work x weekly work duration).

EViews allows to specify the first four types, using the label command, and produces automatically the last two.

For example, a series called GDP can be defined through the sequence:

```
GDP.label(c)

GDP.label(d) Gross Domestic Product at constant prices

GDP.label(u) In 2014 Euros

GDP.label(s) from the Excel file accounts.xls produced by the Statistical Office

GDP.label(r) 2019Q4 is provisory
```

Which clears the contents, gives the definition, describes units, the source, and adds remarks.

- In addition, from version 8, EViews allows to introduce one’s own labels, for instance the country for a multinational model, the agent for an accounting one, or the fact that a series belongs to a particular model.

For instance, you can use:

```
HI.label(agent) Households

MARG.label(agent) Firms
```

- If the workfile window screen is in "Display+" mode, you can sort the elements according to their characteristics. In addition to the name, the type and the time of last modification (or creation) you have access to the description.

Moreover, if you right click on one of the column headings, and choose "Edit Columns" you can display additional columns for any of the label types, including the ones you have created.

This can prove quite useful, as it allows you to filter and sort on any criterion, provided you have introduced it as a label.

This criterion can be for instance:

- The agent concerned
- The country
- The association with a given model
- The formula in the model
- The formula used to create the series (if any)[37]
- The type within this model (exogenous, endogenous, identity, behavior…)
- The sub-type: for exogenous it can be policy, foreign, structural. For endogenous it can be behavior or identity.

Once the display is produced, it can be transferred to a table, which can be edited (lines, fonts…) and used for presentations.

For instance, one can produce a table for a model, with columns for type, agent, units, source, identity / behavior…. This table can be sorted using any of the criteria.

These new functions allow table production to be integrated in the modelling process, a very powerful information tool for both model development and documentation.

For instance, you could use:

```
F_HDI.label(d) Disposable income

U_MARG.label(d)  Margins

F_HDI.label(model) France small

U_MARG.label(model)  USA small

F_HDI.label(agent) Households

U_MARG.label(agentl) Firms
```

[37] You can also use the "source"

One of the main interests of this feature is to create a table(using "freeze"). This table can then be sorted according to any of the criteria.

These definitions follow the series as they are moved through the workfile, or even to an external file.

5.2.9 CONSEQUENCES ON WORK ORGANIZATION

Let us now give some specific considerations on data management.

In the general case, the model builder will be confronted with a large set of series of more or less various origins. Optimal management strategy might appear to vary with each case, but in fact it is unique in its main feature: one must produce a file, in the standard of the model building software, and containing the series having a chance to be useful for the model.

This is true even if the global set of necessary series is produced and managed on the same computer or computer network, using the same software (the task of transfer will be simply made easier): it is essential that the model builder has control over the series he uses, and especially that he manages changes (in particular updates of series in current use). Interpreting a change in model properties (simulations, estimations), one must be able to dismiss a change in the data as a source, except if this change has been introduced knowingly by the model builder himself[38].

Such an organization also makes the management of series easier. In particular, limiting the amount of series in the bank, apart from the fact that it will save computer time and space, will make the set easier to handle intellectually.

Concerning the scope of the series, two extreme options can however be considered:

- Transferring in the model bank the whole set of series that have a chance (even if a small one) to become useful at one time to the development of the model39.
- Transferring the minimum, then adding to the set according to needs.

If a median solution can be considered, the choice leans strongly in favor of the first solution. It might be more expensive initially, in human time, and in size of files, but it will prove generally a good investment, as it avoids often a costly number of limited transfers, and gives some stability to the bank as well as to its management procedures.

5.2.10 THE PRACTICAL OPTIONS

For models managed by institutions or research groups, the most frequently found organization is a team working through computers connected through Internet, where storage and synchronization services like Google Drive allow to share files inside a project. A rigorous work organization is needed to manage the elements of the project, between work in progress for which current information must be provided, and documented final versions, which can

[38] This remark is a particular application of the general principle « let us avoid potential problems which can prove expensive in thinking time ».

[39] Even if they are not considered for actual model variables. For instance, one can be interested in comparing the capital output radio of the modelled country with those of other countries.

only be unfrozen by a controlling manager. Not meeting these principles can lead very quickly to a confusing and unmanageable situation.

The final version can be made available online to followers, along with the associated documentation and examples. If the follower has access to the relevant modeling software, direct access to the files can be provided.

In the case of an operational project (like allowing Government economists to produce forecasts and studies) access can be provided through a front end, which does not require any knowledge of the model management software. This is the case for EViews.

As to the original data, it can come from distant sources like the website of the World Bank, or of the statistical office of the model's country. One might in some cases access directly the data sets of the provider from inside a model-building session (this is the case in EViews for the World Bank's WDI). The producers of modelling packages are giving a high priority to this type of option.

One must however pay attention to format incompatibilities, especially if the operating system is different (Windows and its versions, Linux, UNIX, Macintosh...)[40].

5.3 ACCESSING INTERNATIONAL DATA BASES

In this chapter, we describe the use of data provided by international organizations and institutes.

We will not try to be comprehensive: this is a formidable task, certainly more time consuming than the production of the present book.

This is done much better by various institutes, which the reader can find easily with a simple Google search. One very good instance is:

https://www.economicsnetwork.ac.uk/data_sets

Also, a portal giving access to all UN-related sites is:

https://data.un.org/

where you can get access to all the main sites, and also some specific domains, like agriculture and rural development.

Our purpose will rather be the following: to help the producer of a new model to complete the data base he needs for this task.

So we will rather focus on the technical process, limited to the most promising options in our opinion.

Two cases must be considered:

- The model applies to a single country, which represents the interest of the builder.

[40] Most modelling packages work actually under Windows, except freeware like R.

This is particularly relevant if he belongs to an official organization, and is responsible for providing studies (maybe forecasts) applied to the country's economy. In this case, he will have access to the official data for this country, and his model must conform to it, if only to provide results in the official format for data and concepts.

But this is also true if the builder is a local independent (maybe a PHD student). He will be more familiar with the context and have more direct access to local resources.

However, he will need some foreign information, if only to produce some assumptions on foreign demand and prices.

In this case, any source of information is adequate. Detail will only be useful for building detailed scenarios on the evolution of the world economy, and its consequences on local growth. For instance, a Vietnamese modeler might require a description of external trade identifying Chinese growth, to establish the model assumptions.

But the classifications, base years and accounting systems can remain different.

- The model applies to a group (like the European Union) or the world (a set of groupings covering the whole world).

In this case compatibility between models requires access to a single source. The choice must be made at the start, based on a detailed study of the advantages of all solutions available.

We will rather select the most and focus on the technical access to the main sources, with practical examples. A very comprehensive list (giving access to the various sites) will be found in:

5.3.1 THE WORLD BANK

In our opinion, the main sources for international data are:

- The World Bank World Development Indicators
- The OECD Economic Perspectives.

And as a potential complement:

- The IMF and ILO databases.

We will focus on the World Bank "World Developments Indicators" under EViews.

5.3.1.1 The main criteria for the choice

The two main options provide enough information to build a single product model. We have used both to build two multi country models OECD data for the MacSim project detailed above, and the World Bank WDI for a fifteen country ECOWAS model, presently in the final stage of development.

One can wonder why we did not use a single source, which would have simplified the process, in particular the programming.

The reasons are practical, and the choice has been obvious:

The OECD data uses a quarterly periodicity, essential to describe the dynamics of the MacSim developed economies and applying sophisticated econometric techniques such as cointegration.

The countries in the ECOWAS model are not considered developed enough to be described in the OECD data set.

As to the additional sources:

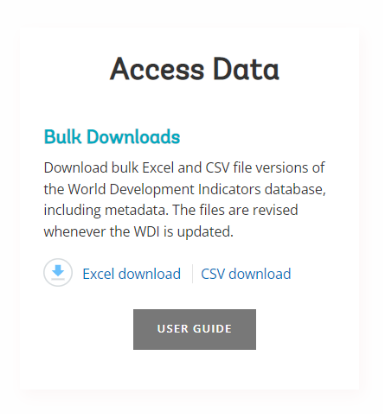
The ILO data set is extremely detailed, but limited to the field of labor: employment, unemployment, revenue and costs. However, if fills gaps on wages, a problem for the WDI and less for OECD.

The IMF data set is logically more detailed on financial series, although many of them can be found in the WDI, with an annual periodicity, however.

5.3.2 THE ORIGINAL SERIES

At this time (22nd June, 2020) The World Bank makes available, on the site

https://datatopics.worldbank.org/world-development-indicators/



The compressed file:

The series are contained in the page “Data” unfortunately as a single page, with all the series (1442 of them) for all the countries (266 of them, including a number of groupings) and the periods 1960 to 2020 (in principle).

As shown below, Columns A to D contain:

- The country name
- The 3-letter acronym
- The definition
- The name using “.” as separator

If used as such, EViews will replace (conveniently) the dots by underscores (“_”).

The following columns contain the data, starting in 1960.

	A	B	C	D	E	F
1	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961
2	Arab World	ARB	Access to clean fuels and technologies for cooking (% of population)	EG.CFT.ACCS.ZS		
3	Arab World	ARB	Access to electricity (% of population)	EG.ELC.ACCS.ZS		
4	Arab World	ARB	Access to electricity, rural (% of rural population)	EG.ELC.ACCS.RU.ZS		
5	Arab World	ARB	Access to electricity, urban (% of urban population)	EG.ELC.ACCS.UR.ZS		
6	Arab World	ARB	Account ownership at a financial institution or with a mobile-money-service provider (% of population ages 15+)	FX.OWN.TOTL.ZS		
7	Arab World	ARB	Account ownership at a financial institution or with a mobile-money-service provider, female (% of population ages 15+)	FX.OWN.TOTL.FE.ZS		
8	Arab World	ARB	Account ownership at a financial institution or with a mobile-money-service provider, male (% of population ages 15+)	FX.OWN.TOTL.MA.ZS		
9	Arab World	ARB	Account ownership at a financial institution or with a mobile-money-service provider, older adults (% of population ages 25+)	FX.OWN.TOTL.OL.ZS		

The list of the countries and series are given in annex, as separate Excel tables.

Obviously, many of the observations are missing, sometimes entirely.

5.3.3 CREATING A SINGLE COUNTRY WORKFILE

If your purpose is to model a single country, no true programming is necessary.

You just have to:

- Locate the country series in the page “Data” of the WDI data set.
- Copy the set in a separate Excel file.
- Create a one-page EViews workfile.
- Import the data into the page by menu or program.

In addition, you can attach the definitions to the series, using the elements of column C of the same page. One can use the first 1431 cells, which apply to all countries. This calls for a little editing.

5.3.4 WORKING ON SEVERAL (OR ALL) COUNTRIES)

We will now present our method to produce a set of EViews workfiles, in which each page contains the whole set of 1341 WDI series associated with a single country or group.

Each page will:

- Use the name of the three-letter acronym for the country or group, used by the World Bank.
- Contain all the WDI series for that country or region.
- Give access to the definition of the variables (both short and long ones).

This is what we have done, and the provided programs will do. As we have considered that the size of a complete file was too big, we divided the countries into 8 regions, following the World Bank’s own partition (column H of sheet Countries).

1 : EAP : EAST ASIA AND PACIFIC

2 : ECA : EUROPE AND CENTRAL ASIA

3 : LAC : LATIN AMERICA COUNTRIES

4 : MEA : MIDDLE EAST AND AFRICA

5 : NAM : NORTH AMERICA AND MEXICO

6 : SAS : SOUTH ASIA

7 : SSA : SUB SAHARIAN AFRICA

8 : OTH : OTHERS (such as OECD members…)

In addition, the region groupings (like the page for South Asia as a whole) will be found in the corresponding regional data set.

The list of countries and sub-groups will be found in the related annex.

5.3.4.1 The method

Although this is not really needed (after all, the files are there), we shall briefly describe the method.

- We create a file called indic.xls with 263 lines (for each country or group) and two columns: the acronym and the number. We read it as a matrix.
- We modify the original World Bank WDI Excel file by separating the "Data" sheet into 6 parts, to meet a constraint on the number of lines. At this moment, only 65536 (=2^{16}) lines can be read, which means 45 countries. Our pages will contain only 40.
- We chose a region.
- We create a workfile for the region.
- We run an EViews program which checks for the presence of the region number, in each of the 6 pages in sequence.
- If the acronym meets the number in **indic**:
  - We create a page with the acronym name.
  - We read the following 1343 series into the page.
- We repeat the process until the end of the sheet is met.

This sequence is available to any user, after editing the program for any changes in the number of countries and series.

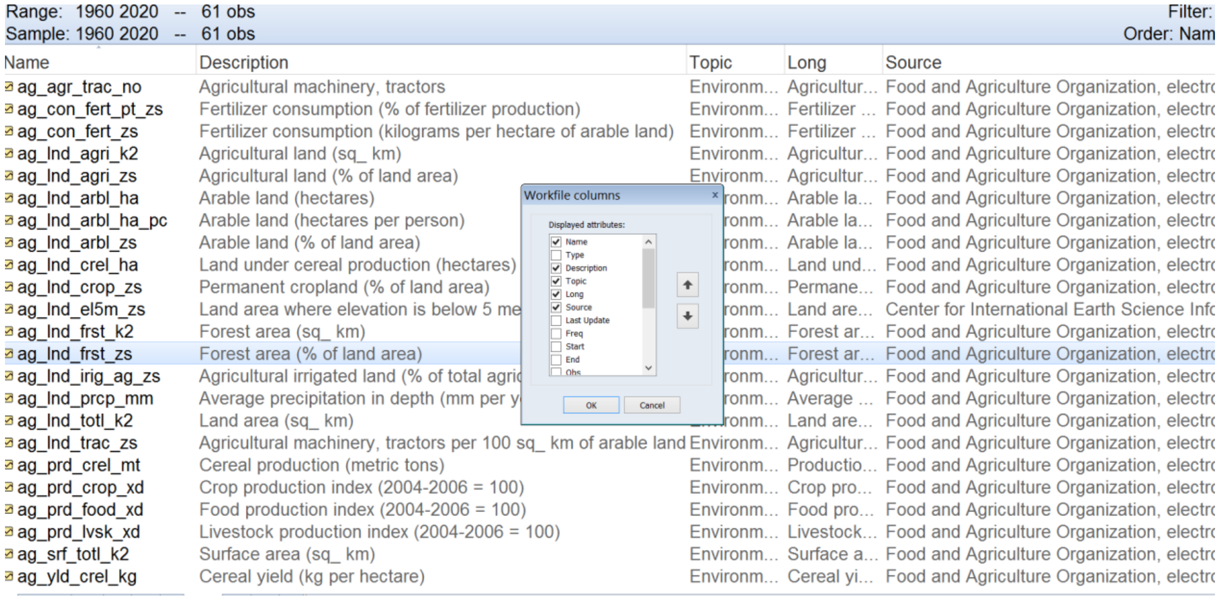
5.3.4.2 the definitions

At the same time as a series is created, a full description is attached. It contains:

- A short description, useful for tables
- The topic of the series, useful for selections
- A longer description, to fully understand the concept
- The source, including the organization and the publication.

The four elements appear when a single series is displayed (using the "sheet" option).

The "Display+" option presents one series per line, including the short description. The other elements can be displayed too, by right-clicking on the top bar, selecting "Edit columns" and clicking in the appropriate boxes (the "Type" and "Last Update" columns can be dropped at the same time).



5.3.4.3 The problems

Although the World Bank provides a lot of information, sometimes in great detail, some very important elements are not present. In particular, these elements are clearly requested if you want to build a general econometric model.

They are, in order of decreasing importance:

- Employment and wages.

These series are clearly required, as they enter the wage-price loop, the production function and the households and firms accounts.

- Capital.

This is required too, but not readily available in most data banks. There are ways to compute it, depending on the related information available.

- Intermediate consumption.

This is needed to compute total demand (which defines imports) and the production price (which defines the trade prices and competitiveness).

- Housing investment.

This is a part of demand.

- Non-wage revenue of households.

This enters household revenue and influences consumption.

- Social contributions.

This affects the revenue of all agents, and the cost of labor (thus the value-added deflator and the capital labor ratio in case of substitution).

5.3.4.4 The program

```
'   A program for creating a workfile with one page
'   for each of the "East Asia and Pacific" countries

cd "d:\eviews\__world_bank_2020"

'     We close the workfiles

'    This workfile contains the single page "countries" with artificial series for all the acronyms (each with the "NA" value);
'   This  can be done easily using the following list

close wb_all
close wb_eap_2020

'    We open the base workfile
'    and save it as wb_eap (for "East Asia and Pacific"

open wb_all
save wb_eap_2020

'     We include a file with a subprogram for creating the series characteristic (short definition, long definition, topic, source)

include def_2020

' We create the acronym series again

pageselect countries
delete(noerr) *
group g_countries
for %1 ABW AFG AGO ALB ANO ARB ARE ARG ARM ASM ATG AUS AUT AZE BDI BEL BEN BFA BGD BGR BHR BHS BIH BLR BLZ BMU BOL BRA BRB BRN BTN BWA CAF CAN CEB CHE CHI CHL CHN CIV CMR COD COG COL COM CPV CRI CSS CUB CUW CYM CYP CZE DEU DJI DMA DNK DOM DZA EAP EAR EAS ECA ECS ECU EGY EMU ERI ESP EST ETH EUU FCS FIN FJI FRA FRO FSM GAB GBR GEO GHA GIB GIN GMB GNB GNQ GRC GRD GRL GTM GUM GUY HIC HKG HND HPC HRV HTI HUN IBD IBT IDA IDB IDN IDX IMN IND IRL IRN IRQ ISL ISR ITA JAM JOR JPN KAZ KEN KGZ KHM KIR KNA KOR KWT LAC LAO LBN LBR LBY LCA LCN LDC LIC LIE LKA LMC LMY LSO LTE LTU LUX LVA MAC MAF MAR MCO MDA MDG MDV MEA MEX MHL MIC MKD MLI MLT MMR MNA MNE MNG MNP MOZ MRT MUS MWI MYS NAC NAM NCL NER NGA NIC NLD NOR NPL NRU NZL OED OMN OSS PAK PAN PER PHL PLW PNG POL PRE PRI PRK PRT PRY PSE PSS PST PYF QAT ROU RUS RWA SAS SAU SDN SEN SGP SLB SLE SLV SMR SOM SRB SSA SSD SSF SST STP SUR SVK SVN SWE SWZ SXM SYC SYR TCA TCD TEA TEC TGO THA TJK TKM TLA TLS TMN TON TSA TSS TTO TUN TUR TUV TZA UGA UKR UMC URY USA UZB VCT VEN VGB VIR VNM VUT WLD WSM XKX YEM ZAF ZMB ZWE


if @isobject(%1)=0 then
genr {%1}=na
g_countries.add {%1}
endif
pagedelete(noerr) {%1}
next

  '    We read the file indic
  '    creating a vector with the category for each acronym
  '    from to 1 to 8

  '   Here 1 = East Asia and Pacific

 '    This file can be adapted to any case
 '    For instance using 1 for relevant countries and 0 for others
 '    one can restrict the workfile to a subset

vector(263) indic
indic.read(type=excel,b1,s=indic) indic.xls
```

```
'   We

!j=1

'   The number of countries and groups

!n=263

'    The number of countries in each Excel page (1 to 7)

!p=40
!l=1
while !n>0

'    The starting ilne for a country's series
'    We initialize by “2” if there is an items definition line
‘    1 otherwise

!k=    2

'    We process the "p" next countries
'    in page number "j"

'    "l" is the number of the country
'    and %1 the associated acronym

'

for !i=1 to !p
pageselect countries
%1=g_countries.@seriesname(!l)

'   We delete the page preventively

pagedelete(noerr) {%1}

'   If the country belongs to the set

if indic(!l)=1 then

'    We create the page

pagecreate(page={%1}) a 1960 2020

'    We read the 1431 series from the EXcel sheet number "j"
'    starting in cell e{!k}

smpl 1960 2020
read(type=excel,e{!k},s=Data{!j},t) wdiexcel20.xls 1431

'    Once data is read we associate definitions

 call def

 '    We close the page

close {%1}
endif

'    If the country does not belong we do nothing

pageselect countries

'     We increment the starting line in the sheet
```

```
!k=!k+1442

'    and the number of countries treated (relevant or not)

!l=!l+1
next

'    We increment the sheet number

!j=!j+1

'    "n" is he number of remaining countries

!n=!n-!p

'    If it is lower than the number of countries in a sheet
'    we have reached the end of the countries set
if (!n<!p) then

    '    The number of remaining countries is now "n"
!p=!n
endif
wend

'    We save the workfile

wfsave wb_eap_2020
```

5.3.5 THE ALTERNATE SOURCES

The simplest and most accurate way to fill the gaps in the WDI data is the access to an alternate source. We shall concentrate on the first case, clearly the most important.

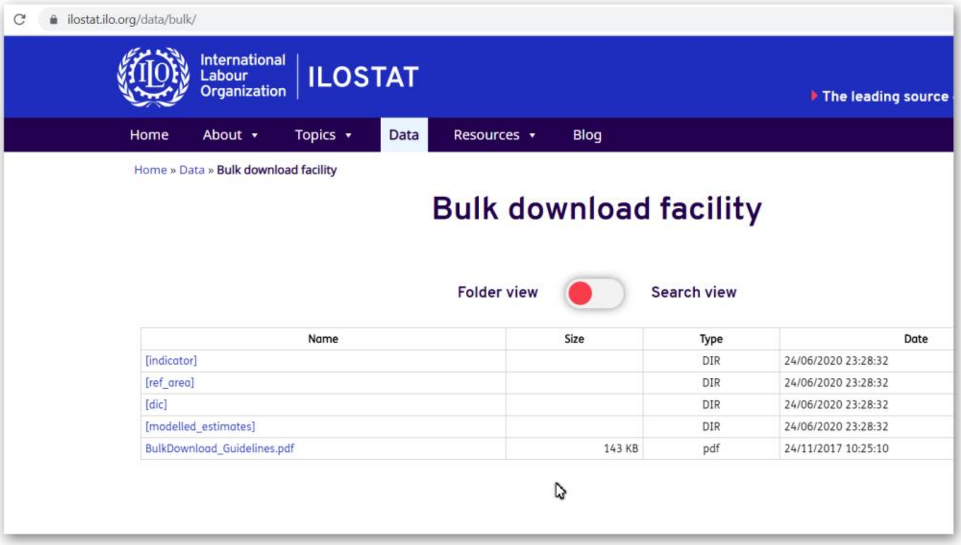
Wages and employment can be obtained from:

5.3.5.1 The International Labor Organization.

This institution (www.ilo.org) looks like the most logical source. It does provide information for 234 countries (almost all countries in the world) and 433 indicators.

There are two main options:

- Downloading a whole set of data, using "Bulk download".



The explanations are given in the last file:

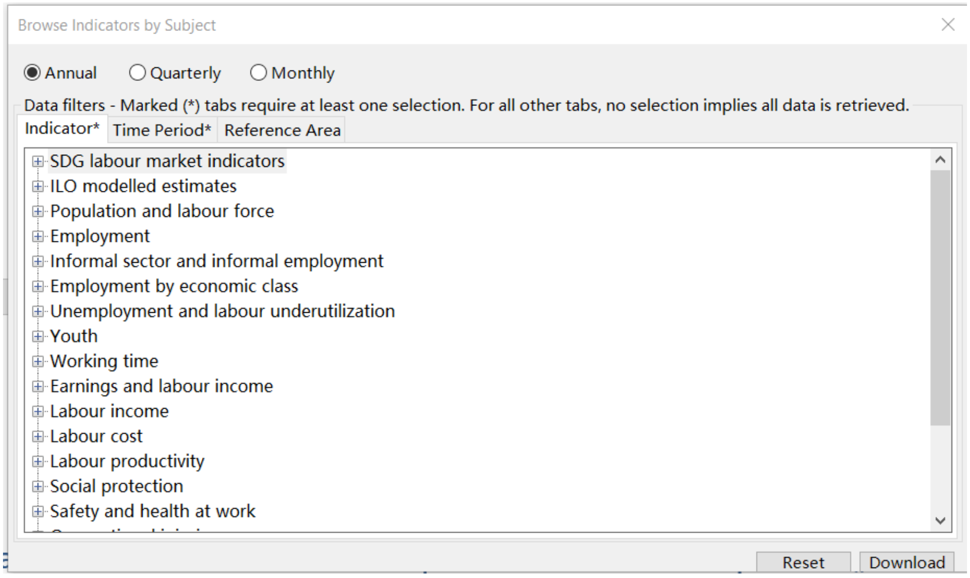
https://www.ilo.org/ilostat-files/Documents/ILOSTAT_BulkDownload_Guidelines.pdf

- Downloading the data through Excel.

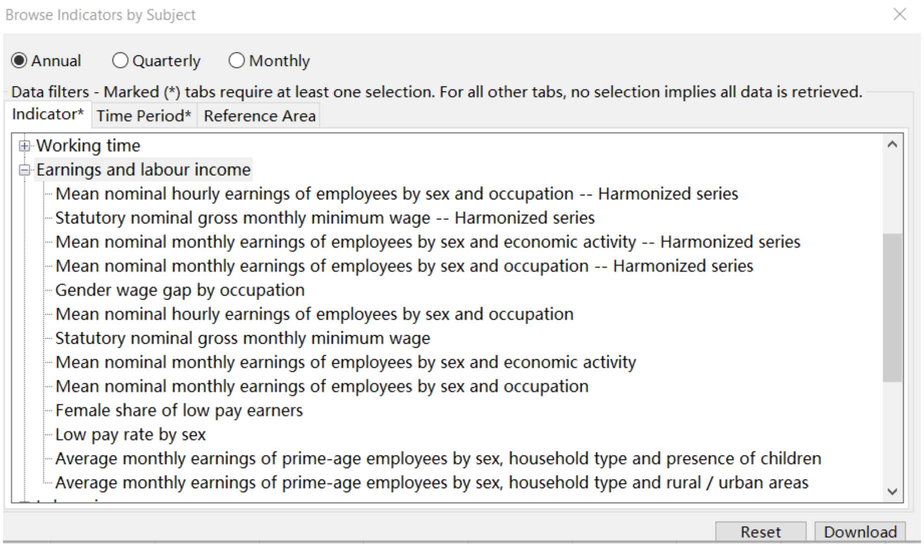
If you are currently accessing the ILO data by Internet, the Excel menu will be modified automatically to include a "ILOSTAT" item, as follows (sorry for the French):

As you can see, you can access data for a country or a subject.

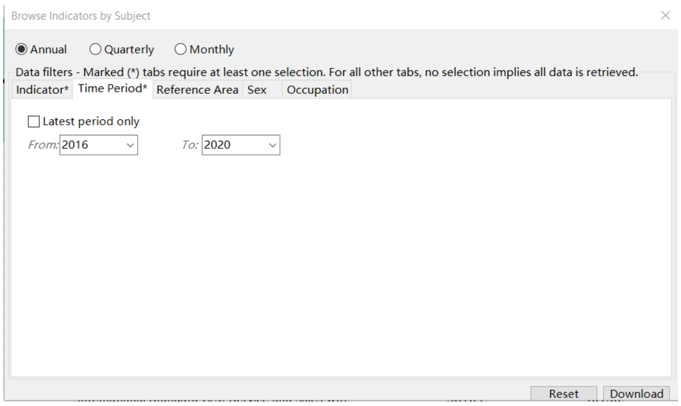
If you chose "Subject" you get the following menu:



which leads to the menu:



You can choose the time period and periodicity:



This will start a download for all available series, for all countries for which at least one series is present.

The results can appear as:

Mean nominal monthly earnings of employees by sex and occupation (local currency)

http://www.ilo.org/ilostat/faces/oracle/webcenter/portalapp/pagehierarchy/Page27.jspx?indicator=EAR_XEES_SEX_OCU_NB&subject=EAR&collection=A&locale=en

Downloaded from ILOSTAT SDMX service on 25juin2020

Country	Source	Sex	Occupation version	Occupation	2015	2016	2017	2018	2019	2020	
Albania	Insurance records	Total	International Standard Cl	Total	46829	47522	48967	50589			Notes
Albania	Insurance records	Total	International Standard Cl	1. Managers	94912	95961	95380	99013			Notes
Albania	Insurance records	Total	International Standard Cl	2. Professionals	67070	64628	66297	67888			Notes
Albania	Insurance records	Total	International Standard Cl	3. Technicians and assoc	50270	52145	54088	56647			Notes
Albania	Insurance records	Total	International Standard Cl	4. Clerical support worke	46278	47826	50878	53123			Notes
Albania	Insurance records	Total	International Standard Cl	5. Service and sales wor	34218	34832	36844	37822			Notes
Albania	Insurance records	Total	International Standard Cl	6. Skilled agricultural, for	23540	27739	29910	31026			Notes
Albania	Insurance records	Total	International Standard Cl	7. Craft and related trade	32303	31447	33604	34887			Notes
Albania	Insurance records	Total	International Standard Cl	8. Plant and machine ope	33756	32086	34123	35400			Notes
Albania	Insurance records	Total	International Standard Cl	9. Elementary occupatioı	25440	27231	28990	30502			Notes
Albania	Insurance records	Total	International Standard Cl	0. Armed forces occupat	56126	54049	58075	59383			Notes
Albania	Insurance records	Total	International Standard Cl	X. Not elsewhere classified		0	0	0			Notes
Albania	Insurance records	Male	International Standard Cl	Total	48358	50084	51392	53181			Notes
Albania	Insurance records	Male	International Standard Cl	1. Managers	89319	96813	95420	99309			Notes
Albania	Insurance records	Male	International Standard Cl	2. Professionals	70537	76779	77638	78799			Notes
Albania	Insurance records	Male	International Standard Cl	3. Technicians and assoc	51801	54999	56561	59922			Notes
Albania	Insurance records	Male	International Standard Cl	4. Clerical support worke	44779	46549	50354	52366			Notes
Albania	Insurance records	Male	International Standard Cl	5. Service and sales wor	36826	38164	40136	41393			Notes
Albania	Insurance records	Male	International Standard Cl	6. Skilled agricultural, for	24102	28171	30439	31611			Notes
Albania	Insurance records	Male	International Standard Cl	7. Craft and related trade	33272	34900	37643	38931			Notes
Albania	Insurance records	Male	International Standard Cl	8. Plant and machine ope	39708	36722	39277	40355			Notes
Albania	Insurance records	Male	International Standard Cl	9. Elementary occupatioı	25562	27991	29745	31483			Notes
Albania	Insurance records	Male	International Standard Cl	0. Armed forces occupat	57256	54027	58074	59551			Notes
Albania	Insurance records	Male	International Standard Cl	X. Not elsewhere classified		0	0	0			Notes
Albania	Insurance records	Female	International Standard Cl	Total	45033	44354	46018	47490			Notes

5.3.5.2 The OECD (Organization for Economic Cooperation and Development.

First, access to this information requests registration and logging in, at least in some cases. But this is free, and advises you periodically (not too often) on various updates. And by stating your preferences, you can direct this information.

This is a very interesting source (stats.oecd.org). It provides data from 2000 to 2018, sometimes 2019 (at present).

The only drawback is the limitation to the 40 following countries:

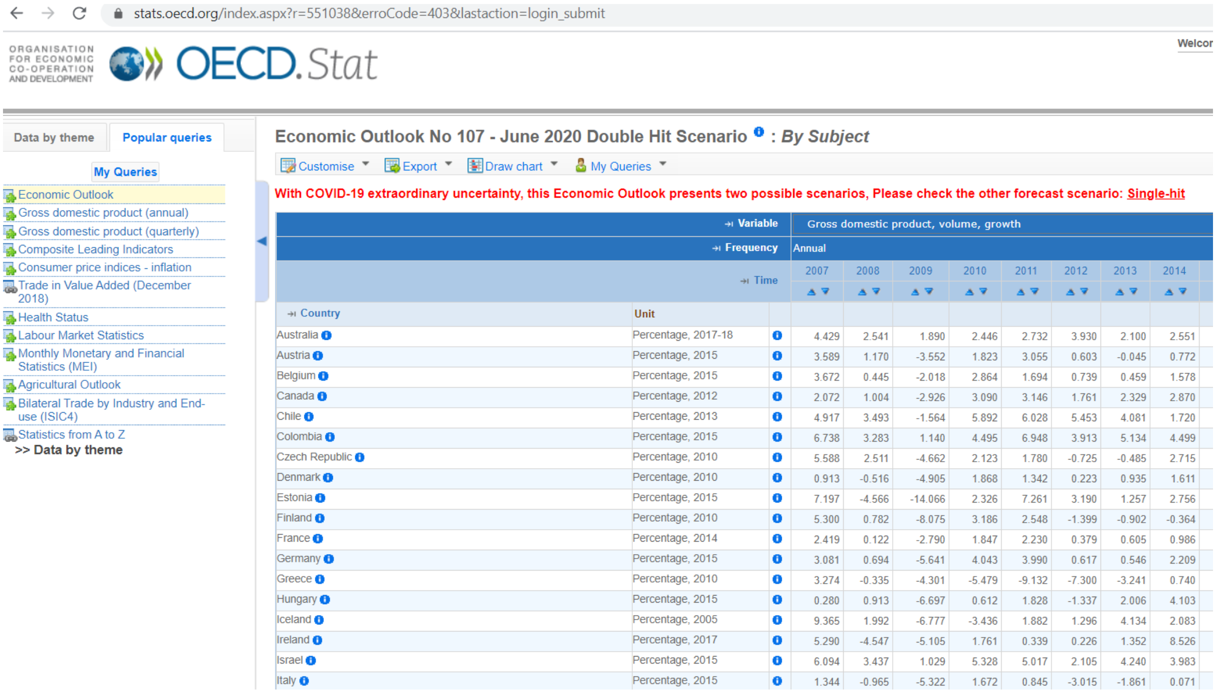
Australia	Estonia	Italy	Portugal
Austria	Finland	Japan	Russian Federation
Belgium	France	Korea	Slovak Republic
Brazil	Germany	Latvia	Slovenia
Canada	Greece	Lithuania	South Africa
Chile	Hungary	Luxembourg	Spain
China, People's Republic of	Iceland	Mexico	Sweden
	India	Netherlands	Switzerland
Colombia	Indonesia	New Zealand	Turkey
Czech Republic	Ireland	Norway	United Kingdom
Denmark	Israel	Poland	United States

(excluding for instance all African countries except South Africa).

After logging in, you will have access to the following menu, with the Economic Outlook the most interesting elements formacromodelers.

You can also access it directly busing:

https://stats.oecd.org/viewhtml.aspx?datasetcode=EO107_INTERNET_2&lang=en

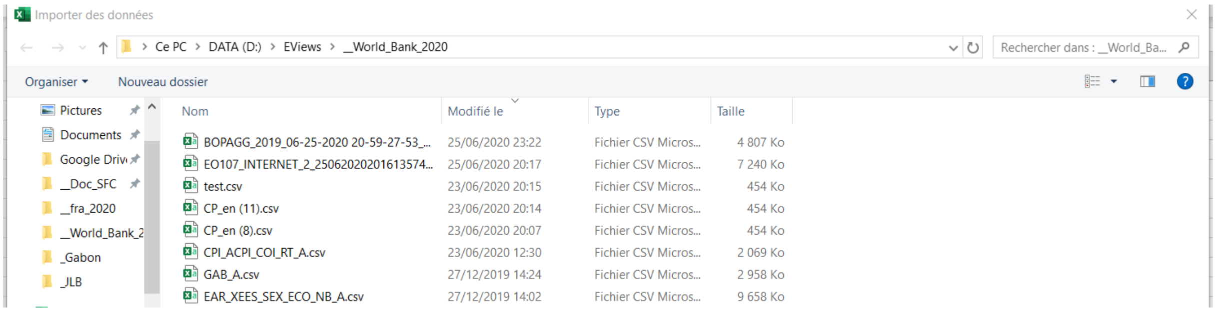


Using "Customize" you can decide on the countries, the topics and the time span.

Then you can ask from the selected elements to be downloaded as and Excel or CSV file.

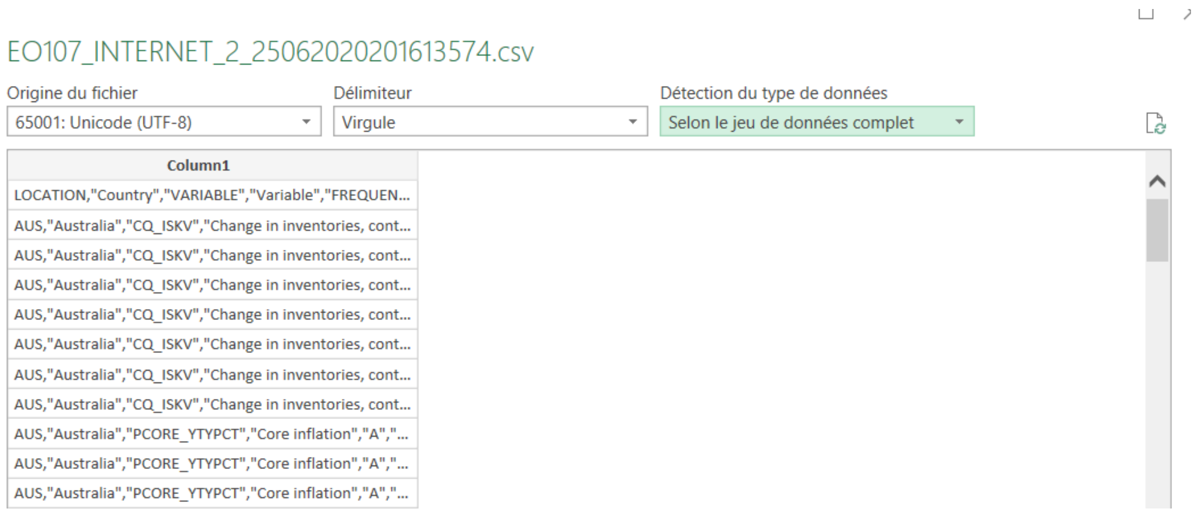
Remember that to read a CSV file in Excel you should not try to open it directly, but rather open a blank sheet and use:

Data >Obtain data>From file>From Text/CSV.



Select the OECD EO107 (or later) downloaded file.

and select "Comma" as a delimiter ("Virgule' in French).to organize the information into colums.

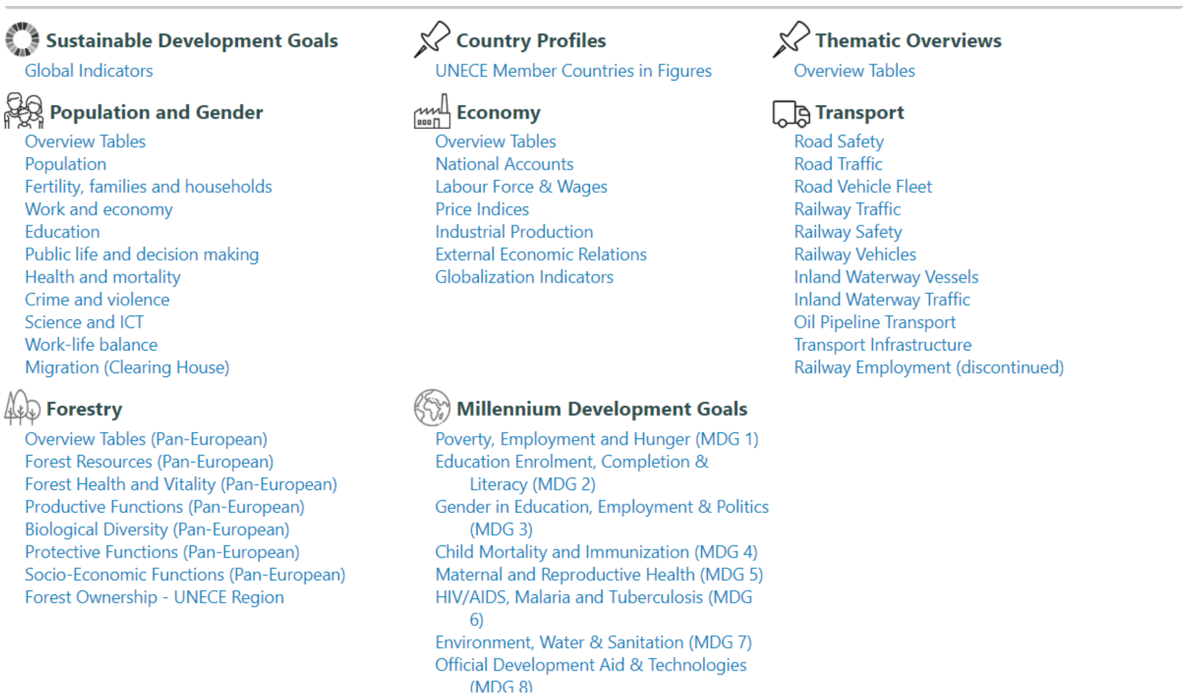


5.3.5.3 The United Nations, in particular the Economic Commission for Europe.

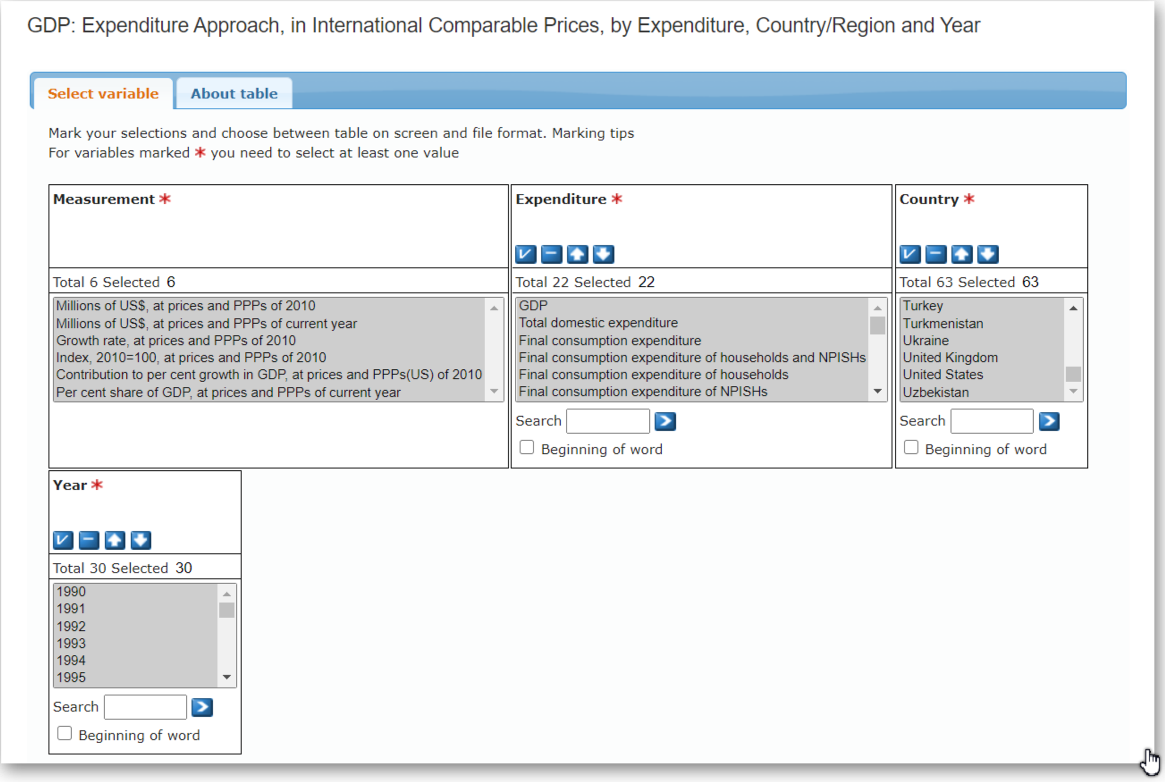
The site:

https://w3.unece.org/PXWeb/en

gives you access to the following menu:



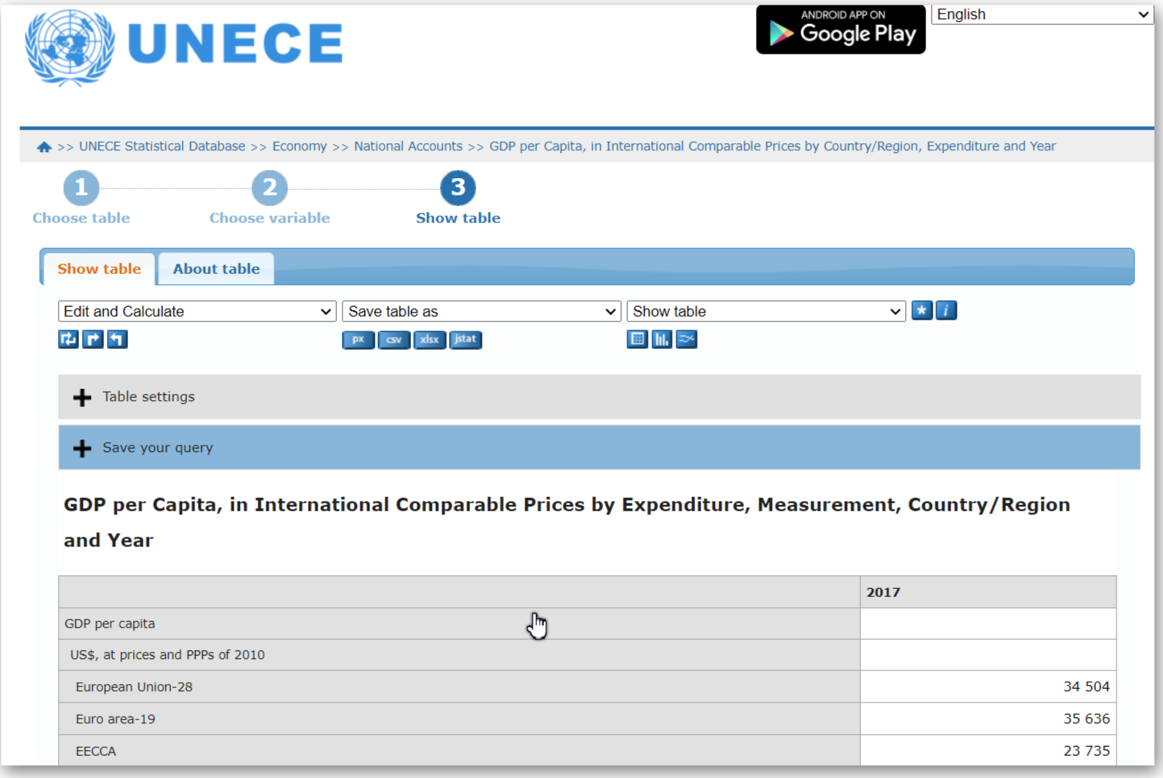
If you chose “National Accounts” you get:



The countries are (63 including groupings):

European Union-28, Euro area-19, EECCA, CIS-11, North America-2, UNECE-52, Western Balkans-6, Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Canada, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Kazakhstan, Kyrgyzstan, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Republic of Moldova, Romania, Russian Federation, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Tajikistan, Turkey, Turkmenistan, Ukraine, United Kingdom, United States, Uzbekistan.

Once you have made your selection, you can save it in different format (through "Save as"), the most practical being probably Excel .xlsx.



5.3.5.4 The International Monetary Fund.

After signing in (free) on the site:

https://www.imf.org/en/Data

you are given access to the following files, for 193 countries and 11 groups:

External Sector

1. Balance of Payments and International Investment Position
2. Coordinated Direct Investment Survey
3. Coordinated Portfolio Investment Survey
4. Currency Composition of Official Foreign Exchange Reserve
5. International Reserves and Foreign Currency Liquidity

Fiscal Sector

1. Fiscal Decentralization
2. Fiscal Monitor
3. Government Finance Statistics
4. Historical Public Debt
5. Public Sector Balance Sheet (PSBS)
6. Revenue Administrations Fiscal Information Tool
7. World Revenue Longitudinal Data

Financial Sector

1. Financial Access Survey
2. Financial Development
3. Financial Soundness Indicators
4. Monetary and Financial Statistics

Real Sector

1. Commodity Terms of Trade
2. Consumer Price Index
3. Direction of Trade Statistics
4. Primary Commodity Price System

Cross-domain

1. Asia and Pacific Economic Outlook
2. Gender Budgeting and Gender Equality
3. International Financial Statistics
4. Middle East and Central Asia Economic Outlook
5. Sub-Saharan Africa Economic Outlook
6. Western Hemisphere Economic Outlook

The reference will be created as something like:

https://data.imf.org/?sk=388DFA60-1D26-4ADE-B505-A05A558D9A42&sId=1479329328660

and propose the files (which we will not present individually):

You can ask for the download of a given file. It will be sent using you registered e-mail address, for instance as:

Bulk export

do_not_reply@imf.org
À : jlbrillet@yahoo.fr

Your dataset BOPAGG_2019 (format: CSV) has been exported.
Download .zip archive file
It will be available under My Data for 7 days.

If you want data for a specific country (like Gabon) and table (like the Balance of Payments), you can use:

Refresh Autoupdate Export | Controls | Find

Report | Tools and panels | Search

Annual | Quarterly

Country: Gabon   Date: (21 of 580 selected)

Gabon: Balance of Payments Analytic Presentation by Country

Millions of U.S. Dollars

Click here for country specific metadata

	2005	2007	2008	2009	2010	2011	2012	2013	2014	2015
Current account (excludes reserves and related items)	1,983.0	2,748.2	3,829.6	804.1	2,453.2	2,604.4	1,977.6	1,462.6	1,112.1	141.0
Goods, credit (exports)	5,456.9	6,835.6	9,103.0	6,375.8	8,857.7	10,150.7	9,782.6	9,358.6	8,290.4	5,112.6
Goods, debit (imports)	1,358.8	1,890.2	2,143.3	2,576.9	3,125.8	3,655.9	3,645.4	4,049.3	3,852.1	3,171.4
Balance on goods	4,098.0	4,945.4	6,959.6	3,798.9	5,731.9	6,494.8	6,137.3	5,309.3	4,438.4	1,941.3
Services, credit (exports)	145.5	179.2	139.4	241.3	482.6	325.7	264.6	271.7	145.4	277.3
Services, debit (imports)	1,034.0	993.4	1,638.9	1,650.0	1,797.6	2,558.7	2,401.9	2,701.2	2,379.3	1,877.9
Balance on goods and services	3,209.6	4,131.2	5,460.1	2,390.2	4,416.8	4,261.9	3,999.9	2,879.8	2,204.5	340.7
Primary income, credit	36.5	105.6	38.8	250.7	102.0	184.5	117.4	40.2	122.1	111.2
Primary income, debit	994.1	1,234.2	1,327.1	1,476.9	1,839.2	1,450.7	1,684.0	882.6	656.1	215.5
Balance on goods, services, and primary income	2,252.0	3,002.6	4,171.8	1,164.0	2,679.5	2,995.6	2,433.4	2,037.3	1,670.5	236.4
Secondary income, credit	17.7	12.1	13.3	8.8	75.7	74.3	50.6	39.4	23.8	91.1
Secondary income, debit	286.7	266.6	355.4	368.8	302.0	465.5	506.4	614.1	582.3	186.4
Capital account (excludes reserves and related items)	0.0	0.0	296.8	-2.5	1.2	-0.4	...	...	...	3.8
Capital account, credit	0.0	0.0	296.8	...	1.5	...	...	...	...	3.9
Capital account, debit	0.0	...	...	2.5	0.3	0.4	...	...	...	0.0
Balance on current and capital account	1,983.0	2,748.2	4,126.4	801.5	2,454.4	2,604.0	1,977.6	1,462.6	1,112.1	144.8
Financial account (excludes reserves and related items)	1,342.0	1,633.3	2,509.6	564.3	1,244.6	1,113.4	1,289.8	586.6	473.9	962.9
Direct investment, assets	159.2	58.9	-128.1	243.5	377.1	25.4	910.7	236.3	480.1	16.5
Equity and investment fund shares	-15.6	59.3	5.2	328.7	265.4	-2.1	9.9	0.9	17.1	-113.8
Debt instruments	174.8	-0.5	-133.3	-85.2	111.8	27.5	900.8	235.4	463.1	130.3
Direct investment, liabilities	326.2	654.8	693.5	635.7	523.9	1,123.5	677.3	324.0	1,263.1	41.7
Equity and investment fund shares	320.6	482.2	695.7	607.8	650.6	553.4	825.9	185.3	242.4	-125.2
Debt instruments	5.5	172.6	-2.2	27.9	-126.6	570.1	-148.6	138.7	1,020.7	167.0

………

which you can export to Excel (not CSV!) using the corresponding menu item.

External Sector

1. Balance of Payments and International Investment Position
2. Coordinated Direct Investment Survey
3. Coordinated Portfolio Investment Survey
4. Currency Composition of Official Foreign Exchange Reserve
5. International Reserves and Foreign Currency Liquidity

Fiscal Sector

1. Fiscal Decentralization
2. Fiscal Monitor
3. Government Finance Statistics
4. Historical Public Debt
5. Public Sector Balance Sheet (PSBS)
6. Revenue Administrations Fiscal Information Tool
7. World Revenue Longitudinal Data

Financial Sector

1. Financial Access Survey
2. Financial Development
3. Financial Soundness Indicators
4. Monetary and Financial Statistics

Real Sector

1. Commodity Terms of Trade
2. Consumer Price Index
3. Direction of Trade Statistics
4. Primary Commodity Price System

Cross-domain

1. Asia and Pacific Economic Outlook
2. Gender Budgeting and Gender Equality
3. International Financial Statistics
4. Middle East and Central Asia Economic Outlook
5. Sub-Saharan Africa Economic Outlook
6. Western Hemisphere Economic Outlook

5.3.6 BACK TO OUR EXAMPLE

Now that we know the principles, let us see how to apply them to the case we have defined earlier. To avoid switching between distant pages, we shall repeat its presentation.

1. In the example, our economist has decided to build a very simple model of a country's economy, which includes the following elements: Based on their production expectations and productivity of factors, firms invest and hire workers to adapt productive capacity. However, they exert some caution in this process, as they do not want to be stuck with unused elements.
2. Productive capital grows with investment but is subject to depreciation.
3. The levels actually reached for labor and capital define potential GDP.
4. They also need intermediate products (proportional to actual GDP), and adapt inventories, from the previous level.
5. Households obtain wages, based on total employment (including civil servants) and a share of Gross Domestic Product. They consume part of this revenue and invest another (in housing).
6. Final demand is the sum of its components: consumption, productive investment, housing investment, inventories, and government demand. Total demand includes also intermediate consumption. Final and total demand are the sum of their components
7. Imports are a share of local total demand, final or intermediate. But the fewer capacities remain available, the more imports will be called for.
8. Exports follow world demand, but the priority of local firms is satisfying local demand. They are also affected by capacity constraints.
9. Supply is equal to demand.

We have voluntarily kept the framework simple, as our purpose is only explanatory at this time. However, the model we are building has some economic consistency, and can actually represent the nucleus for further extensions which we shall present later.

We shall also suppose that the following data is available in an Excel file called FRA.XLS, selected from OECD's Economic Perspectives data set. Series are available from the first quarter of 1962 to the last of 2010. However, the set contains a forecast, and the historical data ends in 2004.

A note: the reason for using older data is not laziness in updating the statistics. The period we are going to consider is the most interesting as it includes years of steady growth (1962 to 1973), followed by uncertainty following the first oil shock. This justifies too using French data: the main point here is using actual data for a non-exceptional country. When we move to a more operational case our data set will include later periods.

The reason for the "FRA" prefix is to identify series for France in a large set of countries, representing all the OECD members as well as some groupings.

They use the following units:

Values: Euros

Deflators: base 100 in 1995.

Volumes (or quantities): Millions of 1995 Euros

Populations: persons

FRA_CGV	Government Consumption, Volume
FRA_CPV	Private Consumption, Volume
FRA_EG	Employment, Government
FRA_ET	Total Employment
FRA_CGV	Government Consumption, Volume
FRA_CPV	Private Consumption, Volume
FRA_EG	Employment, Government
FRA_ET	Total Employment
FRA_FDDV	Final Domestic Demand, Volume
FRA_CGV	Government Consumption, Volume
FRA_CPV	Private Consumption, Volume
FRA_EG	Employment, Government
FRA_ET	Total Employment
FRA_FDDV	Final Domestic Demand, Volume
FRA_GAP	Output Gap
FRA_GDPTR	Potential Output, Total Economy at Current Prices
FRA_GDPV	Gross Domestic Product (Market prices), Volume
FRA_IBV	Gross Fixed Cap Form, Business Sector, Volume(Narrow Definition)
FRA_ICV	Intermediate consumption, Volume
FRA_IGV	Government Investment, Volume
FRA_IHV	Investment in Housing, Volume
FRA_ISKV	Increase in stocks. volume
FRA_KBV	Capital Stock, Business
FRA_MGSV	Imports Goods and Services, N.A. Basis, Volume

FRA_PCP	Deflator, Private Consumption
FRA_TDDV	Total Domestic Demand, Volume
FRA_WSSS	Compensation of Employees
FRA_XGSV	Exports Goods and Services, N.A. Basis, Volume
FRA_XGVMKT	Exports Goods and Services, Market Potential, Volume
FRA_YDRH	Real Household Disposable Income

Applying the principles, we have defined above calls for:

- Creating the model specifications
- Identifying the variables in the model.
- Separating them into endogenous and exogenous.
- Writing down the full identities.
- Establishing each behavioral equation as an identity, presenting in the simplest way the variable it defines, and the explanatory elements.
- Creating the associated series, from the available data.
- Transferring the elements already available into model series using the names allocated to them.
- Specifying formulas computing the remaining elements

Now that we have obtained the data, we can move to the two tasks: transform it to fit the model needs, start specifying the model equations.

It should be clear that this will have to be done through a set of stored statements in a readable language (a program). This option will allow:

- Establishing an apparently consistent set of statements, which can be controlled visually.
- Locating errors and introducing corrections as simply and clearly as possible.
- Storing subsequent versions, including the last and most correct one, until a satisfying version is established.
- Replicating this process with the smallest amount of work.
- Displaying the steps in the process as clearly as possible, introducing comments.
- Once a satisfying stage has been reached, memorizing the actions for later use (especially if the modeling project faces breaks, short or long).
- Allowing external users to master the current state of operations, to evaluate the present stage of development of the project.

The program can be inserted with comments, making the sequence of tasks and the role of individual commands clearer, and allowing to warn of the presence of local problems and the way they have been processed. This is especially useful for a team project, for which the name of the author should also be included.

Under EViews, two other methods are available:

- Using a sequence of menu and sub-menu functions,
- Typing commands without saving them, directly from the command window.

These two methods fail on all criteria. The record of the tasks is not available, which means errors are difficult to detect. Reproducing the task, whether to correct errors or to update specifications or data, calls for a new sequence of menu selections or typed statements[41].

[41] However, one can copy the sequence of statements entered in the command window into a program file.

The obvious choice is even comforted by three features provided by recent versions:

- You can run part of a program, by selecting it with the mouse (in the usual Windows way), clicking on the right button, and choosing “Run Selected”.

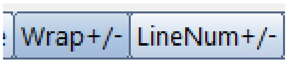
This is generally more efficient than the previous method of copying the selected part into a blank program, and running it. However, the new method does not allow editing, useful when one wants to run a selected AND modified set.

- Symmetrically one can exclude temporarily from execution part of a program, by “commenting it out”. To do this, one should select the relevant part, click on the right button, and choose “Comment Selection”. To reactivate the statements, one should select them again and use “Uncomment Selection”. One can also type a single quote (‘) before the statement.

This can be a little dangerous, especially if you (like myself) have the reflex of saving the program before each execution. To avoid destroying the original, one can save first the modified program under another name[42].

- Finally, one can ask a column of numbers to be displayed left of the program lines. This is particularly efficient if you use the “Go To Line” statement[43].

This is done by first unwrapping the command lines. Each command will use one line regardless of its length, which can be a little annoying for very long ones. Then (and only then) one can ask for the line numbers.



- So actually, the only option is the one we proposed above: defining a program producing all the necessary information, and the framework of equations which is going to use it. But the ordering of the tasks can be questioned, as we have started explaining earlier. Until both are completed, the job is not done, but they are technically independent: one does not need the physical model to create the data, or series filled with values to specify the equations. This means that one can consider two extreme methods:

- Creating all the data needed by the model, then specifying the model.
- Specifying all the model equations, and then producing the associated data.

The criterion is the intellectual feasibility of the ordered sequence of tasks.

Clearly the first option is not realistic, as writing down the equations will surely evidence the need for additional elements. The second is more feasible, as one does not need actual series to write an equation. But as the definition of the equation processes, one has to check that all the addressed elements are or will be available in the required form, either as actual concepts (GDP) or transformations of actual concepts (the budget deficit in GDP points calls for the deficit and GDP series). If a concept appears to be lacking, one will have to: use an alternate available element (a “proxy”), establish an assumption, look in alternate bases not yet accessed, or simply eliminate the element from the model.

This shows that if producing both sets can be done in any order, there is a preference for specifying the equations first, with a general knowledge on data availability. If the data set is not ready, but its contents are known, it is possible to write down the equations and ask the software to proceed the text. The user will be told about possible

[42] Only once of course.

[43] However, you have to be careful to update the numbers when the program changes.

syntax errors, about the nature of the variables (endogenous / exogenous), and the architecture of his model. This will lead to early model corrections, allowing to gain time and avoiding taking wrong directions later. And if the model specifications are still discussed, it is possible to build a first version of the associated data set, which will be updated when the model is complete.

In practice, especially in the simplest cases, one can also start defining the program with two blank paragraphs, and fill them with data and equation creating statements until both are complete. The eight original paragraphs in our model specifications can be treated one by one (not necessarily in the numerical order) filling separately the data and equation generating blocks with the associated elements.

Actually, among the above proposals we favor two alternate techniques:

- **Model then data:** Specifying first the full model, checking that all elements used can be produced either directly or through a formula. Then producing the full set of data, preferably through a direct transfer or a transformation.
- **Model and data**: Producing the equations in sequence, or related block by related block, and establishing simultaneously the statements which create all the series they need.

5.3.6.1 Application to our example

Let us now show on our example how the process can be conducted using the second method, probably more adapted to such a small model (one cannot expect to switch between the two processes too many times).

We shall first present the process in general (non-EViews) terms, treating each case in sequence, and presenting both the equations and the statements generating the associated variables. To make thinks clearer, the equations will be numbered, and the creation statements will start with “>>”.

Also, the endogenous variable will use uppercase characters, the exogenous lowercase. This has no impact on treatments by EViews, but will make interpretation clearer for the model builder and especially for his

(1) *Based on their production expectations and productivity of factors, firms invest and hire workers*.

This defines two behavioral equations for factor demand, in which employment (let us call it LE) and Investment (called I) depend on GDP, called Q.

```
(1) LE=f(Q)

(2) I=f(Q)
```

We need:

```
>> IP=FRA_IBV

>> Q =FRA_GDPV
```

But for LE, we face our first problem. Private employment is not directly available. However, we have supposed that total employment contained only public (government) and private. This means we can use:

```
>> LE=FRA_ET-FRA_EG
```

In another case, private and public employment could have been available, but not the total, which would have been computed as a sum. This highlights the fact that computation and economic causality need not be related.

(2) *Productive capital grows with investment but is subject to depreciation*.

Capital K, measured at the end of the period, is defined by an identity. Starting from the initial level, we apply a depreciation rate (called dr) and add investment. The equation is written as:

```
(3) K(t)= K(t-1).(1-dr(t)) + I(t)
```

Defining K at the end of the period would only change notations.

We need the data for K

```
>> K=FRA_KBV
```

And we get dr by inverting the formula:

```
>> dr=((K(-1) + IP) – K) / K(-1)
```

In other words, dr will be the ratio, to the initial capital level, of the difference between two levels of capital: the value we would have obtained without depreciation, and the actual one.

(3) The levels actually reached define potential production.

Capacity (called CAP) depends on factors LE and K

```
(4) CAP(t)=f(LE(t), K(t))
```

It can be computed directly as:

```
>> CAP=FRA_GDPVTR
```

which rather represents a “normal” GDP value considering the present level of factors.

The direct availability of this concept as a series represents the best case, not often met in practice. Later in the text we shall address the alternate techniques available in less favorable situations.

(4) They need inputs, and also build inventories.

Intermediate consumption can be defined as proportional to GDP, using the actual value. This means that at any level of production, each unit produced will need the same amount of intermediary products.

```
(5) IC = r_icq . Q
```

For inventories, we will estimate its change:

```
(6) CI=f(Q)
```

For this, we need to compute:

```
>> IC=FRA_ISKV

>> r_icq=IC/Q (or FRA_ISKV/FRA_GDP)

>> CI=FRA_CIV
```

(5) Households obtain wages, based on total employment (including civil servants) and a share of Gross Domestic Product. They consume part of this revenue.

Now we need to define total employment, by adding government employment (called lg) to LE.

(7) LT=LE+lg

The new series are obtained by:

```
>> LT=FRA_ET

>> lg=FRA_EG
```

Now we have to compute household revenue, which we shall call R_HI. We shall suppose that the same wage applies to all workers, and that the non-wage part of Household revenue is a given share of GDP, a series called r_rhiq. This gives:

(8) RHI = wr . LT + r_rhiq . Q

Actually the above assumption, while simplistic, is probably not too far from the truth. The sensitivity to GDP of the elements included in this heterogeneous concept can be low (such as pensions, or interests from long-term bonds), high (the revenue of small firm owners, with fixed costs and variable output), or medium (self-employed working in the same capacity as wage earners).

Household consumption is given by applying to RHI the complement to 1 of a savings rate which we shall call sr. For the time being, the savings rate is exogenous:

(9) CO = RHI . (1 – sr)

Housing investment is also a share of RHI, which we shall call r_ih.

(10) IH= r_ih . RHI

The new variables are RHI, wr, r_rhiq, sr, IH and r_ih.

RHI is given simply by:

```
>> RHI = FRA_YDRH
```

Let now compute the real wage rate wr. This is done through the following computation.

Dividing FRA_WSSS by FRA_ET gives the individual nominal value, which we divide again by FRA_CPI/100[44] to get the real value[45].

```
>> wr = (FRA_WSS/FRA_ET)/(FRA_CPI/100)
```

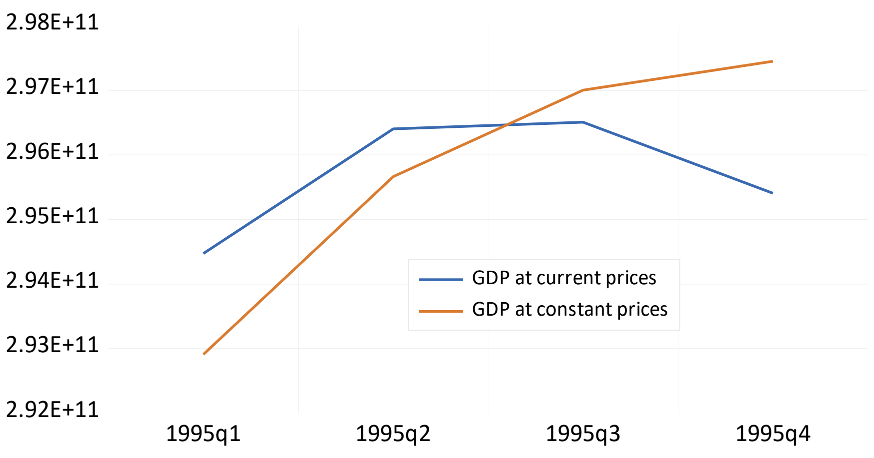
(parenthesizes are added for clarity).

r_rhi will be obtained as the ratio to GDP of household revenue minus wages

```
>> r_rhi = (RHI – wr . LT) /Q
```

---

[44] The OECD deflators are measured as 100 in the year 1995. This means that 1995 the average of values and volume is the same.



[45] Considering the above list of available series, one can observe other optiosn aren possible.

Consumption and housing investment will be obtained directly:

```
>> CO=FRA_CPV

>> IH=FRA_IHV
```

Computing the savings rate and r_ih will use the inversion of the associated equation:

```
sr =(RHI-CO)/RHI
```

or

```
sr=(FRA_YDRH-FRA_CPV)/FRA_YDRH
```

(savings divided by revenue)

```
>> r_ih=IH/RHI
```

Or

```
>>  FRA_IHV/FRA_YDRH
```

(6) Final demand is the sum of its components: consumption, productive investment, housing investment, inventories, and government demand. Total demand includes also intermediate consumption.

```
(11)  FD=IP+CO+IH+gd+CI

(12)  TD = FD + r_ic . Q
```

We need to compute gd as the sum of FRA_IGV and FRA_CGV.

```
>> gd = FRA_IGV+FRA_CGV

>> FD = FRA_TDDV

>> r_ic = FRA_ICV/FRA_GDPV
```

(7) Imports are a share of local demand («domestic demand»). But the less capacities are still available, the more an increase in demand will have to be imported.

This calls for:

```
(13)  UR=Q/CAP

(14)  M=f(FD+IC,UR)
```

We need to compute:

```
>> UR=Q/CAP (its definition)

>> M=FRA_MGSV
```

(8) Exports will essentially depend on World demand. But we shall also suppose that if tensions appear (through UR) local firms will switch some of their output to local demand, and be less dynamic in their search for foreign contracts.

```
(15) X=f(WD, UR)
```

We need:

```
>> X=FRA_XGSV

>> WD=FRA_XGVTR
```

(9) Supply is equal to demand.

The supply-demand equation will for the moment use the following implicit form:

(16) Q + M = FD + X

(all variable values are obtained earlier)

We can now reorder the framework of our model into the following elements:

[1] LE =f(Q)

[2] IP=f(Q)

[3] K= K_{-1} (1-depr) + IP

[4] CAP=f(LE, K_{-1})

[5] IC=r_icq . Q

[6] CI=f(Q)

[7] LT=LE+lg

[8] RHI = wr . LT + r_rhiq . Q

[9] CO = (1-sr) . RHI

[10] IH = r_ih . RHI

[11] FD = CO + IH + IP + CI + gd

[12] TD = FD + r_ic . Q

[13] UR = Q/CAP

[14] M=f(TD, UR)

[15] X=f(wd, UR)

[16] Q + M = FD + X

Endogenous variables

I	Firms investment.
LE	Firms employment.
K	Firms (productive) capital
CAP	Potential output
LT	Total employment.
CI	Change in inventories
IC	Intermediate consumption
IH	Housing investment.
CO	Household consumption.
FD	French final demand
TD	French total demand
M	French Imports.
RHI	Household real income.
UR	Rate of use of capacities
X	French Exports.
Q	Gross Domestic Product

exogenous **variables**

depr	Depreciation rate of capital
gd	State consumption and investment.
lg	Public employment
r_ih	Share of Housing investment in Household revenue.
r_rhiq	Share of GDP transferred to Households, in addition to wages
wd	World demand normally addressed to France.

r_icq	Ratio of intermediate consumption to GDP
wr	Real average wage rate

One observes:

- That we have indeed as many equations as variables to compute.
- That we have separated behavioral equations and identities.
- That accounting identities are completely defined.
- That on the other hand the form of behavioral equations is still vague, although the explanatory elements are known (at least as a first guess).

This distinction is normal. As we have already indicated, identities generally represent a mandatory formal connection, while conforming behavior equations to economic theory is not so restrictive.

- Computing formulas

By considering the formulas we have obtained, we can see that most of the data needed is available directly, so a simple transfer should be enough. We might even have considered using the original names. But as our model will apply only to France, there is no reason to keep the prefix, which helped to identify the French data inside a much larger multi-country file. And one might decide (rightly in our sense) that our names are clearer.

The correspondences are:

```
Q       = FRA_GDPV
CAP     = FRA_GDPVTR
CI      = FRA_ISKV
LT      = FRA_ET
LG      = FRA_EG
FD      = FRA_TDDV
CO      = FRA_CPV
RHI     = FRA_YDRH
I       = FRA_IBV
IH      = FRA_IHV
WD      = FRA_XGVMKT
X       = FRA_XGSV
M       = FRA_MGSV
```

Only eight elements are lacking, seven of them exogenous variables:

gd	=FRA_IGV+FRA_CGV	Government demand
UR	=Q/CAP	Rate of use of capacities (endogenous)
depr	= ((K(t-1) + IP) – K(t))/K(t-1)	Depreciation rate of capital
r_ic	=IC/Q	Ratio of intermediate consumption to GDP
r_ih	=IH/RHI	Share of Housing investment.
r_rhiq	= (RHI – wr . LT) /Q	Non-wage households revenue: share of GDP
sr	=(RHI-CO)/RHI	Savings rate
wr	= (FRA_WSS/FRA_ET)/(FRA_CPI/100)	Real average wage rate

In real cases, this kind of computation will be used often. One must be aware of **one important issue**:

The use of these formulas is logically distinct from the definition of model equations. The only reason we need them is to produce the **historical** values of series not yet available. If the statisticians had made a comprehensive job (and if they knew the requirements of the model) they would have provided the full set, and no computation would have been necessary (just a changes in names).

So these two types of formulas have completely different purposes

- Applying the **computation statements** ensures that all the requested data is available. By associating formulas to missing elements, they allow to produce the full set required for simulation and estimation. If the data was already available in the right format, and the names given to the variables were acceptable, no statement would be necessary. And one can check that in our case, most of the computations are actually direct transfers, which allow to create a model element while retaining the original series.

Actually, one could question the necessity of having a full set of historical values for endogenous variables. These will be computed by the model, which will be simulated on the future anyway. The reasons for producing a full set are the following:

- Estimation will need all the elements in the associated equations.
- Controlling the consistency of identity equations with the data is a prerequisite before any simulation; otherwise, we may start with a flawed model set.
- Checking that the model gives accurate simulations on the past will need all the historical elements.
- Many equations use lagged values. This requires actual values preceding the starting simulation date.

These formulas can include original data, transformed data computed earlier in the program, or simply assumptions. For instance:

- GDP has been drawn directly from the original set.
- The depreciation rate is computed using the sequence of capital values, and investment.
- In the absence of other information, the target for inflation can be set to 2%.

- The **model equations** establish a logical link between elements, which will be used by the model to produce a consistent equilibrium. This means that if the formula for computing variable A contains variable B, variable A is supposed to depend on B, in economic terms.

This is obviously true for estimated equations. For instance, the wage rate can depend on inflation, or exports on world demand. But this is also true for identities:

Household revenue is the sum of its elements. If one grows, revenue changes in the same way (ex-ante, of course). Basically, we suppose that some behaviors apply in the same way to every element of revenue, whatever its source.

If household consumption is estimated, savings are the difference between revenue and consumption.

It is extremely important to understand this issue, at the start of any modeling project.

It is quite possible however that the same formula is present in both sets. For instance, we might not have values for FD, and we believe that CO, I, IH and gd represent the whole set of its components. In this case the formula:

```
FD = CO + IP  + IH + gd
```

will be used both to compute historical values of FD and to define FD in the model.

This introduces an obvious problem: if we make a mistake in the formula, or we use the wrong data, there is no way to detect it.

5.3.6.2 The EViews program

Let us now consider how the above task can be produced. We want to create:

- A workfile for all model elements
- An image of the model, with fully defined identities, and indications as the intended estimated equations,
- The associated data.

5.3.6.2.1 The workfile

- First, we need a work file. In EViews, all tasks are conducted in memory, but they apply to the image of a file which will contain all the elements managed at a given moment.

We can create the file right now (as a memory image) or start from a pre-existing one, in which case the file will be transferred from its device into memory.

Some precautions have to be taken.

- First, only one version of the file must be open in memory. As we state elsewhere, EViews allows the user to open a second version (or even third, and fourth…) of a file already opened. Then changes can be applied only to one of the memory versions, such as series generation and estimations.

This is obviously[46] very dangerous. At the least, one will lose one of the set of changes, as there is no way to transfer elements from an image to the other. Of course, each file can be saved under a different name, but this does not allow merging the changes[47]. At the worst, one will forget the allocation of changes to the files, and one or both will become inconsistent, the best option being to avoid saving any of them, and to start afresh.

This means one should:

- In command mode, check that no file of the same name is opened, and close it if necessary.
- In program mode (the case here) make sure that no file is open at first. This calls for an initial "CLOSE" statement, which will not succeed most of the time[48] but will guarantee that we are in the required situation.
- Second, a new project must start from a clean (or empty) workfile. For an initial file to contain elements is at best confusing, at worst dangerous. For instance, series with the same name as elements in our project can already be present with a different meaning (GDP for a different country?), and available for a larger period. Allowing EViews to estimate equations over the largest period available will introduce in the sample irrelevant values.

A simple way to solve the problem is to delete any existing element, through the statement:

[46] This is only a personal opinion.

[47] Providing this option does not look impossible.

[48] With fortunately no error message.

```
DELETE *
```

which will destroy any pre-existing item, except for the generic C (generic vector of coefficients) and RESID (generic series of residuals) which are created automatically with the work file and cannot be deleted.

There is only one acceptable case for pre-existing elements: if the work file contains some original information, provided to the user by an external source. But even in this case the file has to be saved first, to allow tracing back the steps to the very beginning in which only this original information was present, in its original form.

In any case, in EViews, the possibility to define separate pages (sheets) inside the work file solves the problem. As we have seen earlier, one can just store the original data in one page and start building the transformed data in a blank one, logically linked to the original.

First principle of modeling: always organize your work in such a way that if step n fails, you can always get back to the result of step n-1.

First principle of modeling (alternate version): Always organize your programs in such a way that you can produce again all the elements associated with the present situation.

This (long) discourse leads to the following statements:

```
CLOSE small

WFCREATE(page=model) small Q 1970Q1 2005Q4

DELETE *
```

Applying them guarantees:

- That the file small.wf1 is open in memory with the needed characteristics, for a page called “model”.
- That only one version is open (provided no more than one was open previously, of course, but we shall suppose you are going to follow our suggestions).
- That the page is empty (actually it contains only C and RESID).

5.3.6.2.2 The data

Now that we have a work file, we must fill it with the necessary information.

The original information is represented by the 72 series in the FRA.XLS[49] Excel file. We shall import them using the IMPORT statement. This statement is quite simple (see the User’s manual for detailed options):

[49] EViews allows also to read Excel 2010 .xlsx files (but not to produce them).

```
READ fra.xls 72
```

But beware: even if the Excel file contains dates (in the first column or line) this information is not taken into account. What is used is rather the current sample, defined by the last SMPL statement. Fortunately, in our case, the current sample, defined at workfile creation, is the same as the one in the Excel file. But this will not always be the case: better to state the SMPL before the READ.

```
SMPL 1962Q1 2010Q4

READ fra.xls 72
```

Second principle of modelling: if introducing a (cheap) statement can be useful, even extremely seldom, do it now.

One also has to be careful about the orientation of series: normally they appear as columns, and data starts from cell B2 (second line, second column). Any other case has to be specified, as well as the name of the sheet for a multi-sheet file.

An alternate (and probably better) option

If the follow the above method, all the data will be transferred to the "model" page. This makes things easier in a way, as all information will be immediately available. But

- The separation between original and model data will not be clear.
- The stability of the original data is not guaranteed.
- As the set original series is probably larger, most of the screen will be occupied by elements no longer useful.

Of course, one can separate original and model data by using a prefix for the first type. But it is even better to separate the two sets physically. This can be done through the "link" EViews function.

Instead of loading the original series in the model page, a specific page is created (named for instance "oecd") in which the data is imported.

Then in the model page the model variables are declared as "linked", and a link is defined with the original series in the "OECD" page.

The associated syntax will be presented later.

5.3.6.2.3 The model

- Now, we need to define the model on which we shall work. Producing a model starts with the statement:

```
MODEL modelname
```

Let us call our model _fra_1.

A trick: starting the name of important elements by an underscore allows them to be displayed at the upper corner of the workfile screen, avoiding a tedious scrolling if the number of elements is large. For very important elements (like the model itself) you can even use a double underscore.

The statement

```
MODEL __fra_1
```

defines __fra_1 as the “active” model.

Two cases can be considered:

- The model does not exist. It is created (with no equations yet).
- The model exists. It is opened, with its present equations.

The second option is dangerous in our case, as we want to start from scratch. To make sure of that, the most efficient (and brutal) technique is to delete the model first, which puts us in the first case.

```
DELETE _fra_1

MODEL _fra_1
```

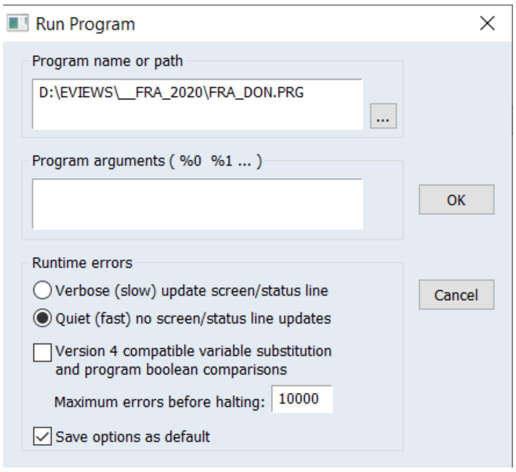
This introduces a slight problem, however. In most cases (including right now) the model does not exist, and the DELETE statement will fail. No problem, as what we wanted is to make sure no model preexisted, and this is indeed the situation we obtain. But EViews will complain, as it could not perform the required task. And if the maximum number of accepted errors is 1 (the default option) the program will stop.

- It is better to specify the “noerr” option, which accepts failure of the statement without error message.

```
DELETE(noerr) _fra_1


MODEL _fra_1
```

- We can change the default number of accepted errors.

Another way to avoid this situation is obviously to set the maximum number of errors to more than 1. This is done by changing the number in the "Maximum errors before halting" box in the "Run program" menu. If you want this option to apply to all subsequent runs, you have to tick in the "Save options as default" box.



Actually, if you have followed the principle above, there is no risk in proceeding in a program which produced error messages, even valid ones. You have saved the elements associated to the initial situation, and even if you forgot to do that, you can always repeat the steps which led to it.

The advantage of this option:

- The program will continue after irrelevant error messages.
- You can produce artificial errors, which can be quite useful as flags[50].
- The messages can be associated to several logically independent errors, which can be corrected simultaneously, leading faster to a correct version.

Now, which number should we specify? In my opinion, depending on the model size, from 1000 to 10000. The number has to be higher than the number of potential errors, as you want to get as close as possible to the end of the program. Of course, you will never make 10000 logical errors. But the count is made on the number of error messages. And in a 2000 equations model, if you have put all the endogenous to zero and you compute their growth rates, this single mistake will generate 2000 messages.

The only drawback is that if your program uses a loop on the number of elements of a group, and this group could not be created, the loop will run indefinitely with the message:

```
Syntax error in "FOR !I=1 TO G.@COUNT"
```

You will have to wait for the maximum number to be reached.

---

[50] The message associated with a real error will locate it between the preceding and following artificial errors.

- Introducing the equations.

Now that we have a blank model, we can introduce the equations one by one. The text of these equations has already been defined; we just need to state the EViews commands.

This is done through the APPEND statement.

The first one will define investment:

```
_fra_1.append IP=f*(Q)
```

Clearly the syntax

- Contains the statement “append”
- Uses the model name on the left, with a dot.
- Adds the text of the equation on the right, with a separating blank.

We must now explain the syntax of our equation.

- At this moment, we expect the model to explain the decision on investment by the evolution of GDP. This seems quite logical, but we have not decided between the possible forms of the theoretical equation, and we have not checked that at least one of these equations is validated by all required econometric tests.
- But at the same time we want EViews to give us as much information as possible on the structure of our model: simultaneities, exogenous parts...
- The best compromise is clearly to produce a model which, although devoid of any estimated equation, nevertheless presents the same causal relationships as the (future) model we consider.

The simplest choice should be, as if we were writing model specifications in a document or on a blackboard, to state:

```
IP=f(Q)
```

Unfortunately, EViews does not accept an equation written in this way. It will consider we are using a function called f, with the argument Q. As this function does not exist, the equation will be rejected.

The trick we propose is to put an asterisk between “f “and the first parenthesis, which gives

```
IP=f*(Q).
```

And state f as a scalar (to avoid confusion with an additional exogenous).

If more than one explanatory variable is used, such as in the productive capacity equation, we would like to write:

```
CAP=f*(LE,K)
```

Again, this is not accepted by EViews, and we can write instead:

```
CAP=f*(LE+K)
```

One just has to state his conventions, and you are welcome to use your own.

However, dropping the f is dangerous, such as in:

```
M=FD+TD
```

This will work too, but the equation can be confused with an actual identity, quite misleading in this case.

The complete set of equation statements is:

```
_fra_1.append LE =f*(Q)

_fra_1.append I=f*(Q)

_fra_1.append K = K(-1)*(1-depr) + I

_fra_1.append CAP=f*(LE+ K)

_fra_1.append IC=r_icq * Q

_fra_1.append CI=f*(Q)


_fra_1.append LT=LE+lg

_fra_1.append RHI = wr * LT + r_rhiq .*Q

_fra_1.append CO = (1-sr) * RHI

_fra_1.append] IH = r_ih * RHI
```

```
_fra_1.append FD = CO + IH + IP + CI + gd

_fra_1.append TD = FD + r_ic * Q

_fra_1.append UR = Q/CAP

_fra_1.append M=f*(TD+UR)

_fra_1.append X=f*(wd+UR)

_fra_1.append Q + M = FD + X
```

They produce a 16 equations model called _fra_1. After running the statements, an item will be created in the workfile, with the name "_fra_1" and the symbol "M" (in blue).

Double-clicking on this item will open a special window, with the list of equations:

- Text (with the icon "TXT" on the left),
- Number (in the order of introduction in the model).
- Economic dependencies: the dependent variable on the left, the explanatory on the right, using actually the syntax we could not apply earlier. Lags are not specified, as we shall see later. So K is presented as depending on K.

Actually, three other display modes are available, using the "View" button:

- Variables: shows the variables (endogenous in blue with "En", exogenous in yellow with "X"). For the endogenous, the number of the equation is given. This allows locating the equation in the model text, which is useful for large models.

The "dependencies" button gives access to a sub-menu, which allows to identify the variables depending on the current one (Up) and which influences it (Down).

For instance, for FD, "Up" will give TD and Q, "Down" will give CO, I, G, and IH

Of course, exogenous will only be allowed the "Up" button.

The "Filter" option will allow selecting variables using a specific "mask". For instance, in a multi-country model the French variables can be identified with FRA_*, provided one has used such a convention.

- Source text: this is basically the text of the model code. We shall see that this changes with estimated equations.
- Block structure: this gives information on the logical structure of the model (detailed later in Chapter 7).

We get:

  - The number of equations.
  - The number of blocks, separated into simultaneous and recursive.
  - The contents of each block.

For the time being, let us only say that a simultaneous block contains interdependent elements. For any couple of elements in the block, a path can lead from the first to the second, and vice-versa. Of course, this property does not depend on the ordering of equations inside the block.

EViews gives also number of feedback variables (this will be explained later too).

On the contrary, a recursive block can be ordered (and EViews can do it) in such a way that each variable depends only (for the present period) on previously defined variables.

This information is useful to improve the understanding of the model, to locate inconsistencies and to correct technical problems.

EViews can detect errors if:

- A variable is defined twice
- The syntax of an equation is wrong (a parenthesis is lacking for instance)

and allow the user to observe errors himself if:

- Normally endogenous elements appear as exogenous: the equation for the variable has been forgotten, or written incorrectly.
- Elements foreign to the model appear: variables have been misspelled.
- A loop appears where there should be none.
- Or (more likely) an expected loop does not appear: for instance a Keynesian model is described as recursive, or a model for two countries trading with each other can be solved as two independent blocks.

All these errors can be detected (and corrected) without calling for the data. This can speed up the building process, especially if the data is not yet produced.

For the production of series, there are two options.

If the original and model series share the same page, one will simply use the "genr" statement, in the sequence.

```
genr  Q=FRA_GDPV

genr  CAP= FRA_GDPVTR

genr  CI =FRA_ISKV

genr  IC=FRA_ICV

genr  LT =FRA_ET

genr  LG =FRA_EG

genr  FD = FRA_TDDV

genr  CO =FRA_CPV

genr  RHI =FRA_YDRH
```

```
genr  IP  = FRA_IBV

genr  IH  = FRA_IHV

genr  WD  = FRA_XGVMKT

genr  GD  = FRA_IGV+FRA_CGV

genr  X =FRA_XGSV

genr  M =FRA_MGSV

genr r_icq= FRA_IC/FRA_Q

genr r_ih=IH/RHI (or FRA_IHV/FRA_YDRH)

genr r_rhiq=(RHI-WR*LT)/Q

genr sr=(RHI-CO)/RHI

genr UR=Q/CAP

genr wr=FRA_WSSS//FRA_LT (FRA_PCP/100)

genr rdep=((K(-1)+IP)-K)/K(-1)
```

If the original series are managed in their own page (a better option in our opinion), one will use:

```
for %1 Q CAP CI IC LT LG FD CO RHI I IH WD X M

link {%1}

next

 Q.linkto oecd\FRA_GDPV

CAP.linkto oecd\FRA_GDPVTR

CI.linkto oecd\FRA_ISKV

IC.linkto oecd\FRA_ICV

LT.linkto oecd\FRA_ET

LG.linkto oecd\FRA_EG

FD.linkto oecd\FRA_TDDV

CO.linkto oecd\FRA_CPV
```

```
RHI. oecd\FRA_YDRH

I.linkto oecd\FRA_IBV

IH.linkto oecd\FRA_IHV

WD.linkto oecd\FRA_XGVMKT

X.linkto oecd\FRA_XGSV

M.linkto oecd\FRA_MGSV

GD .linkto oecd\FRA_IGV+FRA_CGV
```

However, a problem remains for GD, the sum of the two original variables FRA_IGV and FRA_CGV. The LINK function allows to refer to single variables and not functions (as Excel does). Until EViews 8 you had two options.

Creating links to the original elements in the model page.

```
LINK FRA_IGV

LINK FRA_CGV

FRA_IGV.linkto oecd\FRA_IGV

FRA_CGV.linkto oecd\FRA_CGV

genr gd=FRA_IGV+FRA_CGV
```

Or computing a FRA_GDV variable in the original page.

```
genr FRA_GDV=FRA_IGV+FRA_CGV
```

And linking it

```
LINK GD

GD.linkto oecd\FRA_GDV
```

But it is also possible to refer to variables in a different page, as

```
page_name\variable name
```

This means you can use the much simpler method :

```
genr GD= oecd\FRA_IGV+oecd\FRA_CGV
```

Of course, the same method could have been used for single variables.

```
genr Q=oecd\FRA_GDPV
genr CAP=oecd\FRA_GDPVTR
genr CI=oecd\FRA_ISKV
genr IC=oecd\FRA_ICV
genr LT=oecd\FRA_ET
genr LG=oecd\FRA_EG
genr FD=oecd\FRA_TDDV
genr CO=oecd\FRA_CPV
genr RHI. oecd\FRA_YDRH
genr I=oecd\FRA_IBV
genr IH=oecd\FRA_IHV
genr WD=oecd\FRA_XGVMKT
genr X=oecd\FRA_XGSV
genr M=oecd\FRA_MGSV
genr GD =oecd\FRA_IGV+oecd\FRA_CGV
```

it all depends on if you want changes in the original series to be applied automatically, or to control the process through GENR[51]. But if the series is not present (like GD) in the original data, a GENR statement is called for anyway.

Now we have produced a first version of the model, and the associated data. As the behaviors have not been established, we obviously cannot solve it. But we can check two important things:

- The data required for estimation is present.
- The data is consistent with the identities.

These conditions are needed to start estimation, the next stage in the process. The first one is obvious, the second less so. But inconsistencies in identities can come from using a wrong concept for a variable, of computing it wrongly. If this variable is going to be used in estimation, whether as dependent or explanatory, the whole process will be based on wrong elements.

- The time spent in estimation will be lost.
- This time will probably be longer than usual, as it is generally more difficult (sometimes impossible) to find a good fit based on wrong data (fortunately?).
- If a good fit is found, the associated equation can remain in the model for a long time (if not indefinitely), and all the subsequent results will be invalidated. If one is honest, discovering the error later means that a lot of work will have to be done again, including possibly published results.

This test can be conducted through a very simple technique: the **residual check**

5.3.6.3 A first test: checking the residuals in the identities

At this point, asking for a solution of the model cannot be considered. However, some controls can be conducted, which do call for a very specific "simulation". This technique is called "residual check".

This method will compute each formula in the model using the historical values of the variables. This can be done by creating for each equation a formula giving the value of the right-hand side expression (using the GENR statement in EViews). However, there is a much simpler method, provided by EViews.

If we consider a model written as:

$$y_t = f(y_t, y_{t-1}, x_t, \hat{\alpha})$$

with y and x the vectors of endogenous and exogenous variables.

We can perform a very specific "simulation", in which each equation is computed separately using historical values.

Technically this means:

- Breaking down the model into single equation models, as many as there are equations.

[51] Of course, this will also increase the size of the workfile.

- Solving each of these models at the same time but separately, using as explanatory values the historical ones. If we call these historical values y_t^0

It means we shall compute:

$$y_t = f(y_t^0, y_{t-1}^0, x_t, \hat{\alpha}) + e_t$$

This method will control:

- For identities, the consistency between data and formulation.
- For the behavioral equations, the availability of the variables requested by the contemplated estimations. But one gets no numerical information (actually the method we are proposing will give a zero value).

Actually, EViews allows the use of an expression on the left-hand side. This applies also here; the comparison being made between the left and right expressions.

The interest of this method is obvious: if the residual in the equation is not zero, it means that there is *at least* one error in that *particular* equation. Of course, the problem is not solved, but its location is identified. We shall see later that this method is even more efficient for a fully estimated model, and we shall extend our discussion at that time.

It would be illusory, however, to hope to obtain a correct model immediately: some error diagnoses might have been badly interpreted, and corrections badly performed. But even if the error has been corrected:

- There could be several errors in the same equation
- The correcting process can introduce an error in another equation that looked previously exact, but contained actually two balancing errors. Let us elaborate on this case.

Let us consider our example. If we had used for housing investment the value at current prices:

```
genr IH=FRA_IH
```

Then the equation for FD

```
_fra_1.append FD = CO + IH + IP + CI + gd
```

will not hold true, but the one for IH

```
_fra_1.append IH = r_ih . RHI
```

will, as the computation of r_ih as the ratio of IH to RHI will compensate the error by another error.

If we correct the error on IH without correcting r_ih, the IH equation will now appear as wrong, while its actual number of errors has decreased from 2 to 1.

This means achieving a set of all zero residuals might take a little time, and a few iterations, but should converge regularly until all errors have disappeared[52].

5.3.6.3.1 The types of error met

The residual check allows diagnosing the following errors

- Failure to solve
  - syntax error (call to a non-existent function, unbalanced parentheses).
  - series with the right name, but unavailable, either completely (they have not been obtained), or partially (some periods are lacking).
  - bad spelling (call to a non-existent series)
- Non-zero residuals
  - bad spelling (call to the wrong series).
  - errors of logic. This can be more or less serious, as it can come from a purely technical error: forgetting a term for example, or from a conceptual error: stating an unverified theoretical identity.
  - data error: badly entered information, badly computed series, information coming from non-coherent sources, or from different versions of the same bank.
- Non-verified behavioral equations (or with erroneous residual). This issue will be applicable (and addressed) later.

Observing error values can give clues as to their origin:

- If some periods give a correct result:
  - At the base year (where elements at constant and current prices are identical): the price indexes could be mistaken for one another, or values could be mistaken for volumes.
  - If a variable in the formula is null for these periods, it could be responsible.
  - Otherwise, it could come from a typing error (made by the user or the data producer).
  - Or if it appears in the last periods, the provisory elements could be inconsistent.
- Observing the magnitude of the error also can be useful: a residual exceeding the normal economic magnitude (1000% for example) should come from a specification error: bad operator, inversion of

[52] Unless the modeler allows some identities to hold true only approximately.

coefficients, mistaking values for values per capita. A low residual will often come from confusion between two close concepts (the consumption price deflator excluding or including VAT).

- For additive equations, a missing or extra element may be identified by comparing the residual to the actual values of variables. For instance, if the error on final demand for 2010Q1 is 56734 and this is the actual value of housing investment.

- If the sign of the error is constant (and especially if the order of magnitude is similar across periods), the error could come from the absence of an element, a multiplication by a wrong factor, or a missing positive influence.

- If several errors have identical values, they should have the same origin. This is the case when values are mistaken for volume, if the share the same deflator.

- If two variables show roughly identical errors with the opposite sign, this can come the fact that one of them has erroneous values and explains the other.

For instance, if historical values for Q are overestimated, the relative error on UR and Q will be similar with different signs.

```
UR = Q/CAP

Q + M = FD + X
```

5.3.6.3.2 Processing errors

Diagnosing errors in the residual check phase can lead back to different phases of the modelling process:

- Data management: the data obtained from external producers is not consistent, for a full series or for specific observations (this happens!).
- Production of model data: using the wrong original series, using a wrong computation.

Example: using a variable at current prices instead of constant or forgetting an element in a sum.

- Specification of the model (badly written equations).

Example: forgetting the term for Housing investment in the definition of demand. But if the same error was made when computing the series, the two errors will compensate each other.

- Estimation (modified series since estimation, bad coefficients).

Example: an error in the imports equation shows that the explanatory series for domestic demand has been changed since estimation.

Applying this process a number of times will be necessary to produce a coherent model.

5.3.6.3.3 Back to the example

Producing a residual check is quite easy in EViews: one just has to specify the option “d=f” in the SOLVE statement:

```
_fra_1.solve(d=f)
```

Of course, as all equations will be computed separately, all information must be available on the current sample period, including the values of the endogenous variables (which should be explanatory somewhere else). Contrarily to computations and estimations, EViews does not adapt the simulation process to the feasible period (this seems rather logical).

As the model is recursive (super-recursive?) computation gives the result directly, and no element describing the solving method is needed (we shall see them later).

However:

- One should specify the name given to the computed variables.

Every time EViews has to solve a model, the name given to the results will be built from the original name of the variable, with the addition of a suffix (a prefix is also possible but less manageable in our opinion). This avoids destroying the original information, and allows comparing alternate solutions.

The prefix is specified using the statement:

```
modelname.append @all suffix
```

(Remember: **append** adds text to the model, an identity equation is only a special case of text).

In our case, applying the suffix “_C” calls for:

```
_fra_1.append @all _C
```

The equation for FD will give FD_C, which we can compare with the actual values of FD.

Computing the differences between actual and computed values can be done in a loop, using the syntax described later. The elements in the loop can be defined “by hand” but it is more efficient to use the “**makegroup”** statement.

```
_fra_1.makegroup(a,n) groupname @endog
```

```
_fra_1.makegroup(a,n) groupname @exog
```

In our case:

```
_fra_1.makegroup(a,n) g_vendo @endog

_fra_1.makegroup(a,n) g_vexo @exog
```

Two remarks:

- You surely wonder about the reason for the (a,n). This modifies the default options of the "**makegroup"** statement, which would produce a group with the baseline names (in our case with _C added) and leave out the actual names. Stating (a,n):

  - Introduces the actual names (a for actual)
  - Eliminates the baseline ones (n for no baseline)

It would be best to restrict the computations to the identities. The residuals on the "estimated" have no meaning: as the "f" scalar is null, the right-hand side will be computed as zero, and the percentage error as 100% as 100*(value - 0)/value. But being able to compute the whole model proves that estimations can be conducted on that period.

One can create two sub-groups by

```
group g_vbeha CI I LE M X

group g_viden CAP CO FD IC IH K LT Q RHI TD UR
```

Or

```
group g_vbeha CI I LE M X

_fra_1.makegroup(a,n) g_viden @endo

g_viden.drop CI I LE M X
```

This creates first a full group g_viden, then eliminates the estimated from it.

This last technique is clearly inefficient here, but will be much more with a 500 equations model with 50 estimated ones (a more usual situation).

However, both techniques call for a user-defined list, which will have to be updated each time the variable set is modified, something we want to avoid: we propose using a more complex, but automatic one.

A tip: A visual check is made difficult by the relative imprecision of EViews, which often produces small residuals for exact equations. In scientific format, these residuals appear as numbers with high negative exponents, which are hard to identify. One solution is to move to a fixed decimal presentation, by selecting a zone (in the "spreadsheet" view) then using the right mouse button to access "display format" then "fixed decimal".

A simpler solution to observe if there is no error is to display all the residuals as a single graph, and look, not at the series (they should move around in some Brownian motion) but at the scale: both maximum and minimum must be very small.

Another idea is to transfer the residuals to Excel and sort the sheet (unfortunately EViews does not sort a sheet across series on the values at a given period). The non-negligible elements should appear at the top and the bottom according to their sign and the sorting order. Then one can delete the small errors in the middle (less than 0.001%?). As error correction progresses, the number of remaining lines should decrease.

This technique takes more time but allows to identify immediately and fully the faulty elements.

5.3.6.4 Group members view

Managing groups (very useful in modelling) is more flexible and organized. The Preview function (see above) applies to the group as a set.

However, for very simple tasks (like adding and element to a group) using the command window (in which the previous group specification is available) can actually prove faster.

5.3.6.4.1 A trick: generating the groups of identities and behavioral

You certainly have realized by now (and you knew it probably before anyway) that one should avoid as much as possible having to edit the text of modeling programs, each time changes have been made earlier in the process. This represents at best extra work, at worst a source of error. We have just violated this principle, by separating by ourselves the endogenous into behavioral and identity.

This will introduce problems, in particular in large models: the initial process will be tedious and error prone, and one will have to remember to update the list every time the model structure changes.

We propose a simple technique to avoid this, and make the initial separation and its updating automatic. It is based on the presence of the "f" scalar in the behavioral equations.

We just have to:

- Simulate the model with the option "d=f" and f=1, saving the results under a given suffix
- Set f to 2 and update the model (this is necessary for EViews to take into account the change).
- Simulate the model again with f=2 and another suffix.
- Create empty groups of estimated and identity variables.
- Produce a loop over the whole group of endogenous, and test each time if the results of the two simulations are different.

  - If they are, add the variable to the list of estimated elements.

* If not, to the list of identity elements.

Note: when we move to actual estimated formulas, we will introduce a residual appending the suffix "_ec" to the name of the variable. We will use the same technique applying a change to this element.

We can use the following program (for the period 2000 – 2002). We suppose that any percentage error higher than 0.00001 denotes an error.

```
_fra_1.makegroup(am) g_vendo @endog

_fra_1.makegroup(an) g_vexo @exog

group g_varia g_vendo g_vexo

group g_vbeha    'creates an empty group

group g_viden     'creates an empty group

smpl 1986Q1 2004Q4

_fra_1.append assign @all _c

scalar f=0

solve(d=f) _fra_1

scalar f=1

_fra_1.update

_fra_1.append assign @all _d

solve(d=f) _fra_1

for !i=1 to g_vendo.@count

%1=g_vendo.@seriesname(!i)

series pf_{%1}=100*({%1}_d-{%1}_c)/({%1}_c+({%1}_c=0))

if @max(@abs(pf_{%1}))>1e-5 then

g_vbeha.add {%1}

else

g_viden.add {%1}

endif

next
```

This sequence calls for some explanation.

- The loop ("for" to "next") is reproduced for each variable in the list g_vendo. The number of these variables is g_vendo.@count (For EViews, x.@count is an integer scalar containing the number of elements in group x).

  - !i is the rank of the variable in the group g_vendo (from 1 to g_vendo.@count).
  - %1 receives as a character string the contents of g_vendo.@seriesname(!i) , the name of the variable in group g_vendo, with rank !i.
  - The subsequent formulas replace %1 by its string value, and brackets are dropped leaving the characters in the statement.

For regular users of EViews, or people familiar with programming, the above was probably clear. For others, this is the time to give very basic information about EViews programming (even if this is not the purpose of this book).

5.3.7 USING LOOPS AND GROUPS IN EVIEWS

In the programs we are going to present, intensive use is made of two elements: groups and loops.

5.3.7.1 Groups

Groups are named elements which refer to a set of objects (which can be series, series expressions but also other objects**)**, allowing to treat them either as a whole or in sequence.

The statement creating a group is

```
group name-of-the group list-of-elements
```

For instance

```
group g x y z x/y
```

will create a group named g containing the three series x, y and z and the ratio of x to y.

The element must be series of expression, but one can cheat by creating artificial series with the name of the requested element.

One can:

- Group groups
- Add and drop elements from groups:

```
g.add a
```

will add the series a to the group g.

```
g.drop x
```

will drop the series x from the group g.

Two useful elements can be associated with the group:

g.@count is a scalar which contains the number of elements of group g.

g.@seriesname is a character vector which contains the names of the series in group g.

Finally, groups can be created through a mask:

group g_fra fra_* will create a group from all the elements starting with fra_ an underscore.

group g_GDP ???_GDP will create a group from all the GDPs of OECD countries (using three characters as a label).

group g_3 ???_* will create a group from all the elements starting with three characters, then

Groups can be used to display a list of series, as spreadsheet or graph, by double-clicking on its name in the workfile window (where they appear as a blue "G" symbol) or calling for it.

The default display is a spreadsheet format, but one can move to graphs using the "View" button then "graph", or even editing the list of elements by "View" + "group members".

Managing groups has been made more flexible and organized in the last versions. The Preview function (see later) applies to the group as a set.

However, for very simple tasks (like adding and element to a group) using the command window (in which the previous group specification is available) can actually prove faster.

5.3.7.2 Loops

EViews allow two kinds of loops:

- By element
- s (a list or a group)

The syntax is:

```
for %parameter list-of-variables or group-name

block of statements including {%parameter} or %parameter

next
```

The block of statements will be repeated in sequence for each element in the list, which will then replace the parameter.

The presence of brackets around the parameter changes its status. With brackets the associated characters are included in the statements, then the brackets are dropped. Without brackets the parameter is considered as a character string variable.

For instance, with

```
%1="111"
```

The statement

```
genr xxx={%1}
```

will give to the series xxx the value 111,

while

```
xxx=%1
```

will create a character string with the value “111”

The statement

```
genr xxx=%1
```

will be illegal as it tries to transfer a character string to a series.

We get the message:

```
can not assign string expression to numeric variable in "GENR XXX="111""
```

On the other hand, the statement:

```
%2=%1+"333"
```

Will create a “111333” string, while

```
%2={%1}+"333"
```

will be illegal as it mixes strings and values:

```
Scalar assigned to string in "%2=111+"333""
```

- By integer number.

The syntax is:

```
for !name=first-integer to second-integer by third integer

block of statements including {!parameter}

next
```

The block of statements will be repeated in sequence from first-integer to second-integer, incrementing if necessary by third-integer, the value replacing the parameter.

Integers can be negative. If third-integer is omitted, the increment will be 1.

This type of loop can also be applied to a group

```
for !integer=1 to group-name.@count

%1=group-name.@seriesname(!integer)

block of statements including !integer, %1, {%1}

next
```

- group-name.@count is the number of elements in the group group-name.
- %1 receives the contents of group-name.@seriesname(!i) , the name of the variable in group group-name, with rank !integer.

5.3.8 COMPARING WORKFILES : THE WFCOMPARE COMMAND

During the modelling process, you often have to compare two sets of information.

In particular, you might want to:

- Make sure that two sets of data are identical. This applies to the results of a program you are running again, maybe after a long delay.
- Control the evolution of historical values for a model data set, showing for instance which equations will have to be estimated again.
- Summarize the results of a residual check, showing for which equations the right-hand side (using historical values of the explained variable) is different from the right hand side (the result of the computation). By setting a tolerance level slightly higher than zero (for instance 0.0001) one can restrict the display to the errors deemed significant.
- Or you just might want to know which elements of a set are present in another set, for instance which available series are actually used by one model.

This can be done easily, using the **wfcompare** command.

You can compare elements between workfiles and pages inside the same workfile. EViews will display one line per element, in which will be stated its relation, between: unchanged, modified (numerically), added, deleted, replaced (logically, the last case applies for instance to a linked variable have been modified). A filter can be applied.

For series, a tolerance level can be set, under which the series are not considered modified. The display will tell how many periods show a higher difference.

By default, all elements will be displayed, but one can restrict the case (for instance, to all variables present in both pages with a difference higher than the criterion).

Equations and models are not compared but appear in the list.

The syntax of the **wfcompare** command is:

```
wfcompare(tol=criterion,list=comparison_type) list_of_compared_series list_of_reference series
```

For more details you should refer to the EViews Help.

For instance, if you want to compare all French series (starting with "FRA_") between the pages "base" and "updated", for a tolerance level of 0.00001 one will state:

```
wfcompare(tol=1E-5,list=m) updated\fra_* base\fra_*
```

CHAPTER 6 THE ESTIMATION OF EQUATIONS

We now have

- A full description of the framework of the model, in which all the identities are completely specified, and the intents in terms of behaviors are described as clearly as possible.
- A full database containing all the series in the present model, endogenous and exogenous, with their description.

We have also checked that:

- The specification of identities is consistent with the available data.
- The information obtained on the structure of the model (causalities, interdependencies) is consistent with our economic ideas.

Both the list of variables and equations are available as printable documents.

The next stage is obviously to replace each of the tentative behaviors by actual ones, validated both by economic theory and statistical criteria.

6.1 THE PROCESS OF ESTIMATION

What we are proposing is not a book on econometrics, and anyway we will never be as knowledgeable, by far, as the EViews team and collaborators, both in terms of theory and ability to teach it (remember that one of them is Robert Engle…).

This means we will not approach the theoretical aspects of the subject, leaving the reader to the use of the books we propose in our bibliography, or even to the EViews Help manuals, which can be actually used as teaching tools, as they are both comprehensive and very progressive in their approach.

But once the modeler is familiar with the concepts, their application to an actual case[53] is not straightforward at all. This means we think this book can bring a very important contribution: showing how these methods can be used in the process of building our models. The reader will learn how, in very specific cases, first very basic then more operational econometrics can be used (or not used), considering the information he has and the goal he is pursuing.

We shall also show the role econometrics take in the process, not as a single task between data gathering and simulations, but as a recurrent partner in the iterative process of building a working model.

We shall not only give examples working smoothly from the start, but show also how econometrics can be set aside, and how, in some cases, an initial failure can be transformed into success, with some imagination[54].

[53] One in which he is not playing with data, but actually oblided to succeed.

[54] Remember David Hendry's four golden rules of econometrics: 1.Think brilliantly, 2.Be infinitely creative, 3.Be outstandingly lucky, 4.Otherwise, stick to being a theorist.

6.2 SPECIFIC ISSUES

Nevertheless, we feel it will be useful to start with two cases, which are not generally treated by manuals, and can lead to wrong decisions, or wrongly evaluating the results of tests.

We shall use a very practical approach.

6.2.1 THE R2 OR R-SQUARED

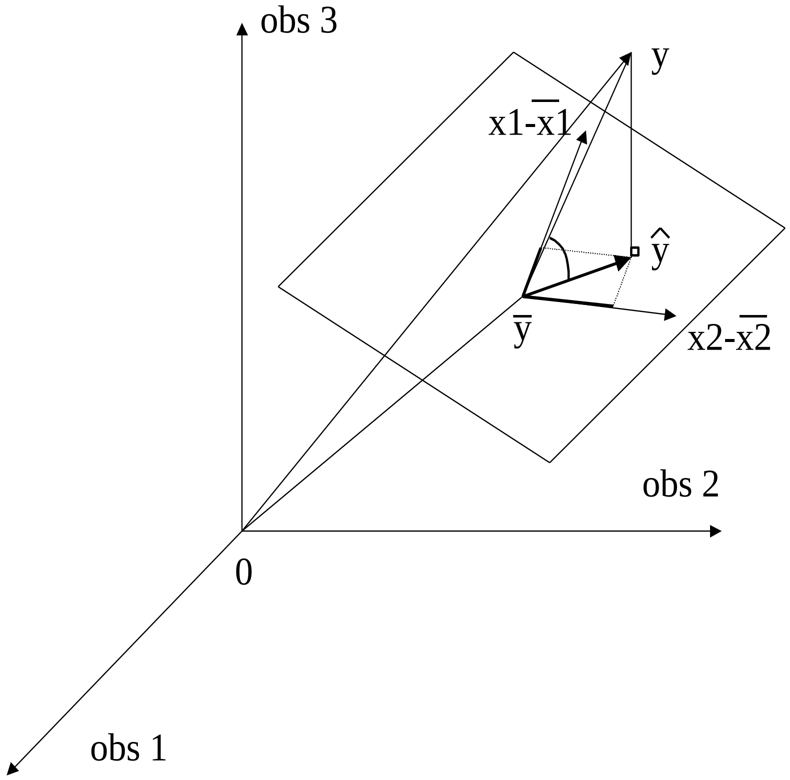
The statistic called "R2" or "R-squared" is the most commonly used to judge the global quality of an estimation. It is defined by the following formula.

$$R^2 = \sum_{t=1}^{T}(\hat{x}_t - \bar{x})^2 \,/\, \sum_{t=1}^{T}(x_t - \bar{x})^2$$

This statistic can therefore be interpreted as the share of the variance of the observed variable x explained by the estimated formula.

A geometrical explanation also can be used: if we consider the space of variables (dimension T = number of observations), the estimation method will consist in minimizing the distance between the explained variable and the space (the plane or hyper plane) generated by the vectors of explanatory series, using combinations of parameter values.

Especially, if the formula is linear relative to estimated parameters and contains a constant term, we can consider the estimation is based on the difference of variables (explained and explanatory) to their means. In this case, minimizing the Euclidian distance will lead (as can be seen on the graph) the vector $(\hat{y}_t - y_t)$ to be orthogonal to the space and therefore to the vector$(\hat{y}_t - \bar{y})$. These two elements represent the non-explained and explained part of$(y_t - \bar{y})$, the variance of which is the sum of their squares. The R^2 can be interpreted as the square of the cosine of the angle between the observed and adjusted series: the closer the R^2 is to 1, the smaller the angle will be and the higher the share of the estimated variable in the explanation of the total variance. The explanation will be perfect if $y-\bar{y}$ belongs to the space, and null if the perpendicular meets the space at the origin.



If the equation presents no constant term, the same reasoning can be applied, but this time the mean is not subtracted. However, the R^2 no longer has the same meaning: instead of variances, direct sum of squares will be used.

We will not go further in the explanation of this test, concentrating instead on its practical properties.

6.2.1.1 Questioning the R-squared

One must be very careful when using the R2 statistic.

- it is favored by trends in variables, independently from economic significance.

The R^2 statistic will be all the higher as the explained variable and at least one of the explanatory variables present a time trend according to the rank of the observation. Thus components of each of these variables on axes of observations will grow in the same or opposite direction (from highly negative to highly positive or the reverse), and give associated vectors very close orientations. In the above graph, the components of variables on the axes will be more or less ordered according to the numbering of the axes themselves. The first observations will be the most negative, then values will grow through zero and reach the most positive ones in the end. The same goes if the ordering follows opposite directions: the estimation will evidence a negative link.

In this case, even in the absence of a true causal relationship between variables, the orientation of the vectors will be similar to a given multiplicative factor, and the R^2 test will seem to validate the formulation. And most time series (like values, quantities or prices), generally present a growing trend, giving this phenomenon a good chance to happen. For example, if we apply the previous equation for French imports:

$$(1) \quad Log(M_t) = a \cdot Log(TD_t) + b + u_t$$

Replacing TD by any steadily growing (or decreasing) variable[55] will give a "good" R^2, better maybe than actual French demand.

Actually, it can be shown that testing for each OECD country the estimation of its imports as a function of the demand of any country, the "true" equation does not come always as the best, although it is never far from it.

- It gives misleading diagnoses when comparing estimations explaining different elements.

This happens in particular when we explain the same concept using a different transformation.

Let us consider our equation 14, as

$$(14) \qquad \Delta Log(M_t) = a \cdot \Delta Log(TD_t) + b + v_t$$

We can see that the time trend has disappeared from both series, and any correlation will come from common deviations around this trend (or rather common changes in the value from one period to another). This is of course a much better proof of a link between the two elements (independently from autocorrelation).

To put the two formulations on equal grounds, they must explain the same element. For this, one can just modify the new equation into:

$$Log(M_t) = Log(M_{t-1}) + a \cdot \Delta Log(TD_t) + b + v_t$$

Compared to the initial formula, this transformation will not change the explanation[56], as obviously the minimization of the sum of squared residuals represents the same process. The only modified statistic will be the R^2, which will increase a lot, as an identical element with a high variance (compared to that of $\Delta Log(M_t)$) has been added on both sides.

[55] Like Australian demand, or the price of a pack of cigarettes in Uzbekistan.

[56] Before estimation EViews will move the lagged term to the left.

The choice between the two formulations should not rely on the R^2 but on the autocorrelation of the residual: if u_t is not correlated one should use (1), if it is one should try (2). But in any case the issues will be solved by error correction models and cointegration, which we shall address later.

The following two formulations are equivalent an indeed give exactly the same results, except for the R-Squared statistic.

Dependent Variable: DLOG(M)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/21/20 Time: 19:25
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments
DLOG(M)=C_M(1)*DLOG(FD)+C_M(2)

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(1)	1.941998	0.092941	20.89494	0.0000
C_M(2)	0.001169	0.001359	0.860131	0.3910

R-squared	0.725731	Mean dependent var	0.015300
Adjusted R-squared	0.724068	S.D. dependent var	0.028994
S.E. of regression	0.015231	Akaike info criterion	-5.519130
Sum squared resid	0.038275	Schwarz criterion	-5.481788
Log likelihood	462.8473	Hannan-Quinn criter.	-5.503974
F-statistic	436.5986	Durbin-Watson stat	1.720498
Prob(F-statistic)	0.000000		

Dependent Variable: LOG(M)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/21/20 Time: 19:23
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments
LOG(M)=LOG(M(-1))+C_M(1)*DLOG(FD)+C_M(2)

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(1)	1.941998	0.092941	20.89494	0.0000
C_M(2)	0.001169	0.001359	0.860131	0.3910

R-squared	0.999457	Mean dependent var	24.31343
Adjusted R-squared	0.999454	S.D. dependent var	0.651772
S.E. of regression	0.015231	Akaike info criterion	-5.519130
Sum squared resid	0.038275	Schwarz criterion	-5.481788
Log likelihood	462.8473	Hannan-Quinn criter.	-5.503974
F-statistic	303832.9	Durbin-Watson stat	1.720498
Prob(F-statistic)	0.000000		

6.2.2 THE CONSTANT TERM

When observing the validity of individual influences, one element plays a very specific role: the constant term.

This element can have two purposes:

- To manage the fact that the equation does not consider elements as such, but the deviations from their means. In ordinary least squares, even if the final result is a linear formulation of the variables and a constant term, the process actually
    - computes the mean
    - substracts it from the variable
    - uses the deviations to estimate a formula with no constant[57]
    - recombines estimated coefficients and means into a constant

This constant is an integral part of the process. It should be included every time at least one of the explanatory elements does not have a zero mean.

- To describe an economic mechanism.

Let us give an example for the first case: if imports have a constant elasticity to demand, we will estimate:

$$\Delta M_t / M_t = a \cdot \Delta TD_t / TD_t$$

Or

$$Log(M_t) = a \cdot Log(TD_t) + b$$

but the estimation process will first use the difference to the average to get "a"

$$Log(M_t) - Log(\bar{M}) = a \cdot (Log(TD_t) - Log(\overline{T}D))$$

or

$$Log(M_t / \bar{M}) = a \cdot Log(TD_t / \overline{T}D)$$

Then the constant

[57] As all elements in the formula have zero mean, the sum of the residuals will also

$$b = Log(\overline{M}) - a \cdot Log(\overline{TD})$$

We can see in particular the consequences of a change in the units (thousands, millions, billions...). The constant term will absorb it, leaving "a" unchanged. In the absence of "b", "a" will get a different value, for no economic reason.

Of course, the more "b" is significant, the more its absence will be damaging to the quality of the estimation (and the more "a" will be affected). But this is no reason to judge "b". We can see this as weighting an object with a balance: the two platters never have the same weight, and even if the damage decreases with the difference, it is always useful to correct it. And in our case there is no cost (actually it makes things cheaper, as the cost of the decision process disappears).

It is not frequent for the constant term to have a theoretical meaning. The majority of such cases come from a formula in growth rates or variations, where the constant term will be associated with a trend.

The only justification for the absence of a constant term is when this theoretical constant is null. In this case, observing a significant value becomes a problem, as it contradicts the theory. We shall give an example soon.

6.3 APPLICATIONS: OUR MODEL

Let us now apply the above principles to our sample model.

In our model, we have to estimate five equations, for which we have already ideas about their logic:

- The change in inventories, employment and investment should depend on GDP
- Exports and imports should depend on the associated demand (world and domestic) and availability of potential supply.

We shall use each of these equations to illustrate a specific aspect of estimation.

- The change in inventories: general elements, homoscedasticity, presence of a constant term.
- Employment: stationarity, error correction models.
- Investment: the necessity to establish a consistent theoretical equation prior to estimation.
- Exports: autoregressive processes, cointegration, long-term stability.
- Imports: going further on cointegration and long-term stability.

Each of our formulations will be based on very simple economic ideas, and we shall select a specification which complies with both econometric tests and economic consistency. They are also chosen in a way which should allow them to merge harmoniously into the model we are building. However, it should be clear that

- Other simple formulations could probably be built on the same sample, with equivalent or maybe better quality. having
- Using another sample (another country for instance) the same economic ideas could lead to different formulations (not only different coefficient values).
- Other economic ideas could be applied, with the same complexity.
- To produce a truly operational model, the present framework would have to be developed in a large way. We will present such developments later.

However the model we are building represents in our sense a simplified but consistent summary of the general class of models of this type. Reading descriptive documents for any operational structural model, one will meet many of the ideas we are going to develop.

6.3.1 THE CHANGE IN INVENTORIES

We shall use this simplest estimation to present the basic features of EViews estimation, and also stress the necessity for homoscedasticity.

Our formulation will suppose simply that firms want to achieve a level of stocks proportional to their production (or GDP). For a particular producer, this should be true both for the goods he produces and for the ones he is going to use for production. For instance, a car manufacturer will allow for a given delay between production and sale (maybe three months, which will lead to an inventory level of 1/4th of annual production). And to be sure of the availability of intermediary goods (like steel, tires, electronic components and fuel for machines in this case) he will buy the necessary quantity (proportional to production) sometime in advance.

We shall suppose that firms have achieved, at the previous period, an inventory level IL representing a number of semesters of production:

$$\boldsymbol{LI_{t-1} = a \cdot Q_{t-1}}$$

And they want to keep this level at the present period:

$$\boldsymbol{LI_t^* = a \cdot Q_t}$$

$$\boldsymbol{LI_t^* = IL_t}$$

Then the change in inventory will represent:

$$\boldsymbol{CI_t = (LI_t - LI_{t-1}) = a \cdot \Delta Q_t}$$

This means that contrary to the general case this equation should not include a constant term. Its presence would call for a trend (and a constant) in the equation in levels, with no economic justification. It would also introduce a problem: adding a constant to an explanation in constant Euros would make the equation non-homogenous.

Even then, the equation faces a problem, concerning the residual: between 1963 and 2004, French GDP has been multiplied by 4. We can suppose the level of inventories too (maybe a little less with economies of scale and improved management techniques).

It is difficult to suppose that the unexplained part of the change in inventories is not affected by this evolution. As the variable grows, the error should grow. But to apply the method (OLS), we need the residual to have a constant standard error. Something must be done.

The simplest idea is to suppose that the error grows at the same rate as GDP, which means that if we measure the change in inventories in proportion to GDP, we should get a concept for which the error remains stable. Of course, we shall have to apply the same change to the right-hand side, which becomes the relative change in GDP.

To avoid causality problems (for a given semester, demand for IC is partly satisfied by Q) we shall use the previous value of Q.

The equation becomes:

$$CI_t/Q_{t-1} = a \cdot \Delta Q_t/Q_{t-1}$$

6.3.1.1 The basic EViews estimation features

As this is our first example, when shall use it to present the basic estimation features.

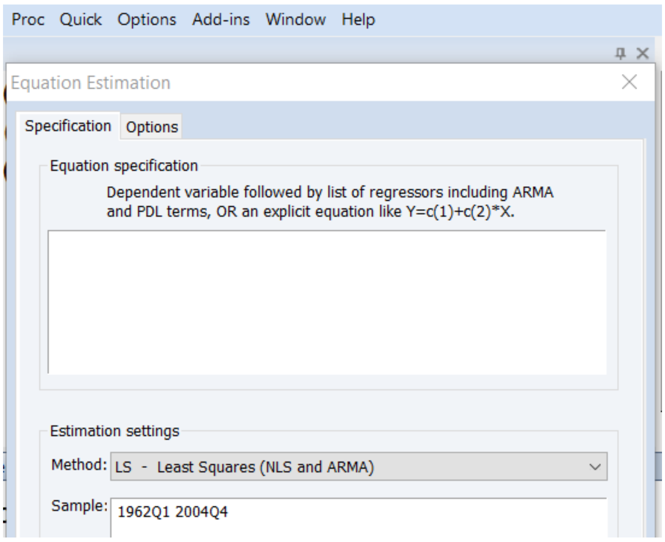
Actually, the technique will be different according to the stage in the estimation process: whether we are exploring several individual formulations, looking for the best option both in statistical and economic terms, or we have already elected the best one, and want to merge it into our model.

We shall start with the first situation.

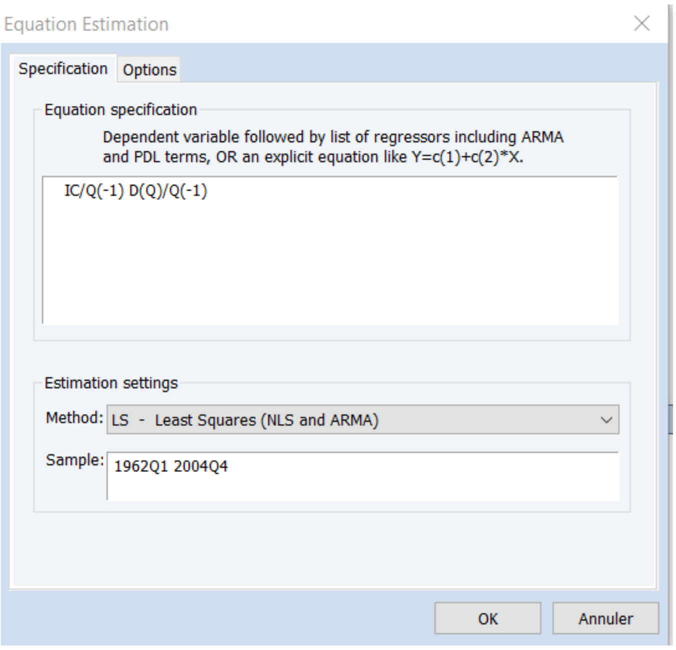
The simplest way to estimate an equation under EViews is through the menus, using in succession:

Quick > Estimate equation

A window appears, in which one has to type the formula.



In the case of ordinary least squares, this can be a list of elements separated by blanks, in our case:



We can also use

```
IC/Q(-1)=c(1)*D(Q)/Q(-1)
```

Of course, the two methods give exactly the same results (in the first case, the “c” vector will also be filled with the estimated coefficient).

The default method will be Least Squares, appropriate in our case. If the equation was not linear in the coefficients, the second presentation would be automatically called for.

One will note that

- A constant term has to be introduced explicitly (as an additional element called “C”).
- EViews allows to specify a sample, which will be applied only to the particular equation (the current sample is not modified). This is quite useful if some periods have to be excluded from the estimation. This will happen for instance if they are deemed not to follow the estimated behavior (like pre – transition data for Central European countries, China or Vietnam), or observations are also provided over the future (OECD’ Economic Perspectives completes the historical data with the results of its forecasts over the next three years).
- On the contrary, one does not have to care about leaving in the sample periods for which estimation is not possible, due to missing elements or the impossibility to compute a term (for instance the logarithm of a negative value). EViews will eliminate by itself the corresponding periods (and tell you about the reduced sample).

In our case we can use the sample:

```
smpl 1962q1 2004q4
```

which means that we consider data from the first quarter of 1962 to the last of 2004.

If the equation is linear in coefficients, EViews recognizes this property, and does not try to iterate on the coefficients, as it knows the values found are the right ones.

Using the "Ok" button gives the following results

Dependent Variable: CI/Q(-1)
Method: Least Squares
Date: 03/21/20 Time: 20:46
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(Q)/Q(-1)	0.183685	0.034776	5.281999	0.0000

R-squared	0.057924	Mean dependent var	0.002150
Adjusted R-squared	0.057924	S.D. dependent var	0.006806
S.E. of regression	0.006606	Akaike info criterion	-7.195844
Sum squared resid	0.007243	Schwarz criterion	-7.177173
Log likelihood	601.8529	Hannan-Quinn criter.	-7.188266
Durbin-Watson stat	0.695764		

We can see that EViews gives the sample used (the relevant periods of our sample). Estimation starts in 1963q2, (with data starting in 1963Q1) and ends in 2004q4.

- We get also the number of periods, and the time and date.
- The other elements are the usual statistics, described earlier. The most important are:
  - The R-squared, the Durbin-Watson test and the Standard Error of regression for global elements.
  - The coefficient, the t-Statistic and the probability for reaching the coefficient value if the true coefficient is null with the estimated standard error.

In our case:

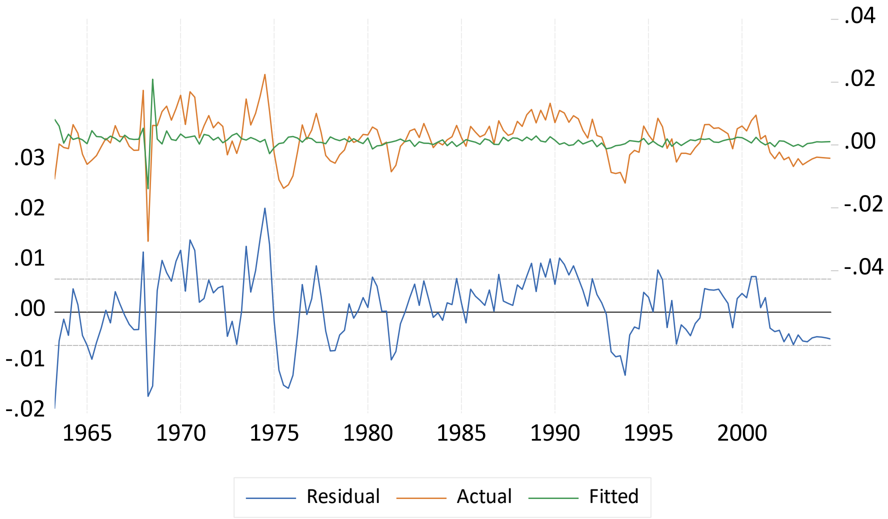
- The R-Squared is very low, even if the extreme variability and the absence of trend of the left-hand element plead in favor of the explanation[58].

However, as with almost all homogenous estimations, a simple interpretation is available, through the standard error: as the explained variable is measured in points of GDP, this average error represents 0.68 points.

[58] If we knew the values for IL, its estimation would get a better R^2 (due to the colinearity of LI and Q). But we would be led to estimate an error correction model on IL, anyway. We have seen the advantage of this formulation, but for the quality to extend to the whole model, all equations must be of this type.

- The coefficient is very significant. The probability of reaching 0.107 for a normal law with mean 0 and standard error .00126 is measured as zero. Of course, it is not null, but it is lower than 0.00005, and probably much so.

- But the Durbin-Watson test takes an unacceptable value, even if the absence of a constant term (and the subsequent non-zero average of residuals) makes its use questionable.
- The graph of residuals is the second important element for diagnosis. It shows the evolution of actual and estimated series (top of the graph, using the right-hand scale) and of the residual (bottom, using the left hand scale, with lines at + and – 1 standard error). This means that inside the band residuals are lower than average, and higher outside it. Of course, it gives only a relative diagnosis.



The graph shows (in our opinion) that the equation provides some explanation, but some periods (1975-1980 in particular) present a large and persistent error

In addition to the display of estimation results and graph of residuals, EViews creates several objects:

- A vector of coefficients, contained in the "C" vector. The zero values or the results from the previous regression are replaced59 .
- A series for the residuals, contained in the "RESID" variable. The "NA" values or the results from the previous regression are replaced60.
- A tentative equation, called "Untitled" for the moment, and containing the developed formula, with "C" as the vector of coefficients, with numbers starting from 1. In our case, the formula is obviously

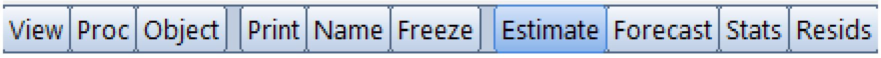
```
IC/Q(-1)=c(1)*D(Q)/Q(-1)
```

[59] But if the present regression contains fewer coefficients than the previous ones, the additional elements are not put to zero.

[60] But this time, residuals from previous equations are given either computed values or « NA ».

Any subsequent estimation will replace this equation by the new “Untitled” version.

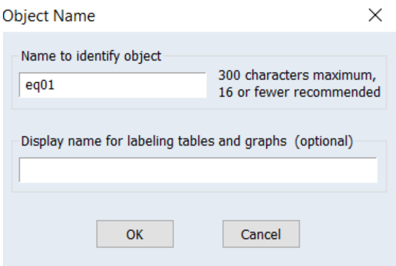
- EViews provides also several options, accessed from the menu, and which can be useful:



- **View** gives three representations of the equation: the original statement, and two formulas including coefficients as parameters (the above “c” type) or as values.
- **“Print”** allows printing the current window: to a printer, to a text file (using characters, which saves space but reduces readability, especially for graphs), or to a graphics RTF file. This last option might call for a monochrome presentation, which is obtained through the « Monochrome » template (the last of the general Graph options).
- **“Name”** allows creating the equation as a named item in the workfile, with an attached comment. It is important to use it immediately after the estimation, as the temporary equation (named “untitled”) will be replaced by the next estimation.

However, inserting an underscore ("_") before the name proposed will place the equations in the first positions of the working window.

EViews proposes as a standard name “EQ” followed by a two-digit number, following the lowest one unused at the moment.



There are two options:

* Give a name representative of the equation (like “EQ_X3U” for the third equation estimating X as influenced by the rate of use).
* Accept the EViews suggestion and rely on the attached comment for the explanation.

Personally we favor the second option:

- It is simpler and more natural to use.
- It allows placing all the equation in the same workfile (and window) location.
- It avoids defining a complex and maybe unclear naming method.
* The comment zone is much wider and can follow any format, including blanks and special characters.

Actually the item saved is more complex than the actual formula. Double-clicking on it shows that it contains the full representation, including the residual (and actually the standard errors of the coefficients, even if they are not displayed).

- **Forecast** produces a series for the estimated **variable** (or the estimated left-hand expression, generally less interesting), and an associated graph with an error band (and a box with the statistics).

6.3.1.2 An alternate technique: using the command window

Instead of using Quick>Estimate, one can work directly through the command window. One just has to add "ls" before the formula.

```
ls IC/Q(-1)=c(1)*D(Q)/Q(-1)
```

This has several advantages:

- By copying and editing the current equation on the next line of the command box, entering changes is made much easier.
- After a session of estimations, the set can be copied into a program file and reused at will. Management of a set of alternate versions is much easier.
- One can control the size of characters. This is quite interesting when working with a team, or making a presentation, as the normal font is generally quite small.
- The only drawback is sample definition: it has to be entered as a command, not as an item in the "estimate" panel.

6.3.1.3 Other possible specifications

Let us go back to our estimated formula. If we are not satisfied with the previous results, we can try alternate options, without changing the economic background.

First, one can observe a very clear outlier for the second quarter of 1968. Economists familiar with French history are aware that this corresponds to the May 1968 "revolution" when student demonstrations turned into general strikes, paralyzing the economy as not only factories closed but also transportation came to a standstill, and available goods could not be delivered to the firms which needed them.

As production was much more reduced than demand, in an unexpected way, satisfying it had to call for inventories, and even a lower production needed intermediate goods (in particular oil and coal) which were not available due to these transportation problems.

The situation came back to normal in the third quarter.

This will lead us to introduce a "dummy" variable, taking the value 1 in the second quarter only. We can observe already the gain from our "time" variable: we can introduce it explicitly, without having to create a specific element which will have to be managed in every simulation, including forecasts.

Dependent Variable: CI/Q(-1)
Method: Least Squares
Date: 03/21/20 Time: 23:00
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(Q)/Q(-1)	0.141900	0.037159	3.818766	0.0002
T=1968.25	-0.019887	0.007058	-2.817550	0.0054

R-squared	0.101169	Mean dependent var	0.002150
Adjusted R-squared	0.095722	S.D. dependent var	0.006806
S.E. of regression	0.006472	Akaike info criterion	-7.230859
Sum squared resid	0.006911	Schwarz criterion	-7.193517
Log likelihood	605.7767	Hannan-Quinn criter.	-7.215703
Durbin-Watson stat	0.614546		

s

The results do improve.

6.3.1.4 Introducing distributed lags

One could assume that the change in inventories does not depend on the present change in GDP, but rather on the sequence of past changes, due both to inertia on firms behavior, and technical difficulties in implementing decisions.

Let us first consider the last five periods leaving the coefficients free. We get;

Dependent Variable: CI/Q(-1)
Method: Least Squares
Date: 03/21/20 Time: 23:22
Sample (adjusted): 1964Q2 2004Q4
Included observations: 163 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@PCH(Q)	-0.015392	0.043290	-0.355545	0.7227
@PCH(Q(-1))	0.058955	0.036838	1.600411	0.1115
@PCH(Q(-2))	0.166449	0.034885	4.771425	0.0000
@PCH(Q(-3))	0.177331	0.037380	4.743971	0.0000
@PCH(Q(-4))	0.073800	0.035119	2.101456	0.0372
T=1968.25	-0.037652	0.006491	-5.801017	0.0000

R-squared	0.386670	Mean dependent var	0.002280
Adjusted R-squared	0.367137	S.D. dependent var	0.006800
S.E. of regression	0.005409	Akaike info criterion	-7.565217
Sum squared resid	0.004594	Schwarz criterion	-7.451337
Log likelihood	622.5652	Hannan-Quinn criter.	-7.518983
Durbin-Watson stat	0.753603		

Not only most of the explanations are not significant, but the value of the first one (maybe the most important) takes the wrong sign.
To make the set of lagged coefficients smoother, we can constraint them to follow a polynomial on the lags.
The syntax for this element is :

PDL(variable, number of lags, degree of the polynomial, conditions).

Here we shall use :

PDL(@pch(Q),4,3,2)

which implies

- A maximum lag of 4.
- A polynomial of degree 3.
- A zero value for the last coefficient (just beyond the last lag).

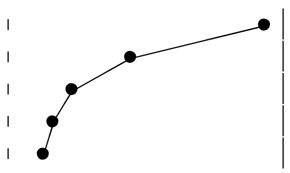
We have also introduced:

- The lagged value of the dependent variable (representing additional inertia).
- A constant term associated with a trend in the level of inventories.

Dependent Variable: CI/Q(-1)
Method: Least Squares
Date: 03/21/20 Time: 23:46
Sample (adjusted): 1964Q2 2004Q4
Included observations: 163 after adjustments

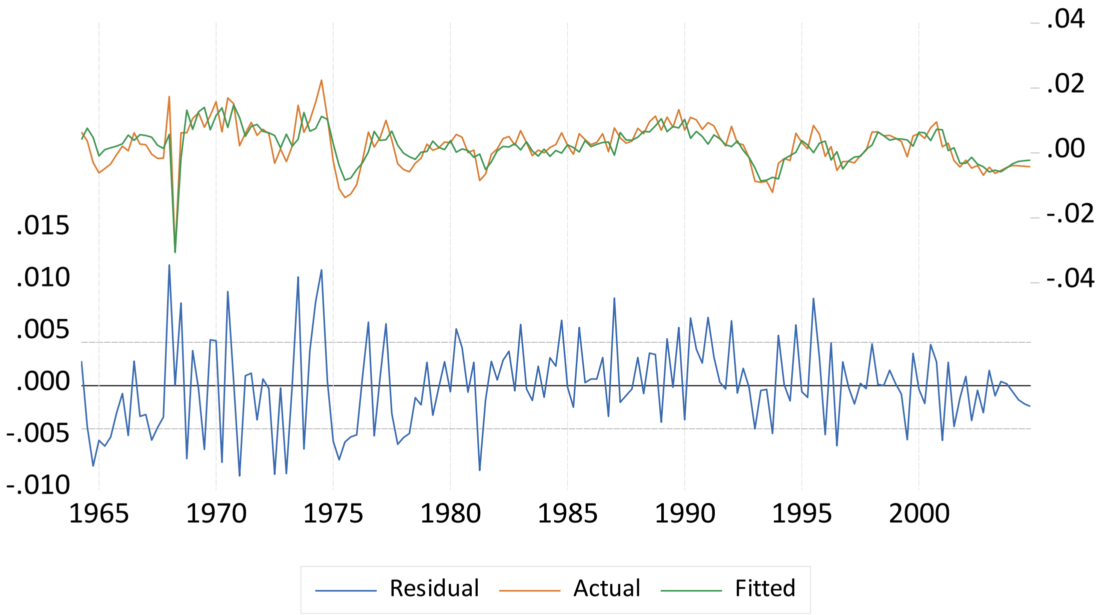
Variable	Coefficient	Std. Error	t-Statistic	Prob.
CI(-1)/Q(-2)	0.612895	0.063632	9.631865	0.0000
T=1968.25	-0.026807	0.005099	-5.256900	0.0000
C	-0.002149	0.000591	-3.636950	0.0004
PDL01	0.054922	0.027450	2.000823	0.0471
PDL02	-0.028519	0.022629	-1.260305	0.2094
PDL03	0.017477	0.009394	1.860457	0.0647

R-squared	0.636943	Mean dependent var	0.002280
Adjusted R-squared	0.625380	S.D. dependent var	0.006800
S.E. of regression	0.004162	Akaike info criterion	-8.089560
Sum squared resid	0.002720	Schwarz criterion	-7.975679
Log likelihood	665.2991	Hannan-Quinn criter.	-8.043325
F-statistic	55.08772	Durbin-Watson stat	2.022998
Prob(F-statistic)	0.000000		



Lag Distribution of...	i	Coefficient	Std. Error	t-Statistic
	0	0.21940	0.03934	5.57655
	1	0.10561	0.02824	3.73939
	2	0.05492	0.02745	2.00082
	3	0.03919	0.02780	1.40989
	4	0.03026	0.02738	1.10538
	Sum of Lags	0.44938	0.07327	6.13293

The results are rather satisfactory, with a nice profile for the reconstructed coefficients, and generally significant explanations. The dummy element provides also a much better explanation.



6.3.1.5 Preparing the equation for the model

Once an equation has been selected for introduction in the model, a different strategy should be used.

If we use the estimated formula, we will face several problems:

- It is not simple to link the equation name with its purpose, which makes the process unclear and forbids to use any automated and systematic process.
- The C vector is used by all equations is only consistent with the last estimated one.
- The residuals cannot be managed simply
- Instead, we propose the following organization, deriving all elements from the name of the dependent variable, though a systematic transformation:
- Naming the equation after the estimated variable.

For instance we can call our equation EQ_CI.

- Using the developed specification, with explicit coefficients.
- Naming the coefficient vector after the estimated variable.

For instance we can call it C_CI. Of course this calls for its creation, with a high enough dimension:

```
coef(10) c_ci
```

(we chose 10 as a round number which we know we shall never reach).

- Introducing an additive explicit residual, named after the estimated variable. The reason is the following.

- It is essential for a model to estimate and simulate the same equation. Of course two versions can be maintained, one being copied into the other after each new estimation. This is:

* Tedious.
* Difficult to manage.
* Error-prone.

It is much better to use a single item. However this faces a problem: one wants access to the residual, in particular for forecasts as we shall see later. And the estimation calls for no residual.

The solution is quite simple: introduce a formal residual, but set it to zero before any estimation.

- Work through a program

This allows:

- Visual control over the specification.
- Easy replication of the estimation (for instance if the data has changed).
- Easy introduction of marginal changes.
- Documentation of the economic context (by introducing comments in the program).

In our case we shall use:

```
coef(10) ec_ci
  genr ec_ci=0
  equation eq_ci ci/q(-1)=c_ci(1)*@pch(q)+c_ci(2)*@pch(q(-1))+ec_ci
  genr ec_ci=resid
```

6.3.2 INVESTMENT: THE NECESSITY TO ESTABLISH A CONSISTENT THEORETICAL EQUATION PRIOR TO ESTIMATION

In this estimation, we shall stress the importance of establishing a sound economic framework before any estimation.

The basic economic idea is quite simple: the purpose of investment is

- To replace capital discarded from wear or obsolescence.
- To allow a higher level production, facing an increase of demand.

Without proceeding further on theory, many formulations can be considered. For instance, investment could have a constant elasticity to GDP, maybe with an error correction term including capital...

In our sense, trying for the best estimation without considering the economics behind the formula, and especially its consequences for model properties, is rather irresponsible. For instance, using the logarithm of investment is quite dangerous. Its value can change in very high proportions, and if we go back to the microeconomic foundation of this behavior, its value could very well be negative, as some firms are led to disinvest from time to time, by selling more capital than they buy.

For instance, the following equation seems to work quite well:

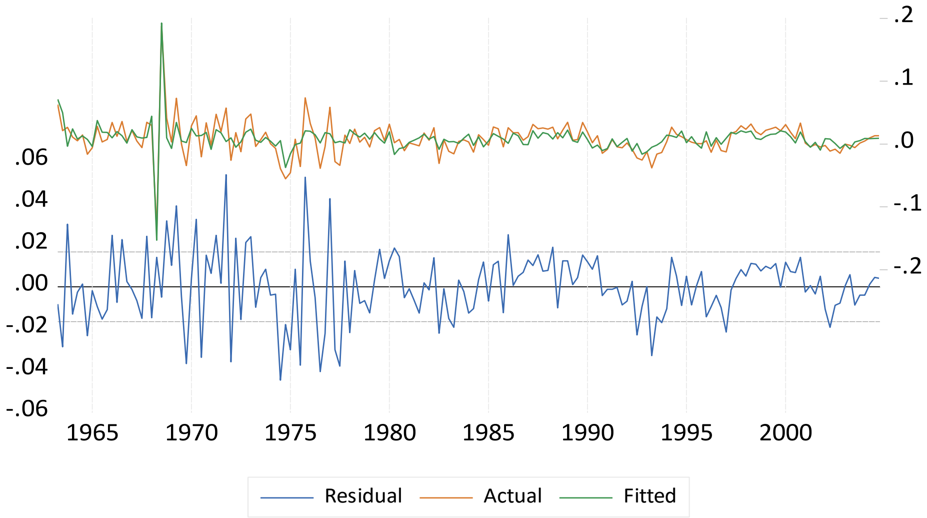
$$\Delta Log(I_t) = a \cdot \Delta Log(Q_t) + b \cdot Log(I_{t-1}/Q_{t-1}) + c \cdot t + d$$

The results are:

Dependent Variable: DLOG(I)
Method: Least Squares
Date: 03/22/20 Time: 11:49
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(Q)	1.835512	0.106843	17.17945	0.0000
LOG(I(-1)/Q(-1))	-0.046829	0.020085	-2.331615	0.0209
C	-0.931254	0.301928	-3.084362	0.0024
T	0.000415	0.000138	3.012286	0.0030

R-squared	0.651895	Mean dependent var	0.008590
Adjusted R-squared	0.645488	S.D. dependent var	0.028010
S.E. of regression	0.016677	Akaike info criterion	-5.325856
Sum squared resid	0.045336	Schwarz criterion	-5.251174
Log likelihood	448.7090	Hannan-Quinn criter.	-5.295544
F-statistic	101.7498	Durbin-Watson stat	2.152872
Prob(F-statistic)	0.000000		

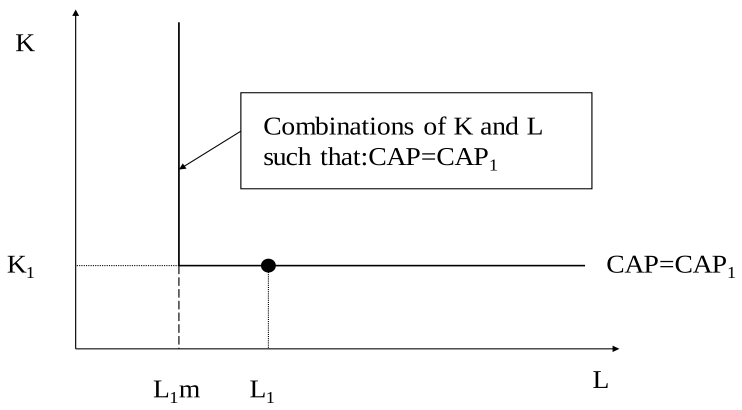


Everything seems to go well: the statistics are quite good (even the Durbin-Watson test), the signs are right, the graph shows a really strong fit. However, when we merge the equation into a model, its simulation properties will be affected by the base solution: even a very high increase in GDP will have a low impact of the absolute level of investment if it this level was very low in the previous period. And investment can show huge variations as it represents a change in a stable element, capital.

One can guess that although linking investment (a change in capital) to the change in production seems a natural idea, jumping to the above formulation was moving a little too fast. In general, one should be naturally reticent in taking the logarithm of a growth rate, itself a derivative.

The right starting approach is to clarify the economic process through a full logical formalization.

Let us suppose that production follows a "complementary factors" function, which means that to reach a given level of productive capacity, fixed levels of capital and employment are required, and a reduction in one factor cannot be compensated by an increase in the other. This means obviously that the less costly process (optimal) is the one which respects exactly these conditions.



With "pk" productivity of capital, and "pl" productivity of labor, we get :

$$\boldsymbol{CAP_t = min(pk_t \cdot K_{t-1}, pl_i \cdot L_t)}$$

(the "t-1" means that we shall use the level of capital reached at the end of the previous period).

Actually, for a given level of employment, there is always some short-term leverage on production, at least at the macroeconomic level. Temporarily increasing labor productivity by 1% can be easily achieved through extra hours, less vacations, cancelled training courses...

This means capital will be the only limiting factor in the short-term.

The capacity equation can be simplified into:

$$\boldsymbol{CAP_t = pk_i \cdot K_{t-1}}$$

Now let us define the rate of use of capacities:

$$\boldsymbol{UR_t = Q_t / CAP_t}$$

$$\boldsymbol{CAP_t = pk_t \cdot K_{t-1} = Q_t / UR_t}$$

Now let us suppose firms actually want to reach a constant target utilization rate UR*, and expect a production levelQ^a_{t+1}. Then by definition:

$$\boldsymbol{K^*_t = CAP^*_t / pk^*_t = Q^a_{t+1} / UR^* / pk_{t+1}}$$

$$\boldsymbol{K_t = CAP_t / pk_t = Q_t / UR_t / pk_t}$$

And defining tx(z) as the growth rate of z:

$$tx^*(K_t) \approx tx^*(CAP_t) - tx(pk_t) \approx tx^a(Q_t) - tx^*(UR_t) - tx(pk_t)$$

This means that the target growth rate of capital can be decomposed as the sum of three terms, one with a positive influence:

- The expected growth rate of production

and two negative ones:

- The target growth rate of the rate of use: if the firms feel their capacities are 1% too high for the present level of production, they can reach the target by decreasing capital by 1% even if production is not expected to change.
- The growth rate of capital productivity: if it increases by 1%, 1% less capital will be needed.

But the element we need is investment. To get it we shall use the definition.

$$K_t = K_{t-1} \cdot (1 - dr_t) + I_t$$

which can be written as

$$tx(K_t) = -dr_t + I_t/K_{t-1}$$

This gives finally:

$$I_t^*/K_{t-1} = tx(K_t) = dr_t + tx^*(K_t) = dr_t \ + tx^a(Q_t) - tx^*(UR_t) - tx(pk_t)$$

In other words:

- If firms expect production to grow by 2.5%, capacities should adapt to that growth
- But if they feel their capacities are under-used by 1%, their desired capacity will only increase by 1.5%.
- If capital productivity is going to increase by 0.5%, they will need 0.5% less capital.
- But once capital growth has been defined, they also have to compensate for depreciation (5% is a reasonable value).

In summary, the accumulation rate (the ratio of investment to the previous level of capital) would be:

2.5 - 1 - 0.5 + 5 = 6%.

If we suppose

- That the depreciation rate is constant, as well as the rate of growth of capital productivity,
- That production growth expectations are based on an average of the previous rates,

And we consider as the rate of use the ratio of actual GDP to a value obtained under normal utilization of factors, which leads to a unitary target, we get the simplified formula:

$$I_t^*/K_{t-1} = a + \sum_{i=0}^{n} \alpha_i \cdot tx^a(Q_{t-i}) - tx^*(UR_{t+1})$$

With

$$\sum_{i=0}^{n} \alpha_i = 1$$

finally, we can suppose, as we shall do also for employment, that the desired growth of capital is only partially reached in practice, either because firms react cautiously to fluctuations of demand, or because they are constrained by investment programs covering more than one period, from the decision to the actual installation of investment goods.[61]

And we shall leave free the coefficients:

$$I_t^*/K_{t-1} = b \cdot I_{t-1}/K_{t-2} + (1-b) \cdot (a + c \cdot \sum_{i=0}^{n} \alpha_i \cdot tx^a(Q_{t-i}) - d \cdot tx^*(UR_t))$$

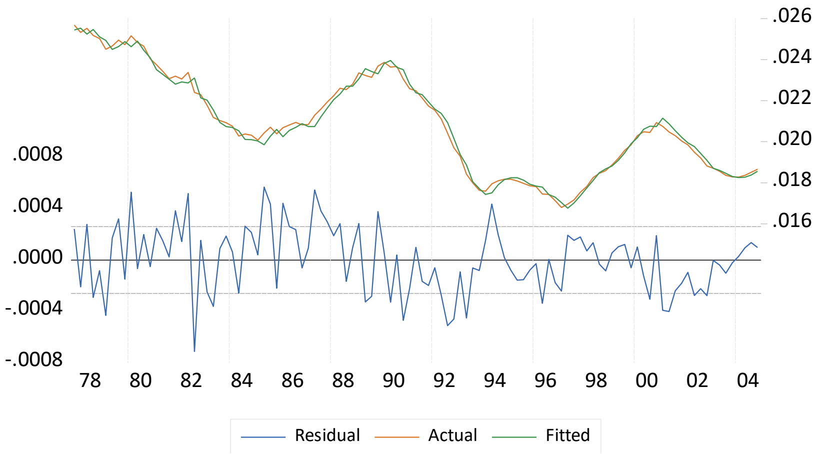
The results are rather satisfactory, with the right sign and acceptable statistics for all explanatory elements. This was not obvious, as their strong correlation (both use Q in the numerator) could have made it difficult for the estimation process to separate their role.

Dependent Variable: I/K(-1)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/22/20 Time: 13:01
Sample (adjusted): 1977Q4 2004Q4
Included observations: 109 after adjustments
I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*(UR_STAR-UR)/UR+C_I(3)*.125*Q/Q(-8)
+C_I(4)

	Coefficient	Std. Error	t-Statistic	Prob.
C_I(1)	0.941668	0.011760	80.07071	0.0000
C_I(2)	-0.002311	0.000863	-2.679352	0.0086
C_I(3)	0.045795	0.013397	3.418178	0.0009
C_I(4)	-0.004818	0.001653	-2.915711	0.0043

R-squared	0.988698	Mean dependent var	0.020830
Adjusted R-squared	0.988375	S.D. dependent var	0.002417
S.E. of regression	0.000261	Akaike info criterion	-13.63100
Sum squared resid	7.13E-06	Schwarz criterion	-13.53224
Log likelihood	746.8897	Hannan-Quinn criter.	-13.59095
F-statistic	3061.832	Durbin-Watson stat	1.649042
Prob(F-statistic)	0.000000		

[61] In this model, we suppose there is no delay between the acquisition of investment (with impact on demand and the supply-demand equilibrium) and the participation of this investment to the production process.



The graph of residuals shows that the quality of the explanation grows with time, and is especially good for the last periods. This is rather important for simulations over the future, and one can wonder what we would have done if the sample had been reversed, and the initial residuals had applied to the last periods.

We will deal with this problem of growing errors on recent periods when we address forecasts.

The equation we have built is not only satisfactory by itself, but we can expect it to provide the model with adequate properties. In particular, the long-term elasticity of capital to production is now unitary by construction. Starting from a base simulation, a 1% permanent shock on Q will leave the long run value of UR unchanged[62]. This gives the same relative variations to production, capacity and (with a constant capital productivity) capital.

The coefficients “a” and “b” determine only the dynamics of the convergence to this target.

Actually we have estimated a kind of error-correction equation, in which the error is the gap between actual and target capacity (the rate of use).

We hope to have made clear that to produce a consistent formulation, in particular in a modelling context, one must start by establishing a sound economic background.

6.3.3 EMPLOYMENT: STATIONARITY, ERROR CORRECTION MODELS, BREAKPOINT TEST.

6.3.3.1 The economic framework

Of course, the employment equation should follow also a complementary factors framework.

In the previous paragraph, we have shown that in this framework the element determining capacity is the sole capital, while firms could ask from workers a temporary increase in productivity, high enough to ensure the needed level of production[63]. Adapting employment to the level required to obtain a “normal’ productivity target will be done by steps.

[62] As the left-hand side represents the (fixed) long-term growth rate of capital.

[63] This is true in our macroeconomic framework, in which the changes in production are limited, and part of growth is compensated by increases in structural productivity (due for instance to more capital-intensive processes). At the firm

This means estimating employment will allow us to apply elements on error correction models in a very simple framework.

We shall suppose that firms:

- Know the level of production they have to achieve.
- Know also the level of production which should be achieved by each worker under normal circumstances (in other term his normal productivity).

From these two elements they can determine the normal number of workers they need.

But they do not adapt the actual employment level to this target, and this for:

- Technical reasons: between the conclusion that more employees are needed and the actual hiring[64], firms have to decide on the type of jobs called for, set up their demands, conduct interviews, negotiate wages, establish contracts, get authorizations if they are foreign citizens, maybe ask prospective workers to train... Of course this delay depends heavily on the type of job. And this goes also for laying out workers.
- Behavioral reasons: if facing a hike in production, firms adapt immediately their employment level to a higher target, they might be faced later with over employment if the hike is only temporary. The workers they have trained, maybe at a high cost, have no usefulness at the time they become potentially efficient. And laying them out will call generally for compensations....

6.3.3.2 The formulas: stationarity and error correction

We should realize that we are facing an error correction framework, which we can materialize as:

"Normal" labor productivity does not depend on economic conditions. It might follow a constant trend over the period, such as:

$$Log(pl_t) = a + b \cdot t$$

Firms use this target to define "normal" employment:

$$LE_t^* = Q_t / pl_t^*$$

They adapt actual employment to this target with some inertia:

---

level, employment can produce bottlenecks. This will be the case if a sudden fashion appears for particular goods requiring specialized craftspeople, even if the tools and machines are available for buying.

[64] But not the start of actual work: what we measure is the number of workers employed, even if they are still training for instance.

$$\Delta Log(L_t) = \alpha \cdot \Delta Log(L_t^*) + \beta \cdot Log(L_{t-1}^*/L_{t-1}) + \gamma + \varepsilon_t$$

We recognize here the error correction framework presented earlier, which requires:

$Log(L_t^*/L_t)$ to be stationary.

But α does not have to be unitary. However, if we follow the above reasoning, its value should be between 0 and 1, and probably significantly far from each of these bounds.

To estimate this system we face an obvious problem: pl* is not an actual series (LE* either, but if we know one we know the other).

But if we call "pl" the actual level of productivity (Q/LE) we can observe that:

$$Log(L_t^*/L_t) = Log((Q_t/pl_t^*)/(Q_t/pl_t)) = -Log(pl_t^*/pl_t)$$

The stationarity of $Log(L_t^*/L_t)$ is equivalent to that of $Log(pl_t^*/pl_t)$

Now it should be obvious that if pl* and pl have a trend, it must be the same, actually the trend defining completely pl*. If not, they will diverge over the long run, and we will face infinite under or over employment. So target productivity can be identified using the trend in the actual value, if it exists.

This means we can test the stationarity of the ratio as the stationarity of actual productivity around a trend, a test provided directly by EViews.

We can expect a framework in which actual productivity fluctuates around a regularly growing target, with cycles which we do not expect to be too long, but can last for several periods[65].

6.3.3.3 The first estimations

First, we compute actual labor productivity

```
genr PROD = Q / LE
```

[65] Which will create (acceptable) autocorrelation in the difference to the trend.

and regress it on time:

```
ls log(PROD) c t
```

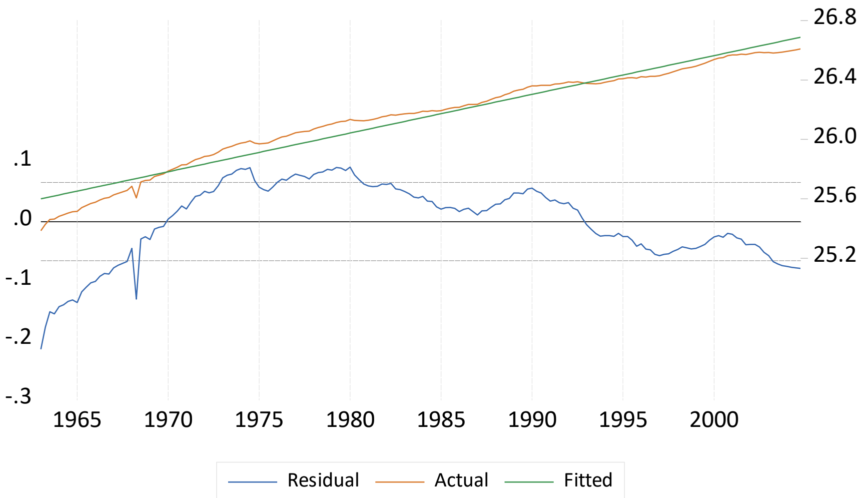
to get the structural productivity trend.

Results are quite bad. Of course productivity shows a significant growth, but the standard error is quite high (more than 5 %). More important, the graph of residuals and the auto-correlation test show that we are not meeting the condition we have set: that observed productivity fluctuates around a trend, with potential but not unreasonably long cycles.

Dependent Variable: LOG(Q)
Method: Least Squares
Date: 03/22/20 Time: 13:16
Sample (adjusted): 1963Q1 2004Q4
Included observations: 168 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-25.54456	0.829672	-30.78877	0.0000
T	0.026054	0.000418	62.30088	0.0000

R-squared	0.958986	Mean dependent var	26.14374
Adjusted R-squared	0.958739	S.D. dependent var	0.323533
S.E. of regression	0.065719	Akaike info criterion	-2.595031
Sum squared resid	0.716947	Schwarz criterion	-2.557841
Log likelihood	219.9826	Hannan-Quinn criter.	-2.579938
F-statistic	3881.400	Durbin-Watson stat	0.036326
Prob(F-statistic)	0.000000		



The problem apparently lies in the fact that the average growth rate is consistently higher in the first part of the period, and lower later. Seen individually, each sub-period might seem to meet the above condition.

From the graph, we clearly need one, and probably two breaks. One will observe that the first period follows the first oil shock, and the beginning of a lasting world economic slowdown. The reason for the second break is less clear (some countries like the US and Scandinavia show a break in the opposite direction).

For choosing the most appropriate dates, we can use two methods:

- A visual one: 1973 and 1990 could be chosen, possibly plus or minus 1 year.
- A statistical one: the most appropriate test is the Chow breakpoint test, which diagnoses if the introduction of one or more breaks improve the explanation. To make our choice automatic, we shall consider three intervals, and apply the test to all reasonably possible combinations of dates from those intervals. As we could expect, all the tests conclude to a break. But we shall select the combination associated to the lowest probability (of no break), which means the highest likelihood ratio[66]. Of course, this criterion works only because the sample and the number of breaks remain the same.

The best result corresponds actually to 1972q3 and 1992q4, as shown in this table of log-likelihood ratios.

The equation for structural productivity is

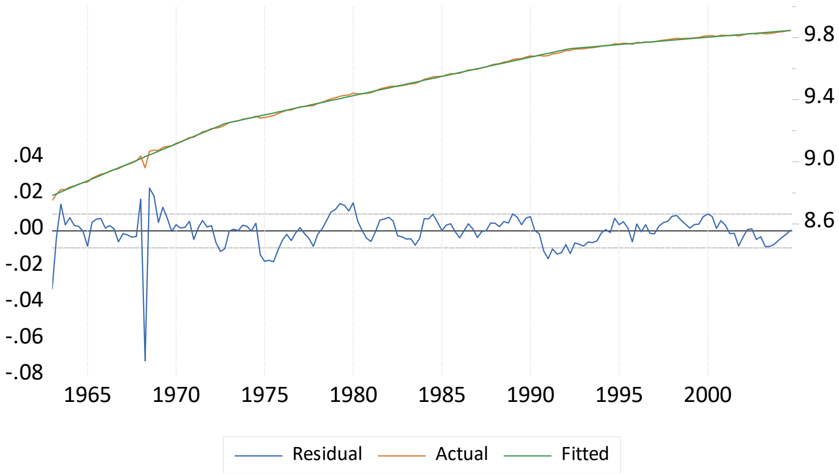
log(prle)=c_prle(1)+c_prle(2)*(t-2004)+c_prle(3)*(t-1972 .50)*(t<1972 .50)+c_prle(4)*(t-1992.75)*(t<1992.75)

Dependent Variable: LOG(PRLE)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/22/20 Time: 15:32
Sample (adjusted): 1963Q1 2004Q4
Included observations: 168 after adjustments
LOG(PRLE)=C_PRLE(1)+C_PRLE(2)*(T-2004)+C_PRLE(3)*(T-T1)
(T<T1)+C_PRLE(4)(T-T2)*(T<T2)

	Coefficient	Std. Error	t-Statistic	Prob.
C_PRLE(1)	9.839846	0.002212	4447.971	0.0000
C_PRLE(2)	0.009382	0.000262	35.76099	0.0000
C_PRLE(3)	0.022782	0.000457	49.89498	0.0000
C_PRLE(4)	0.015405	0.000361	42.64593	0.0000

R-squared	0.999093	Mean dependent var	9.466339
Adjusted R-squared	0.999077	S.D. dependent var	0.307411
S.E. of regression	0.009341	Akaike info criterion	-6.485251
Sum squared resid	0.014310	Schwarz criterion	-6.410871
Log likelihood	548.7611	Hannan-Quinn criter.	-6.455064
F-statistic	60233.59	Durbin-Watson stat	1.553017
Prob(F-statistic)	0.000000		

[66] the highest F gives the same conclusion



One will note:

- That we have introduced no residual, contrary to our usual practice.
- That we have introduced reversed trends, which stop after a while instead of starting inside the period.

  - Target productivity is not a behavior.

The first element is quite logical: what we are estimating for the model is not actual productivity (this is given in the model by an identity, dividing actual GDP by employment). We are looking for the **exact** formula for target productivity, prone to error only because we have not enough information to produce the true value. If the sample grew, or the periodicity shortened, the precision would improve constantly. The residual might not decrease, but it does not represent an error, rather the gap between the actual and "normal" values of labor productivity. Whereas, in a normal behavioral equation, the residual corresponds to an error on the variable, and cannot be decreased indefinitely, as the identification of the role of explanatory elements becomes less reliable with their number.

  - Partial trends should apply to past periods.

The reason here is purely technical. Our model is designed to be used on the future. So it is essential to make the forecasting process as easy as possible.

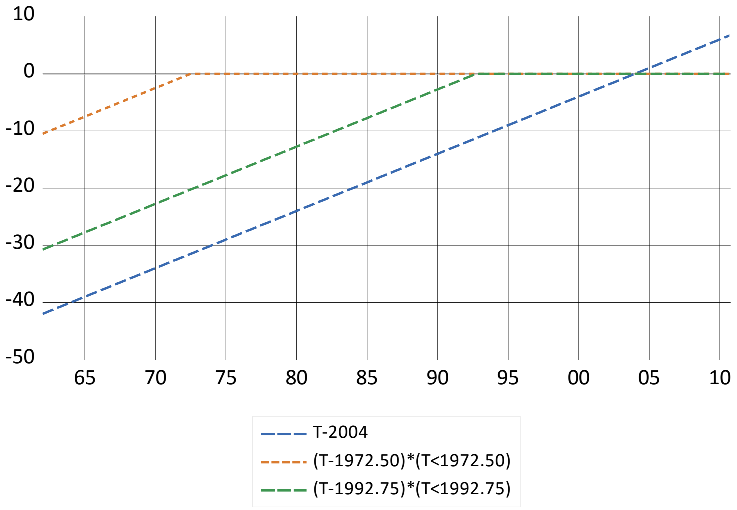
If the partial trends are still active in the future, we shall have to manage them simultaneously. We can expect that we want to control the global trend of labor productivity, if only to make it consistent with our long-term evolutions of GDP (which should follow world growth) and employment (which should follow population trends). Obviously, controlling a single global trend is easier than a combination of three trends.

Also, the last trend is the most important for interpretation of model properties, and it is better to make it the easiest to observe.

On the other hand, our technique has no bad points, once it has been understood.

Finally, the reason for breaking the trend in 2004 is also associated with handling of its future values. If the global coefficient is changed, this will be the period for a new break, and this is the best period to introduce it.

We can summarize the above in the following graph.



We can see that in the beginning three trends apply, then two, then after 2004 only the global (blue) trend is maintained.

The results look quite good, both in the validation of coefficients and the graphs. We are presenting the program version, which will be introduced in the model (as a identity)[67].

However, we observe a very high residual in the second quarter of 1968. As we are estimating a trend, this is not the place for considering a one-period outlier. We will come back to this problem when we estimate employment itself.

Now we must test the stationarity of the residual. We shall use the Dickey Fuller test (or Phillips – Perron).

First we need to generate from the current RESID a variable containing the residual (the test is going to compute its own RESID, so it is not possible to test on a variable with this name).

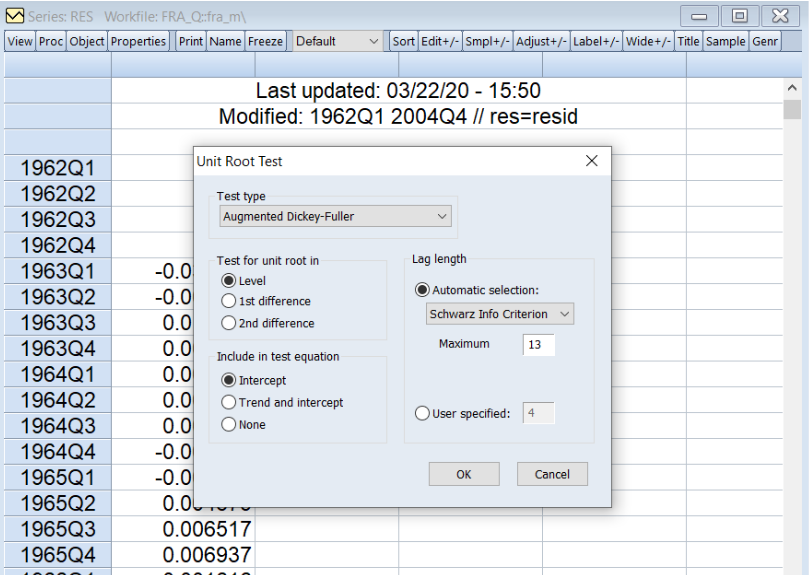
- In program form, the test is conducted by:

```
genr res_prle=resid

uroot(1,p) res_prle

uroot(h,p) res_prle
```

- Using menus, one has to

  - Display the variable
  - Select **View>Unit root test**
  - Choose the method.

[67] This is not absolutely needed, as a variable depending only on time can be considered exogenous and computed outside the model. But we want to be able to change the assumption in forecasts, and this is the easiest way.



Null Hypothesis: RES has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=13)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.029197	0.0000
Test critical values:	1% level	-3.469933	
	5% level	-2.878829	
	10% level	-2.576067	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RES)
Method: Least Squares
Date: 03/22/20 Time: 15:55
Sample (adjusted): 1963Q3 2004Q4
Included observations: 166 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RES(-1)	-0.675886	0.096154	-7.029197	0.0000
D(RES(-1))	-0.163434	0.074306	-2.199455	0.0293
C	0.000183	0.000677	0.270373	0.7872

R-squared	0.420548	Mean dependent var	2.58E-05
Adjusted R-squared	0.413438	S.D. dependent var	0.011383
S.E. of regression	0.008718	Akaike info criterion	-6.628887
Sum squared resid	0.012389	Schwarz criterion	-6.572646
Log likelihood	553.1976	Hannan-Quinn criter.	-6.606059
F-statistic	59.15010	Durbin-Watson stat	2.024797
Prob(F-statistic)	0.000000		

The test concludes very strongly the stationarity of the residual.

The values of target productivity and desired employment are given by:

```
genr log(prle_t)= c_prle(1)+c_prle(2)*t+c_prle(3)*(t-1972.50)*(t<1972.50)+c_prle(4)*(t-1992.75)*(t<1992.75)

genr led=q/prle_t
```

Now, as to the estimation of employment itself, LE will be estimated (using here the developed form) by:

```
equation eq_le.ls dlog(LE)=c_le(1)*dlog(LED)+c_le(2)*log(LED(-1)/LE(-1))+c_le(3)
```

where LED is equal to Q/prle_t, the trend obtained in the previous equation.

It is now time to consider the 1968 residual. With such a high value, there should be some economic explanation.

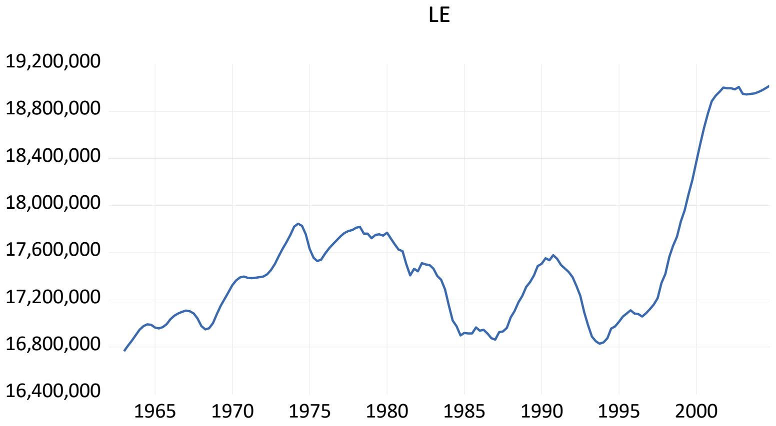
Indeed, the behavior of firms did not follow normal lines. They believed (rightly) that the decrease in production they were facing was quite temporary. For them, laying out the corresponding number of workers was not reasonable, as it would cost severance payments, and when things went back to normal there was no reason they would find as efficient and firm-knowledgeable workers as before.
This means labor productivity decreased, then increased to get back to normal.

Dependent Variable: DLOG(LE)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/22/20 Time: 16:35
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments
DLOG(LE)=C_LE(1)*DLOG(LED)+C_LE(2)*LOG(LED(-1)/LE(-1))
+C_LE(3)+C_LE(4)*((T=1968.25)-(T=1968.50))+EC_LE

	Coefficient	Std. Error	t-Statistic	Prob.
C_LE(1)	0.295420	0.030602	9.653533	0.0000
C_LE(2)	0.199058	0.024568	8.102254	0.0000
C_LE(3)	0.000475	0.000190	2.495656	0.0136
C_LE(4)	0.017596	0.003294	5.342659	0.0000

R-squared	0.443499	Mean dependent var	0.000756
Adjusted R-squared	0.433257	S.D. dependent var	0.003228
S.E. of regression	0.002430	Akaike info criterion	-9.178280
Sum squared resid	0.000962	Schwarz criterion	-9.103598
Log likelihood	770.3864	Hannan-Quinn criter.	-9.147968
F-statistic	43.30049	Durbin-Watson stat	1.087096
Prob(F-statistic)	0.000000		

The results are rather significant, except for the **last** coefficient.



Following the reasoning made earlier, c_le (3) (or rather c_le(3)/c_le(2)) should represent the logarithm of the long-term gap between the target employment and the level reached. This gap will be significant if both:

- Employment shows a trend (the target is moving), which means that GDP and target productivity show different trends.
- A difference between the growths of GDP and target productivity is not compensated immediately (the value of c_le(1) is different from one)

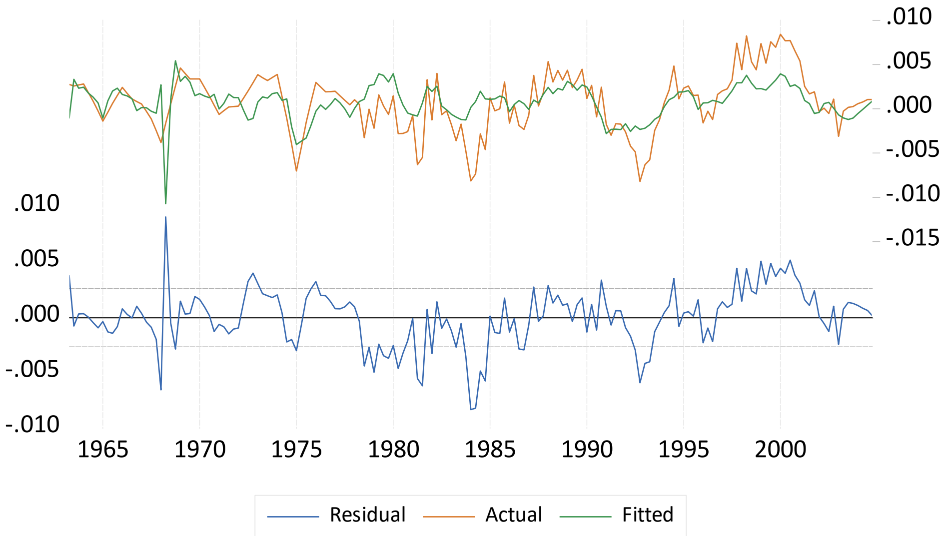
The second condition is clearly met, but for the first the answer is dubious. Instead of trend, one rather observe a break in the level. Nevertheless, the coefficient diagnosed as significant.

As to **the** first coefficients, they are quite significant, maybe lower than expected.

Dependent Variable: DLOG(LE)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/22/20 Time: 16:59
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments
DLOG(LE)=C_LE(1)*DLOG(LED)+C_LE(2)*LOG(LED(-1)/LE(-1))
+C_LE(3)+EC_LE

	Coefficient	Std. Error	t-Statistic	Prob.
C_LE(1)	0.166334	0.020297	8.195005	0.0000
C_LE(2)	0.216267	0.026322	8.216159	0.0000
C_LE(3)	0.000598	0.000204	2.931151	0.0039

R-squared	0.346046	Mean dependent var	0.000756
Adjusted R-squared	0.338071	S.D. dependent var	0.003228
S.E. of regression	0.002626	Akaike info criterion	-9.028889
Sum squared resid	0.001131	Schwarz criterion	-8.972877
Log likelihood	756.9122	Hannan-Quinn criter.	-9.006155
F-statistic	43.39114	Durbin-Watson stat	0.894407
Prob(F-statistic)	0.000000		



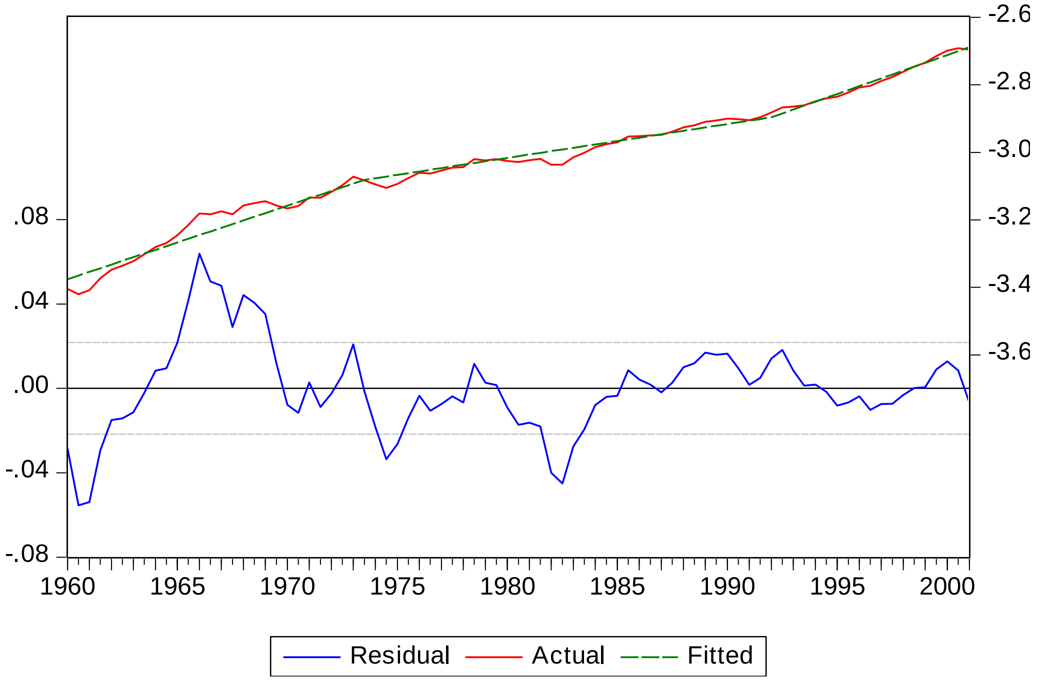
This will not be true in the US case. We will use another data base, this time bi-yearly.

First, the "1992" break exists too, but it is now positive, as shown here:

Dependent Variable: LOG(PRLE)
Method: Least Squares
Date: 11/07/12 Time: 12:48
Sample (adjusted): 1960S1 2001S1
Included observations: 83 after adjustments
LOG(PRLE)=C_PRLE(1)+C_PRLE(2)*T+C_PRLE(3)*(T-1973.5)
(T<1973.5)+C_PRLE(4)(T-1992)*(T<1992)

	Coefficient	Std. Error	t-Statistic	Prob.
C_PRLE(1)	-48.83508	2.620014	-18.63924	0.0000
C_PRLE(2)	0.023062	0.001313	17.55985	0.0000
C_PRLE(3)	0.011769	0.001133	10.38824	0.0000
C_PRLE(4)	-0.013038	0.001639	-7.954501	0.0000

R-squared	0.985989	Mean dependent var	-3.025035
Adjusted R-squared	0.985457	S.D. dependent var	0.180258
S.E. of regression	0.021738	Akaike info criterion	-4.772521
Sum squared resid	0.037331	Schwarz criterion	-4.655951
Log likelihood	202.0596	Hannan-Quinn criter.	-4.725690
F-statistic	1853.184	Durbin-Watson stat	0.252303
Prob(F-statistic)	0.000000		

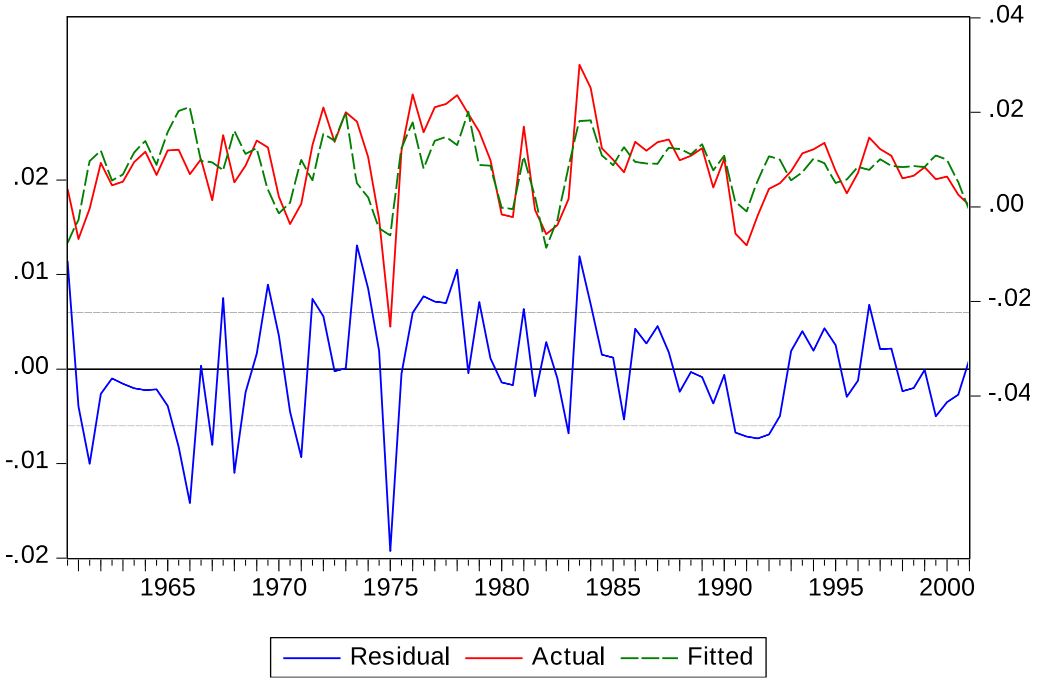


Second, employment has grown substantially over the sample period, which means that a constant term is called for:

Dependent Variable: DLOG(LE)
Method: Least Squares
Date: 11/07/12 Time: 12:50
Sample (adjusted): 1960S2 2001S1
Included observations: 82 after adjustments
DLOG(LE)=C_LE(1)*DLOG(LED)+C_LE(2)*LOG(LED(-1)/LE(-1))+C_LE(3)

	Coefficient	Std. Error	t-Statistic	Prob.
C_LE(1)	0.420082	0.045549	9.222675	0.0000
C_LE(2)	0.097051	0.031296	3.101115	0.0027
C_LE(3)	0.004784	0.000774	6.182176	0.0000

R-squared	0.532205	Mean dependent var	0.008455
Adjusted R-squared	0.520362	S.D. dependent var	0.008682
S.E. of regression	0.006013	Akaike info criterion	-7.353944
Sum squared resid	0.002856	Schwarz criterion	-7.265893
Log likelihood	304.5117	Hannan-Quinn criter.	-7.318593
F-statistic	44.93870	Durbin-Watson stat	1.488317
Prob(F-statistic)	0.000000		



6.3.3.4 A virtuous dummy element

Reverting to the French case, the outlier observed earlier (and in the equation for the change in inventories) remains. The year 1968 presents a strong negative residual in the first semester, and a negative one for the last. As we are now considering a dynamic behavior, this is the time for treating this problem.

As stated earlier, French people (and people familiar with French post-war history) will certainly recall the May 1968 "student revolution" which lasted roughly from March to June. During that period, the French economic process was heavily disturbed, in particular the transportation system, and GDP decreased (by 7.6% for the quarter). If the equation had worked , employment would have decreased too, especially as productivity growth was quite high. On the contrary, it remained almost stable.

The explanation is obvious: firms expected the slump to be purely temporary, and activity to start back after a while (actually they were right, and GDP grew by 7.5% in the next semester, due in part to the higher consumption allowed by "Grenelle[68]" wage negotiations, very favorable to workers). They did not want to lay out (at a high cost) workers whom they would need back later with no guarantee to find the same individuals, familiar with the firms' techniques. So the employment level was very little affected.

This means that the global behavior does not apply here, and the period has to be either eliminated from the sample, or rather treated through a specific dummy variable, taking the value 1 in the first semester and –1 in the second (when employment increased less than the growth in GDP would call for).

This case is rather interesting: some economists could be tempted to introduce dummies just because the equation does not work, and indeed the results will be improved (including in general the statistics for the explanatory variables). This can probably be called cheating. On the contrary, **not** introducing the present dummy can be considered incorrect: we know that the local behavior did not follow the formulation we have selected, so it has to be modified accordingly.

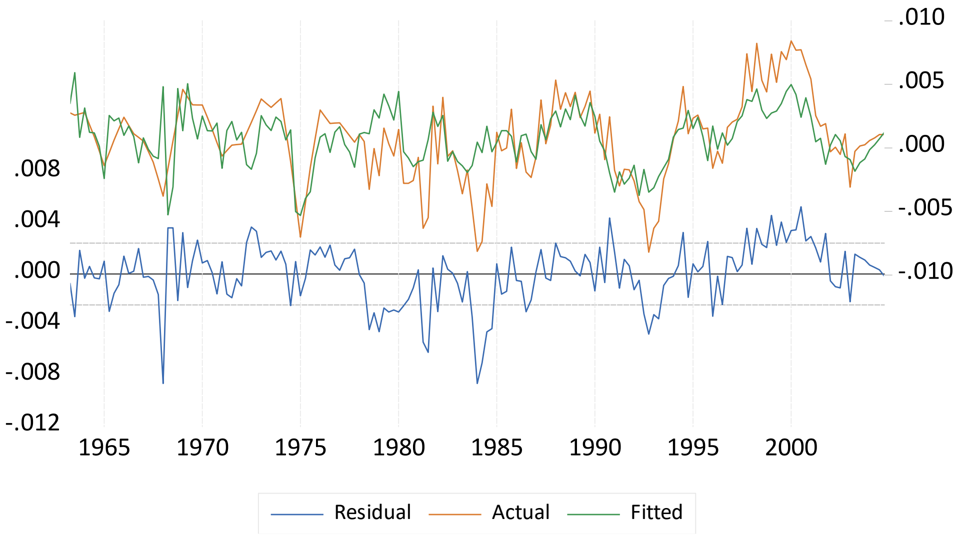
[68] From the location of the Ministry of Employment where negotiations where conducted.

The global results are slightly improved, and the first coefficient increases significantly, meaning that the adaptation of employment to the target is more effective at first. The introduction of the element was not a negligible issue.

Dependent Variable: DLOG(LE)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/22/20 Time: 17:15
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments
DLOG(LE)=C_LE(1)*DLOG(LED)+C_LE(2)*LOG(LED(-1)/LE(-1))
+C_LE(3)+C_LE(4)*((T=1968.25)-(T=1968.50))+EC_LE

	Coefficient	Std. Error	t-Statistic	Prob.
C_LE(1)	0.295420	0.030602	9.653533	0.0000
C_LE(2)	0.199058	0.024568	8.102254	0.0000
C_LE(3)	0.000475	0.000190	2.495656	0.0136
C_LE(4)	0.017596	0.003294	5.342659	0.0000

R-squared	0.443499	Mean dependent var	0.000756
Adjusted R-squared	0.433257	S.D. dependent var	0.003228
S.E. of regression	0.002430	Akaike info criterion	-9.178280
Sum squared resid	0.000962	Schwarz criterion	-9.103598
Log likelihood	770.3864	Hannan-Quinn criter.	-9.147968
F-statistic	43.30049	Durbin-Watson stat	1.087096
Prob(F-statistic)	0.000000		



We shall use:

```
equation eq_prle.ls(p) log(prle)=c_prle(1)+c_prle(2)*t+c_prle(3)*(t-1973)*(t<1973)
+c_prle(4)*(t-1992)*(t<1992)
equation eq_le.ls(p) dlog(le)=c_le(1)*dlog(led)+c_le(2)*log(led(-1)/le(-1))+c_le(3)
+c_le(4)*((t=1968.5)-(t=1968))+ec_le
```

```
genr ec_le=resid
```

and as we have to introduce two new variables

```
g_vendo.drop led prle_t

g_vendo.add led prle_t
```

Note: the reason for the initial "drop" statement is to avoid the duplication of the elements inside the group, in case the procedure is repeated. If the elements are not present, nothing happens.

6.3.4 EXPORTS: AUTOREGRESSIVE PROCESS, COINTEGRATION, LONG-TERM STABILITY.

Estimating exports will be simpler from the theoretical side. We shall use it as an example for introducing first autoregressive processes, then cointegration.

Let us first start with the simplest idea: exports show a constant elasticity to world demand. In other words:

$$\Delta X/X/(\Delta WD/WD) = a$$

or by integration:

$$Log(X) = a \cdot Log(WD) + b$$

Estimation should give to a value close to unity, as world demand is measured as the normal demand addressed to France by its clients, and takes into account:

- The expansion of international trade.
- The types of goods France exports,
- The structure of the countries to which France naturally exports.

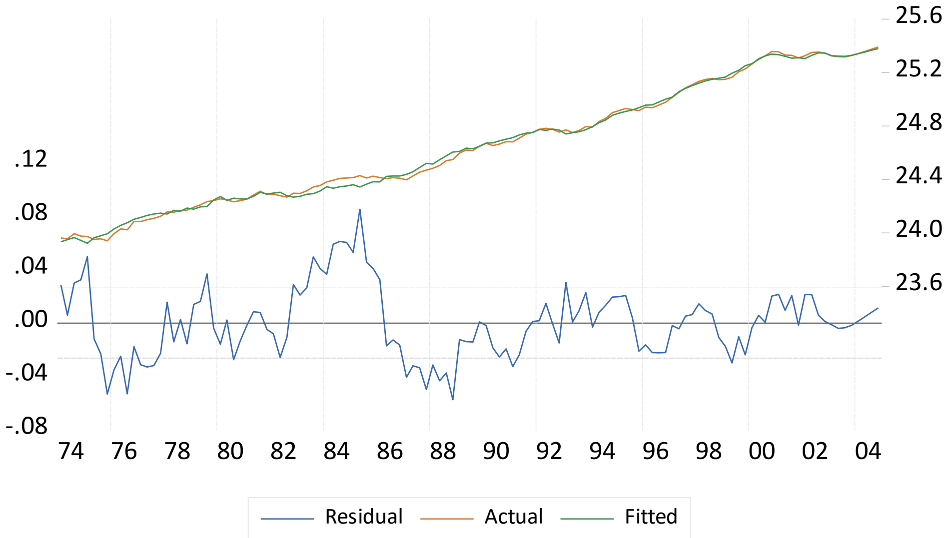
For instance, both luxury goods and Germany have a higher share in this indicator, compared to the global world market. As a weighting of normal imports, it takes also into account the growing importance of international trade.

Indeed the coefficient we obtain is close to unity (but significantly different).

Dependent Variable: LOG(X)
Method: Least Squares
Date: 03/22/20 Time: 17:39
Sample (adjusted): 1974Q1 2004Q4
Included observations: 124 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(WD)	0.887012	0.004704	188.5550	0.0000
C	2.940895	0.115151	25.53944	0.0000

R-squared	0.996580	Mean dependent var	24.64867
Adjusted R-squared	0.996552	S.D. dependent var	0.446469
S.E. of regression	0.026216	Akaike info criterion	-4.428916
Sum squared resid	0.083846	Schwarz criterion	-4.383427
Log likelihood	276.5928	Hannan-Quinn criter.	-4.410437
F-statistic	35552.98	Durbin-Watson stat	0.480527
Prob(F-statistic)	0.000000		



However the low value of the Durbin-Watson test indicates a strongly positive autocorrelation of residuals, and invalidates the formulation. The graph shows indeed long periods with a residual of the same sign, even though the two variables share quite often common evolutions.

Let us try to eliminate auto-correlation, supposing that the residual is actually:

$$e_t = \rho \cdot e_{t-1} + u_t$$

where ρ should be significant (positive here), and u(t) independent across time.

The simplest idea is to transform the equation

$$Log(X_t) = a \cdot Log(WD_t) + b + e_t$$

which is also true in the previous period.

$$Log(X_{t-1}) = a \cdot Log(WD_{t-1}) + b + e_{t-1}$$

We can multiply the second equation by ρ, and subtract it from the first:

$$Log(X_t) - \rho \cdot Log(X_{t-1}) = a \cdot (Log(WD_t) - \rho \cdot Log(WD_{t-1})) + b \cdot (1 - \rho) + e_t - \rho \cdot e_{t-1}$$

The residual for the new equation is the uncorrelated u:

$$Log(X_t) - \rho \cdot Log(X_{t-1}) = a \cdot (Log(WD_t) - \rho \cdot Log(WD_{t-1})) + b \cdot (1 - \rho) + u_t$$

6.3.4.1 Introducing an autoregressive process

To estimate the above formula, it is not necessary to establish the full equation (which calls for a full non–OLS specification, as it is not linear in the coefficients).

One can very well use the same presentation as for ordinary least squares, introducing in the estimation window the additional term AR(n), n representing the autocorrelation lag, in our case 1:

```
ls Log(X) log(WD) c ar(1)
```

But the application to the developed formula is also quite simple:

```
equation eq_lx.ls Log(X)=c_x(1)*log(WD)+c_x(2)+[ar(1)=c_x(3)]
```

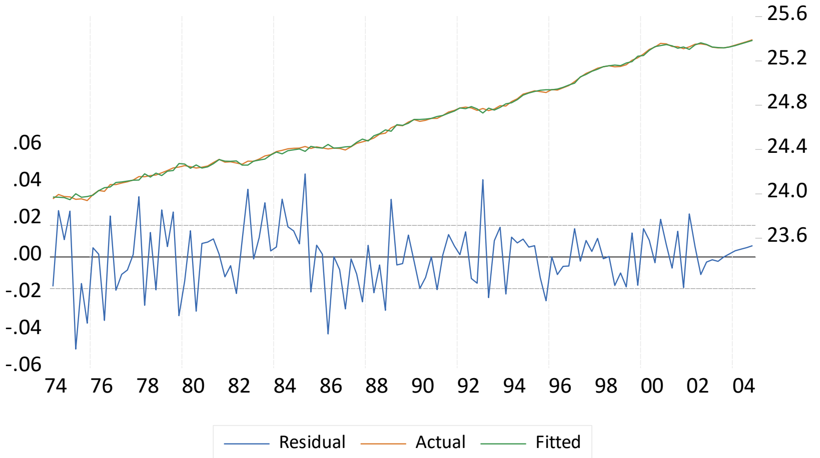
Dependent Variable: LOG(X)
Method: ARMA Conditional Least Squares (Marquardt - EViews legacy)
Date: 03/22/20 Time: 17:50
Sample (adjusted): 1974Q2 2004Q4
Included observations: 123 after adjustments
Convergence achieved after 1 iteration
LOG(X)=C_X(1)*LOG(WD)+C_X(2)+[AR(1)=C_X(3)]

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.884418	0.012700	69.64166	0.0000
C_X(2)	3.003843	0.311465	9.644243	0.0000
C_X(3)	0.757250	0.059233	12.78433	0.0000

R-squared	0.998533	Mean dependent var	24.65426
Adjusted R-squared	0.998509	S.D. dependent var	0.443924
S.E. of regression	0.017142	Akaike info criterion	-5.270440
Sum squared resid	0.035263	Schwarz criterion	-5.201850
Log likelihood	327.1320	Hannan-Quinn criter.	-5.242579
F-statistic	40847.79	Durbin-Watson stat	2.304652
Prob(F-statistic)	0.000000		

Inverted AR Roots	.76

The results are rather satisfactory: the first coefficient retains the theoretical value, the new coefficient is significant, the global precision is much improved (see also the graph) and the DW test is closer to satisfactory.



However our formulation is a little too simplistic. We want exports to decrease with the rate of use of capacities, representing the fact that if firms are already using selling most of their potential production, they will be to be less dynamic in their search for foreign markets (more on this later).

Let us introduce the rate of use UR in the formula.

We get:

Dependent Variable: LOG(X)
Method: ARMA Conditional Least Squares (Marquardt - EViews legacy)
Date: 03/22/20 Time: 18:00
Sample: 1978Q1 2004Q4
Included observations: 108
Convergence achieved after 13 iterations
LOG(X)=C_X(1)*LOG(WD)+C_X(2)*LOG(UR)+C_X(3)+[AR(1)=C_X(4)]

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.381166	0.088426	4.310555	0.0000
C_X(2)	0.219096	0.083380	2.627682	0.0099
C_X(3)	21.58771	36.98678	0.583660	0.5607
C_X(4)	0.998913	0.006440	155.1013	0.0000

R-squared	0.998621	Mean dependent var	24.74243
Adjusted R-squared	0.998581	S.D. dependent var	0.399745
S.E. of regression	0.015058	Akaike info criterion	-5.517431
Sum squared resid	0.023582	Schwarz criterion	-5.418092
Log likelihood	301.9413	Hannan-Quinn criter.	-5.477153
F-statistic	25100.00	Durbin-Watson stat	2.144309
Prob(F-statistic)	0.000000		

Inverted AR Roots	1.00

The relevant coefficients are significant, the average error lower[69] (1,5%), the Durbin-Watson test acceptable, but there is a problem: the sign for the new element is wrong (and unsurprisingly, the coefficient for World demand is now much too low.

Let us not despair. If this old fashioned tool did not work, let us try a more up to date: cointegration.

6.3.4.2 Applying cointegration under EViews

Just as for stationarity, we will not develop the theory behind the method, leaving it to actual econometricians, an excellent source of information being actually the EVIews manual. We sill also rely on basic cointegration theory, and not on any of the recent developments.

Let us just say that cointegration is actually a simple extension of stationarity to a set of two or more variables. To establish cointegration between two elements, one has to prove that in the long run these elements move together, maintaining a bounded "distance" (or rather that a linear combination is bounded), while the value of each of the two elements is unbounded (a necessary condition).

For a group of more than two elements to be cointegrated, no subset of this group must have this property (stationarity of no single element and cointegration of no subset).

If we want to go beyond intuition, the reason for the last condition is that if a cointegrating relation is evidenced between elements, some of which are already cointegrated, one can always recompose the encompassing equation into the true cointegrating equation (considered as a new stationary variable) and other variables.

[69] As the logarithm measures relative evolutions, an absolute error on a logarithm is equivalent to a relative error on the variable itself.

For instance,

if

$$a \cdot x + b \cdot y + c \cdot z$$

is tested as a cointegrating equation, but:

$$a \cdot x + b' \cdot y$$

is too (we can use the same a as a cointegrating equation is known to a given factor), then

$$a \cdot x + b \cdot y + c \cdot z$$

is equivalent to:

$$(a \cdot x + b' \cdot y) + (b' - b) \cdot y + c \cdot z$$

three new elements, one of which is stationary, which forbids us to test cointegration on the three.

So the two properties must be checked: moving together means both "moving" and "together".

Using images rather related to stationarity (as they apply to the actual difference of two elements, without weighting coefficients) we can illustrate the concept as

- Astral bodies moving in outer space and linked together by gravity. Their distance is bounded but their position relative to each other is unknown within those bounds, and we do not know if one is leading the other.
- Human beings: if they are always close to each other, they can be decided to be related (love, hate, professional relationship). But only if they move: if they are in jail, a small distance means nothing.

In our example, the first idea could be to test cointegration between X, WD and UR. But to ensure the stability of our long-term simulations, we need exports to have a unitary elasticity to WD. If this is not the case, when X reaches a constant growth rate, it will be different from that of WD: either France will become the only exporter in the world (in relative terms) or the role of France in the world market will become infinitely negligible. Both prospects are unacceptable (the first more than the second, admittedly).

This constraint can be enforced very easily by considering only in the long run (cointegrating) equation the ratio of X to WD, which we shall link to the rate of use. We will test cointegration between these two elements.

Let us first test their stationarity. We know how to do it (from the estimation of employment).

In a program, this can be done through:

```
UROOT(1,p) Log(X/WD)

UROOT(1,p) Log(UR)
```

Note: if we use menus, we should first display the group (either by selecting elements in the workfile window or creating a group). The default display mode is “spreadsheet” but the “View” item proposes other modes, among them “Unit root test” and “cointegration test” (if more than one series is displayed at the same time).

Null Hypothesis: LOG(X/WD) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on SIC, maxlag=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.515780	0.3201
Test critical values:	1% level	-4.046072	
	5% level	-3.452358	
	10% level	-3.151673	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(X/WD))
Method: Least Squares
Date: 03/22/20 Time: 19:06
Sample: 1978Q2 2004Q4
Included observations: 107

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(X(-1)/WD(-1))	-0.157329	0.062537	-2.515780	0.0134
D(LOG(X(-1)/WD(-1)))	-0.265307	0.092520	-2.867562	0.0050
C	0.038142	0.016339	2.334424	0.0215
@TREND("1978Q2")	-0.000260	0.000117	-2.217517	0.0288

R-squared	0.172932	Mean dependent var	-0.001515
Adjusted R-squared	0.148842	S.D. dependent var	0.017713
S.E. of regression	0.016341	Akaike info criterion	-5.353556
Sum squared resid	0.027505	Schwarz criterion	-5.253637
Log likelihood	290.4152	Hannan-Quinn criter.	-5.313050
F-statistic	7.178750	Durbin-Watson stat	1.937887
Prob(F-statistic)	0.000201		

Null Hypothesis: LOG(UR) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 10 (Automatic - based on SIC, maxlag=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.399050	0.3778
Test critical values:	1% level	-4.054393	
	5% level	-3.456319	
	10% level	-3.153989	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(UR))
Method: Least Squares
Date: 03/22/20 Time: 19:08
Sample (adjusted): 1980Q3 2004Q4
Included observations: 98 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(UR(-1))	-0.091478	0.038131	-2.399050	0.0186
D(LOG(UR(-1)))	1.464814	0.094658	15.47473	0.0000
D(LOG(UR(-2)))	-1.665424	0.170251	-9.782189	0.0000
D(LOG(UR(-3)))	1.785757	0.212274	8.412515	0.0000
D(LOG(UR(-4)))	-1.737132	0.246525	-7.046472	0.0000
D(LOG(UR(-5)))	1.618277	0.251123	6.444148	0.0000
D(LOG(UR(-6)))	-1.586227	0.255319	-6.212736	0.0000
D(LOG(UR(-7)))	1.412498	0.236345	5.976417	0.0000
D(LOG(UR(-8)))	-1.047248	0.214050	-4.892542	0.0000
D(LOG(UR(-9)))	0.719400	0.157728	4.561021	0.0000
D(LOG(UR(-10)))	-0.292147	0.096418	-3.030009	0.0032
C	-0.019947	0.008365	-2.384508	0.0193
@TREND("1978Q2")	3.72E-05	3.38E-05	1.101800	0.2737

R-squared	0.793872	Mean dependent var	0.000238
Adjusted R-squared	0.764772	S.D. dependent var	0.018172
S.E. of regression	0.008813	Akaike info criterion	-6.502075
Sum squared resid	0.006603	Schwarz criterion	-6.159171
Log likelihood	331.6017	Hannan-Quinn criter.	-6.363378
F-statistic	27.28050	Durbin-Watson stat	2.006954
Prob(F-statistic)	0.000000		

These first tests show that both UR and the ratio of exports to world demand cannot be considered stationary, even around a trend: the T statistic is too low, and the estimated probability for the coefficient to zero is too high[70].[71]

[70] Not extremely high, however.

Let us now see if the two elements are co integrated, using the Johansen test.

For EViews this calls for:

coint(*option*,p) *list-of-variables* or *group-name*

Option represents the type of cointegration tested:

a	No deterministic trend in the data, and no intercept or trend in the cointegrating equation.
b	No deterministic trend in the data, and an intercept but no trend in the cointegrating equation.
c	Linear trend in the data, and an intercept but no trend in the cointegrating equation.
d	Linear trend in the data, and both an intercept and a trend in the cointegrating equation.
e	Quadratic trend in the data, and both an intercept and a trend in the cointegrating equation.
s	Summarize the results of all 5 options (a-e).

In our case, we shall use option d (trend in the cointegrating equation, no trend in the VAR)

```
coint(d,p) log(x/wd) log(ur)
```

[71] We can observe that in a traditional least squares estimation, the same T value would give the opposite diagnosis.

Date: 03/22/20 Time: 19:09
Sample (adjusted): 1979Q1 2004Q4
Included observations: 104 after adjustments
Trend assumption: Linear deterministic trend (restricted)
Series: LOG(X/WD) LOG(UR)
Lags interval (in first differences): 1 to 4

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.211670	31.14976	25.87211	0.0100
At most 1	0.059815	6.414546	12.51798	0.4095

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.211670	24.73521	19.38704	0.0076
At most 1	0.059815	6.414546	12.51798	0.4095

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b'*S11*b=I):

LOG(X/WD)	LOG(UR)	@TREND(62Q2)
51.27274	34.10694	0.080146
-18.80755	21.73616	-0.041619

Unrestricted Adjustment Coefficients (alpha):

D(LOG(X/WD))	-0.005934	0.002273
D(LOG(UR))	-0.003137	-0.001939

1 Cointegrating Equation(s): Log likelihood 622.6250

Normalized cointegrating coefficients (standard error in parentheses)

LOG(X/WD)	LOG(UR)	@TREND(62Q2)
1.000000	0.665206	0.001563
	(0.12689)	(0.00013)

Adjustment coefficients (standard error in parentheses)

D(LOG(X/WD))	-0.304258
	(0.07800)
D(LOG(UR))	-0.160833
	(0.05275)

Obviously, cointegration is accepted (a message says so). But we can also:

- Understand the logical process:

EViews tests first if there is no cointegration. If this is accepted (if it shows a high enough probability), the process stops. But here this is refused, as the probability (that there is no cointegration) is too low.

In this case, there is at least one cointegrating equation, and EViews proceeds to testing if there is more than one. This cannot be refused here, as this time the probability is too high.

We have at least one relation, and at most one: we have one.

If the second assumption (at most 1 relation) had not been rejected, there would be at least two and we would have to continue (there cannot be more relations than variables, however).

Evidencing more than one relation is problematic, maybe worse for the model builder than finding no relation (in which case we can always proceed with adding new elements). Even if a cointegrating equation has no implications on causality between elements, we generally intend to include it in a single dynamic formula (a VAR), which does explain a given variable. With two equations, we are stuck with a parasite one, which will be difficult if not impossible to manage in the context of the model (if we stop at econometrics, the problem is smaller).

- Look at the probabilities:

We can observe if the existence (or the rejection) of at least one relation is barely or strongly accepted (and also of only one for that matter).

- Observe the coefficients in the cointegrating equation.

The equation introduces a tradeoff between several concepts (here the share of French exports in world demand and their rate of use). We have always an idea on the sign of the relationship, and also on an interval of economic validity. There is no guarantee that the value will follow these constraints. It can even happen that the right sign obtained by Ordinary Least Squares will become wrong when cointegration is tested on the same relation.

Here it is not too difficult to judge on the soundness of the explanation.

First, the sign is right: exports go down when the rate of use goes up (the sign is positive but both elements are on the same side of the formula).

The size of the coefficient is more difficult to judge. The derivative of the equation relative to Q gives:

$$\Delta(X)/X = -0.665\ \Delta(UR)/UR = -0.665\ (\Delta(Q)/Q - \Delta(CAP)/CAP.$$

In the short run, CAP does not change:

$$\Delta(X)/X = -0.665\ \Delta(UR)/UR = -0.665\ \Delta(Q)/Q$$

$$\Delta(X) = -0.665\ X/Q\ \Delta(Q)$$

Let us suppose an increase in Q of 1 billion Euros coming only from local demand FD. In 2004, the share of exports in French GDP was 32%. The exports target will decrease by

```
∆ X = 0.665 * 0.32 = 213 millions of Euros.
```

The substitution effect looks quite reasonable.

Of course:

- In the long run, capacities will build up, UR will get back to its base value (we know that from the investment equation) and the loss will disappear.
- The changes in Q can come also from X, introducing a loop. This means UR might not be the best representative element. Perhaps we should restrict UR to the satisfaction of local demand (but this looks difficult to formulate).

6.3.4.3 Once cointegration has been evidenced

Two things have to be done:

- Storing the cointegrating equation and its parameters.
- Estimating the VAR (the dynamic equation) and creating the associated element.

The first task should be easy, as EViews does display the requested equation, with its values. However:

- This equation is not available as an item.
- The coefficients are not available as a vector (or as scalars).

There is a **trick**, however, which solves this problem. One can estimate a VAR, in other terms a system which includes both a dynamic equation and a cointegrating equation. This is not directly useful for us, as

- This requires to use the same variables in both forms, in other words to extend the unitary elasticity assumption to the dynamic equation, which is neither needed statistically nor realistic from an economic point of view (as we have seen when estimating employment).
- The output does not provide information on the quality of the cointegration, an essential element in our process.

But the good point is that by estimating a VAR using the same elements as the tested cointegration, we get the same cointegrating coefficients, and this time they are stored in a vector! We can then specify the cointegrating equation using these elements.

This has the extremely high advantage of allowing to establish a program which will adapt automatically to any change in the data, an essential element in all operational modelling projects.

For this particular example, after:

```
coint(d,p) log(x/wd) log(ur)
```

we shall use

```
vary _var_x.ec(d,p) 1 1 log(x/wd) log(ur)[72]
```

to create and estimate the VAR, and store the coefficients in a vector by accessing the first line of matrix _var_x.b (not displayed in the workfile window):

```
vector(10) p_x

p_x(1)=_var_x.b(1,1)

p_x(2)=_var_x.b(1,2)

p_x(3)=_var_x.b(1,3)
```

The cointegrating equation will be:

```
0= p_x(1)*log(x/wd)+p_x(2)*log(ur)+p_x(3)*@trend(60S1)
```

Actually the first parameter is not really needed, as it is equal by construction. We think using it makes the equation clearer.

Estimating the dynamic equation calls for the computation of the residual in the cointegrating equation:

```
genr res_x= p_x(1)*log(x/wd)+p_x(2)*log(ur)+p_x(3)*t
```

Then we estimate the VAR, releasing the constraint on the unitary elasticity of X to WD. In principle, the coefficient should be positive for WD, negative for UR, and we can introduce a lag structure for both elements.

Unfortunately, the coefficient for UR is not significant here, maybe because of a delay in its influence.

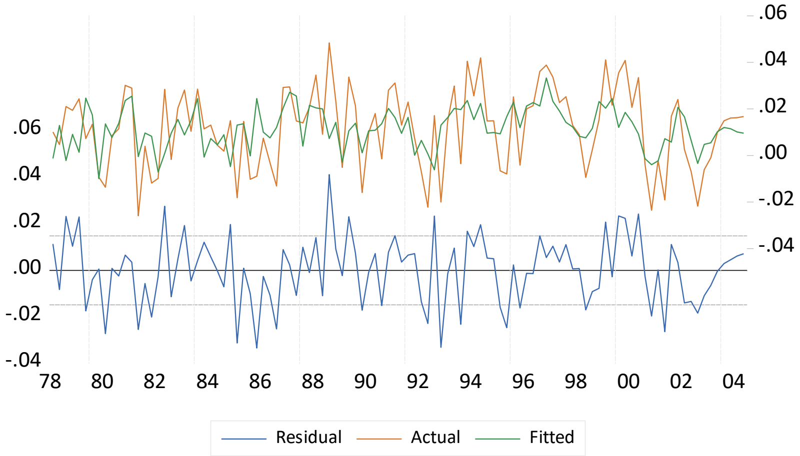
---

[72] The figures 1 and 1 indicate the scope of lagged variations of the left-hand side variable in the VAR which will be added to the right-hand side. Here it will be 1 to 1 (or 1 with lag 1). The options used are consistent with the ones we have used in coint (which have been determined automatically by EViews).

Dependent Variable: DLOG(X)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/22/20 Time: 19:18
Sample (adjusted): 1978Q3 2004Q4
Included observations: 106 after adjustments
DLOG(X)=C_X(4)*DLOG(WD)+C_X(6)*RES_X(-1)+C_X(7)

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(4)	0.494017	0.084894	5.819245	0.0000
C_X(6)	-0.091424	0.038151	-2.396342	0.0184
C_X(7)	0.019601	0.006141	3.191585	0.0019

R-squared	0.266535	Mean dependent var	0.011661
Adjusted R-squared	0.252293	S.D. dependent var	0.017116
S.E. of regression	0.014800	Akaike info criterion	-5.560452
Sum squared resid	0.022562	Schwarz criterion	-5.485071
Log likelihood	297.7039	Hannan-Quinn criter.	-5.529900
F-statistic	18.71468	Durbin-Watson stat	2.032482
Prob(F-statistic)	0.000000		



6.3.5 IMPORTS: GOING FURTHER ON COINTEGRATION AND LONG-TERM STABILITY.

In a single country model, the rest of the world is exogenous, and imports and exports have to be estimated separately, following of course the same guidelines:

- Imports will depend on demand and capacity utilization, through constant elasticities.

However, the definition of demand is not straightforward. For exports, we just had to consider global imports from each partner country in each product, and compute an average using a double weighting: the role of these partners in French exports, and the structure of products exported by France.

This was possible because we considered the rest of the world as exogenous, and did not try to track the origin of its imports.

Now imports can come from three endogenous elements:

- Local final demand, such as foreign cars.
- The intermediate goods necessary to local firms to satisfy this local demand. For cars it will be electronics, steel, energy to run the machines…
- Identically, intermediate goods necessary to produce exported goods.
- But not finished goods used to satisfy foreign demand: France does not re-export significantly goods without transformation (contrary to Hong Kong for instance).

Basically, two methods can be considered:

- We can consider final demand and exports. Obviously, they do not have the same impact on imports (due to the absence of re-exports). We can generate a global demand by applying to exports a correcting factor. Under rather standard assumptions (same import share in all uses, unitary ratio of intermediate consumption to value added), this factor can be set at 0.5.
- We can also define intermediate consumption, and add it to final demand to get the total demand of the country, a share of which will be imported. This method is obviously more acceptable from the economic point of view. Unfortunately it relies on the computation of intermediate consumption, a variable less accurately measured, and sensitive to categories and the integration of the productive process[73].

We have chosen this last method nevertheless, favoring economic properties over statistical reliability.

Dependent Variable: LOG(M)
Method: Least Squares
Date: 04/25/20 Time: 12:16
Sample (adjusted): 1962Q1 2004Q4
Included observations: 172 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(TD)	1.986960	0.013282	149.6030	0.0000
C	-30.78521	0.378155	-81.40902	0.0000

R-squared	0.992462	Mean dependent var	25.78331
Adjusted R-squared	0.992417	S.D. dependent var	0.723716
S.E. of regression	0.063021	Akaike info criterion	-2.679148
Sum squared resid	0.675173	Schwarz criterion	-2.642549
Log likelihood	232.4067	Hannan-Quinn criter.	-2.664299
F-statistic	22381.05	Durbin-Watson stat	0.069582
Prob(F-statistic)	0.000000		

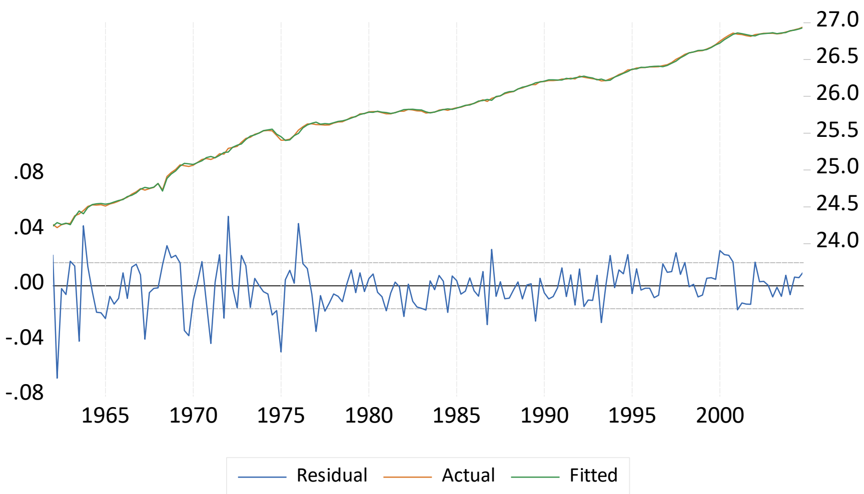
[73] For instance, if good A (say cotton) is used to produce good B (unprinted fabric) which gives good C (printed fabric), both A and B will be counted as intermediary consumption. If the fabric is printed as the same time it is produced, only A will be counted. If we consider value added, the total amount will not change, just the number of elements.

Dependent Variable: LOG(M)
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/25/20   Time: 12:18
Sample: 1962Q1 2004Q4
Included observations: 172
Convergence achieved after 8 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(TD)	2.027497	0.063354	32.00281	0.0000
C	-31.90108	1.786973	-17.85202	0.0000
AR(1)	0.976587	0.020587	47.43628	0.0000
SIGMASQ	0.000271	2.29E-05	11.81244	0.0000

R-squared	0.999480	Mean dependent var	25.78331
Adjusted R-squared	0.999471	S.D. dependent var	0.723716
S.E. of regression	0.016642	Akaike info criterion	-5.312913
Sum squared resid	0.046529	Schwarz criterion	-5.239715
Log likelihood	460.9105	Hannan-Quinn criter.	-5.283215
F-statistic	107738.0	Durbin-Watson stat	1.836241
Prob(F-statistic)	0.000000		

Inverted AR Roots	.98



The fit is quite good including the sensitivity to fluctuations (which was the least we could expect), and the autocorrelation eliminated. But we face a problem: the high value of the coefficient.

Now the question is: should the growth of demand be the only explanation for the growth of imports? In other words, is there not an autonomous force increasing the weight of foreign trade, independently from growth itself? Or: if demand did not grow for a given period, would imports stay stable, or keep part of their momentum? The associated formula would present:

- A more or less unitary elasticity of imports to demand.
- A positive trend representing the regular growth of world trade.

Applying this idea, we get:

Dependent Variable: LOG(M)
Method: Least Squares
Date: 04/25/20 Time: 12:32
Sample (adjusted): 1962Q1 2004Q4
Included observations: 172 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(TD)	1.608779	0.056271	28.58972	0.0000
C	-42.37624	1.719498	-24.64454	0.0000
T	0.011273	0.001640	6.872870	0.0000

R-squared	0.994108	Mean dependent var	25.78331
Adjusted R-squared	0.994039	S.D. dependent var	0.723716
S.E. of regression	0.055878	Akaike info criterion	-2.913993
Sum squared resid	0.527683	Schwarz criterion	-2.859095
Log likelihood	253.6034	Hannan-Quinn criter.	-2.891719
F-statistic	14257.73	Durbin-Watson stat	0.098040
Prob(F-statistic)	0.000000		

Dependent Variable: LOG(M)
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/25/20 Time: 12:33
Sample: 1962Q1 2004Q4
Included observations: 172
Convergence achieved after 7 iterations
Coefficient covariance computed using outer product of gradients

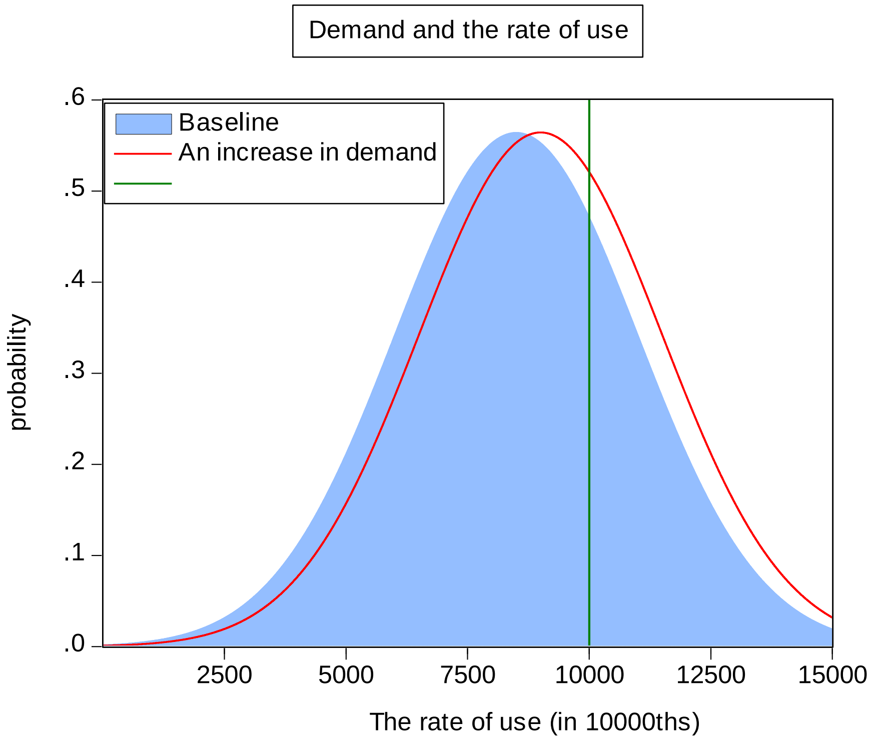
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(TD)	2.116669	0.100512	21.05895	0.0000
C	-26.91516	7.657330	-3.514954	0.0006
T	-0.003789	0.004770	-0.794421	0.4281
AR(1)	0.980164	0.021535	45.51418	0.0000
SIGMASQ	0.000269	2.30E-05	11.72939	0.0000

R-squared	0.999483	Mean dependent var	25.78331
Adjusted R-squared	0.999471	S.D. dependent var	0.723716
S.E. of regression	0.016653	Akaike info criterion	-5.305020
Sum squared resid	0.046312	Schwarz criterion	-5.213523
Log likelihood	461.2317	Hannan-Quinn criter.	-5.267897
F-statistic	80700.22	Durbin-Watson stat	1.843649
Prob(F-statistic)	0.000000		

This does not work. Either we get autocorrelation, or a negative (but not significant) trend.

The problem with our formulation is actually very clear: in terms of model properties, it is reasonable to suppose that in the short run, an increase in final demand will increase imports beyond their normal share, by generating local bottlenecks on domestic supply. But the explanatory element should be the rate of use of capacities. At it is fixed in the long run, the share should go back to normal with time.

A rate of use of 85% (a normal value over the whole economy) does not mean that all firms work at 85% capacity, in which case they can easily move to 86% if needed. It means that the rates of use follows a given distribution, some at lower than 85%, some higher, some at 99%, and a finite number at 100% (see graph).



An increase in demand will move the curve to the right, and more firms will face the limit: an increase of 1% will meet it halfway for firms starting from a 99.5% rate of use. The additional demand, if clients do not accept local substitutes, will have to be supplied by imports.

However, local firms will react to this situation, and try to gain back lost market shares by increasing their capacities through investment: this is the mechanism we have described earlier. In our small model, the long-term rate of use is fixed: the sharing of the additional demand will come back to the base values. These values can increase with time due to the expansion of world trade.

Our formula will make imports depend on total demand and the rate of use:

$$\Delta M/M = a \cdot \Delta TD/TD + b.\Delta UR/UR$$

Where

$$IC = tc\ .\ Q$$

$$TD = FD + IC$$

And tc is the quantity of intermediary consumption units required to produce one unit of GDP.

By integration, we get

$$Log(M) = a \cdot Log(TD) + b \cdot Log(UR) + c$$

Introducing the rate of use, we get:

Dependent Variable: LOG(M)
Method: Least Squares
Date: 04/25/20 Time: 12:41
Sample (adjusted): 1963Q2 2004Q4
Included observations: 167 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(TD)	1.982239	0.013327	148.7350	0.0000
LOG(UR)	1.118627	0.187822	5.955783	0.0000
C	-30.65432	0.379702	-80.73265	0.0000

R-squared	0.993071	Mean dependent var	25.82899
Adjusted R-squared	0.992987	S.D. dependent var	0.683591
S.E. of regression	0.057248	Akaike info criterion	-2.865037
Sum squared resid	0.537488	Schwarz criterion	-2.809025
Log likelihood	242.2306	Hannan-Quinn criter.	-2.842303
F-statistic	11752.39	Durbin-Watson stat	0.099373
Prob(F-statistic)	0.000000		

Dependent Variable: LOG(M)
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/25/20 Time: 12:43
Sample: 1963Q2 2004Q4
Included observations: 167
Convergence achieved after 7 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(TD)	2.012993	0.082841	24.29940	0.0000
LOG(UR)	0.160941	0.127756	1.259755	0.2096
C	-31.50111	2.352651	-13.38962	0.0000
AR(1)	0.972438	0.023201	41.91354	0.0000
SIGMASQ	0.000245	2.23E-05	10.96716	0.0000

R-squared	0.999473	Mean dependent var	25.82899
Adjusted R-squared	0.999460	S.D. dependent var	0.683591
S.E. of regression	0.015881	Akaike info criterion	-5.400439
Sum squared resid	0.040859	Schwarz criterion	-5.307085
Log likelihood	455.9366	Hannan-Quinn criter.	-5.362549
F-statistic	76850.10	Durbin-Watson stat	1.820626
Prob(F-statistic)	0.000000		

Dependent Variable: LOG(M)
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/25/20 Time: 12:45
Sample: 1963Q2 2004Q4
Included observations: 167
Convergence achieved after 9 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(TD)	1.931196	0.201264	9.595350	0.0000
LOG(UR)	0.240297	0.208748	1.151137	0.2514
C	-33.98000	7.679693	-4.424657	0.0000
T	0.002421	0.006392	0.378777	0.7054
AR(1)	0.967930	0.024792	39.04209	0.0000
SIGMASQ	0.000245	2.26E-05	10.85188	0.0000

R-squared	0.999473	Mean dependent var	25.82899
Adjusted R-squared	0.999457	S.D. dependent var	0.683591
S.E. of regression	0.015935	Akaike info criterion	-5.388806
Sum squared resid	0.040881	Schwarz criterion	-5.276782
Log likelihood	455.9653	Hannan-Quinn criter.	-5.343338
F-statistic	61066.95	Durbin-Watson stat	1.818857
Prob(F-statistic)	0.000000		
Inverted AR Roots	.97		

No formulation is acceptable on all counts, with no autocorrelation, a significant explanation of the rate of use, and a positive significant trend (to say nothing of the demand coefficient).

Actually, we are facing a usual problem : the two explanations (deviations of demand from its trend, and output gap) are strongly correlated.

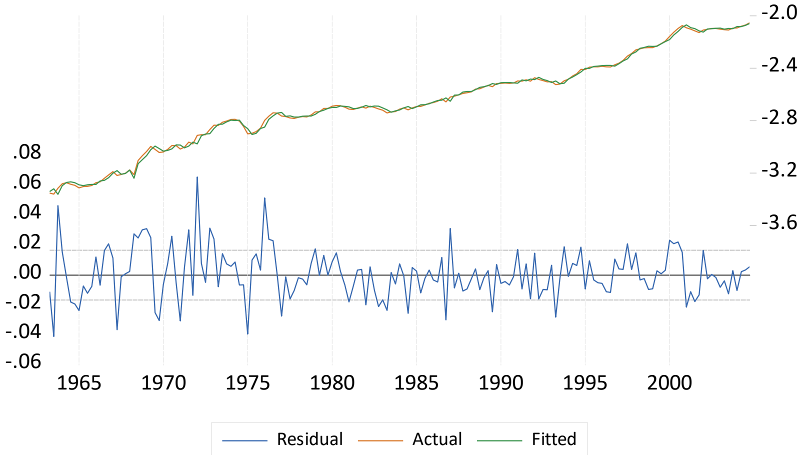
One idea can be to force the first coefficient to unity, and estimate the ratio of imports to demand.

Dependent Variable: LOG(M/TD)
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/25/20 Time: 13:07
Sample: 1963Q2 2004Q4
Included observations: 167
Convergence achieved after 7 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(UR)	1.129681	0.096782	11.67241	0.0000
C	-60.03575	5.793368	-10.36284	0.0000
T	0.028911	0.002922	9.893121	0.0000
AR(1)	0.969395	0.018275	53.04455	0.0000
SIGMASQ	0.000268	2.33E-05	11.47345	0.0000

R-squared	0.997749	Mean dependent var	-2.664699
Adjusted R-squared	0.997693	S.D. dependent var	0.345836
S.E. of regression	0.016610	Akaike info criterion	-5.311368
Sum squared resid	0.044692	Schwarz criterion	-5.218015
Log likelihood	448.4993	Hannan-Quinn criter.	-5.273478
F-statistic	17951.06	Durbin-Watson stat	1.721525
Prob(F-statistic)	0.000000		

Inverted AR Roots	.97



This works quite well on all counts, and ensures that the growth of imports and demand will converge in the long run, once the trend has been suppressed, and the rate of use has stabilized through the investment equation.

We could stop here. But as in the exports case, we can try to separate the behavior into short-term and long-term, in other words to apply a n error correction framework. To represent this process, we need a formula which:

- Enforces a long-term unitary elasticity of imports to the total demand variable, with a positive additional effect of the rate of use.
- Allows free elasticities in the short run.

We shall start by testing the cointegration between the share of imports in demand: M/TD and the rate of use.

Before, we test the stationarity of M/TD, or rather its logarithm (UR has already been tested):

It is strongly contradicted by the Dickey Fuller test:

Null Hypothesis: LOG(M/TD) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.921357	0.1582
Test critical values:	1% level	-4.007084	
	5% level	-3.433651	
	10% level	-3.140697	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(M/TD))
Method: Least Squares
Date: 03/22/20 Time: 20:09
Sample (adjusted): 1963Q3 2010Q4
Included observations: 190 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(M(-1)/TD(-1))	-0.061830	0.021165	-2.921357	0.0039
D(LOG(M(-1)/TD(-1)))	0.271652	0.067991	3.995397	0.0001
C	-0.188041	0.066624	-2.822416	0.0053
@TREND("1962Q1")	0.000421	0.000149	2.828711	0.0052

R-squared	0.108959	Mean dependent var	0.007611
Adjusted R-squared	0.094587	S.D. dependent var	0.019241
S.E. of regression	0.018308	Akaike info criterion	-5.142088
Sum squared resid	0.062347	Schwarz criterion	-5.073730
Log likelihood	492.4984	Hannan-Quinn criter.	-5.114397
F-statistic	7.581507	Durbin-Watson stat	2.018032
Prob(F-statistic)	0.000082		

Now we test cointegration of LOG(UR) and LOG(M /TD)

It fails!

Date: 03/22/20 Time: 20:12
Sample (adjusted): 1979Q1 2010Q4
Included observations: 128 after adjustments
Trend assumption: Linear deterministic trend (restricted)
Series: LOG(M/TD) LOG(UR)
Lags interval (in first differences): 1 to 4

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None	0.056385	11.83231	25.87211	0.8234
At most 1	0.033818	4.403551	12.51798	0.6832

Trace test indicates no cointegration at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None	0.056385	7.428760	19.38704	0.8694
At most 1	0.033818	4.403551	12.51798	0.6832

Max-eigenvalue test indicates no cointegration at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b'*S11*b=I):

LOG(M/TD)	LOG(UR)	@TREND(62Q2)
-15.37135	-8.656178	0.136527
27.20696	-15.09145	-0.171007

Unrestricted Adjustment Coefficients (alpha):

D(LOG(M/TD))	0.001320	-0.001626
D(LOG(UR))	0.004203	0.001371

1 Cointegrating Equation(s): Log likelihood 731.9615

Normalized cointegrating coefficients (standard error in parentheses)

LOG(M/TD)	LOG(UR)	@TREND(62Q2)
1.000000	0.563137	-0.008882
	(0.43358)	(0.00089)

Adjustment coefficients (standard error in parentheses)

D(LOG(M/TD))	-0.020290
	(0.01466)
D(LOG(UR))	-0.064610
	(0.02652)

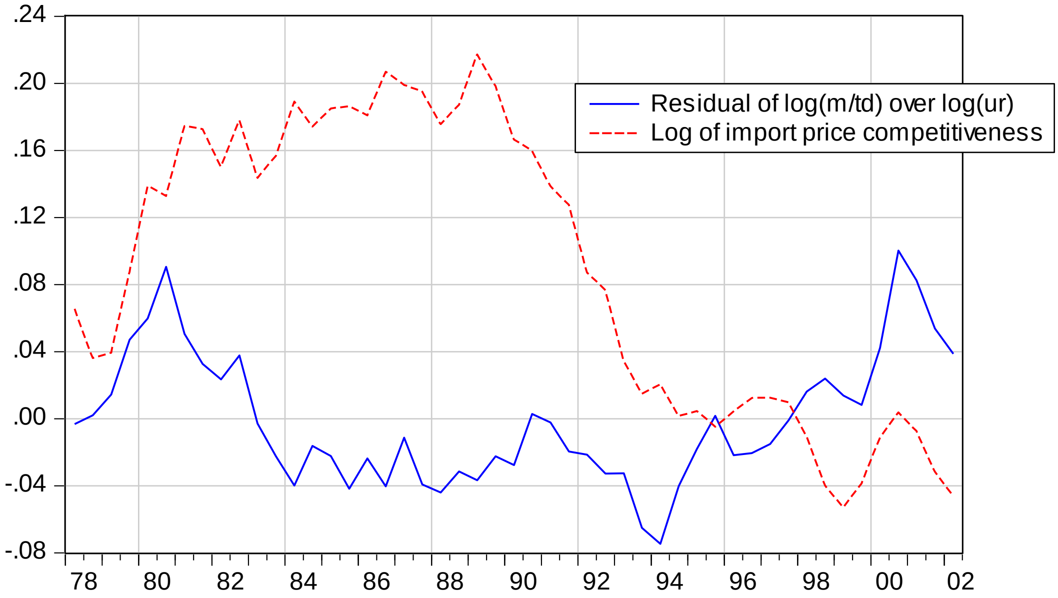
Let us not despair. Diagnosis an absence of cointegration is not such bad news[74], as it allows us to proceed further. If a set of two variables does not work, why not a set of three?

Now which additional element could we consider? The natural candidate comes both from theory and from the data:

[74] Identifying two equations would be much worse, especially in a modelling framework.

If demand is present but local producers have no capacity problems, how can foreign exporters penetrate a market? Of course, through price competitiveness, in other words by decreasing the import price compared to the local one.

This observation is confirmed by the data. Let us regress the import-demand ratio over the rate of use, consider the residual (the unexplained part) and compare it to the ratio of import to local prices: we observe a clearly negative relation.



We observe a clear correlation, especially if we consider variations around a negative trend.

After having tested of course:

- Non-stationarity of Log(COMPM)
- Non- cointegration of Log(COMPM) with both Log(UR) and log(M/TD) individually,

We can test the cointegration of the three elements:

It works!!

Date: 03/22/20 Time: 20:17
Sample (adjusted): 1979Q3 2004Q4
Included observations: 102 after adjustments
Trend assumption: Linear deterministic trend (restricted)
Series: LOG(M/(FD+CT*Q)) LOG(UR) LOG(COMPM)
Lags interval (in first differences): 1 to 3

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.240307	46.08716	42.91525	0.0233
At most 1	0.122323	18.05342	25.87211	0.3403
At most 2	0.045452	4.744804	12.51798	0.6333

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.240307	28.03374	25.82321	0.0252
At most 1	0.122323	13.30861	19.38704	0.3038
At most 2	0.045452	4.744804	12.51798	0.6333

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b'*S11*b=I):

LOG(M/(FD+...	LOG(UR)	LOG(COMPM)	@TREND(62Q2)
-38.34645	58.18449	-21.17524	0.221992
40.54223	-13.27895	14.45687	-0.274847
2.500665	4.998751	15.34349	0.024427

Unrestricted Adjustment Coefficients (alpha):

D(LOG(M/(F...	0.002033	-0.003559	-0.000202
D(LOG(UR))	-0.004216	-0.002450	-0.000251
D(LOG(COM...	-0.000361	0.000809	-0.001906

1 Cointegrating Equation(s): Log likelihood 982.2613

Normalized cointegrating coefficients (standard error in parentheses)

LOG(M/(FD+...	LOG(UR)	LOG(COMPM)	@TREND(62Q2)
1.000000	-1.517337	0.552209	-0.005789
	(0.16825)	(0.08279)	(0.00025)

Adjustment coefficients (standard error in parentheses)

D(LOG(M/(F...	-0.077941
	(0.04357)
D(LOG(UR))	0.161669
	(0.04152)
D(LOG(COM...	0.013852
	(0.03721)

2 Cointegrating Equation(s): Log likelihood 988.9156

Normalized cointegrating coefficients (standard error in parentheses)

LOG(M/(FD+...	LOG(UR)	LOG(COMPM)	@TREND(62Q2)
1.000000	0.000000	0.302737	-0.007052
		(0.14837)	(0.00043)
0.000000	1.000000	-0.164414	-0.000832
		(0.11811)	(0.00035)

Adjustment coefficients (standard error in parentheses)

D(LOG(M/(F...	-0.222221	0.165519
	(0.05989)	(0.06405)
D(LOG(UR))	0.062329	-0.212769
	(0.05869)	(0.06277)
D(LOG(COM...	0.046648	-0.031760

Of course we have to consider the coefficients in the equation. They describe:

- An apparently high sensitivity of imports to the rate of use (but remember the investment equation will stabilize it in the end).

However the true effect is not so high. If the equation was applied to the short-term, with a share of imports in total demand of 0.15 (the value for 2004), we would get:

$$\Delta(M) = 1 \times 0.15 \ \Delta(TD) + 1.52 \times 0.15 \ \Delta(Q)$$

$$\Delta(M) = 0.15 \ \Delta(TD) + 0.23 \ \Delta(Q)$$

And if we do not consider the change in exports:

$$\Delta(FD - Q) = 0.15 \ \Delta(FD + ct . Q) + 0.23 \ \Delta(Q)$$

$$\Delta(Q) = 0.85 / 1.52 \ \Delta(FD) = 0.56 \ \Delta(FD)$$

A quite acceptable multiplier for France.

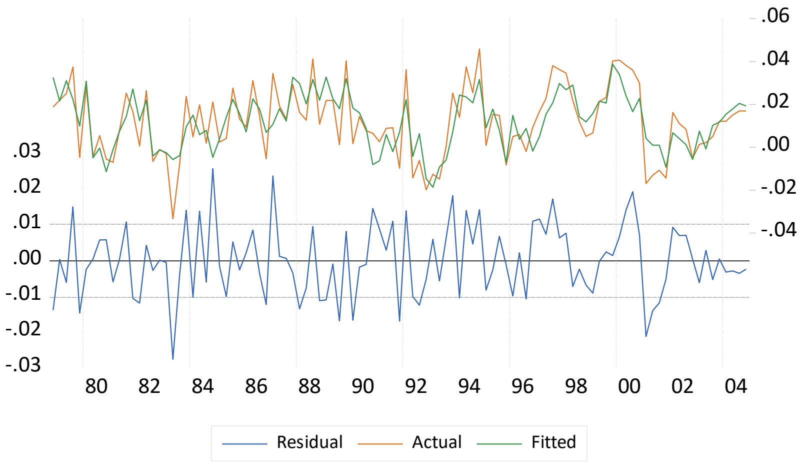
- A much higher sensitivity than in the autoregressive formula.

We can now test the VAR:

Dependent Variable: DLOG(M)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/22/20 Time: 20:26
Sample (adjusted): 1978Q4 2004Q4
Included observations: 105 after adjustments
DLOG(M)=C_M(5)*DLOG(FD+CT*Q)+C_M(6)*DLOG(UR)+C_M(7)
*DLOG(COMPM)+C_M(8)*RES_M(-1)+C_M(9)+EC_M

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(5)	1.676415	0.236019	7.102896	0.0000
C_M(6)	0.213352	0.065300	3.267272	0.0015
C_M(7)	-0.105752	0.078397	-1.348931	0.1804
C_M(8)	-0.102809	0.025468	-4.036803	0.0001
C_M(9)	-0.275761	0.068639	-4.017567	0.0001

R-squared	0.614643	Mean dependent var	0.011899
Adjusted R-squared	0.599229	S.D. dependent var	0.016169
S.E. of regression	0.010236	Akaike info criterion	-6.279400
Sum squared resid	0.010477	Schwarz criterion	-6.153021
Log likelihood	334.6685	Hannan-Quinn criter.	-6.228189
F-statistic	39.87492	Durbin-Watson stat	2.059633
Prob(F-statistic)	0.000000		



The results are rather acceptable, including the graph. By the way, it shows that rejecting a low R-Squared is not justified when the dependent element shows a high variability.

6.3.6 BACK TO THE RESIDUAL CHECK

The method we are using for storing equations has an additional advantage. Now that the residuals have been introduced with their estimated values, all the equations should hold true. The checking process can now be extended to all the endogenous variables.

Theoretically the estimated equations should be consistent with the data, as merging with the model the actual estimated equations ensures the consistency. However:

- If the package does not provide this direct storing, or if the equation had to be normalized by the modeler, editing the formulation could introduce errors.
- The storing of coefficients may have been done badly.
- The text, series or coefficients may have been modified by the user after estimation

- One could have accessed other series or coefficients than the ones used by the estimation (for example one can seek them in another bank including series of similar names).

The reasons for a non-zero residual are less numerous than for identities. They can come only from the fact that equation elements have changed since the last estimation.

Obviously, the main suspect is the data. New data series are due to be inconsistent with the previous estimation, whether it has been updated (moving to a more precise version) or corrected (suppressing an error).

Actually, in EViews, applying a new version of an equation to a model requires, in addition to its estimation, to actually **merge** it again into the model. This will create a new compiled version, without need to explicitly update the model.

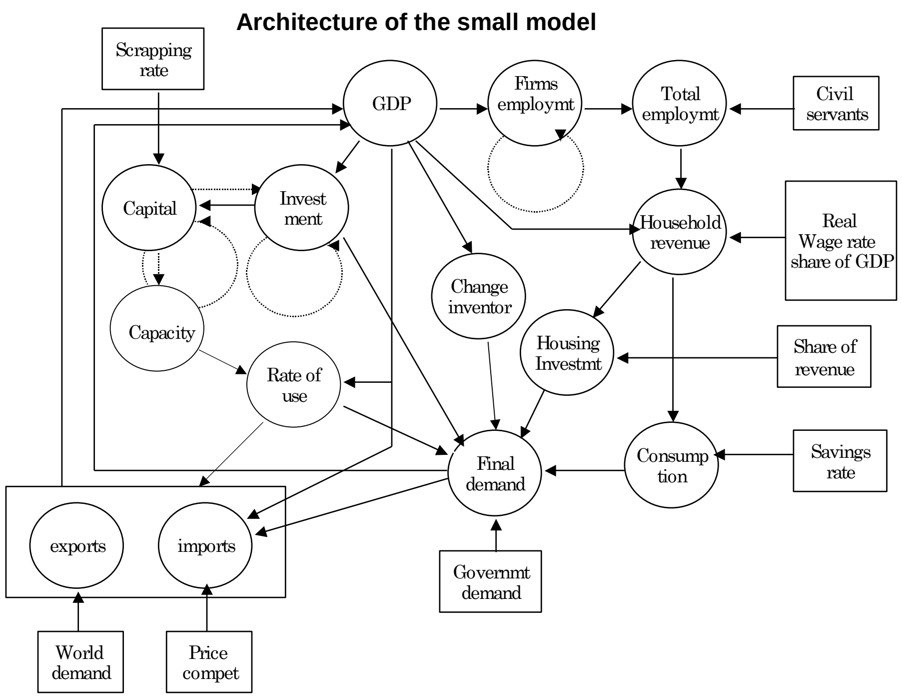
Anyway, in our opinion, applying a new estimation should call for a full model re-creation. This is the only way to guarantee a clear and secure update.

For our model, the statements for the residual check will be the following:

```
'   We check the residuals
smpl 1980Q1 2002S1
__fra_1.append assign @all _c
solve(d=f) _fra_1
for !i=1 to g_vendo.@count
%2=g_vendo.@seriesname(!i)
genr dc_{%2}={%2}-{%2}_c
genr pc_{%2}=100*dc_{%2}/({%2}+({%2}=0))
next
```

The solution series will have the suffix “_c”, and the residuals the prefix “dc_” for the errors in levels, and “pc_” for the relative errors.

We can now present the framework of the model.



6.3.7 THE PRESENT MODEL

In EViews notations, its specifications are:

[1] cap = pk * k(-1)

[2] ur = q / cap

[3] q + m = fd + x

[4] i/k(- 1) = 0.944*I(- 1)/K(- 2) - 0.00256*(ur-ur_star)/ur + 0.0428*.125*q/q(- 8) - 0.00448+ ec_i

[5] ic = ct * q

[6] log(prle_t) = 9.8398+ 0.009382* (t - 2002) + 0.02278* (t - t1) * (t<t1) + 0.01540* (t - t2) * (t<t2)

[7] led = q / prle_t

[8] DLOG(le) = 0.2954*DLOG(led) + 0.1990*LOG(led(- 1)/le(- 1)) + 0.0004748* + 0.01759**((T = 1968.25) - (T = 1968.50)) + ec_le

[9] lt = le + lg

[10] rhi = wr * lt + r_rhiq * q

[11] co = rhi * (1 - sr)

[12] ih = r_ih * rhi

[13] ci/q(- 1) = -0.02680*(T = 1968.25) + 0.6128*ci(- 1)/q(- 2) - 0.0021490 + 0.2193*@PCH(q) + 0.1056*@PCH(q(- 1)) + 0.05492*@PCH(q(- 2)) + 0.03918*@PCH(q(- 3)) + 0.03026*@PCH(q(- 4)) + ec_ci

[14] fd = co + i + gd + ci + ih

[15] td = fd + ic

[16] res_m = log(m / (fd + ct * q)) ‘ 1.517 * log(ur) +-0.552 * log(compm) + 0.00579 * (@trend(60:1) * (t<=2004) + @elem(@trend(60:1) , "2004q4") * (t>2004))

[17] DLOG(m) = 1.676*DLOG(td) + 0.2133*DLOG(ur) - 0.1057*DLOG(compm) - 0.1028*res_m(- 1) - 0.2757+ EC_M

[18] res_x = p_x(1) * log(x / wd) + 0.9815 * log(ur) + 0.001441 * (@trend(60:1) * (t<=2004) + @elem(@trend(60:1) , "2004q4") * (t>2004))

[19] DLOG(x) = 0.4940*DLOG(wd) - 0.09142*res_x(- 1) + 0.01960+ ec_x

[20] k = k(-1) * (1 - dr) +i

CHAPTER 7: TESTING THE MODEL THROUGH SIMULATIONS OVER THE PAST

We have now achieved the production of a model, for which:

- All the identities are consistent with the data.
- All estimations are in our opinion acceptable from a statistical point of view.
- All the coefficient values and equation specifications seem consistent with economic theory[75].
- All the necessary connections between model elements are present.
- The estimated equations have been built in such a way that they should provide a steady state growth, and a long-term solution for the full model.

This does not mean our model is acceptable.

- The consistency of the data might hide errors compensating each other.
- Some complex equations might have hidden wrong individual properties.
- Connections between concepts might have been forgotten or wrongly specified.
- The growth rates provided naturally by equations could be the wrong ones.
- Some error correction processes could be diverging instead of converging.
- Assembling individually acceptable formulations might create a system with unacceptable properties.
- Exogenous elements might be linked with each other.
- Theoretical balances might not have been specified.

And finally:

Another modeler would have obtained a different model, with potentially different properties.

Let us just give an example for a problem: if in the short run increasing government demand by 1000 creates 800 consumption and 600 investment, while exports do not change and imports increase by 300, the individual equations might look acceptable, but the model will diverge immediately through an explosive multiplier effect (800 +600-300 = 1100).

Our next goal will be to control:

- That the model gives a numerical solution
- That it can be used for forecasts.
- That it can be used for policy analysis.

Actually, our first tests will rather answer the reverse question:

- Is there some indication that the model is unsuitable for forecasts, and for policy analysis?

In our opinion, it is only later, by simulations over the future (its normal field of operation, actually) that we can really validate the use of a model. But as usual, problems should be diagnosed as soon as possible. And the availability of actual data increases strongly the testing possibilities.

---

[75] In some cases we might have been obliged to calibrate the values.

Finally, the errors evidenced at this stage might help to build a better forecasting structure, before any attempt on the future.

Let us first address the process of solving the model.

7.1 THE SOLUTION

To solve the model we need to apply a method (an “algorithm”). Let us present the different options.

7.1.1 GAUSS-SEIDEL

This is the most natural algorithm: one often uses Gauss-Seidel without knowing it, like M. Jourdain (the Bourgeois Gentilhomme) makes prose.

The method starts from initial values. They can be the historical values on the past, on the future the values computed for the previous period or for an alternate base simulation. The whole set of equations will be applied in a given order, using as input the most recent information (computations replace the initial values). This gives a new set of starting values. The process is repeated, using always the last information available, until the distance between the two last solutions is small enough to be considered negligible. One will then consider that the solution has been reached.

Let us formalize this process.

Considering the model

$$y_t = f_t(y_t, y_{t-1}, x_t, \hat{\alpha})$$

Where y_t is the vector of endogenous variables at period t.

As only present values will change during computation, we will not consider the other elements, and will drop the time index.

$$y = f(y)$$

We will use it to define the particular endogenous, and an exponent to define the iteration count.

a - We start from y^0, value before any computation, setting the number of iterations to zero.

b - We add 1 to the number of iterations (which we shall note k); this gives to the first iteration the number 1.

c -We compute y_i^k from i = 1 to n, taking into account the i-1 values we have just produced. This means we compute:

$$y_i^k = f(y_1^k, \ldots, y_{i-1}^k, y_i^{k-1}, \ldots., y_n^{k-1})$$

(at the first iteration, explanatory elements will take the starting value y^0 if their index is higher than the computed variable)[76].

d – At the end of the process, we compare y^k and y^{k-1}: if the distance is small enough for every element (using a criterion we shall present) we stop the process, and use the last value as a solution. If not, we check if we have reached the maximum number of iterations, in which case we accept the failure of the algorithm, and stop. Otherwise we resume the process at step b.

Clearly, this algorithm requests an identified model, with a single variable on the left hand side (or an expression containing a single simultaneous variable).

7.1.2 RITZ-JORDAN

The Ritz-Jordan method is similar to the one above: it simply abstains from using values computed at the current iteration:

$$y^k = f(y^{k-1})$$

Refusing to consider the last information, it looks less efficient than Gauss-Seidel. In our opinion, its only interest appears when the model faces convergence problems: it makes their interpretation easier by reducing the interferences between variables.

This method is not provided by EViews.

7.1.3 NEWTON AND ITS VARIANTS

Contrary to the two above, the Newton method applies naturally to non-identified formulations. It represents actually a generalization to an n-dimensional case of the well-known method using a sequence of linearizations to solve a single equation.

Let us consider the model:

$$f_t(y_t, y_{t-1}, x_t, \hat{\alpha}) = 0$$

[76] This means only variables which are used before they are computed must be given values for initialization. We shall come back to this later.

that we will simplify as above into:

$$f\ \ (y) = 0$$

The linearization of f around a starting solution gives, by calling “fl” the value of f linearized:

$$(\partial f/\partial y)_{y=y^0} \cdot (y - y^0) = fl(y) - f(y^0)$$

Solving the system for fl (y) = 0 leads to:

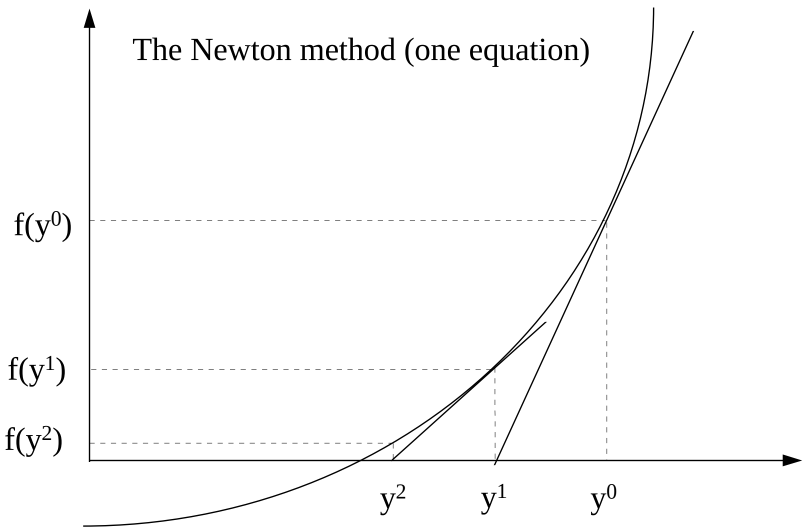
$$y = y^0 - (\partial f/\partial y)^{-1}_{y=y^0} \cdot f(y^0)$$

or with an identified system:

$$y - f(y) = 0$$

we would get naturally:

$$y = y^0 - (I - \partial f/\partial y)^{-1}_{y=y^0} \cdot (y^0 - f(y^0))$$



Linearizing the model again, around the new solution y^1, and solving the new linearized version of the model, we define an iterative process which, as the preceding, will stop when the distance between the two last values gets small enough. Implementing this method is more complex: in addition to inverting a matrix, each iteration involves the computation of a Jacobian. This can be done in practice in two ways:

- Analytically, by determining from the start the formal expressions of derivatives. At each iteration, we shall compute them again from the present values of variables. This method supposes either undertaking the derivation "by hand" with a high probability of errors, or having access to an automatic formal processor, a program analyzing the text of equations to produce the formal expression of their derivatives. To a high initial cost, this method opposes a simplification of computations during the iterative process[77].

EViews allows both methods[78].

- By finite differences, determining separately each column of the Jacobian by the numerical computation of a limited first order development, applied in turn to each variable. One computes the y vector using the base values, then for starting values differing only by a single variable, and compares the two results to get a column of the Jacobian. One will speak then of a method of secants, or pseudo-Newton.

The derivate formulation becomes in this case:

[77] However, changing some model specifications calls for a new global derivation (or a dangerous manual updating).

[78] Unfortunately, the associated code is not apparently available to the user, which would allow interesting computations.

$$\sum_{j=1}^{n}\left\{\left[\mathrm{f}_{\mathrm{i}}(y^{k}+e_{j}\Delta y_{j})-\mathrm{f}(y^{k})\right]/\Delta y_{j}\right\}(y_{j}-y_{j}^{k})=fl_{i}(y)-f_{i}(y^{k})$$

where e_j is a vector of dimension n (number of endogenous), with 1 in position i and 0 otherwise.

In other words, the element of the Jacobian:

$\partial \mathrm{f}_{\mathrm{i}} / \partial \mathrm{y}_{\mathrm{j}}$ will be approximated by $\mathrm{f}_{\mathrm{i}}(\mathrm{y}^{\mathrm{k}}+\mathrm{e}_{\mathrm{j}}\Delta \mathrm{y}_{\mathrm{j}})-\mathrm{f}(\mathrm{y}^{\mathrm{k}})\big| / \Delta \mathrm{y}_{\mathrm{j}}$

One will have only to compute the y vector n+1 times: one time with no modification and one time for each of the endogenous variables.

The expensive part of this algorithm being clearly the computation of the Jacobian and its inversion, a variant will consist in computing it only each m iterations. The convergence will be slower in terms of number of iterations, but the global cost might decrease.

EViews provides another alternative: Broyden's method, which uses a secant method and does not require to compute the Jacobian at each step. As we shall see later, this method proves often very efficient.

7.1.3.1 The identified case

If the model is in "identified" form:

$$y = f(y)$$

the Newton algorithm will be applied to

$$y - f(y) = 0$$

and the Newton formula becomes:

$$y = y^0 - (I - \partial f/\partial y)^{-1}_{y=y^0} \cdot (y^0 - f(y^0))$$

or

$$y = (I - \partial f/\partial y)^{-1}_{y=y^0} \cdot (f(y^0) - (\partial f/\partial y)_{y=y^0} \cdot y^0)$$

This does not change the technical process.

7.1.4 BROYDEN'S METHOD

Broyden's method (also called secant method) computes the Jacobian only once, in the same way as Newton's, and computes a new value of the variable accordingly.

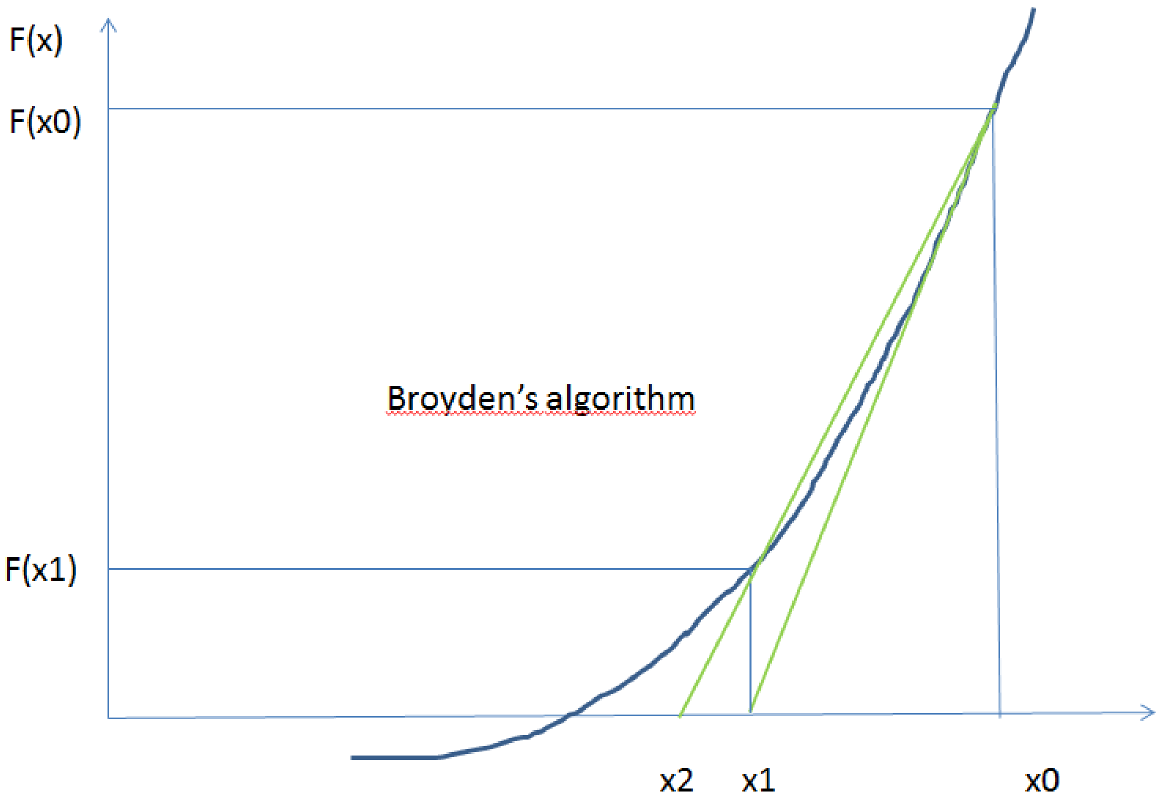
After that, it updates the Jacobian, not by derivation, but by considering the difference with the previous solution, and the direction leading from the previous solution to the new one.

The formula for updating the Jacobian is:

$$J_{t+1} = J_t + (F(x_{t+1}) - F(x_t) - J_t \Delta x_t).\Delta x'_t/(\Delta x'_t.\Delta x_t)$$

where J is the Jacobian, F the function which should reach zero, and x the vector of unknown variables.

Let us clarify all this with a graph based on the single equation case.



We can see that the direction improves with each iteration, less than Newton but more than Gauss-Seidel (for which it does not improve at all).

Otherwise the method shares all the characteristics of Newton's, in particular its independence to equation ordering. It takes generally more iterations, but each of them is cheaper (except for the first).

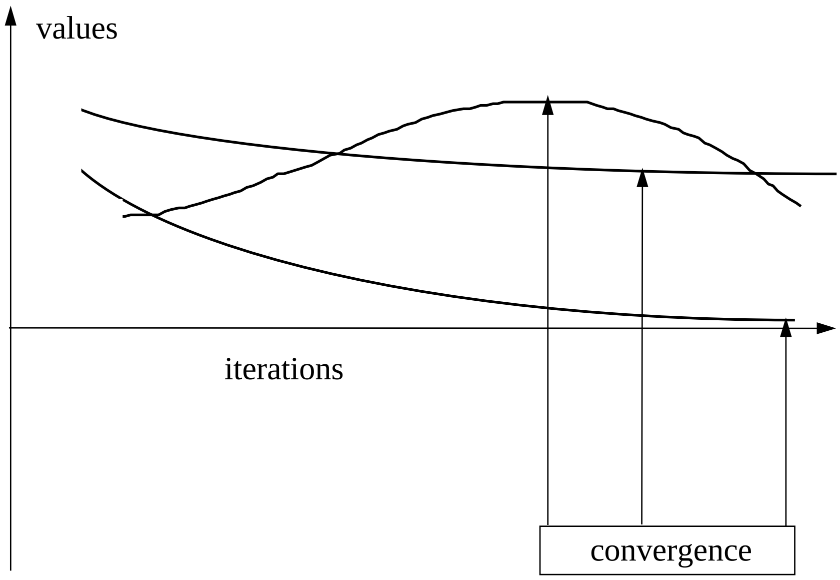
We shall see on a set of practical examples that on average it looks like the most efficient option on the whole, both in terms of speed and probability of convergence[79]. But the diagnosis is not so clear cut.

7.1.5 ITERATIONS AND TEST OF CONVERGENCE

Methods described above have a common feature: starting from initial values, they apply formulations to get a new set. The process is repeated until the two last sets are sufficiently close to be considered as the solution of the system.

One cannot identify the difference between two iterations with the precision actually reached (or the difference to the solution). This is valid only for alternate processes. For monotonous ones, it actually can be the reverse: the slower the convergence, the smaller the change in the criterion from one iteration to the other, and the higher the chance that the criterion will be reached quite far from the solution. As to cyclical processes, they can reach convergence mistakenly at the top or bottom of a cycle.

[79] The most important feature in our opinion.



So one could question the use of this type of method, by stressing that the relative stability of values does not mean that the solution has been found. However, one can observe that if the values do not change, it means that the computation which gave a variable would give the same result with the new values of its explanatory variables: it means also that the equation holds almost true (but that very different values have the same property).

In this case, it is clear that we do not get the exact solution. This criticism should not be stretched too much: the precision of models is in any event limited, and even supposedly exact algorithms are limited by the precision of computers.

7.1.5.1 The general options

For the algorithm to know at which moment to stop computations, we shall have to establish a test.

In fact, the only criterion used in practice will consider the variation of the whole set of variables in the solution, from an iteration to the other.

It can be measured, for each variable:

in relative values:

$$d_i = \left|\left(y_i^k - y_i^{k-1}\right)/y_i^{k-1}\right|$$

or in levels:

$$d_i = \left| y_i^k - y_i^{k-1} \right|$$

As to the condition for accepting convergence, it can be defined:

by variable: $d_i < c_i, \forall i$

on the whole set: $d_i < c_i, \forall i$

or sometimes through a global measure: $f(d) < c$

Generally one will choose a criterion in relative value, each error being compared with a global criterion. This value will have to be small compared to the expected model precision (so that the simulation error will not really contribute to the global error), and to the number of digits used for results interpretation.

The most frequent exception should correspond to variables which, like the trade balance, fluctuate strongly and can even change sign: here the choice of a criterion in level seems a natural solution, which will avoid a non-convergence diagnosis due to negligible fluctuations of a variable around a solution (by pure chance) very small.

For example, if the convergence threshold is 0.0001 in relative value, convergence will be refused if solutions for the US trade balance alternate by chance between - 1 billion current US Dollars and - 1.0002 billion[80], while a difference of 200 000 Dollars, at the scale of US foreign trade, is obviously very small. And this difference, which represents less than one millionth of US exports and imports, might never be reduced if the computer precision guarantees only 8 significant figures[81].

In practice we shall see that the test could be restricted to a subset of variables in the model, the convergence of which extends mathematically to global convergence.

The value given to the criterion can depend:

- On the algorithm used:
  - In case of Gauss-Seidel, each additional digit bears roughly the same cost. The convergence is qualified as linear.
  - In case of Newton, the number of digits gained increases with the iterations: beyond the minimum level (say 0.01%) a given gain is cheaper and cheaper (this will be developed later). The convergence is called quadratic

[80] There is no risk for this in present times.

[81] Exports and imports will be precise to the 8th digit, but the difference, a million times smaller, to the 2nd only.

- On the type of simulation:

    - For a forecast, one will not be too strict, as we all know the precision is quite low anyway. Forecasting growth of 2.05% and 2.07% three years from now delivers the same message, especially as the actual growth might materialize as 1% (more on forecast precision later).
    - For a shock analysis, especially if the shock is small, the evaluation of the difference between the two simulations is obviously more affected by the error: decisions increasing in GDP by 0.07% and 0.09% will not be given the same efficiency.

- And perhaps on the stochastic character:

In a stochastic simulation, it is essential that the consequence for the solution of introducing small random residuals is precisely associated with the shock, and not on the simulation process.

As to the number of iterations, it will be used as a limit, after which we suppose that the model has no chance to converge. In practice one never considers stopping an apparently converging process, just because it has taken too much time. So the only relevant case is when the process is not progressing, because it is oscillating between two or more solutions, and the deadlock has to be broken. Reverting to the use of damping factors (described later) should solve the problem in the Gauss-Seidel case.

7.1.5.2 The EViews options

Testing convergence under EViews is not very flexible: the only option allowed is the level of the (relative) convergence criterion, and it will apply to all variables.

One can also decide on the maximum number of iterations. For most models, after 1000 iterations, convergence becomes rather improbable. But just to make sure, one can set an initial high number. Observing the actual number required can allow to improve the figure.

7.1.6 STUDY OF THE CONVERGENCE

We are now going to show how the choice of the algorithm affects the convergence process.

Let us begin by stating the problem, and introducing some definitions.

7.1.6.1 The incidence matrix

The *incidence matrix* of an n-equation model

$f(y, \ldots) = 0$ (n endogenous variables, n equations)

will be defined as the Boolean matrix A (dimension n by n) such that

- $A_{i,j} = 1$ if the variable y_j appears formally, through its unlagged value, in the equation of rank i.
- $A_{i,j} = 0$ otherwise.

We will suppose the model to be normalized, therefore put under the form:

$$y - f(y) = 0$$

where the variable y_i will appear naturally to the left of the equation of rank i: the main diagonal of the matrix will be composed of 1s.

The definition of the incidence matrix, as one can see, depends not only on the model, but also on the ordering of equations, actually the one in which they are going to be computed.

The formal presence of a variable in an equation does not necessarily mean a numerical influence: it could be affected by a potentially null coefficient, or intervene only in a branch of an alternative. Often we will not be able to associate to a model a unique incidence matrix, nor a matrix constant with time, except if one considers the total set of potential influences (the matrix will be then the Boolean sum of individual Boolean matrices).

One will also notice that defining the incidence matrix does not require the full formulations, or the knowledge of variable values. We simply need to know the list of variables which appear in each explanation, as well as their instantaneous or lagged character[82].

Application to our model

To apply this technique to our model, we can rely on the block structure provided by EViews, through access to:

(double-click)>View>Block structure,

which gives in our case:

```
Number of equations: 20
Number of independent blocks: 3
Number of simultaneous blocks: 1
Number of recursive blocks: 2
Block 1: 3 Recursive Equations
         cap(1)                     prle_t(6)                  x(19)
Block 2: 14 Simultaneous Equations (1 feedback var)
         ic(5)                      ci(13)                     led(7)                     le(8)
         td(15)                     m(17)                      q(3)
Block 3: 3 Recursive Equations
      res_m(16)                  res_x(18)                  k(20)
```

All these elements are consistent with the model graph.

We can use these elements to improve the ordering.

[82] Following our methodology, the incidence matrix can be produced before any estimation.

First, we can use the above separation to move the three predetermined variables at the beginning, and the three post determined at the end, which give the following matrix:

	cap	prle_t	x	q	ur	ic	ci	i	led	le	lt	rhi	ih	co	fd	td	m	res_x	res_m	k
cap	1																			
prle_t		1																		
x			1																	
q			1	1											1		1			
ur	1			1	1															
ic				1		1														
ci				1			1													
i				1	1			1												
led		1		1					1											
le		1								1										
lt									1		1									
rhi				1							1	1								
ih												1	1							
co												1		1						
fd							1	1					1	1	1					
td																1				
m					1										1	1	1			
res_x					1													1		
res_m					1										1				1	
k								1												1

We can see that the model has been divided into three parts:

- A three equation block, with elements which do not depend on the complement, or on subsequent variables in the same block. The variables in this can then be computed once and for all, in a single iteration, at the beginning of the solving process. Actually they do not depend on any variable in the same block, but this is not required.

This property is called recursiveness, and the block is usually named the **prologue**.

We can see that variables can belong to this block for various reasons:

- prle_t depends only on time. The only reason for introducing its equation is to allow easy modification in forecasts.
- cap depends on an exogenous and a predetermined variable.
- x should depend on the rest of the equilibrium (through UR) but this link has not been evidenced statistically, leaving only the instantaneous influence of the exogenous WD.

In practice, however, respecting the convergence threshold will need two iterations, the starting value being different from the solution found, unless the recursivity is known from the start, and the first solution accepted without control[83].

[83] Which is of course the case for EViews.

- A three equations block, in which elements do not affect the rest of the model, and do not depend on subsequent variables in the same block. These variables can be computed after all the others, once and for all in one pass. Again, they do not depend on any variable in the same block, but this is not necessary. The only condition is that they do not depend on elements computed later (or that the matrix is lower-triangular).

In this block we find:

- The residuals for the cointegration equations, which will only be corrected at the next period.
- The end-of-period capital, which obviously cannot affect the equilibrium for the period.

We shall see later another important category: variables with a purely descriptive role, like the government deficit in GDP points

- The rest of the model is simultaneous, and sometimes called the heart. We can check on the model graph that for any given couple of variables in the set, there is at least one sequence of direct causal relationships leading from the first to the second, and vice versa. This means also that exogeneizing any element (at a value different from the model solution of course) all the other elements will be affected.

We can now try to better interpret the simultaneity in the heart. The first stage is observing the presence of loop variables.

The incidence matrix allows defining **loop** variables, as variables that enter in an equation of rank lower than the one that defines them, or will be used before they are computed. In matrix notations, we shall have:

for variable j, $\exists\ A_{i,j} = 1$ such as $i < j$

The variables appearing as an explanatory factor in their own equation of definition also will have to be added to this set. But in practice this case is rather rare.

Let us look at our incidence matrix. Two loop variables are present: FD and M. The reason is that they are used to compute Q, in an equation which appears at the beginning (of the heart).

Actually X should also be present, but as UR appears only through its lagged value, and WD is exogenous, its exact value can be computed immediately, which means it is located in the prologue. In a way it is now technically exogenous (considering only same period relationships).

Of course, a model can contain a sequence of non-recursive blocks. This will happen for instance for two subsequent non-recursive blocks if elements of the second depend on elements in the first, but not vice-versa. Between the two blocks, a recursive one can be introduced.

We shall see examples of this situation when we deal with more complex models.

The definition of the set of loop variables presents the following interest: if this set is empty, the model becomes recursive, which means that the sequential calculation of each of the equations (in the chosen order) gives the solution of the system. Values obtained at the end of the first iteration will satisfy the global set of equations, none of these values being questioned by a later modification of an explanatory element. And a second iteration will not modify the result.

This favorable case is rare enough[84]. However, one can often identify simultaneous subsets (or « blocks ») with a recursive structure relative to each other, such that the p first equations of the model are not influenced by the last n – p (as we have shown on our example). The process of simulation can be then improved, as it will suffice to solve in sequence two systems of reduced size, allowing to gain time as the cost of solution grows more than proportionally with the number of equations. This property is evident for Newton, where the cost of both Jacobian computation and inversion decrease, less for Broyden and even less for Gauss-Seidel, where the only proof comes from practice.

It is obvious that discovering the above properties and performing the associated reordering are interesting for the model builder, as they allow to improve the organization of the solution process, and therefore reduce computation time. This process will also allow to detect logical errors, for example by evidencing the recursive determination of an element known as belonging to a loop (such as investment in the Keynesian loop). Most software packages, including EViews, take care of this search and the associated reorganization, but further improvement may be sought in the solving process by modifying by hand the order of equation computations.

In the light of previous observations, one can look:

- For the best block-recursive structure.
- Inside each block, for the order which permits the fastest convergence, or ensures its highest probability.

The first goal is indisputable, and in fact the easiest to realize from the algorithmic viewpoint. The separation found is unique, but some orderings of blocks can be equivalent (for example equations using only exogenous or lagged elements can be placed in any order).

The second is much less obvious and in any case more complex. One will seek generally to minimize the number of loop variables. The cost of this technique will depend on the ambition: the search for one set of loop variables from which we cannot eliminate an element (Nepomiaschy and Ravelli) is cheaper than the search for all orderings with the smallest possible number of elements (Gilli and Rossier). The first type of set will be called minimal, the second minimum. In fact, minimizing the number of loop variables might not be a good preparation for the use of the Gauss - Seidel algorithm, as we will see later.

EViews determines automatically the block structure of the model (which is de facto optimal, even if other organizations exist). As to reordering inside the simultaneous blocks, if it does not apply an optimization algorithm, it determines the loop variables associated with a given ordering (actually associated to the initial one) and places the associated equations at the end of the block. The efficiency of this last action is questionable, as it means that in a given iteration all computations use the previous value of loop variables, delaying somewhat the impact of “new” information.

For instance, in our model, we can reduce the number of loop variables by transferring the equation for Q to the end of the heart:

[84] And the associated model is probably quite poor.

	cap	prle_t	x	ur	ic	ci	i	led	le	lt	rhi	ih	co	fd	td	m	q	res_x	res_m	k
cap	1																			
prle_t		1																		
x			1																	
ur	1			1													1			
ic					1												1			
ci						1											1			
i				1			1										1			
led		1						1									1			
le		1							1											
lt								1		1										
rhi										1	1						1			
ih											1	1								
co											1		1							
fd						1	1					1	1	1						
td															1					
m				1										1	1	1				
q			1											1		1	1			
res_x				1														1		
res_m				1										1					1	
k							1													1

Now Q is the only loop variable (but a strong one as it appears 6 times in that capacity).

The new ordering is:

Prologue

[1] cap = pk * k(-1)

[2] log(prle_t) = 9.8398+ 0.009382* (t - 2002) + 0.02278* (t - t1) * (t<t1) + 0.01540* (t - t2) * (t<t2)

[3] DLOG(x) = 0.4940*DLOG(wd) - 0.09142*res_x(- 1) + 0.01960+ ec_x

Heart

[4] ur = q / cap

[5] i/k(- 1) = 0.944*I(- 1)/K(- 2) - 0.00256*(ur-ur_star)/ur + 0.0428*.125*q/q(- 8) - 0.00448+ ec_i

[6] led = q / prle_t

[7] DLOG(le) = 0.2954*DLOG(led) + 0.1990*LOG(led(- 1)/le(- 1)) + 0.0004748* + 0.01759**((T = 1968.25) - (T = 1968.50)) + ec_le

[8] ic = ct * q

[9] lt = le + lg

[10] rhi = wr * lt + r_rhiq * q

[11] co = rhi * (1 - sr)

[12] ih = r_ih * rhi

[13] ci/q(- 1) = -0.02680*(T = 1968.25) + 0.6128*ci(- 1)/q(- 2) - 0.0021490 + 0.2193*@PCH(q) + 0.1056*@PCH(q(- 1)) + 0.05492*@PCH(q(- 2)) + 0.03918*@PCH(q(- 3)) + 0.03026*@PCH(q(- 4)) + ec_ci

[14] fd = co + i + gd + ci + ih

[15] td = fd + ic

[16] DLOG(m) = 1.676*DLOG(td) + 0.2133*DLOG(ur) - 0.1057*DLOG(compm) - 0.1028*res_m(- 1) - 0.2757+ EC_M

[17] q + m = fd + x

Epilogue

[18] res_x = p_x(1) * log(x / wd) + 0.9815 * log(ur) + 0.001441 * (@trend(60:1) * (t<=2004) + @elem(@trend(60:1) , "2004q4") * (t>2004))

[19] res_m = log(m / (fd + ct * q)) ‘ 1.517 * log(ur) +-0.552 * log(compm) + 0.00579 * (@trend(60:1) * (t<=2004) + @elem(@trend(60:1) , "2004q4") * (t>2004))

[20] k = k(-1) * (1 - dr) +i

and solving the model will be done in the following way (we shall see later the corresponding statements):

```
Model: _FRA_1
Date: 03/23/20   Time: 11:57
Sample: 2002Q1 2004Q4
Solve Options:
        Dynamic-Deterministic Simulation
        Solver: Gauss-Seidel
        Max iterations = 5000, Convergence = 1e-08
Scenario: Baseline
Solve begin 11:57:47
        2004Q1    Block 1 - 3 eqns     Solved (recursive block)
        2004Q1    Block 2 - 14 eqns    Convergence after 12 iterations
        2004Q1    Block 3 - 3 eqns     Solved (recursive block)
        2004Q2    Block 1 - 3 eqns     Solved (recursive block)
        2004Q2    Block 2 - 14 eqns    Convergence after 12 iterations
        2004Q2    Block 3 - 3 eqns     Solved (recursive block)
        2004Q3    Block 1 - 3 eqns     Solved (recursive block)
        2004Q3    Block 2 - 14 eqns    Convergence after 12 iterations
        2004Q3    Block 3 - 3 eqns     Solved (recursive block)
        2004Q4    Block 1 - 3 eqns     Solved (recursive block)
        2004Q4    Block 2 - 14 eqns    Convergence after 12 iterations
        2004Q4    Block 3 - 3 eqns     Solved (recursive block)
```

Solve complete 11:57:47

It is now possible to compare convergence properties of the different methods, beginning with the Gauss-Seidel algorithm.

One can see that, for the Gauss-Seidel algorithm, only loop variables influence (by their starting values) the result of the current iteration. Let us note y_b the vector of loop variables (complement y_c), associated to the current ordering. One could consider replacing in turn each occurrence of a variable already computed by its current expression (possibly already transformed). This will give a model in which only loop variables appear on the right side:

$$y = g(y_b)$$

And for the convergence process, we have only to consider in the y_b elements of the result:

$$y_b = g(y_b)$$

One has just to solve the new (smaller) model, and y_c will be obtained in one pass at the end of the process.

Our model gives:

```
[4] ur  = f4(q)

[5] ic  = f5(q)

[6] ci = f6(q)

[7] i  = f7(q,f4(q))

[8] led = f8(q)

[9] le = f9(f8(q))

[10] lt=f(f10(f9(f8(q)))

[11] rhi = f11(f(f10(f9(f8(q))),q)

[12] ih  = f12(f11(f(f10(f9(f8(q))),q))

[13] co  =  f13(f12(f11(f(f10(f9(f8(q))),q)))
```

[14] fd = f14(f13(f12(f11(f(f10(f9(f8(q))),q))),f7(q,f4(q)), f6(q), f12(f11(f(f10(f9(f8(q))),q)))

[15] td = f(f14(f13(f12(f11(f(f10(f9(f8(q))),q))),f7(q,f4(q)), f6(q), f12(f11(f(f10(f9(f8(q))),q))),f5(q))

[16] m = f15(f14(f13(f12(f11(f(f10(f9(f8(q))),q))),f7(q,f4(q)), f6(q), f12(f11(f(f10(f9(f8(q))),q))), f5(q), f4(q))

Even for this simple example, producing the total set of explicit relations would have been difficult if not impossible. But if g is formally unknown, its values can be computed for any y_b (in a typical Gauss- Seidel iteration).

Thus, in our example, the successive application (in the given order) of the formulas in the f system associates to an initial value of Q a numerical value of the associated g (but not its solution of course). No other initial value has to be considered.

Again, this shows that for the convergence, one has only to consider the reduced model:

$$\mathrm{y}_b = \mathrm{g}(y_b) \text{ or } \mathrm{y}_b^{\mathrm{k+1}} = \mathrm{g}(y_b^k)$$

and that the control of convergence can be limited to these variables.

Let us use this observation to interpret the convergence of the algorithm. We shall linearize the process around the solution:

$$\mathrm{y}^{k} - \mathrm{y}^* = \mathrm{g}(y^{\mathrm{k-1}}) - \mathrm{g}(\mathrm{y}^*) \approx (\partial \mathrm{g}/\partial y)_{\mathrm{y}=\mathrm{y}^{\mathrm{k-1}}} (y^{\mathrm{k-1}} - \mathrm{y}^*)$$

This approximation, if it can be used in practice (which supposes either that one is near the solution, or that the model is linear enough), shows that the vector of errors is multiplied at each iteration by the Jacobian. If the Jacobian is stable enough in the process, this makes convergence almost linear, and the cost of obtaining an additional solution digit will be independent from its position in the representation of the variable. And in order for the algorithm to converge effectively, the generally necessary and always sufficient condition is that when one raises the Jacobian matrix to a power growing to the infinite, its value goes to zero. An equivalent condition is that the every eigenvalue of the matrix is strictly lower than 1 in module (one will say that the spectral radius is lower than unity). Thus, in our example, convergence will be assured if:

$$\Delta \mathrm{g}(\mathrm{q})/\mathrm{q} < 1$$

One sees that minimizing the number of loop variables, if it reduces the size of the Jacobian, has no specific reason to reduce the spectral radius[85].

Actually it frequently happens that the reordering does not modify in any way the convergence process, as it affects only the point at which it enters in an otherwise unchanged loop. This is the case in our example: by transferring the Q equation to the end, we start with the value for Q in the workfile and not the one which balances the supply-demand equilibrium using the initial values for X, M and FD. But from then on the process is exactly the same. Even more: if the data values meet the equilibrium (they do most of the time), computing Q will not change its value, and the process will be identical even in its figures.

One can even (with no more proof) have the reverse intuition: by concentrating the process on a limited number of variables, one might increase their individual role.

In addition, using an automatic algorithm transfers reorganization control from a logical economist (who has probably chosen a logical choice in terms of causality) to a mathematical blind tool.

Obviously, if the number of loop variables is high, the probability of convergence would be very low if coefficients were chosen randomly. In practice, fortunately, one can associate to the mathematical convergence process a more economic one, the iterative progress to a balanced equilibrium by combining the behaviors of agents (examples can be found in the Keynesian loop, or in the WS-PS wages-prices loop). The probability of success is much improved, and most models converge if they use reasonable assumptions and a reasonable starting point[86].

Let us illustrate this point by simplifying to the extreme our usual model, taking only into account GDP (Q), private demand (C) and government demand (g). Our model does not consider external trade.

(1) C = a Q

(2) Q = C + g

and

$0<a<1$

With the Gauss-Seidel algorithm, this model will always converge to the solution Q*, as for the unique loop variable (Q):

$$Q^k - Q^* = a \cdot (Q^{k-1} - Q^*)$$

[85] In simpler words: the highest modulus found in the set of eigenvalues.

[86] Not to converge for improbable assumption values should be considered as a good point.

and |a| < 1[87]

Let us now invert causalities:

(1) C = Q - g

(2) Q = C / a

This time the model diverges:

$$Q^k - Q^* = 1/a \cdot (Q^{k-1} - Q^*)$$

and

$$abs\left(\frac{1}{a}\right) < 1$$

but this is normal, as to a convergent economic framework (the Keynesian multiplier) we have substituted a framework where, to simplify:

- Households consume what remains of supply once exogenous demand is satisfied.
- Firms seek to maintain supply in a constant ratio (higher than 1) with consumption.

If such a framework does not converge mathematically, it the same logically: although there exists obviously a solution (the same as higher), an initial error on Q will apply identically to C, then will amplify on Q.... And the same instability will be found if we shock public demand (computing the public expenditure multiplier).

One can observe also that with a >1, the first model would diverge (as secondary effects on production would be larger than their source). And the second would converge (as a change in consumption is now reduced on production): the two errors compensate each other.

Formalizing the problem, with the first framework, an initial change in government demand will lead at iteration k, to the cumulated variation:

[87] With a>=1 the model would diverge, which could be expected from an economic point of view.

$$\Delta Q^k = (1 + a + a^2 = \ldots . + \rightleftarrows a^k) \cdot \Delta g = (1 - a^{k-1})/(1 - a) \cdot \Delta g$$

which converges to

$$\Delta Q^k = 1/(1 - a) \cdot \Delta g$$

if a < 1

and with the second

$$\Delta Q^k = (1 + 1/a + (1/a)^2 + \ldots . + (1/a^k)) \cdot \Delta g = (1 - (1/a)^{k-1})/(1 - (1/a))\Delta g$$

which converges only if a > 1

The convergence of one model is equivalent to the divergence of the other.

7.1.6.2 Accelerating Gauss-Seidel: damping factors

An efficient technique to accelerate convergence, or to force naturally diverging models to reach a solution, is to introduce one or more «damping factors». This technique is based on the following elementary observation, inspired by the above.

If y is solution of

$$y = f(y)$$

then it will be also the solution of

$$y = D \cdot f(y) + (\text{I-D}) \cdot y$$

where y represents the vector of endogenous and D an identity matrix.

It goes similarly for g:

if y_b is a solution of

$$y_b = g_b(y_b)$$

then it will be also of:

$$y_b = D \cdot g_b(y_b) + (\text{I-D}) \cdot y_b$$

But the new formulation modifies the Gauss - Seidel solving process, by introducing inertia on starting values y_b.

In practice that will mean initializing the iteration k, not by:

$$y_b^{k\text{-}1}$$

but by

$$D_b \cdot y_b^{k-1} + (\text{I-D}_b) \cdot y_b^{k-2}$$

The change of starting values plays a role only if variables are used before they are computed during the iteration. Therefore damping factors will be applied to loop variables only.

One shows easily that by applying damping factors, we are moving towards a Newton type method, at least in simple cases. Indeed, in a model with a single loop variable, the introduction of damping factors means using:

$$y_b^k = D\, g_b\, (y_b^{k\text{-}1}) + (I - D)\, y_b^{k\text{-}1}$$

or

$$y_b^k - y_b^{k-1} = D\,(g_b(y_b^{k-1}) - y_b^{k-1})$$

This is equivalent to the Newton method if one chooses

$$D = (I - \partial g_b / \partial y_b)^{-1}{}_{y_b = y_b^{k-1}}$$

or the inverse of the Jacobian of:

$$y_b - g_b(y_b) = 0$$

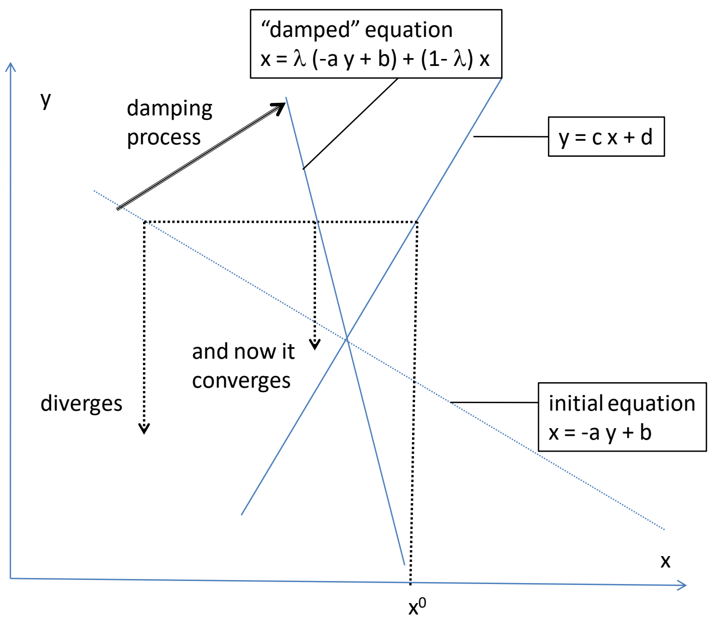
This method, more or less efficient following the stability of the Jacobian, will give the solution in a single iteration in the case of a linear model.

Let us complete our presentation by a graphic example.

We will use the simple formal model

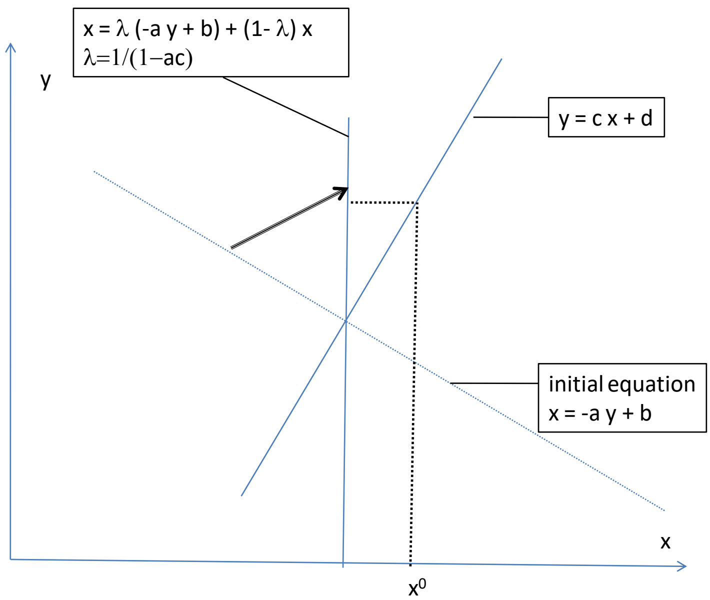
(1) y = c x + d

(2) x = - a y + b



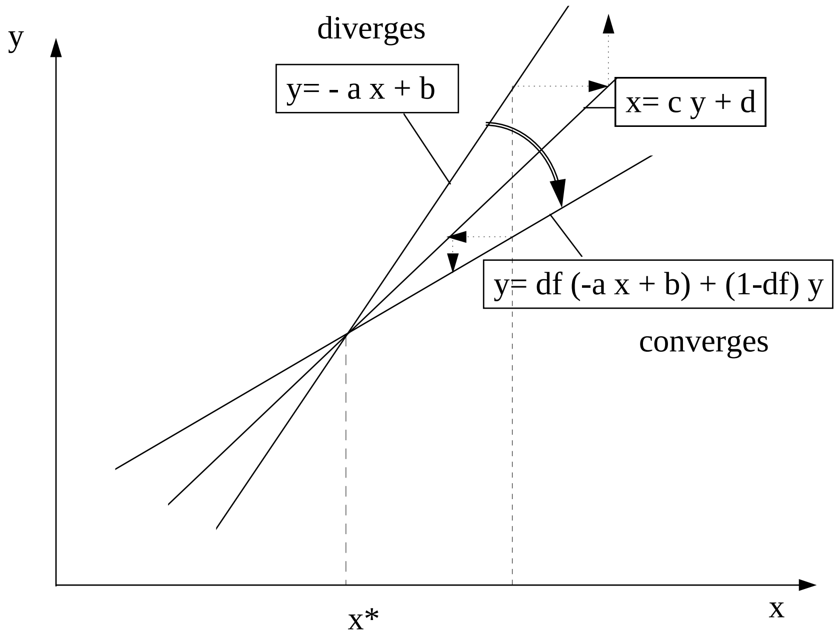
The basic process clearly diverges, but if we apply a convenient damping factor (we have chosen a value close to 0.5) we can make it converge. What we have done is simply replacing one of the equations (y = - ax + b) by another which associates to a given value of y a combination, with constant weights, of the values of x associated to the same y by the two equations. The solution is not changed, but the convergence process is modified.

One could have reached the exact value, by making the new "curve" vertical



Applying damping factors is not however so simple.

- It can be done only through trial and error, excepts if one solves the problem described above, which is actually more complex than solving the model itself by the Newton method[88].
- It is only truly efficient for values between 0 and 1 (alternate convergence). If the two curves have slopes with the same orientation, making convergence monotonous, the method becomes more difficult to handle, as an efficient damping factor must now be chosen outside the interval (although it can work, as the following diagram shows).



-

And when the size of the model grows, the searching process becomes rapidly complex.

---

[88] One can also concentrate on one loop variable, considered as the most important, and measure its changes for three consecutive iterations. A very rough approximation of the convergence factor can be obtained by:

$$a = (x^k - x^{k-1})/(x^{k-1} - x^{k-2})$$

and one can use as a damping factor the value:

$$\lambda = 1/(1 - a)$$

7.1.6.2.1 More complex models

If the equivalence to Newton cannot be extended to the case of several loop variables, it is only because the method restrains D to a diagonal matrix. One could however consider:

A Looking for the diagonal matrix "closest " to the Jacobian, according to a given measure.

B Decomposing the vector y on the eigenvector basis of the Jacobian, which would allow to diagonalize the process.

$$y^{k} - y^{k-1} \approx (\partial g / \partial y)_{y=y^{k-1}} (y^{k-1} - y^{k-2}) = V D V^{-1} (y^{k-1} - y^{k-2})$$

$$V^{-1}(y^{k} - y^{k-1}) \approx DV^{-1} (y^{k-1} - y^{k-2})$$

with D diagonal matrix of the eigenvalues of the Jacobian.

A relative stability of the Jacobian (therefore a good linearity, or an initialization value close to the solution) is then necessary[89].

- Let us see now how the above considerations influence the convergence of the Newton algorithm.

The usefulness of loop variables, and the minimization of their number, is now indisputable, if one decides to compute the Jacobian by finite differences. Indeed one just has to know starting values for y_b to compute $g(y_b)$ and the Jacobian $\partial g_b/\partial y_b$ (this last by Gauss-Seidel iterations),

- Only the part $\partial g/ \partial y_b$ is not identically null.
- And only the part $\partial g_b/\partial y_b$ is going to affect the calculation of the new value of yb.

We just have to apply the Newton formula to the loop variables alone. Once again, only the values of loop variables of the previous iteration (and their influence on the Jacobian of g) will play a role in the iterative process.

This has obviously the advantage of limiting computations, for the Jacobian itself (calculation of the only non-zero columns) as well as for its inversion.

The usefulness of taking into account loop variables looks clear. The reality is less so:

---

[89] Experience shows that the Jacobian of the linear differentiation is much less stable than the relative one:

$$D = (I - \partial g_b/\partial y_b)^{-1}{}_{y_b = y_b^{k-1}}$$

- The Jacobian matrix associated to the initial formulation f is very sparse (the average number of same period explanatory variables per equation does not grow too fast with the size of the model, and stays generally in the order of 3 or 4). This means the Newton method applied to f can use an inversion algorithm specific to sparse matrices, appreciably more effective. The formulation in g concentrates the process of determination on some variables, but the implicit formulation complexity produces a very dense Jacobian matrix, which can be inverted only by a general algorithm.
- As to the calculation of the matrix itself, it will be obviously more expensive for f than for g, if one uses finite differences. But the Jacobian matrix associated to the f vector can also be determined analytically: then one will just have to compute independently each non zero element of the Jacobian, according to a formula generally simpler than the initial equations90. If the average number of instantaneous variables per equation is lower than the number of loop variables, the global cost will decrease. This reasoning is however only valid if one makes abstraction of the (important) cost of the initial formal derivation91, which supposes also a certain stability of the formulation of the model.

As for the comparison with Gauss-Seidel, one can expect a lower number of iterations, but each of them will be clearly much more expensive.

Let us see how the Newton process converges.

For this let us develop to the second order the derivation formula:

$$y^{k+1}\text{-}g(y^{k+1}) - (y^k\text{-}g(y^k)) = (I\text{-}\partial g / \partial y)_{y=y^k}(y^{k+1} - y^k) - {}^T(y^{k+1} - y^k)(\partial^2 g / \partial y^2)_{y=y^k}(y^{k+1} - y^k)$$

plus a third order term.

Using the linearization to determine y^{k+1} eliminates the central terms, leaving:

$$y^{k+1} - g(y^{k+1}) = - {}^T(y^{k+1} - y^k)(\partial^2 g / \partial y^2)_{y=y^k}(y^{k+1} - y^k) + e^3(y^{k+1} - y^k)$$

and as $y^{k+1}\text{-}y^k$is a linear function of $y^k\text{-}g(y^k)$ (again by construction), the error $y^{k+1}\text{-}g(y^{k+1})$ is a quadratic function of $y^k\text{-}g(y^k)$.

Similarly, the distance of y to the solution y* is a quadratic function of the distance at the previous iteration, as the derivation to the first order has given:

90 For instance, when derivation is applied to a sum.

91 EViews is able to compute the analytical derivatives. However, it does not provide the formulas.

$$(y^{k+1}\text{-}g(y^{k+1})) - (y^*\text{-}g(y^*))\ [=0] = (I\text{-}\partial g / \partial y)_{y=y^k}(y^{k+1} - y^*)$$

The convergence of the algorithm is not therefore uniform: in fact it is going to accelerate with the number of iterations, as the linearized form gets nearer to the derivative at the solution. One will say that it is quadratic in the case of a formal derivation, but only supralinear if one uses finite differences, because one does not compute the exact derivative. The Newton algorithm should be the more advisable as the precision requested from the solution gets higher.

7.1.6.3 Comparison of the relative efficiencies

Let us consider first the ability to find the solution. The convergence of the Gauss-Seidel algorithm appears quite doubtful, if it based on several iterative interacting processes (the number of loop variables) all of which must be converging. In other terms, a matrix must have all its eigenvalues lower than unity in modulus. It is however more frequent than one can expect. As we have seen earlier, the numerical process is generally linked to a set of economic causalities, determining a normally convergent process. Actually, if this process does not converge, the model itself should probably be tested for economic consistency.

The Newton algorithm on the contrary puts no condition on the numerical intensity of relationships, but on its variability (in practice the stability of the Jacobian and in a lesser measure the Hessian). This condition is generally verified in the case of econometric models. In particular, the convergence will be immediate for a linear model showing a unique solution. The Newton algorithm will then be advisable for models presenting convergence difficulties, or for model builders wanting to free themselves from this concern, which remains always present and sometimes crucial (for instance if they frequently face deadline problems).

On the other hand, considering the speed of convergence, the choice is much less obvious. If the Newton method converges in less iterations, generally between three and five (remember that convergence accelerates with iterations), each iteration is far more expensive. In fact this supplement can be measured, in the case of a derivation using finite differences: each iteration takes as many Gauss-Seidel iterations as the model contains loop variables (plus one), as well as a matrix inversion (growing with the dimension as a third degree polynomial). One can therefore minor the total cost, whatever the method, by a number of Gauss-Seidel iterations.

In practice one observes that the number of iterations necessary for convergence, whatever the method employed, grows slowly enough with the size of the model. On the other hand the number of loop variables is more or less proportional to it.

We shall give now the time taken by each of the methods (in case of success):

- Gauss – Seidel
- Newton with analytical derivations
- Newton with numerical approximation
- Broyden with analytical derivations
- Broyden with numerical approximation.

The test was conducted on a Toshiba Qosmio F60-10J, using an Intel I5 chip and 4 GB memory.

For 10 models of growing size, all built or contributed by the author during cooperation projects and seminars:

- France : a very small French model, resembling the one we are using here
- France cf : a small French model with complementary factors

- France cd : a small French model with Cobb-Douglas. Both these models will be presented later.
- Vietnam : a small Vietnamese model with one product
- Algeria 1 : a small Algerian model with one product (+oil in some cases)
- Algeria 5 : a five product Algerian model
- Algeria QAM : a 19 product Quasi-Accounting Algerian model
- Vietnam : a 3 product Vietnamese model
- World : a 12 country world model
- China : a 3 product, 4 region model

The first three models are available in the files accompanying the book.

With the following characteristics:

- Prod : number of products
- Coun : number of products or regions
- Per : number of periods
- Equ : number of equations
- Block size : size of the largest non – recursive block
- Feedback : number of feedback (or loop) variables

Countries	prod	coun	per	equ	Block s	Feedba	Gauss-Seid	Newton A	Newton N	Broyden A	Broyden N
France	1	1	q	18	13	2	0,002	failed	failed	0,003	0,004
France cf	1	1	q	87	47	2	0,010	0,018	0,030	0,014	0,014
France cd	1	1	q	88	51	3	0,020	0,025	failed	0,017	0,017
Vietnam	1	1	a	94	55	4	0,030	0,026	0,045	0,019	0,020
Algeria	1	1	a	149	85	7	0,028	0,026	0,036	0,023	0,029
Algeria	5	1	a	630	408	23	failed	0,169	0,237	0,145	0,146
Algeria QAM	19	1	a	1361	922	19	failed	0,312	0,433	0,261	0,234
World	1	12	q	1396	1063	71	0,087	0,084	0,082	0,083	0,083
Vietnam	3	3	a	1858	1260	37	2,446	2,008	2,522	failed	failed
China	3	4	a	2648	1900	45	failed	failed	failed	0,349	0,321

One can observe:

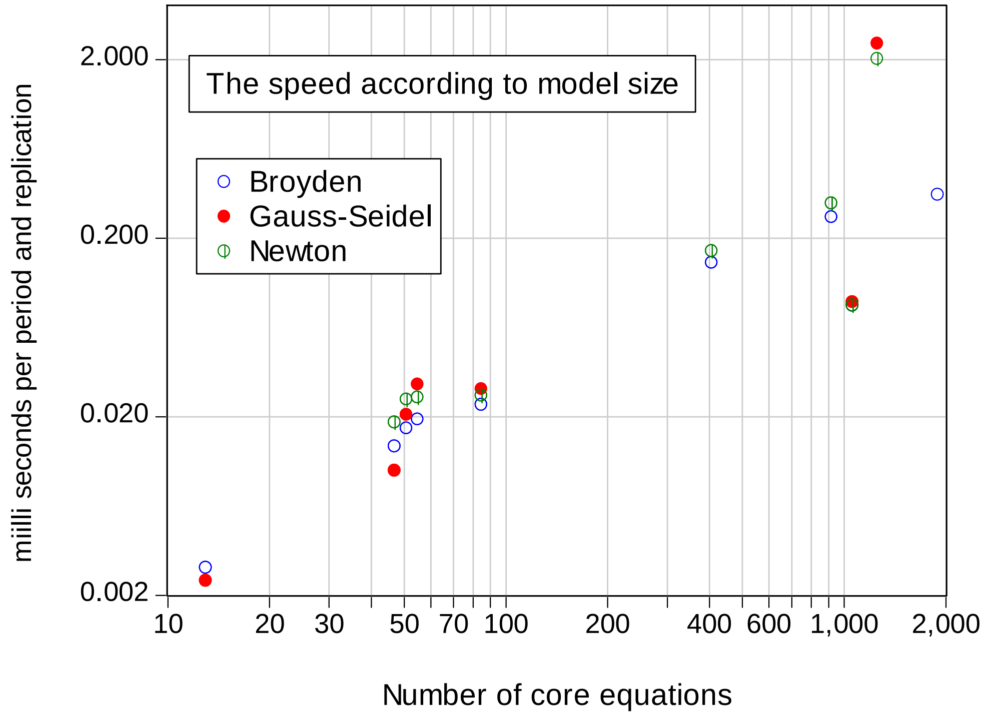
- The bad news is that each method fails at least once, the good that at least one works for any model. If this was true in all cases, convergence would be guaranteed.
- Gauss-Seidel works better for small models. This is quite contrary to expectations, as the cost of inverting a matrix of growing size should become more and more expensive as size grows. There might be a cost of entry for Newton.
- Broyden is globally better for large models, for no particular reason. However, it does fail once. But it is the only one working for the Chinese model.
- For the large Vietnamese model, Broyden fails, and Newton is better than Gauss-Seidel.

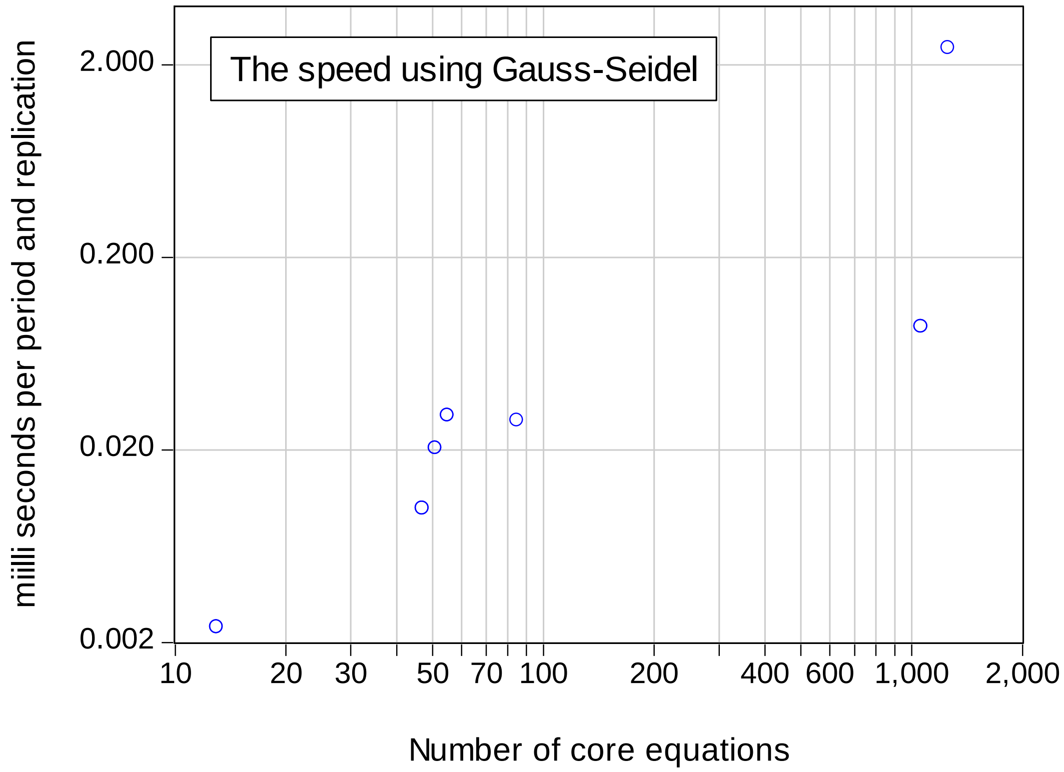
Let us now consider the link between model size and computation cost.

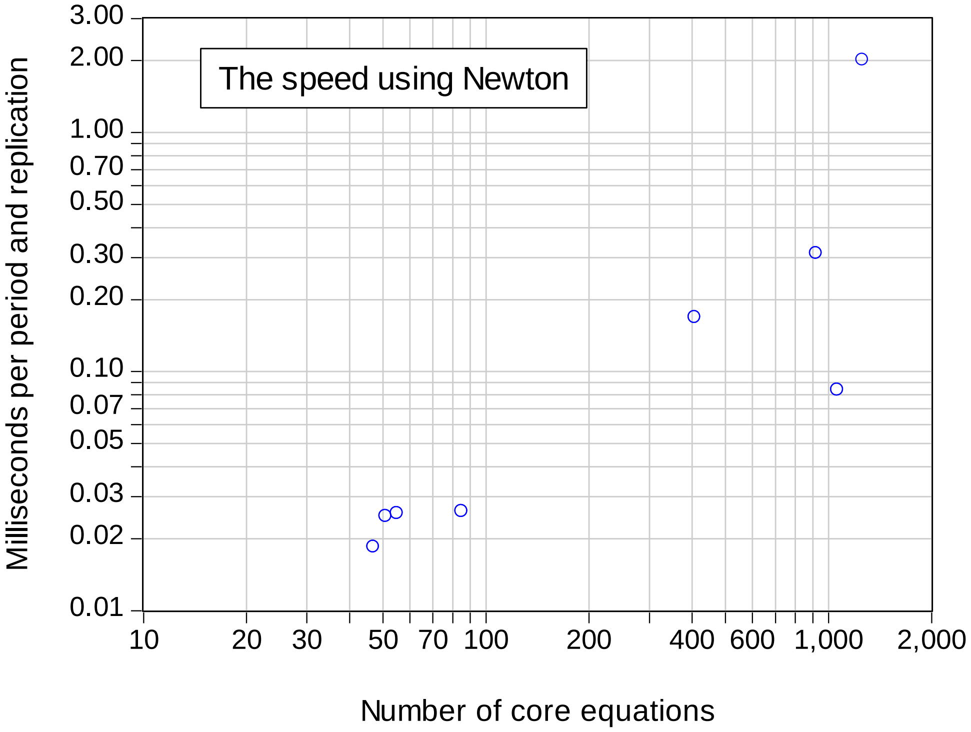
We shall present four graphs, all of them considering the logarithm of both variables:

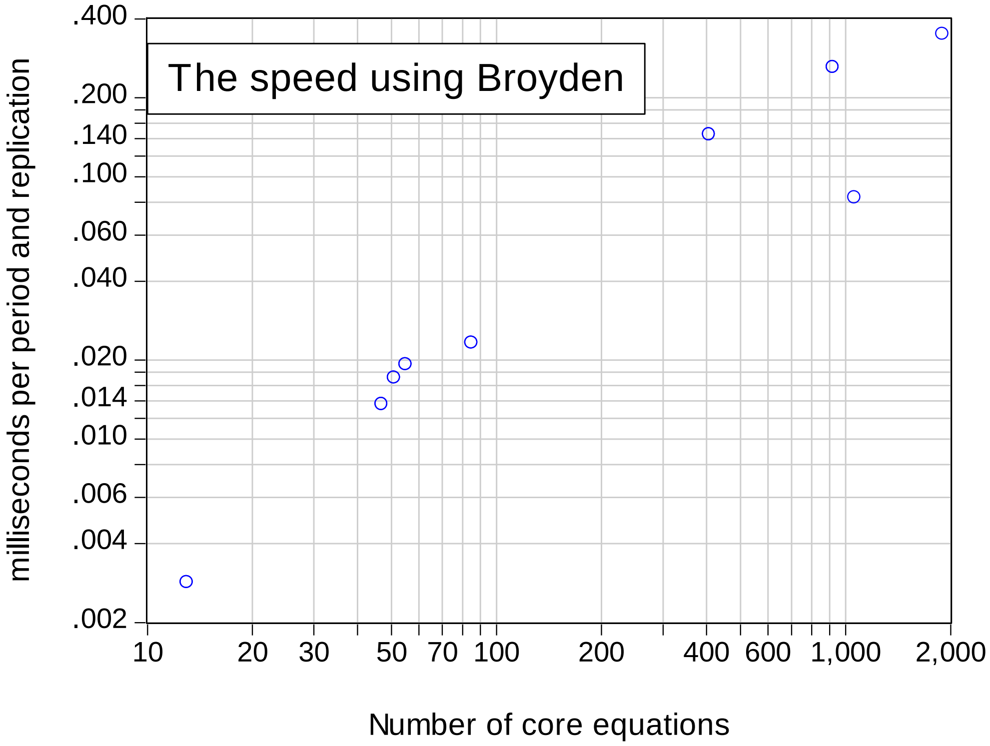
- A global one for all methods
- Three individual ones dedicated to each method.

They show that the link is globally log-linear, with a slight downward inflexion for the "inversion" methods. As to the elasticity itself, a separate estimation gives a value close to 1, and a global one a value of .99 (with a T-stat of 14.7).





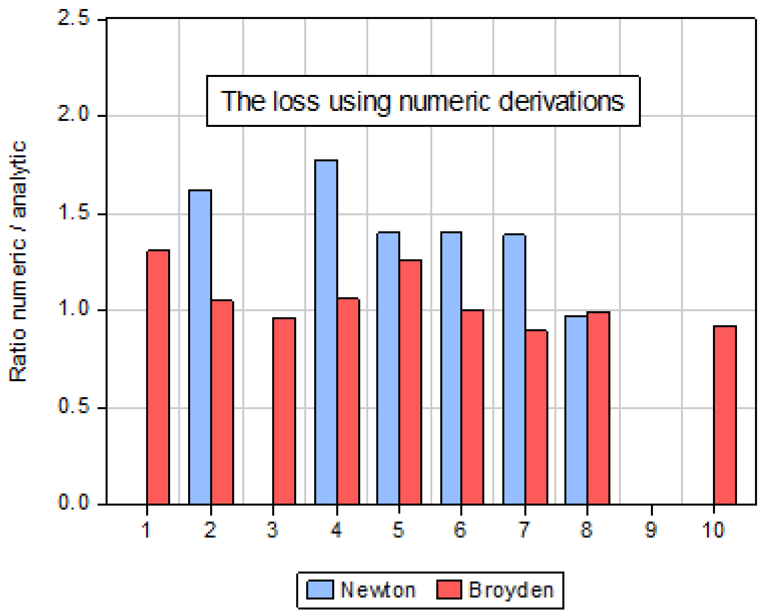




- The two last graphs compare analytical and numerical methods, for both related algorithms.
- The diagnosis is complemented by an histogram presenting the ratio of costs (when both options led to convergence).

They show clearly :

- That establishing analytical derivatives is quite efficient for the Newton method, with an average gain of about 40%.
- That for Broyden (where the derivatives are only computed once) there is no gain, but no loss either.



Countries	Newton analytic	Newton numerical	Newton ratio	Broyden analytic	Broyden numerical	Broyden ratio
France	failed	failed	failed	0,00	0,00	1,31
France cf	0,02	0,03	1,62	0,01	0,01	1,05
France cd	0,02	failed	failed	0,02	0,02	0,96
Vietnam	0,03	0,05	1,78	0,02	0,02	1,06
Algeria	0,03	0,04	1,40	0,02	0,03	1,26
Algeria	0,17	0,24	1,40	0,14	0,15	1,01
Algeria QAM	0,31	0,43	1,39	0,26	0,23	0,89
World	0,08	0,08	0,97	0,08	0,08	0,99
Vietnam	2,01	2,52	1,26	failed	failed	failed
China	failed	failed	failed	0,35	0,32	0,92

7.1.7 SOLVING THE MODEL: BASIC EVIEWS TECHNIQUES

The previous program has made model data and equations consistent. The present one will solve it.

7.1.7.1 The syntax

Again, we shall SOLVE the model, but this time the options are more numerous.

Solving the model will generally be done in a program. However, one can use the menus, first by accessing the model item, then using the “solve” item, which gives access to all the features we are going to present.

In a program, these options can be set before any computation, using the SOLVEOPT statement, or at simulation time (SOLVE).

The syntax is the following:

```
model-name.solveopt(options)
```

Or

```
solve(options) model-name
```

Where the options are (taken from the EViews Help file):

m= (default=5000) Maximum number of iterations for solution (maximum 100,000).

c= (default =1e-8) Convergence criterion. Based upon the maximum change in any of the endogenous variables in the model. You may set a number between 1e-15 and 0.01.

s=arg Solution type: "d" (deterministic, default), "m" (stochastic - collect means only), "s" (stochastic - collect means and s.d.), "b" (stochastic - collect means and confidence bounds), "a" (stochastic - collect all; means, s.d. and confidence bounds).

d=arg Model solution dynamics: "d" (dynamic solution, default), "s" (static solution), "f" (fitted values - single equation solution). Forecasting calls for a dynamic solution, which is the default option. There are few reasons to choose another (except d=f as above).

n=arg NA behavior: "n" (stop on "NA" values), "t" (do not stop when encountering "NA" values). Only applies to deterministic solution; EViews will always stop on "NA" values in stochastic solution.

a=arg Alternate scenario solution: "t" (solve both active and alternate scenario and collect deviations for stochastic), "f" (default, solve only the active scenario).

o=arg Solution method: "g" (Gauss-Seidel), "n" (Newton), "b" (Broyden).

i=arg For the initialization values: “a” (actual), “p” (previous period solution). The default is “a” if the values are available, “p” otherwise. In forecasts, “p” should be applied, but one can initialize the endogenous which gives the two options (with “a” as the default).

For stochastic simulations:

r=integer (default=1000) Number of stochastic repetitions (used with stochastic "s=" options).

b=number (default=.95) Size of stochastic confidence intervals (used with stochastic "s=" options).

7.1.7.1.1 The suffix

In addition to the above options, EViews allows the user to define the name of solution variables.

Of course, one should not consider using the actual names as such. This will destroy the information on the original values. However, the connection must be as easy as possible. The obvious solution is to use the original names, but modify them by some addition: a given prefix or suffix.

This will call for a specific statement:

```
model_name.assign @all suffix
```

This statement will add to the text of the model *model_name* the statement:

```
assign @all suffix
```

where *suffix* is a string.

For instance,

```
_fra_1.assign @all _b
```

will give to the variables obtained by the next model solution a name composed of the original variable followed by the suffix “_b”. In this case, the solution for Q will be called Q_B.

7.1.7.2 Problem processing techniques

Like all packages of this type, EViews does not always reach a solution. We will now provide a typology of techniques one can use for solving convergence problems, focusing on the ones available under EViews, and the way to apply them.

In all cases one must be certain that the residual check shows no error in the first place. This is essential, and very cheap (at least in the observation of errors, and the identification of their location).

7.1.7.2.1 The tools provided by EViews

In addition to changing model specifications, EViews provides a number of specific tools. Some of these can only be specified through menus, but they remain active if one runs a program in the same session.

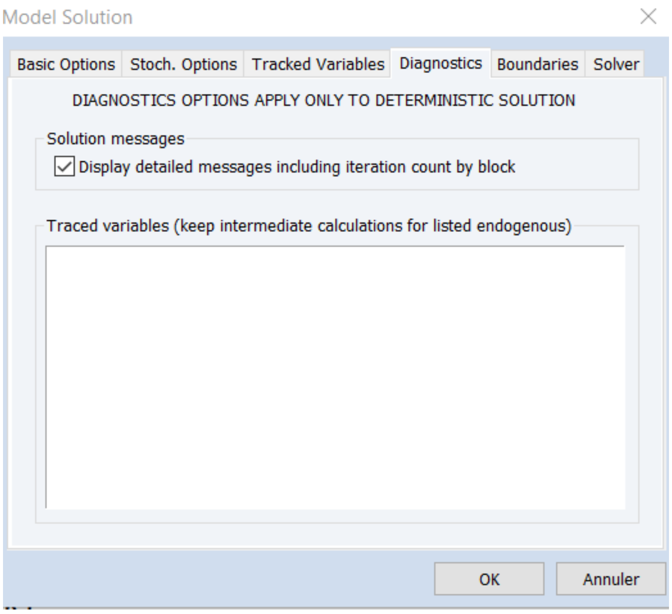
- Displaying the number of iterations necessary for convergence (only useful of course if the model converges on part of the sample). But if the model starts diverging after a few periods, it is interesting to observe if the process is gradual (the number of iterations is growing until breakdown) or sudden (the number is stable until the crash).

This is obtained by

- double-clicking the model item
- using:

solve>diagnostics

The following window opens:



We tick the box for “displaying detailed messages including iteration count by block”.

When we solve the model, a window will open displaying the number of iterations needed for solving each block (several messages per period if the model contains more than one non-recursive block).

This window will remain open after solving, and one can scroll back and forth, to the first message if needed.

Model: _FRA_1
Date: 03/23/20   Time: 11:57
Sample: 2002Q1 2004Q4
Solve Options:

Dynamic-Deterministic Simulation
Solver: Gauss-Seidel
Max iterations = 5000, Convergence = 1e-08
Scenario: Baseline
Solve begin 11:57:47

2004Q1	Block 1 - 3 eqns	Solved (recursive block)
2004Q1	Block 2 - 14 eqns	Convergence after 12 iterations
2004Q1	Block 3 - 3 eqns	Solved (recursive block)
2004Q2	Block 1 - 3 eqns	Solved (recursive block)
2004Q2	Block 2 - 14 eqns	Convergence after 12 iterations
2004Q2	Block 3 - 3 eqns	Solved (recursive block)
2004Q3	Block 1 - 3 eqns	Solved (recursive block)
2004Q3	Block 2 - 14 eqns	Convergence after 12 iterations
2004Q3	Block 3 - 3 eqns	Solved (recursive block)
2004Q4	Block 1 - 3 eqns	Solved (recursive block)
2004Q4	Block 2 - 14 eqns	Convergence after 12 iterations
2004Q4	Block 3 - 3 eqns	Solved (recursive block)

Solve complete 11:57:47

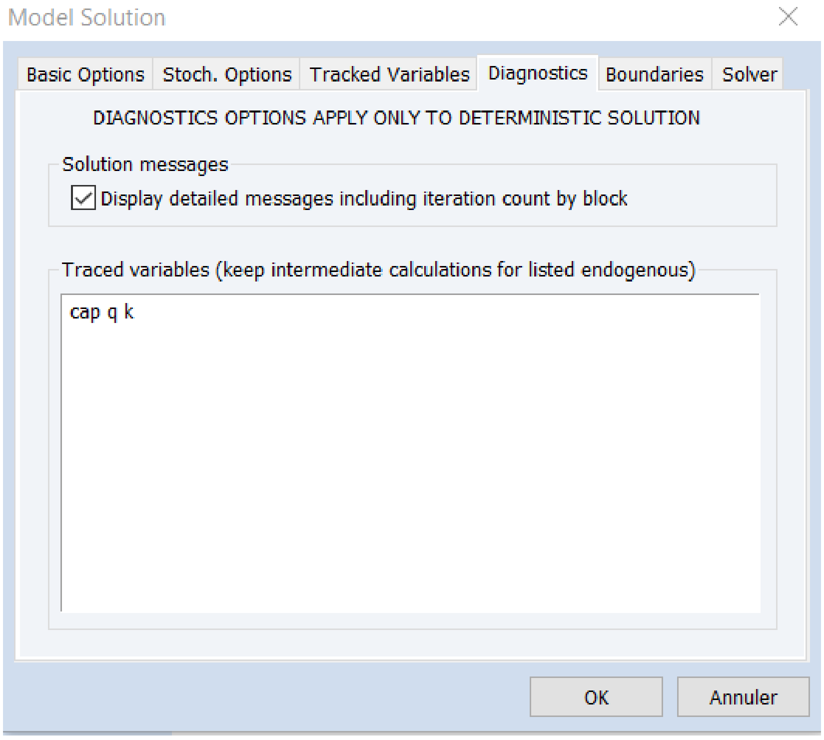
- Displaying the values obtained at each iteration for a set of variables. Of course, following the convergence process is only practical using the Gauss-Seidel method.

This is again obtained by

- double-clicking the model item
- using:

```
solve>diagnostics
```

- Introducing the names of the variables in the window.



Alternatively, the same can be obtained in a program, using:

```
model_name.trace list_of_variables.
```

This can be useful if you select a given set of important elements, which will then be available at all times after any simulation.

- After solving, using:

View>trace output

A window will open with the values of the variables at each iteration, for each period (one column of values per variable).

Date	Block	Iteration	Objective	CAP	Q	K
2004Q1	1	1		4.19E+11		
	2	0			3.54E+11	
		1	0.022360		3.54E+11	
		2	0.020004		3.54E+11	
		3	0.041031		3.54E+11	
		4	0.004830		3.54E+11	
		5	0.000891		3.54E+11	
		6	0.000232		3.54E+11	
		7	3.16E-05		3.54E+11	
		8	3.90E-06		3.54E+11	
		9	4.79E-07		3.54E+11	
		10	5.87E-08		3.54E+11	
		11	7.20E-09		3.54E+11	
	3	1				2.34E+12
2004Q2	1	1		4.18E+11		
	2	0			3.56E+11	
		1	0.020415		3.56E+11	
		2	0.022601		3.55E+11	
		3	0.097650		3.55E+11	
		4	0.010504		3.55E+11	
		5	0.001899		3.55E+11	
		6	0.000479		3.55E+11	
		7	6.31E-05		3.55E+11	
		8	7.52E-06		3.55E+11	
		9	8.90E-07		3.55E+11	
		10	1.05E-07		3.55E+11	
		11	1.25E-08		3.55E+11	
		12	1.48E-09		3.55E+11	
	3	1				2.35E+12
2004Q3	1	1		4.20E+11		
	2	0			3.58E+11	
		1	0.020342		3.58E+11	
		2	0.025339		3.57E+11	
		3	0.137144		3.57E+11	
		4	0.013647		3.57E+11	
		5	0.002419		3.57E+11	
		6	0.000588		3.57E+11	
		7	7.47E-05		3.57E+11	
		8	8.53E-06		3.57E+11	
		9	9.69E-07		3.57E+11	
		10	1.10E-07		3.57E+11	
		11	1.25E-08		3.57E+11	
		12	1.42E-09		3.57E+11	
	3	1				2.37E+12
2004Q4	1	1		4.24E+11		
	2	0			3.60E+11	
		1	0.025946		3.60E+11	
		2	0.028302		3.59E+11	
		3	0.153634		3.59E+11	
		4	0.014515		3.59E+11	
		5	0.002533		3.59E+11	
		6	0.000595		3.59E+11	
		7	7.31E-05		3.59E+11	
		8	8.05E-06		3.59E+11	
		9	8.82E-07		3.59E+11	
		10	9.65E-08		3.59E+11	
		11	1.06E-08		3.59E+11	
		12	1.16E-09		3.59E+11	
	3	1				2.38E+12

If the variable belongs to a recursive block, only one value will be displayed. In our case, the unique values for CAP and X will preceded the iterative process description, then the unique K value.

The first value (iteration 0) is the initial one.

The columns follow the order used in the solving process, not the one in which the list is specified. This is quite useful for following the process.

But the interpretation of the evolutions as a column of figures is often difficult, in particular if one tries to evaluate the presence and nature of cycles. It is better to transform these figures into graphs. This can be done in two ways:

- Selecting with the mouse the requested zone (for instance the values for a given period) and copying it into a graphic software such as Excel.

- The same can be done inside EViews but it is a little more complex. One must:

  - Select with the mouse the requested zone
  - Create variables associated with the displayed elements (with different names, for instance by inserting a common prefix).
  - Display the variables as a table filled with "na" values.
  - Copy the selection into the sheet.
  - Apply View>Graph.

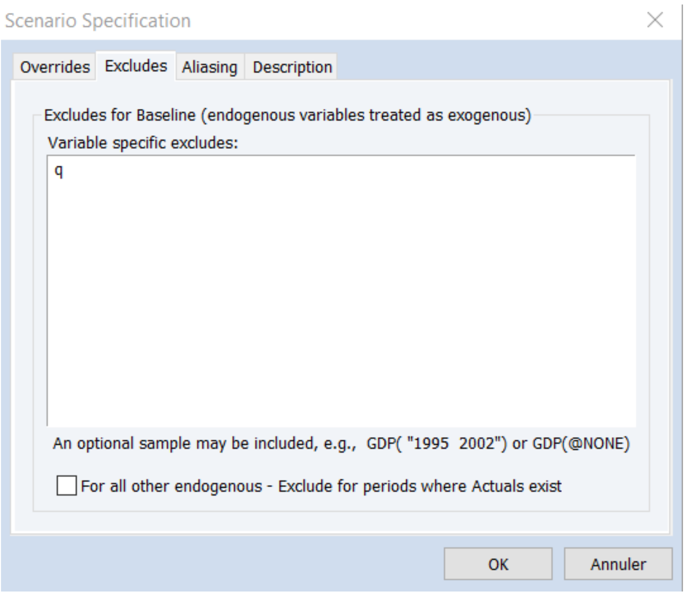
- Excluding equations from the solving process.

This is very easy. Using **menus**, one has to

  - access the model,
  - use:

Basic options>Edit scenario options

In the "exclude series" block, specify to series to exogenize.



Using a **program**, one needs the statement:

```
model-name.EXCLUDE list-of-series
```

The corresponding equations will not be applied, and the variables will retain their initial value.

Note: the list will remain active for all subsequent solutions produced under the same scenario, until it is modified using a different list. Reverting to the full model calls for a statement with a blank list, which should be stated immediately.

One can consider eliminating a single variable, like the change in inventories (if it seems to introduce an explosive process), or a full set, like all the deflators in the model (if the price system seems to explode, or to disturb the computation of the real sector).

However, one must be aware that if an identity variable is excluded, the identity is no longer enforced. Balancing variables should never be excluded. For instance, excluding Q makes supply no longer equal to demand (and the model recursive…).

7.1.7.2.2 Other tools

Some tools can be considered by the user:

7.1.7.2.2.1 USING DAMPING FACTORS

This has been described earlier. Technically, you will have to rewrite the equation. For instance

```
(4) Q + M = FD + X
```

will become:

```
(4) Q  = rel_q * (FD + X - M) +(1-rel_q) * Q
```

where damping is enforced by making rel_q different from 1 (lower in principle).

It might be useful to set the factor as a series, even if it will be constant over time. To apply a change to a scalar, the model needs to be recompiled (by UPDATE).

Also, this method is a little different from the one described earlier. Application of the factors is not done at the end of the iteration, but at the time the variable is computed. If the variable is used later in the ordering, this affects the convergence process.

7.1.7.2.2.2 REORDERING THE MODEL

Using Gauss-Seidel, reordering will change the solving process. Not for Newton as the Jacobian is based on derivation of individual equations.

7.1.7.2.2.3 CHANGING PARAMETER VALUES

Changing the values of parameters will obviously modify the solving process. In particular, a coefficient which looks too high can be reduced or even eliminated. This could be the case for instance if the intensity for the short-term accelerator effect in the investment equation.

This technique can be used to eliminate a specific influence of a variable, leaving unchanged its other roles.

This calls for a good knowledge of the properties of the model.

7.1.7.2.2.4 CHANGING EQUATION SPECIFICATIONS

A mechanism which seems to create problems can be replaced by another. For instance, the purchasing power parity assumption for the exchange rate can be smoothed or replaced by a fixed nominal rate.

7.1.7.2.3 Applying these tools

Let us now propose a methodology in case the model fails in reaching a solution.

In all cases, you should move to menus. This environment will give much more information and provide many more tools.

7.1.7.2.3.1 THREE SIMPLE CASES

First, let us consider three specific cases, requiring simple treatment.

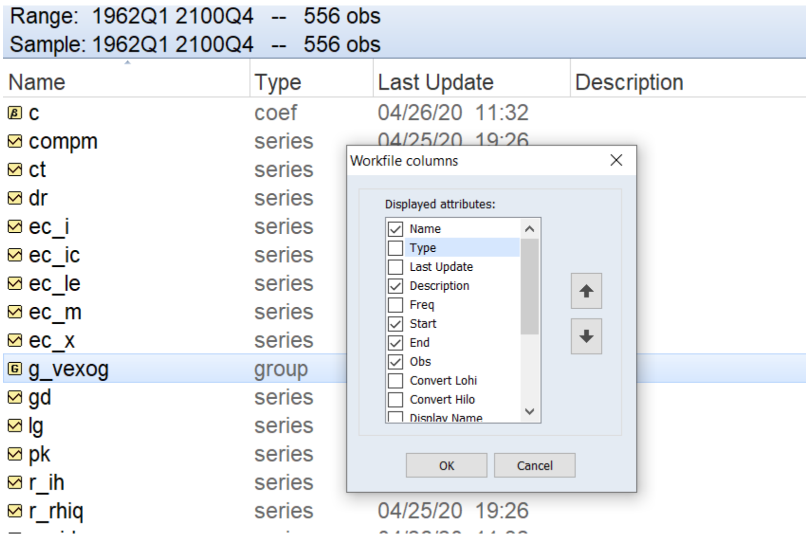
- Some data the process requires is not available.

```
Model: PIC_B
Date: 04/25/20   Time: 19:28
Sample: 1975Q1 2000Q4
Solve Options:
    Dynamic-Deterministic Simulation
    Solver: Gauss-Seidel
    Max iterations = 5000, Convergence = 1e-08

Scenario: Scenario 1
Solve begin 19:28:03
    1975Q1   NA generated for CAP
Solve terminated - Unable to compute due to missing data  in "CAP  = PK  * K(-1)".
```

It should be filled or you should change the period.

One frequent case: you forgot to establish an assumption (compute an exogenous variable) on the future.

If you created a group for the exogenous, you just have to display its values. If there is apparently more than one series missing, the best strategy is to access the group (perhaps copying the series to a specific page) and use "Edit Columns" to call for the starting, ending and number of observations of the series.



Name	Descrip...	Start	End	Obs
c				
compm		1978Q3	2100Q4	490
ct		1963Q1	2100Q4	552
dr		1963Q2	2100Q4	551
ec_i		1965Q1	2100Q4	544
ec_ic		1964Q2	2100Q4	547
ec_le		1963Q2	2100Q4	551
ec_m		1962Q1	2010Q4	196
ec_x		1962Q1	2010Q4	196
g_vexog				
gd		1963Q1	2100Q4	552
lg		1963Q1	2100Q4	552
pk		1977Q3	2100Q4	494
r_ih		1970Q1	2100Q4	524
r_rhiq		1970Q1	2100Q4	524
resid				0
sr		1970Q1	2100Q4	524
t		1962Q1	2100Q4	556
wd		1974Q1	2100Q4	508

You can see immediately that variables EC_M and EC_X are missing.

If there are too many variables to trust eye control, you can

- Freeze the table.
- Sort it according to the column (D in our case).

The shorter series will appear at the top.

- You made technical mistakes.
  - Your model is not consistent with the data. Of course you have performed a residual check, but maybe something has changed since. One should perform it again.
  - The exogenous assumptions have been given abnormal values.

For instance with WD=0 we get:

Solve terminated - Log of non positive number in "DLOG(X) = @COEF(1) * DLOG(WD) + @COEF(2) + @COEF(3) * RES_X(-1) + EC_X"

- The initial values are abnormal, and the model crashes before it reaches an acceptable region.

To solve the problem, it will be necessary in this case to retrace the causal chaining

For instance if WR is 1000 too high due to wrong units, we get:

Model: _FRA_1
Date: 03/23/20 Time: 15:30
Sample: 2000Q1 2002Q4
Solve Options:
    Dynamic-Deterministic Simulation
    Solver: Gauss-Seidel
    Max iterations = 5000, Convergence = 1e-08

Scenario: Baseline
Solve begin 15:30:59
2000Q1 Block 1 - 3 eqns Solved (recursive block)
2000Q1 NA generated for LE - Log of non positive number
2000Q1 Block 2 - 14 eqns No convergence after 4 iterations
Solve terminated - Log of non positive number

And we can follow the evolution of variables in the "Trace output" display:

Date	Block	Iteration	Objective	X	RHI	CO	TD	M	Q
2000Q1	1	1		9.29E+10					
	2	0			2.01E+11	1.83E+11	6.71E+11	8.78E+10	3.35E+11
		1	910.3030		4.33E+13	1.00E+13	1.34E+12	1.23E+11	9.92E+11
		2	2929.029		2.22E+14	2.02E+14	2.21E+14	3.59E+14	-1.39E+14
		3	330.6255		2.11E+14	1.93E+14	7.34E+13	3.59E+14	-1.73E+14
		4	1.452647		2.07E+14	1.89E+14	-1.13E+13	3.59E+14	-1.98E+14

We can observe the following sequence, starting from historical values (Iteration 0):

At iteration 1, household revenue gets much too high, consumption follows, then total demand, imports and GDP.

At iteration 2, imports becomes higher than demand, and GDP becomes negative at the end of the iteration.

At iteration 3, the employment equation uses this negative value in a logarithm, and the process breaks down; as the first message indicated.

- The maximum number of iterations has been reached.

You should specify a higher number (say 1000, or even 100 000) until convergence. You do not want to miss a late convergence, which corresponds to a solution. Time is no longer an issue, except If you are producing stochastic or forward looking estimations, but you can check an other solving algorithm, or a different ordering for .

The absence of convergence with a high number is exceptional. It can only come from oscillations between two very close values of a single variable (purely numerical problem due to the precision of computer computations) in which case the chosen threshold is too small.

For instance, a variable which can change signs, which means that it is computed as a difference (like the change in inventories or the trade balance) takes values much lower than its components, and loses precision digits in the process. The same is true also for growth rates, when the associated element is stable over time.

In practice, this should happen only with Gauss-Seidel. Again, try another algorithm and consider the result. Either it is acceptable, or the model crashes. In that case, the next case applies.

7.1.7.2.3.2 THE SOLUTION CRASHES

Let us now consider the main case : **the algorithm crashes.**

In practice, EVIews will stop computations in only two cases, in addition to reaching the maximum number of iterations.

- The values of a variable become too high.

This will happen essentially with Gauss-Seidel: the spider web process we have presented earlier leads away from the (existing) solution.

However, it can also happen with Newton (and Broyden less clearly), and much faster. These algorithms produce a sequence of linearizations and solutions of the linearized model, supposedly more and more precise until it corresponds to the solution of the model itself. If this linearization does not gain precision it can only lose it, and quite fast (except in the very unlikely case of oscillations) due to the quadratic nature of the algorithm, which applies both to convergence and divergence. The associated solution will fast enter a domain with unacceptable values, most likely as to their sign (negative GDP or unemployment?) which will enter a logarithm when the Jacobian is computed. The algorithm will crash.

But of course reaching unacceptably high values is also possible, for instance when the Jacobian becomes almost singular.

Now, the highest value EViews will consider is at this time is more than 10^{308}. Obviously no economic variable can take this value (even for a country with explosive inflation).

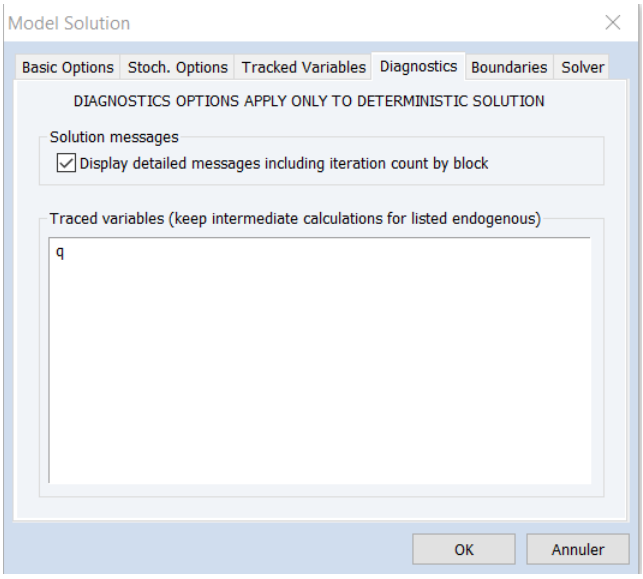
Let us consider the following model (using the names from the small model):

```
Q + M = FD + X

 CO = 0.8 * Q

 I = 0.5 * d(Q)

 FD = CO + I + g

 M = 0.2 * FD
```

With X and g exogenous.

This model does not seem so crazy. But one can observe (as in a previous version of the example) that increasing g by 1 has a secondary impact of 1.1 (0.8 + 0.5 – 0.2). Using Gauss-Seidel the process will explode. This is what will happen in practice.

We will ask for the display of iteration numbers, and the sequence of values for the variable q, using:



And we will apply the Gauss-Seidel algorithm. We get:

Model: __MOD_1
Date: 04/26/20 Time: 16:05
Sample: 2020Q1 2100Q4
Solve Options:
  Dynamic-Deterministic Simulation
  Solver: Gauss-Seidel
  Max iterations = 100000, Convergence = 1e-08

Scenario: Baseline
Solve begin 16:05:57
  2020Q1 NA generated for FD - Overflow
  2020Q1 No convergence after 19977 iterations
Solve terminated - Overflow

And the sequence of values for Q (obtained by clicking on "Trace output")

Date	Iteration	Objective	Q
2020Q1	0		5.19E+11
	1	1.111003	5.23E+11
	2	1.013193	6.20E+11
	3	14.02256	7.20E+11
	4	1.016646	7.99E+11
	5	0.540969	9.05E+11
	6	0.470440	1.02E+12
	7	0.384636	1.13E+12
	8	0.330716	1.26E+12
	9	0.293828	1.38E+12
	10	0.266657	1.51E+12

19971	0.080861	1.2E+308
19972	0.082774	1.3E+308
19973	0.080861	1.3E+308
19974	0.082774	1.4E+308
19975	0.080861	1.4E+308
19976	0.047738	1.4E+308
19977	0.003812	1.4E+308

As you can see, the high value accepted by EViews calls for a very high number of iterations for the algorithm to crash.

We have suggested using another algorithm. With Newton we get:

Date	Iteration	Objective	Q
2020Q1	0	1.48E+11	5.19E+11
	1	0.000774	-1.90E+12
	2	0.000774	-1.90E+12
2020Q2	0	1.16E+12	5.22E+11
	1	0.005610	-2.61E+13
	2	0.005610	-2.61E+13
2020Q3	0	1.33E+13	5.25E+11
	1	0.089759	-2.68E+14
	2	0.089759	-2.68E+14

As the model is linear, it converges in one iteration (the second is needed to check convergence).

Solving the model using another algorithm has proven one very important thing: your model has a solution. But this can have two outcomes, which you will have to separate yourself.

- The solution is acceptable: the algorithm is at fault, or you used it inappropriately. For instance, you used Gauss-Seidel with an illogical ordering. Switching to a successful method can be done indefinitely for this particular model.
- The solution is wrong (as in the present case). You have not solved the problem, but you have the elements to do so, even if some research will be necessary.

Of course, we could have used another of the techniques presented earlier (damping factors, exogeneizing variables) but changing the algorithm was both faster and more informative

- The second error type(s) will call for a slightly modified model.

1. Q + M = FD + X
2. FD = CO + I + g
3. Log(CO) = Log(Q) + res_co
4. I = 0.50 * d(Q)
5. M / FD = r_m

The use of Log in Equation 3 is the main difference from the previous version.

In the same context as before, we get two different error messages, according to the algorithm.

In the example we are going to develop, we get from Gauss-Seidel.

Model: __MOD_1
Date: 04/26/20 Time: 18:33
Sample: 2020Q1 2030Q4
Solve Options:
Dynamic-Deterministic Simulation
Solver: Gauss-Seidel
Max iterations = 100000, Convergence = 1e-08

Scenario: Baseline
Solve begin 18:33:04
2020Q1 Convergence after 64 iterations
2020Q2 Convergence after 59 iterations
2020Q3 Convergence after 67 iterations
2020Q4 Convergence after 60 iterations
2021Q1 Convergence after 76 iterations
2021Q2 Convergence after 61 iterations
2021Q3 NA generated for CO - Log of non positive number
2021Q3 No convergence after 7 iterations
Solve terminated - Log of non positive number

And from Newton:

Model: __MOD_1
Date: 04/26/20 Time: 19:04
Sample: 2020Q1 2030Q4
Solve Options:
Dynamic-Deterministic Simulation
Solver: Newton
Max iterations = 100000, Convergence = 1e-08

Parsing Analytic Jacobian:
12 derivatives kept, 0 derivatives discarded

Scenario: Baseline
Solve begin 19:04:31
2020Q1 Convergence after 2 iterations
2020Q2 Convergence after 1 iterations
2020Q3 Convergence after 2 iterations
2020Q4 Convergence after 2 iterations
2021Q1 Convergence after 2 iterations
2021Q2 Convergence after 2 iterations
2021Q3 No convergence after 16 iterations
Solve terminated - Solver has stalled: starting values may be too far from a solution

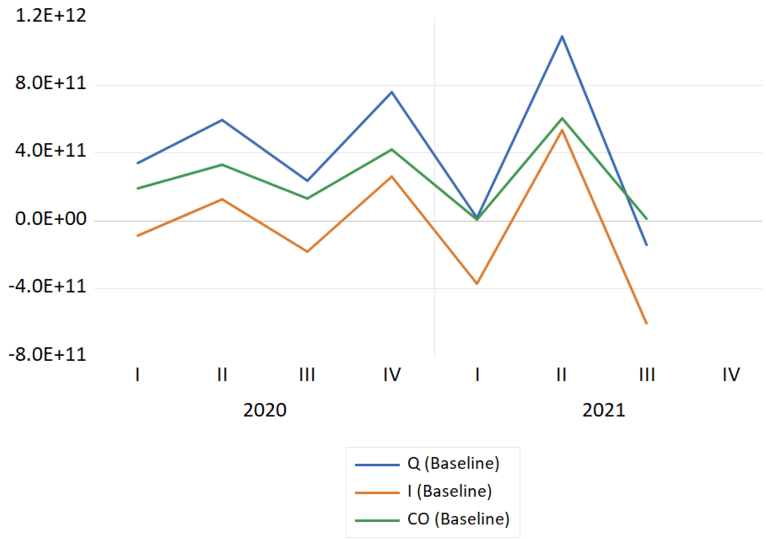
One can see that Newton take much less iterations, but both crash at the same period (after 6 “successes”).

Now the situation is both better and worse.

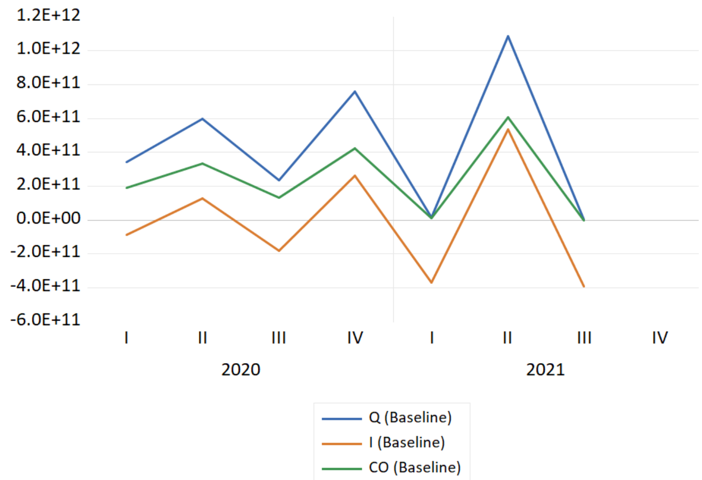
- Both algorithms fail very quickly, so switching algorithms is not relevant.
- But they succeed in the beginning, which gives us additional information..

This information will be put to good use.

Let us draw the graph of solution values, first for Gauss-Seidel:



Then Newton:



As expected, the two algorithms give the same values (the model solution) except for the erroneous period.

Let us consider the links between the elements, by considering derivative of the system.

1. D(Q) + D(M) = D(FD)
2. D(FD) = D(CO) + D(I)
3. D(CO)/CO= D(Q)/Q or D(CO) = 0.56 D(Q).
4. D(I) = 0.50 D(Q) – 0.50 D(Q(-1))
5. D(M) = r_m D(FD) or D(M) = 0.29 D(FD)

In summary

D(Q) = 0.71 D(FD) = 0.71 . (0.55 D(Q) + 0.50 D(Q)) – 0.50 D(DQ(-1))

D(Q) = - 0.50 / 0.25 D(Q) = - 2 D(Q)

We observe the same diverging process as in the graphs. The amplification of the change in Q brings the solution to a negative value.

We can see that the problem does not lie with the consumption equation, rather with investment. This means that facing this kind of problem we might have to retrace the causality process, perhaps by several steps.

For Gauss-Seidel, an efficient method (not provided by EViews) to understand the flaws in the convergence process consists in changing the algorithm to Ritz-Jordan, which takes into account only values of the previous iteration, eliminating interferences and allowing to improve individualizing divergent processes. At the beginning of a process, this method individualizes immediately the faulty equation (less efficiently than a residual check, however).

This example has helped us to understand the problem a little better. Let us now treat the general case.

A first statement: to solve the problem, it is essential to solve the model, even by doing it in a non-operational environment.

Basically, the model builder can face four situations. The model can:

- Never converge, using any algorithm.
- Converge only for a few periods.
- Converge for the whole period considered, but with unacceptable properties.
- Converge for the whole period, with acceptable properties.

The last situation never occurs at the first try. In forty years, it occurred to me once. And I consider myself as a specialist.

The second and third differ only by the number of periods for which the results provide information.

As to the first case, it should be avoided much as possible, using the approach we propose, as we get no results at all (except on the way divergence occurs). But this is easier said than done, of course.

Let us state also that when the model fails, there is a good chance the current model has no solution for that period, as it presents two relatively independent flaws:

- The iterative process is not converging using Gauss-Seidel. If the ordering of the equations reproduces an economic equilibration, this is not working: starting in the vicinity of a given solution does not lead to another solution, but away from it.
- The linearization of the model provides the Newton algorithm with an image with very different properties from what we expect.

Now, to make the model converge, you can apply the following techniques.

- First, you should control extensively the text of your equations (including the parameter values), the status of your variables (endogenous/exogenous), and the block structure. It is better if somebody outside the modelling team can contribute.
- You can drop one or several equations, and try solving the model again. The choice depends on the variables which seem to lead the divergence: either the ones which actually take the worst values, or the elements which influence them. By forcing a variable to take a given value, one can often measure its contribution to the problem (especially if the variable goes through unacceptable values) by observing if the problem disappears. This method has however the same limit as the others: the exogeneity of a variable destroys all influences forwarded by it, and making the error disappear does not mean we have exogenize the faulty equation.

- You can also change the values of parameters, or drop an individual influence by setting its coefficient to zero. This is different from dropping the variable altogether: in our model, you can drop the role of the rate of use in exports but leave it in imports. By applying variations to some strategic coefficients, chosen either because they play an important role in the model (in the case of a general absence of convergence), or because they appear linked to the diagnosed error (in case of a local error), one might establish a connection between these coefficients and the intensity of the problem. However this method, once again, does not locate necessarily the origin of the problem, which can come from far upstream. And it remains to determine how to use the information, if one excludes setting arbitrarily the coefficient to the value ensuring convergence.
- You can try alternate (probably simpler92) formulations. For instance, you could replace a complex CES function using the relative cost of factors in a sophisticated way, by a very simple complementary factors function. Of course, you can start by exogeneizing the relative cost of factors.

If the argument of a logarithm takes a negative value, you can introduce a test avoiding this situation (for example replacing Log(x) by Log (max (x, a)), with a>0. The system might actually converge to an acceptable value (x>0)[93], but reaching a solution with x<0 will give you information anyway[94].

- You can use other starting points. You should try to guess as well as possible the domain in which you solution will reside, and initialize the variables within it. This is easy on historical periods.
- You can check the derivatives of your equations. You can compute them by hand, or by computations (a little harder, but EViews provides the elementary tools). If you start from quasi-solution values but the gradient is wrong, you can go in the wrong direction, perhaps by a wide margin.
- Finally, you can change the simulation period. It is interesting to know if the problem is general, or appears only for some periods. Examining the particular aspects of the faulty periods (slow growth, strong inflation, important unemployment...) can give useful information on the origins of the error. Using it can however be difficult. One specific case is the role of external trade: in forecasts, too fast an expansion compared to local production and demand can drive the multiplier to abnormal and diverging values, which will disappear if we simulate the model on the first periods available.

Once the model has converged, at least for a while, you find yourself in the general case, and you can start studying the results. But the error correction process includes an additional difficulty, as the final version will have to eliminate the artificial changes you introduced, to make it in fully operational. While an immediate convergence, even for a few periods, guaranteed that the full model had no technical flaw.

The methods we have presented can be used again, but the goal is now to proceed to the last case. And for this we have more information. This will of course be easier with a longer simulation, starting from on a regular base trajectory. In other terms, a forecast, or rather a simulation over the future.

We can also concentrate on:

- As we have done on the small model, observing the way the model is diverging, and which variables (or sets) appear to lead the way. For instance, variables at constant prices might follow an acceptable path, and prices

[92] If the model you want to produce is quite complex, it might be a good idea to start with a simpler version, and introduce layers of complexity one by one.

[93] A rare situation in which the studying the error actually solves the problem.

[94] When things go wrong, the potential for solving the problem is of course quite different if you have a (bad) solution or nothing at all. This means it is generally essential to manage to reach a solution, even in a non-acceptable framework.

explode. Or unemployment might explode while employment is acceptable, driving wages and the partition of value added between workers and firms....

We can see if the divergence is monotonous or cyclical, regular or erratic, sudden or regular.

- Shocking assumptions and looking at the consequences. They can be acceptable in the beginning, which means that the problem lies in the dynamics. They can grow wrong in a regular way, or explode suddenly.

The interpretation and solution of a convergence problem are not simple, essentially because of interferences due to the interactive nature of the system. It will be often necessary to use several methods, and to repeat each of them several times, which will not dispense with a thorough reflection on the logical structure of the model. It is therefore essential for the model builder to know his model well, which will be all the more difficult as its size grows.

Your model might describe a sequence of equilibrating processes, which converge initially, but change through dynamics which makes them diverging after a while. For instance, the expanding effect in the Keynesian multiplier might get higher than one, making the process explosive (the same with the price-wage loop). Or the error-correction process can actually become error-increasing.

In an operational model, a growing call for labor in the productive process can make unemployment become negative in time, and it is used as a logarithm in the wage equation[95]. Or as we have seen earlier, an accidentally very low level of investment can make the associated equation explode suddenly. Or exports growing much slower than imports can make GDP become negative.

- Finally, when one uses the Gauss - Seidel, the computation of eigenvalues of

$$(I - \partial g / \partial y)^{-1}$$

can help to understand the convergence process, and evidence the mathematical characteristics of a possible divergence (intensity and number of divergent processes, presence of cycles). If the highest modulus is higher than one, this can justify the problem.

Processing this information can be done by a gradual elimination of loop influences, analogous to the one recommended by Deleau and Malgrange (1978) to study dynamic model properties. In the best of cases, an association will be obtained between each divergent process and a single loop variable (or two for a couple of conjugate eigenvalues, producing a cyclical process).

But in any event the change must not reduce the theoretical validity of the model: one does not try to enforce a solution thanks to an ad hoc formulation, but rather to correct an inconsistency.

7.2 A FIRST VALIDATION

Although the convergence of a model is not without connection to its theoretical validity, it is not enough to ensure it. For this, a certain number of more or less exact methods must be applied.

7.2.1 EX POST SIMULATIONS

[95] Replacing the logarithm by a level does not change the situation, as negative unemployment is of course unacceptable.

Now we want to know if the model is able to perform realistic forecasts. For this, two techniques are available: ex-post simulations (or ex-post forecasts) over the past, and stochastic simulations, generally over the future. Let us start with the first case.

We have a problem: we do not know the future (if we did we would not need the model anyway), so we have to rely on the past, which provides limited information. We are not in the situation of a scientist who can produce relations as exact as he wants through the required number of experiments.

So we shall perform simulations on the past, but place the model in a situation as close as possible to the one it will meet in the future. And there is one element it certainly will not know: the residuals in the estimated equations, in other terms the unexplained part of the behaviors.

What he guesses is their most probable values, the average of their distribution: zero.

So we shall simulate the model on the past, with zero residuals. This is called “ex post simulations”.

They are undertaken on the total set of periods used for the estimation of coefficients: it may seem that the closer simulation results will be to the observed values, the better the model should be.

One will then compare model results to historical values, through tables of residuals, graphs showing the two curves, or statistics.

In practice one shall perform:

- Static simulations, which use historical values for lagged variables. We will get in fact a sequence of independent simulations with a one-period horizon.
- Dynamic simulations, which will use the results from previous periods (except of course for the first l one).

This last criterion should be favored, as it is more representative of the future use of the model. One can also assume that it is more difficult to meet, but nothing truly proves it.

In addition to the visual observation of errors (made easier by a graphic presentation) information can be synthesized using to the following criteria:

- The average absolute error
- the average absolute percentage of error.
- the standard error (square root of the average squared error)
- the standard error divided by the average of observations means.
- the bias

One will generally start with one of the first four, which we shall now present with

$\hat{x}_t$ as the simulated value, and x_t as the observed value.

The root mean square error:

$$\text{RMSE} = \sqrt{1/T \sum_{t=1}^{T} (x_t - \hat{x}_t)^2}$$

The mean absolute error:

$$MAE = 1/T\sum_{t=1}^{T}|(x_t - \hat{x}_t)|$$

The mean absolute percentage error:

$$MAPE = 100/T\sum_{t=1}^{T}|(x_t - \hat{x}_t)/x_t|$$

The root mean square error, normalized:

$$RMSER = \sqrt{1/T\sum_{t=1}^{T}(x_t - \hat{x}_t)^2} \Big/ \bar{x}$$

The choice between a criterion in level or in relative value will come as usual from the nature of the variable. In general the relative criterion will be more relevant (and more telling): an error of 1% on household consumption (CO) is more easily interpreted than an error of 3 billion 2005 US Dollars. The absolute criterion will be used for elements presenting a strong variability, especially if their sign is not determined: the error on the trade balance will be measured naturally in billions of currents units (Euros or US Dollars). But to better understand the figure, it will be more appropriate to divide it by a some normalizing element, like GDP. Levels can be used also if the notion of relativity is already present in the variable itself (growth rates, ratios): the absolute error on the variable «growth rate of wages» will give the relative precision on the salary itself. The error on the profits ratio will also be better measured in points[96].

As for the choice between the two types of moments: first and second order, it will be often forced, packages providing in general only one of the two statistics. One should note that:

- The first moment is always lower than the second, when the formula is of the same type (Cauchy-Schwartz inequality).
- This difference is all the more important as individual errors are of unequal size: a simulation including one or two particularly badly simulated observations will be particularly hurt by the second criterion.

Finally, another question appears, in the case of dynamic simulations of quantities or especially indexes: should one consider the error between simulated and observed levels of the variable, or the additional error introduced by the simulation of the period? In others terms, should one consider the error on the level or on the growth rate? In the case of a price index, both can be meaningful: in economic forecasts, or when judging the economic efficiency of a policy, one generally uses the inflation level, but the cumulated evolution on the price levels is the determinant of price competitiveness and of purchasing power.

And the result can be completely different, as the following example will show.

Let us suppose that the simulation of the price index p, from t = 1 to T (let us suppose T even), uses an econometric formulation in growth rates, where the formula:

[96] We get back to the same notion as usual: to choose the best criterion, one just has to consider the way the variable (and its changes) is normally presented.

$$p_t = p_{t-1} \cdot (1 + f_t) + e_t$$

can be affected by two sets of errors:

(1) $e_1 = a, e_T = -a, e_t = 0$ for other values of t.

(2) $e_t = a$ when t is even, $e_t = -a$ when t is odd

(the sum of errors is zero in both cases).

We shall not discuss the way these results have been obtained: the first estimation looks skewed, at least for small samples, while the second should be corrected for autocorrelation. We have to admit the problems we shall meet come at least partially from faulty estimations.

The criterion "absolute mean relative error" will give the results:

	error on the level	error on the rate
case (1)	a (T - 1) / T	2 a / T
case (2)	a / 2	A

The order is completely inverted according to the criterion used.

In more general terms, the issue is the following: is a simulation which presents a lasting gap to observed reality, but a relatively stable trend parallel to it, better than a simulation which oscillates around true values, with the same error band as the first[97]?

More generally, the difference between the two criteria comes from the dynamic character of the model. At a given period, the error comes from the accumulation of present and past ones, the impact of the latter reduced in principle with time.

The growth in the variance of the global error will be reduced by the fact that the sum of original residuals is zero, so positive and negative ones will alternate. However, this condition is only met on the whole sample. It is quite possible, even with no apparent autocorrelation, that the current sum has a significantly positive or negative value.

[97] We prefer the first case, actually both should lead to a change in the formulation itself.

So it is not surprising with highly dynamic equations that the smallest simulation error is quite often found at the last period, and the largest error close to the middle.

But in some occasions, an accumulation of errors can lead the variable away from its observed value. Let us show this on a very simple example.

Let us suppose that the wage rate depends on the consumption price (in growth rates) and the unemployment rate (unemployment divided by the total active population).

$$\boldsymbol{tx(W_t) = a \cdot tx(PC_t) - b \cdot UN_t/(L_t + UN_t) + c}$$

Unemployment decreases with job creation, with a reduced intensity:

$$\Delta UN_t = -d \cdot \Delta L_t$$

Both equations are estimated, and subject to errors. Let us suppose the second error is not auto correlated. On the series UN the error will be the sum of non-correlated errors: it can very well keep the same sign over the period, for instance if the few first errors are mostly positive. In the end the error will cancel out, but it will have introduced on W itself a cumulative positive error which can make its level diverge markedly from its historical value[98].

7.2.1.1 Reassessing the criterion

Actually this criterion is not so relevant, and its use is highly questionable, for an evident reason: the numerical history one seeks to approach was known at estimation. It has been used to determine values of coefficients, but also to choose among possible formulations. Even a scrupulous model builder will be led sometimes to favor equations that give results close to reality, to the detriment of theoretical quality.

An extreme case is the introduction of dummy variables, taking the value 1 in one or several periods, and zero otherwise: they will allow the equation to go perfectly through a point quite badly explained. One does not see why this introduction, which improves the criterion, can improve the model and its properties.

It is clear in this case that the additional explanation does not add anything to the model, but rather masks a problem (all the more as these variables are attached to the most badly explained periods). They can be accepted only if the period presents a specific feature, impossible to translate into an explanatory element or behavior (as in the May 1968 case we have seen earlier, or the present coronavirus pandemy). But even then one runs the risk, by eliminating the main fluctuations of the explained series, of making the interpretation of its behavior more difficult. And the problem of forecasting these variables will remain.

More generally, there is no strict connection between the proximity of formula results to the observed reality and its intrinsic quality (one can observe that very accurate estimations can provide coefficients values entirely unacceptable by any theory, for example presenting wrong significant signs). A model produced by selecting systematically the

[98] Of course, this would not happen with error correction formulations.

formulations giving the smallest distance to observed values, to the possible detriment of economic meaning, might give good ex-post simulations[99]. By pushing the argument to the extreme, one can end up with a model describing exactly reality, provided enough variables are added[100].

But a model determined according to these principles will probably fail the more difficult test of the analytic shocks, which we shall present later.

It is nevertheless also clear that a model failing to describe past evolutions will have to be rejected or at least corrected (but by theoretical rather than technical contributions), as the quality of its forecasts will be even lower. Passing the present test is therefore necessary, but not sufficient, to guarantee model quality.

Let us consider the formation of simulation errors, originating in the residuals (we shall consider later the consequences of errors on coefficients and the actual form of the equations). They have three sources:

- The original estimation errors: it is clear that the smaller the estimation errors, the better the expected adequacy of simulation results to observed reality.
- The expanding effects of the instantaneous links between variables. Taking into account as explanatory elements variables affected by estimation errors (directly or indirectly) will increase the number of error sources and in most cases the global error (most equations use sums or products, and in the absence of correlation, the variance of a sum is the sum of variances).

This is not guaranteed however, as some errors can compensate by pure chance, and some model mechanisms can play a moderating role. For instance, a positive error on demand will see its effect on production dampened by the increase in imports it will generate. This will happen in our model:

```
[14] FD  = CO + IP + IC + IH + gd

[15] TD  = FD + CI

[16] DLOG(M) = 1.676*DLOG(TD) + 0.2133*DLOG(UR) - 0.1057*DLOG(COMPM)

          - 0.1028*RES_M( - 1) - 0.2757+ EC_M

[17] Q + M = FD + X
```

A positive error on IP and FD will increase M, which will limit the impact on Q.

- The dynamic error (in dynamic simulations), coming from lagged variables which are now the result of a simulation. The error should grow with each period simulated, as again the number of error sources increases. However if we impose to all our estimated equations an error correcting structure, the problem is

[99] Although it is quite possible that in an incoherent model the instantaneous and dynamic links between variables exert a strong expanding effect on originally small estimation errors..

[100] But the small explanation given by the last variables introduced should give them very low acceptability statistics.

partially solved: if a variable moves away from a target value, the behavior of model agent(s) will drive it back to this target (it is not surprising that an "error correcting" model corrects the errors better than others). With these models, medium-long-term simulations can actually give better results than the short-medium-term ones, and ex post simulations can be more accurate at the end than in the middle101.

We shall improve our understanding of the implications of error correction when we address stochastic simulations.

7.2.1.2 Partial tests

To better interpret the simulation errors, some of their sources can be eliminated

- By exogeneizing variables.

The variables will take their exact values, which should reduce the global error. The measure of the decrease gives some idea as to the contribution of the variable.

However this technique will mix two effects, in the case of a behavioral variable: suppressing the estimation error and the model errors transiting through this variable. This will make the results difficult to interpret.

In the case of eliminated identities, the change in the error will show clearly the contribution of the variable. But if you for balancing equations (like the supply-demand equilibrium giving GDP) the constraint will no longer hold true.

One can consider:

- Eliminating a single variable (or a small set) and observing the decrease of the residual compared to the full model.
- Eliminating a large set of variables (for instance all the deflators) and observing the expansion of the residual through model mechanisms.

With EViews this is done quite simply through the EXCLUDE statement.

- By re-introducing residuals (starting from a fully null set) :

This is done by setting the estimation residuals to their observed values, starting in two directions:

- Keeping every residual null, except for one equation. The (probable) reduction of the error will show the gains to be expected from improving the equation.
- Setting every residual to its historical value, except for one equation. The remaining error will show how the model dynamics expand the original estimation error.

Of course, combinations of variables also can be used: for instance observing the global error coming from external trade, or the error remaining if all prices showed no residual error....

With EViews, we can simply put the requested residuals to zero (to avoid losing their estimated value, they can be stored in an alternate variable). However, there is a much better technique, which uses the SCENARIO option. We shall present it shortly.

- by using static simulations, and comparing them to dynamic ones

[101] On the past, this is favored by the fact that the sum of the historical errors is zero.

This allows to separate the impact of the error on the first period, and from its dynamic influences.

7.2.2 EX-POST FORECASTS

To make the above criterion lose its artificial character, it looks more appropriate to simulate the model over a period which has not been used for estimation. This can be done by excluding from the estimation sample the last known periods, and using them to check the quality.

But this introduces several problems:

- It reduces the size of the sample, and therefore limits the significance of estimated coefficients. To reproduce the conditions of the test, it is necessary to remove a high number of periods (5 at least in the case of a medium-term annual model).
- Evaluating the result faces the same subjectivity problems as the previous test.
- But the main problem is that if we select models on the basis of the test, we use actually the same technique as on the full sample, and the same "ad-hoc" critique applies.

It all depends on the way the test is conducted.

- First, it is always applied to a model (or a set of models) which has been estimated initially on the full sample. One never restricts the sample during the whole estimation process, then tests the models on the eliminated periods. This would be quite inconvenient, as these periods are the most representative of the domain on which the model will be use: the future.

So the search for the best formulas in always conducted on the full sample. When we restrict the period, we get other equations which might or might not be satisfying from an economic and statistical point of view. If the equations have remained stable, the residuals are probably of the same size, and it is quite improbable that the multiplying effects of the model properties will change very much, so the ex post error should not be much larger.

But if the results change, we will question the specifications themselves: the last periods are not consistent with the global choice.

This means that the only message from the new test is the stability of the set of equations, which could have been evidenced using a simple Chow breakpoint test, a more objective criterion. What we still get is a more global (and perhaps utilitarian) measure.

Of course, if the modelling process had been applied to the reduced sample, the process would be much more valid, provided one uses the result as an absolute measure of model quality, and not as a way to select the best model in a set. In this case, the process is not much different from estimating on the full sample. The choice relies on a test (obtaining good results on out of sample simulations) which basically requests equation stability, which means that the equations obtained are not very far from the ones from the full sample.

In other words, the success of the ex post forecast requires that the estimations work on the full sample, and that they keep stable if we eliminate the last periods. Of course, stability is a requirement for model reliability. But it is not a proof: an equation inverting a true and strong link between two variables (making demand depend on imports for instance) should remain stable over any sample.

7.2.3 SOLVING THE MODEL: SCENARIOS

To apply the above techniques, an EViews feature is quite useful: the scenarios.

When one wants to change the way the model is solved, it can mean:

- Solving the same model under
    - Different assumptions on exogenous variables
    - Different horizons
    - Different options for formulations allowing alternate versions (for instance the exchange rate equation can propose a constant nominal rate, a constant real rate or uncovered interest rate parity).

- Solving different coexisting versions of the same model.
    - One can be official, having passed all the requested tests and familiar to the operational forecasters, and the other can be more experimental, in the process of being tested.

The **scenario** option provides an efficient way of dealing with these issues.

- It works first through the creation of a new scenario.

In a program, the statement is:

```
model-name.scenario(n) scenario-name
```

This will create a blank scenario even if it exists already. If a previously defined scenario should be reused, one should state:

```
model-name.scenario scenario-name
```

Then one should:

- Assign a suffix to the scenario/model solution:

```
model-name.assign @all suffix
```

From EViews 8 this can actually be obtained in one step:

```
model-name.scenario(a=suffix) scenario-name
```

- Create assumptions associated with the scenario, using the name of the variables followed by the suffix.

- State that the solution should look for the alternate versions of these variables by using the override statement:

```
model-name.override list-of-overriden-exogenous
```

For instance to solve our model **_fra_1** dynamically, with zero residuals for exports and imports (series called m_ec and x_ec) and the current suffix “_s”.

```
 '    We solve over the past
 '    with residuals set to zero
 '    using the suffix « _s »
  genr m_ec_s=0
  genr x_ec_s=0
_fra_1.scenario "scenario 1"
_fra_1.append assign @all _s
_fra_1.solveopt(n=t m=1000,c=1e-6,o=n,d=d)
_fra_1.override ec_m ec_x
  solve _fra_1
```

To override variables, two operations are needed:

- Stating that the variables have to be overridden.
- Creating the alternate assumptions using the scenario suffix.

If one step is missing, the consequences are different.

If a variable is omitted from the list, the override operation will not be applied, without any adverse consequence (but no message). This allows to define a pool of potentially overridden variables in which the choice will be made later.

If a variable in the list has not been created, with the adequate suffix, EViews will refuse to solve the model, just as if the original variable did not exist.

If the program switches to another scenario, the options associated to the previous scenario will be put aside but remain associated with it, until they are modified under this very scenario. A new override statement cancels any previous overrides.

To eliminate all overrides, one can just specify a blank list. This should be done as soon as the overrides are no longer pertinent, as a precaution.

From EViews 8, one can use an alternate option: you can drop overrides individually through the REVERT command:

```
model_name.revert list-of-unoverriden-series
```

But the most important addition is the possibility to specify directly the changes in the assumptions, using ADJUST.

The syntax is:

```
model_name.adjust(init=initial_series) series adjustment
```

For instance the statements:

```
_mod_1.scenario(a=_2) "scenario 2"

_mod_1.adjust(init="scenario 1") gdp =+10000
```

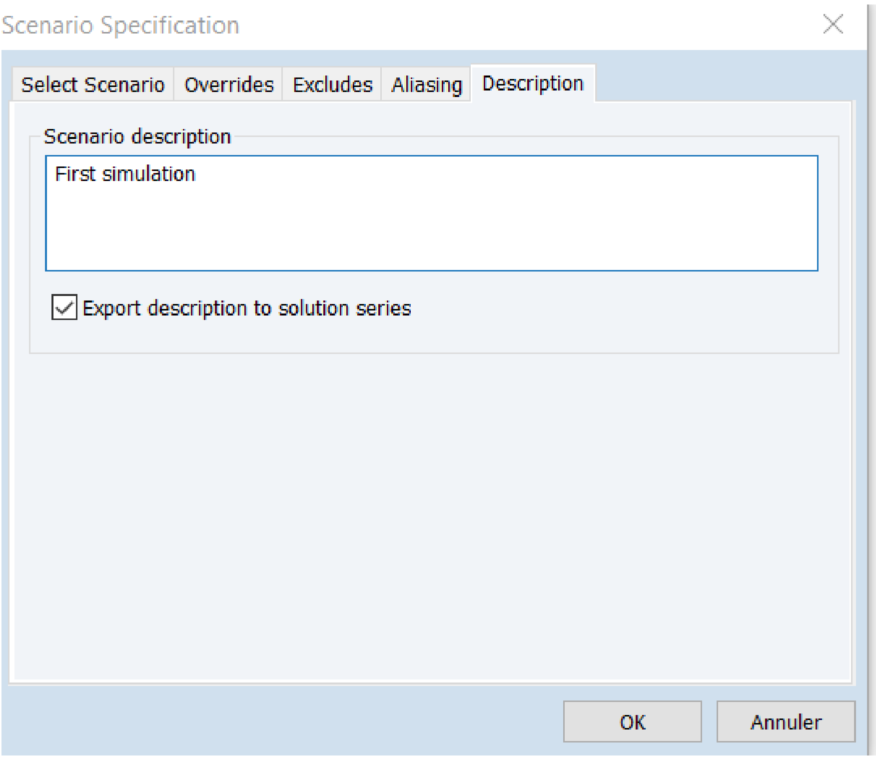
will create a series called gdp_2 with a value 10000 higher than the value from "scenario 1" (maybe called gdp_1). The series will be added to the "override" list, and to the "exclude" list as well if the variable is endogenous.

Be careful to introduce a space before the "=" sign, or gdp will take the value 10000!

Introducing a change involving elements, such as series and parameters, is possible but more difficult. One can prefer using actual expressions, as we have done in our examples.

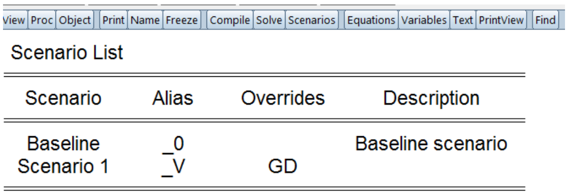
7.2.3.1 Scenario description

A description can be attached to each scenario, and optionally exported to the series it computes.



7.2.3.2 Scenario list view

By using “Scenario view” for a particular model, one will obtain a list of scenarios including any information or comments, as well as possible overrides and excludes.


View Proc Object Print Name Freeze Compile Solve Scenarios Equations Variables Text PrintView Find

Scenario List

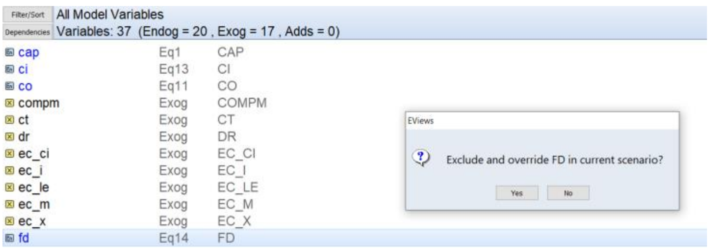
Scenario	Alias	Overrides	Description
Baseline	_0		Baseline scenario
Scenario 1	_V	GD	

7.2.4 OVERRIDING THROUGH MENUS

You have certainly understood that in our opinion menus are not the best way to manage alternate scenarios. However one might want to use this option to make a quick study, for instance to test a simple assumption.

This is done by accessing the model, then its “Variables” presentation, under a specific scenario (excluding “Baseline”).

Right-clicking on a variable name will open a box allowing you to change its override status, using “Edit Override”.

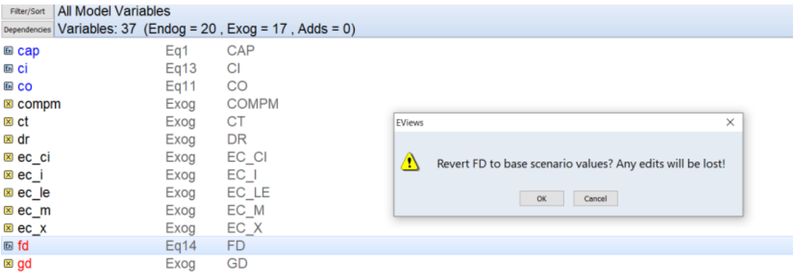


If the variable is not currently overridden, it will be added to the list (endogenous variables will be endogenized first, by adding them to the EXCLUDE list).

Then a window will open, with the variable values in spreadsheet format, which you can edit.

	FD_V	Unadjusted	Delta	Delta%
2006Q4	3.74E+11	3.74E+11		
2007Q1	3.81E+11	3.81E+11		
2007Q2	3.84E+11	3.84E+11		
2007Q3	3.86E+11	3.86E+11		
2007Q4	3.88E+11	3.88E+11		
2008Q1	3.90E+11	3.90E+11		
2008Q2	3.93E+11	3.93E+11		
2008Q3	3.96E+11	3.96E+11		
2008Q4	3.99E+11	3.99E+11		

Just as in programs, clicking on the REVERT item will get the series back to the original values, and the endogenous back to their original status.



For an endogenous EXCLUDEd variable (through any method), REINCLUDE will also change its status back, but will create an add factor for the variable, with null values.

7.2.5 OUR EXAMPLE

We shall now simulate the model with residuals set to zero, on the 1980q1 to 2004q4 period, the longest one for which all estimations could be conducted, considering the availability of exogenous values.

```
'    We solve over the past
'    with residuals set to zero
for !i=1 to g_vbeha.@count
%1=g_vbeha.@seriesname(!i)
genr ec_{%1}_s=0
next
_fra_1.scenario "scenario 1"
_fra_1.append assign @all _s
_fra_1.solveopt(n=t m=1000,o=g,d=d)
_fra_1.override. ec_k ec_ic ec_le ec_m ec_x
solve _fra_1
'   We compute the residuals and their averages
'   - the relative errors
'   - the absolute errors on the growth rates
smpl 1980Q1 2004Q2
 matrix(20,2)  v_psm_b
for !i=1 to g_vendo.@count
%st1=g_vendo.@seriesname(!i)
%st2=g_vendo.@seriesname(!i)+"_s"
series psa_{%st1}=100*({%st2}-{%st1})/({%st1}+({%st1}=0))
series tc_{%st2}=100*({%st2}-{%st2}(-1))/({%st2}(-1)+({%st2}(-1)=0))
```

```
series tc_{%st1}=100*({%st1}-{%st1}(-1))/({%st1}(-1)+({%st1}(-1)=0))
series psd_{%st1}=tc_{%st2}-tc_{%st1}
v_psm_b(!i,1)=(@mean(@abs(psa_{%st1})))
v_psm_b(!i,2)=(@mean(@abs(psd_{%st1})))
smpl 1980Q1 2004Q4
next
```

We get the following results

	Simulation errors	
	Percentage error on the level	Error on the growth rate
CAP	0,755	0,035
CI	104,041	566,870
CO	1,720	0,210
FD	1,246	0,213
I	2,016	0,465
IC	0,989	0,348
IH	1,720	0,209
K	0,755	0,051
LE	2,431	0,261
LED	2,541	0,384
LT	1,871	0,202
M	2,142	0,837
PRLE_T	1,856	0,038
Q	0,989	0,348
RES_M	0,890	0,747
RES_X	23,156	12,928
RHI	1,720	0,210
TD	1,569	0,363
UR	1,600	1,296
X	2,268	1,195

The quality looks rather acceptable, considering the size of the model[102]. The only source or uncertainty comes from the estimation residuals, which have a zero sum over the period. And our equations, except for the change in inventories, follow an error correction framework. This means that any residual appearing at a given period is corrected with time, sometimes very fast. The variance of the global error will grow, but converge in the long run to a finite value, as we shall observe later.

[102] But one could also suggest that the smaller the model, the larger the domain of the assumptions and the smaller the sources of uncertainty.

7.2.6 ANALYTIC SHOCKS

If ex-post simulations give a valued criterion (more or less valid) of the quality of the model in forecasts, analytic shocks allow judging its properties (so the reliability of policy studies) by studying in terms of quantified economic logic the results of model simulations. Thus the model will show if its quantitative properties are consistent with the economic theory that has presided over its construction, and not too atypical as compared to the other models of the same class (the second condition is not of course necessary, but to ignore it begs for further study and strong arguments).

For this, a set of simple changes will be undertaken independently, on the main assumptions, trying to encompass each of the main areas described by the model. Every type of domain must be treated: agents, production function, supply and demand effects. This means we will study the sensitivity of the economy to a variation of exogenous demand (by State investments for example), the behavior of households (by a variation of the income tax), of firms (by a variation of their social contributions), the foreign trade (by the impact of a devaluation or a change in the tariffs rates), the role of capital productivity, of an exogenous price increase. The quantitative and logical analysis of the evolutions of the main economic elements will give indications on model quality, by comparison with economic theory and properties of other models.

If the information gained from this criterion is probably more instructive, it also loses its objective character: a model which makes imports increase when local firms create new productive capacities will be clearly wrong, but an original behavior compared to the family of other models (for example if exports show a particularly strong sensitivity to price competitiveness, or depend on the age of capital) will not be conclusive, in the absence of a general consensus on the theoretical value. This might even, in some cases, affect the sign of the connection: facing a rise of inflation, households can increase their savings (to maintain their purchasing power) or decrease them (to buy immediately what will soon become more expensive).

In practice, to validate an atypical behavior, one will have to identify the source of the originality. Sometimes it will have been expected before any simulation, as it comes from an original feature introduced voluntarily into the model. If the model identifies foreign direct investment, a growth in local GDP can attract it and produce an increase in exports, overcoming the usual substitution effect.

In general, to make the interpretation easier, the change in assumptions will not fluctuate with time. One will distinguish however three types of shocks, according to the nature of the variation:

- One-time shocks, where the modification disappears after one period. For example the State will increase its investment for one budgetary year. In this case one will be often interested, not only in the evolution of effects (which should disappear with time), but also in the cumulation of these effects over time (such as the total gain on production).
- Sustained shocks, where the modification is maintained for a number of periods, which can be the full sample used in the simulation. For example a reduction of the income tax rate will be maintained on the period of simulation.
- Cumulative shocks, where the shock is amplified at each period by an identical amount. Their interpretation will be often difficult, as the global evolution comes from multiple sources.

It is clear that if one wants to correctly interpret these shocks, they must remain logically acceptable

In summary, one must consider two criteria:

- The acceptability of the shock.
- The readability of results.

This allows discarding one-time shocks in general, especially for policy variables. They fail most of the time on both grounds. Except for an additional investment or subsidy, policy decisions are generally taken to last for a while, especially if they apply to tax rates or social security benefits. Structural changes (gains on factor productivity) are also permanent. As to external shocks (like an increase in world demand) they could very well be one-time.

But the most decisive argument is more down to earth: most of the interpretation will be done through graphs, and one-time shocks leads back all the variables to the baseline scenario. After a few periods, the return to zero of variations will make it impossible to separate the lines from each other, creating a high confusion close to the time axis, while in the case of a sustained shock the stabilization of consequences will create a set of more parallel (or slowly evolving) lines. As to cumulative shocks, they will produce expanding graphs and make the interpretation of the initial changes (associated with a smaller scale) more difficult.

So we are left most of the time with sustained shocks, which are indeed the easiest to read, provided the shock is reasonable. This applies first to the intensity: the less reasonable the shock, the more difficult its interpretation. This means the shock must be small enough (for instance affordable in terms of budgetary cost) and large enough (why should the government bother with such small decisions?).

But it applies also to the cumulated cost of the shock, in the case of policy decisions. If the period is long enough, and the shock important enough, one can consider that its cost will become unbearable after a number of periods. And the economic agents know about this. So not only should we stop the shock in time, but the model should adapt the behavior of agents to this future event. This is not possible with our present model, in which our agents look only backwards in time.

This is the most basic principle in the Lucas critique, which we shall approach later, when we address rational expectations.

Of course, this long-term problem will be more acute when we consider forecasts: we could stress that the sample period is small enough to allow sustained policy changes.

7.2.6.1 The choice of modified assumptions

Concerning shocked variables, one will look for a set that asks contributions from all the mechanisms of the model: as the purpose is to validate the whole model, no abnormal behavior should remain. An initial investment in the full examination of these elements will avoid later problems, the identification of which will prove to be far more expensive (provided it can be done).

Especially, one will seek to activate the whole set of agents, and the whole set of domains: production, labor market, prices and wages, external trade, financial mechanisms. Supply and demand oriented influences must all be studied. However, if there was only one test to realize, the most representative would consist in increasing the exogenous part of demand, thus determining the Keynesian multiplier. We shall come back to this in the following paragraph.

7.2.6.2 The choice of the results considered.

All variations of endogenous elements may or may not be examined, according to the size of the model. In most cases, the first analyses will consider aggregated results, resorting only to detailed values if a problem comes to light.

- In the case of a macro-economic model, the main elements will be:

  - The supply and demand equilibrium: gross domestic product, demand and its decomposition, exports and imports (the ratio of the variation of GDP to the increase of exogenous demand defines the Keynesian multiplier).
  - Prices and wages, interest rates.
  - Employment and unemployment.
  - Foreign and budget balances.

- Some ratios: savings ratio, profits, utilization of production capacities, financial balances and debts as a share of GDP.

- If the model also addresses financial issues, we could add

  - The stocks of assets: debts, equities, currency, as a share of GDP.

7.2.7 OUR EXAMPLE

In our example, the number of available assumptions is quite low. The most reasonable choices are:

- Government demand GD
- World demand WD

And maybe also:

- Capital productivity PK
- The savings rate SR
- The depreciation rate DR[103]
- The share of non-wage household revenue in GDP R_RHIQ
- The share of household housing investment in revenue R_HI
- The residuals in each of the five estimated equations.

The main elements are demand oriented ones. We shall select the most widely accepted: government demand GD.

We shall use the following statements, taken from a program (we have left the comments):

```
'    We produce a shock
```

[103] Shocking DR introduces an interesting problem, as it is de facto forbidden. Indeed, the development of our investment equation is based on the fact that DR is constant, which allows to integrate it to the constant term, so changing it violates model specifications. It is the only shock which will actually modify the long-term rate of use, and this for wrong reasons. To allow it to change, one can simply estimate the change in capital instead of investment (the assumption on DR stability disappears) and rearrange the present capital equation so that it now determines investment.

```
'    First, an unshocked simulation

_fra_1.append assign @all _b
_fra_1.solveopt(n=t m=1001,o=g,d=d)
_fra_1.override
solve _fra_1
'     The shocked simulation
_fra_1.solveopt(n=t m=1002,o=g,d=d)
smpl 1980Q1 2004Q4
genr gd_v=gd+.01*q_b*(t>=1981)
genr dv_gd=+.01*q_b*(t>=1981)
_fra_1.scenario "scenario 1"
_fra_1.append assign @all _v
_fra_1.override gd
solve _fra_1
'     We compute the differences
'     in absolute and relative terms
for !i=1 to g_vendo.@count
%1=g_vendo.@seriesname(!i)
series dv_{%1}={%1}_v-{%1}_b
series pv_{%1}=100*dv_{%1}/({%1}_b+({%1}_b=0))
next
```

A few additional comments.

- We have started the shock a little after the start of the simulation. This allows checking that in absence of shock the two simulations give the same result, which should mean that the difference can be interpreted as coming only from the shock. Additional differences can come from:

o A different starting date
o A different model.
o Different assumptions.

This is not an absolute proof, however:

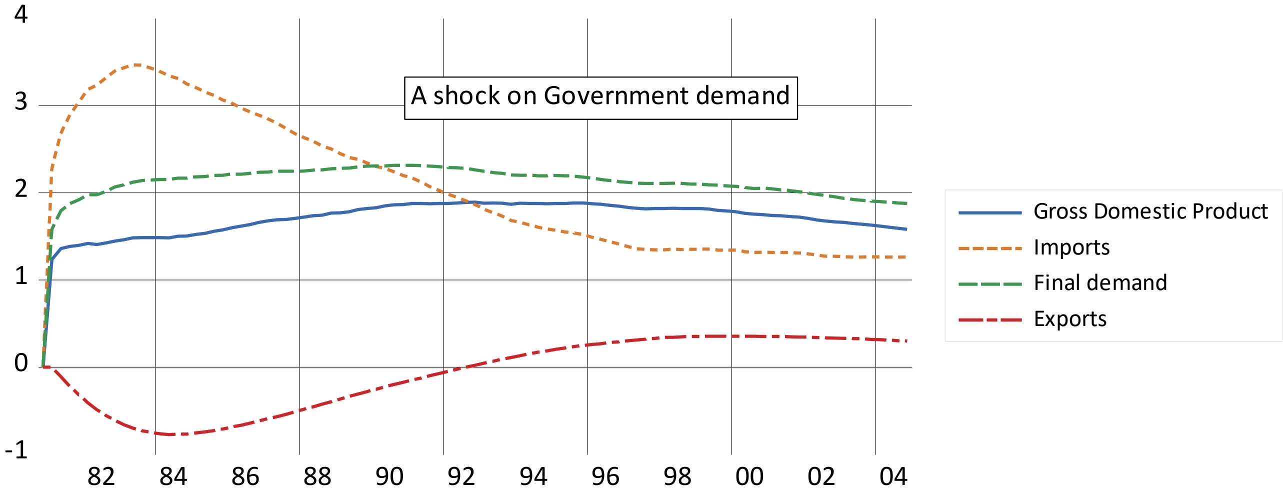
o An additional change in the series might start after 1981.
o The specification of the model might contain changes which apply only after 1981 (for instance a new tax) or only if a condition is met, which will occur later (like a high level of inflation).

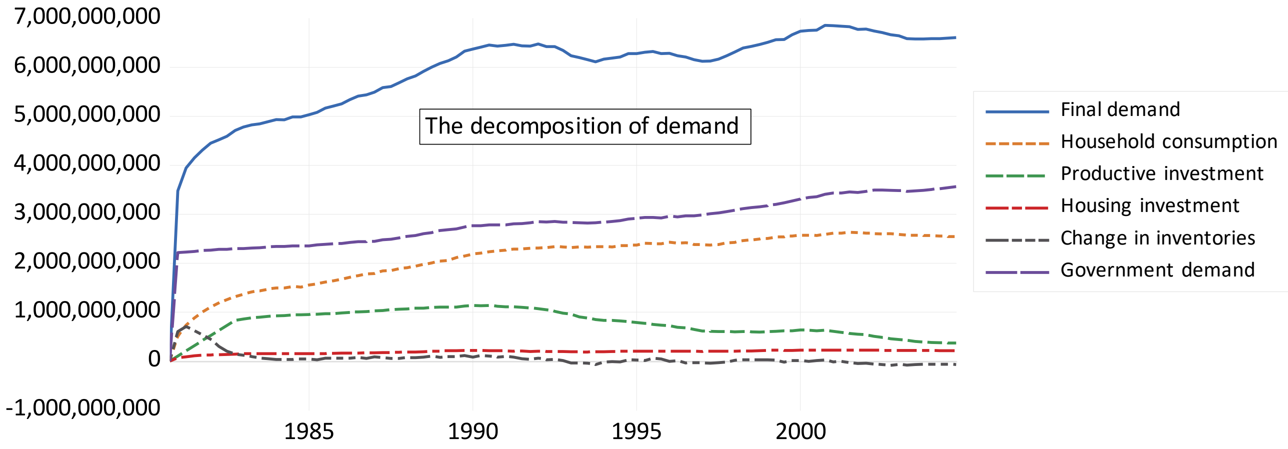
To guarantee the absence of such errors, it is advisable to simulate the baseline scenario just before the shock, even it has been produced earlier. The efficiency of present algorithms and the speed of present computers make the cost negligible in most cases.

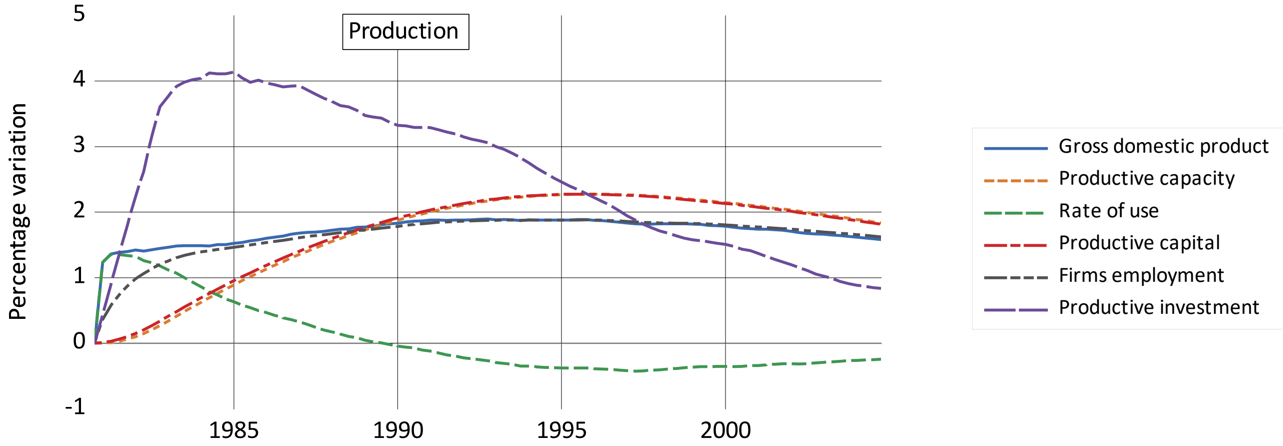
- We produced the alternate assumption in two steps, computing first the shock itself, with the same format ("DV_" plus the name of the variable) as the absolute endogenous deviations. This will allow us to compare the ex-ante and ex post changes in GDP and create a graph presenting all changes.

7.2.7.1 The results

We shall base our commentary on three graphs showing: the evolution of the supply-demand equilibrium in relative terms, the decomposition of final demand in absolute terms (including the initial shock), and the evolution of elements linked to production.







The properties shown by these graphs are quite usual for this type of model (considering the reduced size of our version).

- The ex-ante demand variation is amplified by the increase in production factors (investment and employment), the latter affecting household revenue and consumption. These two elements keep increasing for a while, with some inertia, traditional for an error correcting process.
- Later, both changes become proportional to GDP. But while households obtain as revenue a high share of the increase, the motivation for investment decreases as capital adapts to GDP. In the long run, the only incentive left is the replacement of a higher level of discarded capital.
- Capital and capacity adapt slowly, but this inertia makes them overshoot the target. A cyclical process appears on the rate of use, with a rather fast decreasing intensity.
- External trade plays a simple role.

    - The change in exports follows the rate of use, with a decreasing amplitude as capacities adapt.
    - Imports take a large initial share in the additional demand, when capacities have not adapted. They compensate the ex-post increase in demand, reducing the multiplier value. In following periods, capacities adapt, and the multiplier increases regularly. The overshooting of capital actually favors local producers in the medium-term.

On the whole, although this model remains overly simplistic, it looks rather sound if we consider its size.

7.2.7.2 The drawbacks of shocks over the past

Although producing historic shocks for a model designed for the future does not face the same criticisms as the ex-post tests, they still meet several problems.

- They are conducted on a limited period. If one wants to observe the long-term properties of the model, the available period will generally be too small. The problem will increase if data production has started recently, or if the country has undergone a recent transformation making the use of older data unacceptable. For Central and Eastern Europe countries, acceptable series start in 1995 at best.
- They are conducted on the wrong period. If one wants to interpret the dynamic properties of a model over a sufficiently long-term, it is necessary to start the shocks quite far in the past. At that time, the economy of the country might have presented different characteristics. Even if the economic behavior was the same, some structural parameters were quite different, like the importance of external trade, or the relative roles of labor and capital in the production process
- They are based on a more or less unstable baseline. No model is completely linear (as it will always include both sums and logarithms). This means that the response to shocks will depend on the baseline, which can present irregularities. For example, a given increase in government consumption will generate more GDP if the activity is low, and the country can use more idle capacities to respond to demand.

This means that it will be difficult to decide if irregular variations in the results of the test come from model properties or from a non-smooth baseline.

Obviously, all these problems disappear when the shock applies to forecasts:

- The period is unlimited, allowing to evidence completely convergence processes and in particular cycles.
- As tests are conducted on the same period as future operational applications, the evaluation will be more relevant
- The present context and future issues are more familiar to the modeler
- As the baseline is smooth (especially in the medium and long-term when it matters) any irregularity will be allocated to model properties.

This is why we suggest to limit to a minimum the studies on the past, and proceed to simulations on the future. This is true even for control of convergence: it is on the future and not on the past that the model will have to converge. But a limited series of immediate tests can evidence problems right away, which can help to gain time in the global process by staying on the right path.

CHAPTER 8: TESTING THE MODEL OVER THE FUTURE

The above arguments lead to the natural idea: test the model on the future.

- We shall have initial information on the reliability of spontaneous forecasts, and of their distance to what we expect of the near future.
- The results will me more representative of future use.
- The actual results can be interpreted as the actual consequences of present policy decisions.
- The first periods of the shock will be representative of the present efficiency of present policies.
- The tests can be conducted on a period of any length, allowing to observe convergence and cycles.
- With regularly growing assumptions we can test that the simulation is regular, and that it converges to both a steady state growth and a long-term stable solution.
- Applying to these regular solution constant shocks we can check that we get smooth evolutions, and interpretation is easier. We also have enough observations to treat the Lucas critique.

There are only two drawbacks:

- We cannot check the simulation results against true values. We shall try to prove this is not so important, and can be replaced by other tests.
- We do have to produce a simulation over the future, an unknown domain in which convergence might be more difficult to achieve.

But this convergence will have to be obtained anyway. So our first test will be:

- **Can the model converge at any horizon?**

In our sense, this is a prerequisite for structural econometric models. This is where this type of model has an advantage over CGE and VARs: to produce dynamic simulations which show the way the economy converges to a long-term equilibrium, and allow a full interpretation of the process. If this is missing (if the model converges only over the medium-term) this advantage is lost.

8.1 MAKING THE MODEL CONVERGE IN THE LONG RUN

First, let us stress that this is not the same as producing a long-term forecast. What we are considering is the very long run (the first converging version of the model can take several centuries to stabilize). We are not going to suppose that the model will remain valid for such a period. What we want to achieve is:

<u>A technical reproduction on a long period of the mechanisms associated with the equations we have evidenced, as if they remained valid forever[104]. The period is much longer than the longest on which the model will be used.</u>

For instance, many models present a cyclical behavior, linking investment, capital, the output gap and profitability. The length of this cycle can exceed 30 years. To check that its period is constant, and that it is converging regularly, we need at least three cycles, which means that our simulations have to last until 2100.

So at the beginning of the tests, one should consider a very long period.

[104] But we know this is not the case

This could be compared to testing a car (actually a prototype of a car) on an artificial track, for as long as it can run, exceeding by far its normal life (say several millions of kilometers), at a constant speed. And also changing one of the car elements, or the driving conditions, and observing the consequences, both immediate and after a long period..

What we want to get is information which will help us to build a better model, which will be used later in quite different circumstances (a shorter but less regular environment).

We have to keep this in mind when we undertake our simulation.

8.1.1 THE ASSUMPTIONS

To produce a long-term simulation, we need assumptions on the exogenous elements.

The choice of these assumptions can have several goals:

- Producing short-term evolutions which coincide with what we know already of the near future. This can also be true of the recent past which we have to simulate in the absence of complete information (global GDP but not its product decomposition).
- Producing reasonable medium-term solutions, consistent with the goals of the Government, or of what we expect from the local economy (a high but decreasing Budget deficit for instance).
- Allowing the model to converge regularly to a long-term equilibrium, which we shall examine to understand model properties.

For the time being, we shall concentrate on the last goal. Of course, we shall control that the two first targets are at least partially met, as an unreasonable simulation will react to shocks in an unreliable way, due to the model non linearities, making it more difficult to interpret.

This means we shall not differentiate the assumptions which share the same dimension, but apply to them common growth rates. In practice three rates have to be defined (in addition to a null one):

A rate for populations (including employment).

- A rate for variables at constant prices.
- A rate for deflators.

Of course, the second rate could be replaced by labor productivity. Combining it with employment we would get quantities.

Actually, we shall see that most model assumptions have no dimension. This should be favored as it suppresses the need to make normally independent assumptions grow at the same rate (in our model, Government demand GD and World demand WD).

A tip : in most data sets, the last observed period changes between series. If you want to apply a given growth rate from the first missing observation, you just have to specify the SMPL statement as:

SMPL period if *variable*=na

genr *variable*=*variable*(-1)*(1 + *rate*)

8.1.1.1 A particular case: the estimation residuals

In general (this is the case for all the equations in our model) the estimated expression has been defined without dimension, if only to eliminate the risk of heteroscedasticity. And anyway, the most probable value of a residual is supposed to be zero. But one can suppose that the last value obtained is a better indicator of the future, in particular if the previous values held the same sign. For the time being, the choice is not so important, as we are not overly concerned with short-term solutions. But it will become crucial later, when we approach actual forecasts.

8.1.2 ADAPTING THE FORMULATIONS

To make the model converge in the long run, applying any formula to variables which grow at the theoretical rate should produce a result which grows at its own theoretical rate. This is easily controlled, both visually but also technically as we shall see later.

Identities will provide this property automatically, if they are specified correctly and if the assumptions are well defined.

For estimated equations, the error correction format allows to introduce the necessary constraints in the long-term equation. One does not have to be concerned with the VAR, which includes only growth rates, stable in the long run by definition. But one has to manage (generally suppress) any additional trend.

Let us explain this on our exports equation, using for now the trend provided by EViews. The basic equation is:

```
[3] DLOG(x) = 0.4940*DLOG(wd) - 0.09142*res_x( - 1) + 0.01960+ ec_x

[18] res_x = p_x(1) * log(x / wd) + 0.9815 * log(ur) + 0.001441 * @trend(60:1)
```

Let us suppose the rest of the model (in particular the investment equation) has allowed UR to stabilize in the long run. Then any growth in res_w will come from:

- X growing at a different rate from WD.
- The trend : @trend(60:1)

To suppress the trend at the end of the historical period, we can replace its expression by[105]:

@trend(60:1) * (t<=2004) + @elem(@trend(60:1) , "2004q4") * (t>2004)

Starting in the first quarter of 2005 (represented by t>2004) the trend will be replaced by its 2014q4 value.

[105] Let us recall the syntax for the @elem function:

@elem(series,"date") gives a scalar with the value of the series for that particular date.

This works, but one can find the expression overly complex. To simplify it, we can apply the technique presented earlier, estimating the employment equation.

First, let use replace the trend by our time variable t.

Then, we can replace t by (t-2014)[106] without changing he actual specification. The constant term will adapt and the formula will give the same values.

Finally, we can break down the computation into:

- A trend of (t-2014) for the historical period.
- 0 for the forecast period.

Combining both, we get the formula;

(t-2014)*(t<=2014)+0*(t>2014)

Which simplifies of course into:

(t-2014)*(t<=2014)

On this example, let us see how growth rates of endogenous elements converge to that of exogenous.

We have (with Δlog(WD) and ec_x constant)

$$\boldsymbol{\Delta Log(X_t) = a - b \cdot res\_x_{t-1}}$$

$$\boldsymbol{res}\_\mathrm{x}_{\boldsymbol{t}} = \boldsymbol{Log(X_t/WD_t) + c}$$

giving

$$\boldsymbol{Log(X_t) = Log(X_{t-1}) + a + c - b \cdot (Log(X_{t-1}) - Log(WD_{t-1}))}$$

and

$$\boldsymbol{Log(X_t) = (1 - b) \cdot Log(X_{t-1}) + a + c - b \cdot Log(WD_{t-1})}$$

[106] Actually, we it might be better to use 2004.75, associated with the last historical value.

The growth rate of X will converge to that of WD.

In very special cases, the trend must not be suppressed but replaced by a given value, as it applies to non-constant ratios.

For instance in the employment equation, the structural labor productivity follows an estimated trend, estimated by:

$$\boldsymbol{Log(Q_t/LE_t) = a + b \cdot t}$$

The trend of labor productivity will be replaced in forecasts by

$$\boldsymbol{\Delta\, Log(PRL_t) = Log((1 + txq)/(1 + txn))}$$

Where txq and txn represent (as seen above) the long-term growth rates of quantities and populations.

In the long run, the growth rates of quantities and employment are fixed, which defines the growth rate of labor productivity. The estimated coefficient has to be replaced by this value.

8.1.3 IMPROVING THE CHANCE (AND SPEED) OF CONVERGENCE

We have presented earlier some techniques. They can be applied indifferently to simulations over the past or the future. Few new elements will be specific to forecasts.

Actually the only significant one is associated with the starting values of the endogenous, for the computation process of a given period. EViews allows initializing the process either by "historical" values or those obtained at the previous period. This is done through the option: i=a (actuals) or i=p(previous) in the SOLVE or SOLVEOPT statement. In forecasts, only the last one should be available.

But actually one can define values on the future:

- By computing the theoretical values, applying to the endogenous variables their theoretical growth rate.
- By giving to the actual names the value obtained for a previous solution, if it used comparable assumptions to the current problem.

These values are considered as "actual" by EViews, even if they correspond to future periods[107]. They can be used as starting points, and it is possible (in particular in the second case) that they will improve the speed and probability of convergence. The issue is whether the growth rate of variables is higher than the difference for the same period between the two solutions. Both cases are quite possible, depending on the difference between assumptions in the second case.

---

[107] EViews leaves to the user the management of the present date.

In any case, the general principle applies: if one option does not work and another is available, just try it!

8.1.4 SOLVING PARTIAL MODELS

As presented earlier, an efficient way of diagnosing the origin of convergence problems is to exclude endogenous variable(s) from the solution. This calls for generating values for these variables. Generating starting values for all variables allows to use this feature at no further cost.

8.1.5 CHECKING THE EXISTENCE OF A LONG-TERM SOLUTION

We shall now present a very efficient way of checking that model specifications allow a long-term solution.

For this solution to exist, the following must be true:

<u>All the </u>exogenous<u> assumptions sharing the same dimension must grow in the long run at a common rate.</u>

Of course, their evolution can differ from this rate for the first periods.

As stated earlier, in practice three basic rates will appear:

- variables at constant prices (let us call it txq)
- deflators (let us call it txp)
- populations (let us call it txn)

The other rates will be combinations of the three (in addition to pure stability). The main ones are labor productivity and purchasing power of wages: txq-txn or rather (1+txq)/(1+txn)-1

The trends must be stopped, or set to theoretical values (like the trend in labor productivity).

In estimated formulas, the long-term equations (in principle cointegrating relations) should use elements with zero dimension, and the dynamic equations too. This is not completely necessary, but it makes things much simpler, and is required anyway if we want the equations to be homoscedastic. Stationarity is more easily checked on elements with no dimension.

For dynamic equations, using logarithms (in general their variations) will provide this property naturally.

Under the above conditions, every formula must give a solution growing at the rate associated with the element it defines. For instance, the equation for imports at constant prices must give a result growing at rate q.

It is enough that a single equation does not meet these conditions to forbid the whole model from having a long-term solution, let alone reach it.

In addition, the estimated formulas must lead to the solution and not away from it. In particular the sign of the coefficients associated with error correcting terms must be the right one, and the value lower than one to create a diverging process.

If these conditions are not met, the model will probably explode in the long run, and it is quite possible that this process affects already the medium-term solutions. To be sure this is not the case, the only option is to check the presence of a long-term solution.

Of course, one can start by simulating the whole model over a long enough period. Unfortunately, even if high care has been taken, at least one error will generally remain, propagating to the whole system without generally giving clear clues as to the origin of the problem.

However, a very simple technique can be applied: producing a residual check over the future!

One just has to generate (as proposed earlier) a full set of exogenous and endogenous elements following the above constraints. Then the model is solved using the “fit” option (separate computation of each equation) and the growth rate of the endogenous result[108] is compared to the growth rate of the variable the equation defines. The two must be identical. If not, the elements in the equation are not consistent with its formulation.

The origins of problems are somewhat different from the usual residual check. The most frequent are:

- A trend has been forgotten or miscomputed in the equation, for instance the opening of world trade in the imports equation[109].
- The dimension of one or more element on the right hand side is not consistent with its role. This can apply to identities or estimated formulas.

For instance the exogenous variable associated with social benefits might be computed in purchasing power per head, while the social benefits equation just multiplies this element by the deflator. Or a linear estimation adds up elements with various dimensions.

- The dimension of the endogenous variable is inconsistent with the formula. For instance a CES production function uses the wrong coefficients, used as exponents.
- We have decided to apply different formulas for the past or the future, for instance, blocking the trend after the estimation period, and the second type contains errors of the above types..

For instance, a trend term such as:

```
C(1)*((t-2007)*(t<2007)+0.02*t*(t>=2007))
```

instead of:

```
C(1)*((t-2007)*(t<2007)+0.02*(t-2007)*(t>=2007))
```

will make the element jump from 0 to 0.02*2007 when forecasts start in 2007.

---

[108] Not the growth rate of the left hand side expression itself.

[109] This trend can be continued in the initial forecasting periods, but it must decrease and disappear after a while. In this case, one must check the residuals only after this condition is met.

The obvious advantage of this technique is to identify most of the erroneous equations in the model (hopefully a limited se)in one pass. But as in the more usual application, a zero residual does not prove the absence of error, and a non-zero one can come from several errors. This means that in general reaching an (apparently) cleaned version will take several iterations.

The programs provided at the end of the book will systematically use this technique.

8.2 CONVERGENCE PROBLEMS IN THE SHORT RUN

Success in the above test does not guarantee reaching a consistent solution, however: now the problem lies in the initial periods, and particularly in the very first one.[110]And the above method cannot be applied, as the lagged elements on the right hand side do not yet grow at the theoretical rate.

The most immediate danger is an error in the assumptions: the above program, which sets the growth rates to all the exogenous variables, should avoid this problem. We have checked that their growth rates are correct, and consistent with the formulas in which they are used. The only error could come from values given explicitly, for instance because we have advance information on some elements, or because we are trying to specify endogenous target by manipulating exogenous instruments.

Finally, it is possible that the management of residuals in the first forecast periods brings high variations, and in particular strong cycles, which

- make one of the elements pass through unacceptable values (such as negative levels for goods, which enter behavioral equations by their logarithm)
- or produce a model with diverging properties. This can be true for Gauss-Seidel. If the linearized model is very different from the right one, it might can contain eigenvalues with a modulus higher than one. And the Newton method can make the solution move in the wrong direction.

As to the remedies, they are not different from the ones developed in the previous chapter..

8.3 CONVERGENCE PROBLEMS IN THE MEDIUM RUN

Now if the model has started converging, and contains a converging error-correcting process, it should give solutions over the whole period.

Of course this is not true if the second condition is not met: in that case we know nothing about the long run properties, and divergence can occur at any time. Actually the greatest danger occurs when convergence is still obtained at the end of the forecasting period, but numerical properties are already wrong, not yet bad enough to be noticed.

Otherwise, the only danger which remains is if the cycles are too strong, which can only occur if the model is wrong. But now we have help, just as if we tested a motor which runs for a while before breaking down: it is much easier to understand the reason than if it does not start at all.

The main solutions have been developed earlier:

- Applying shocks to the main assumptions and observing the consequences. Maybe the diverging process will start before the breakdown, and its reasons can be made explicit.

[110] Just as takeoff is the most dangerous part of a plane flight.

- Excluding some variables, or changing the intensity (increasing or decreasing) of local explanations. Now we can observe how these changes affect the horizon of the crash, and the properties: less cyclical or more consistent with economic theory. This is much more informative than just trying to get a solution.

8.4 TESTING THE RESULTS

Let us now suppose that our model is converging in the long run. It means that we have efficiently managed to establish a long-term solution (remember that we are considering a very long period).

8.4.1 OF SIMULATIONS

What we should check at this stage is how the model behaves in the short and medium runs. Of course, we have introduced assumptions too crude to provide a realistic forecast. In particular, setting from the first the growth of word demand to the common value txq will stop the expansion of exports (imports should follow). A more realistic option should be to reduce it gradually. However:

- The impact of this gradual decrease will change as the model evolves, so it will have to be reconsidered with time.
- This will go against the main present goal in simulating the model over the future, which is obtaining trajectories which are stable, thus easier to interpret, in a future as near as possible.

In practice the quality of our simulations has just to be reasonable enough not to question the teachings of sensitivities to shocks. They will represent the most important diagnosis.

8.4.2 OF SHOCKS

For the shocks, we shall use the same « Scenario » feature as in the previous chapter. We have already described the advantages of waiting for this stage to make the tests. The situation has not changed much.

As we have stated earlier, it is on the future that shocks are better performed:

- They can last longer.
- They start from a smoother baseline.
- The period is the same as for future uses.

In particular, one can check the exact long run values and ratios. Long-term convergence itself is guaranteed by the previous production of a long-term scenario, provided the shock is not too large.

Concerning the period itself, it is better to start with a very long one (like two or three centuries). This might look unreasonably high, but it allows:

- To check that indeed the long run is reached numerically.
- To interpret the convergence (or divergence) of potential cycles, which have to be observed in a high enough number. Cycles with a 50 years period are not uncommon (especially if the model slightly overdoes it) so observing 4 instances requires 200 years...

Of course once the model is stabilized, the cycles will probably converge faster, and the test period can be reduced, allowing to concentrate of the short-medium-term properties (which will be much more visible in shortened graphs).

8.5 EVIEWS FEATURES

We will now see how we can apply the above principles to EViews.

8.5.1 PRODUCING SIMULATIONS

We shall now describe in turn the phases is the simulation process.

8.5.1.1 Preparing the workfile

Let us see how we can manage this on our simple model. We shall concentrate on the specific features (the full program is made available later).

The first task is producing a « forecast » file. The period considered should be quite large.

```
close fra_q.wf1
open fra_q.wf1

wfsave _fra_1.wf1
pageselect fra_m2
pagecopy(page=fra_proj)
pagestruct(end=2100Q4)  *
```

8.5.1.2 Producing the assumptions

We can also compte the endogenous (as justified earlier).

Now we will produce the assumptions.

Fortunately, most of the assumptions have no dimension, and are independent from the growth rates.

The only exceptions are:

- Government and world demand, gd and wd, quantities which grow at txq.
- Government employment lg which grows at txn

```
scalar txq=exp(0.006)-1
scalar txn=exp(0.0005)-1
smpl 2015Q1 2100Q4
genr T=t(-1)+0.25
 '  exogenous
for %1 gd wd
smpl 2005Q1 2100Q4 if {%1}=na
genr {%1}={%1}(-1)*(1+txq)
next
for %1  r_rhiq dr sr pk compm ec_i ec_le ec_m ec_x  ec_ci ct r_ih
smpl 2005Q1 2100Q4 if {%1}=na
genr {%1}={%1}(-1)
next
for %1 lg
```

```
smpl 2005Q1 2100Q4 if {%1}=na
genr {%1}={%1}(-1)*(1+txn)
next
for %1 wr
smpl 2005Q1 2100Q4 if {%1}=na
genr {%1}={%1}(-1)*(1+txq)/(1+txn)
next
'   endogenous
for %1 ci i m  x co fd k rhi q cap  ih
smpl 2005Q1 2100Q4 if {%1}=na
genr {%1}={%1}(-1)*(1+txq)
next
for %1 ur res_m  res_x
smpl 2005Q1 2100Q4 if {%1}=na
genr {%1}={%1}(-1)
next
for %1 le lt led
smpl 2005Q1 2100Q4 if {%1}=na
genr {%1}={%1}(-1)*(1+txn)
next
for %1 prle_t
smpl 2005Q1 2100Q4 if {%1}=na
genr {%1}={%1}(-1)*(1+txq)/(1+txn)
next
```

Changing coefficients

For our simulation to provide a steady state solution, we only need to change one coefficient: the trend in labor productivity which must follow txq-txn (in Δlogs) at each period. As it applies to t which grows by 0.25 a each quarter, we must multiply the coefficient by 4. And as we have seen, our management of trends makes the change in coefficient apply only from 2004, without calling into question the previous values of trend productivity.

We will replace c_prle(2) by 4*log((1+txq)/(1+txn))

8.5.2 PRODUCING A BASE SOLUTION

To produce a base solution, we use the same technique as usual. We compute the growth rates of the endogenous to check that they converge to the long-term values (zero, txq, txn or a combination of both).

Two remarks:

- We shall also compute the theoretical growth rates, and the difference between actual and theoretical, which should decrease to zero in the long run.

Actually we shall compute the Δlogs, for which the variables associated with combinations of growth rates (like labor productivity) will nevertheless give round values, making a visual check easier.

As stated earlier, this technique represents a trick, and one can rely on actual growth rates. It is in our sense less efficient, but it is much easier to explain (consider the length of my message, and I am not sure it actually went through).

- The blank “override” statement insures that only base series are used..

- The blank “exclude” statement insures against any previous exclusion associated with the present scenario. It is free, and one cannot be sure that it is not needed.

```
smpl 2006Q4 2100Q4
_fra_1.scenario "Scenario 1"
_fra_1.append assign @all _p
_fra_1.override
_fra_1.exclude
_fra_1.solveopt(n=t m=1000,o=g,d=d)
_fra_1.solve
for !i=1 to g_vendo.@count
%st2=g_vendo.@seriesname(!i)
genr tc_{%st2}=100*log(({%st2}_p+({%st2}_p=0))/({%st2}_p(-1)+({%st2}_p(-1)=0)))
genr tc0_{%st2}=100*log(({%st2}+({%st2}=0))/({%st2}(-1)+({%st2}(-1)=0)))
genr dtc_{%st2}=tc_{%st2}-tc0_{%st2}
next
```

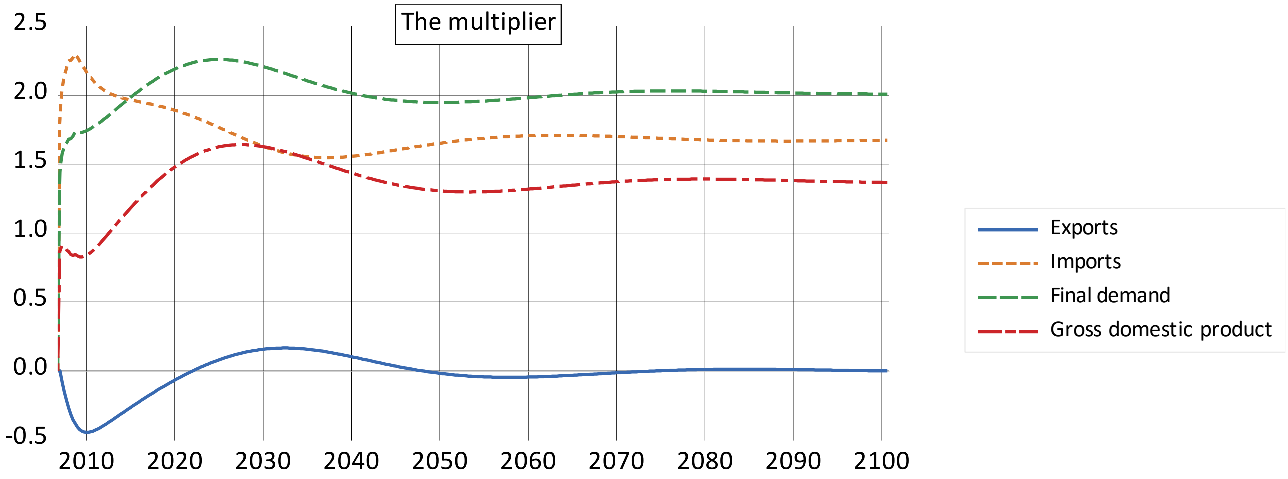
8.5.3 PRODUCING SHOCKS

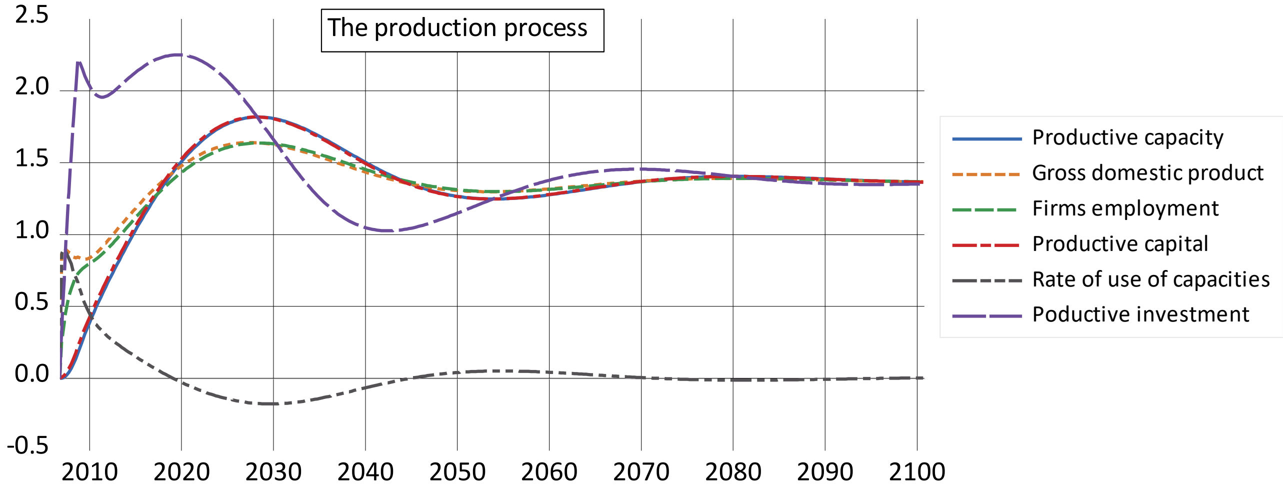
To produce shocks we use the same technique as before. We apply to gd an increase of 1 GDP point (using q_p, the GDP level computed by the baseline simulation). We solve the model using the suffix “_v”. Then we compute the absolute and relative differences (destroying first any variable of this type, to avoid confusions).

Finally we can export the results to Excel (not necessary with the present EViews graphic potential).

A trick: when a simulation breaks down in a program which contains several SOLVE statements for a given model, the EViews message will not let you identify which of them produced the error. A simple way to make this identification is to change the maximum number of iterations in the statement (1000 into 1001, 1002,….), as this statement will be displayed in the message.

```
 genr gd_v=gd+.01*q_p*(t>=2007)
_fra_1.solveopt(n=t m=1002,o=g,d=d)
_fra_1.scenario "Scenario 1"
_fra_1.append assign @all _v
_fra_1.override gd
smpl 2006Q4 2100Q4
_fra_1.solve
for !i=1 to g_vendo.@count
%1=g_vendo.@seriesname(!i)
series dv_{%1}={%1}_v-{%1}_p
series pv_{%1}=100*dv_{%1}/({%1}_p+({%1}_p=0))
next
group chocs  pv_* dv_*
write(t=xls)  varia_p.xls chocs
'      We create groups for the supply-demand equilibrium, and for the production process
group g_v1 pv_x pv_m pv_fd pv_q
group g_v2 pv_cap pv_q pv_le pv_k pv_ur pv_i
```





The results look rather similar to the previous shock. We can only observe that with a more recent period, the role of the trade variables increases, leading to a lower multiplier. The main observation is that now the smooth evolutions make the results much clearer.

In the long run, the value of the multiplier is about 1.4, a somewhat high value. But we do not take into account the loss in competitiveness on both external trade elements, which would come through inflation.

We can observe that the rate of use of capacities gets back to its base value. This leads exports to do the same.

The addition of investment and consumption to the ex ante change in Government demand makes the global variation grow to 2 %. Imports grow less as they depend also on the unchanged exports. And GDP even less, as it depends even more on exports.

The stability of the rate of use gives investment and capital on one side, and GDP and capacity on the other, a constant ratio, and thus the same relative change. As productivity of capital is fixed, all these variables change in the same proportion. And as productivity of labor is also fixed, firms employment follow. The change in inventories also, although with more sustained fluctuations.

But government employment does not change. Thus total employment and wages grow less, as well as household income and consumption, which depend also on GDP

8.5.4 CHANGING MODEL SPECIFICATIONS

To test model properties in the long run, it can be interesting

- To test alternate estimated formulas, the properties of which have to be ascertained.
- To allow some of its equations to follow formulations different from estimated ones, either through a change of coefficient values, or even a change in the actual formula.

This can be done of course by producing a new model creation program. But this will have several drawbacks.

- It will take more time and involve more statements.
- It will not single out an “official” version.
- One will have to maintain several model-building programs, making file management less clear.
- The forecasting program will have to be run anyway.

In our sense it is better:

- To maintain an official version of the model, representing the best tested version according to the present information.
- To introduce the changes in the “forecasting” program (or programs, documenting the differences)..

This can be done in two ways.

- If a new official version has been estimated for some equation, the full set of associated statements should be specified (residual set to zero, defining the equation and estimating it, storing the estimated residual, merging the equation).
- If one is just testing an alternate version with at least one fixed coefficient, with or without re-estimating the rest, we suggest the following procedure, which we shall present using investment as an example:

  - Copy into the “forecasting” program the set of statements, as above.
  - State a vector of parameters (which will not be estimated).

```
vector(10) p_i
```

  - Estimate the equation as usual, with a different name (here: the normal name followed by “1”).

```
coef(10) c_i
smpl 1980Q1 2004Q4
genr ec_i=0
equation eq_i1.ls i/k(-1)=c_i(1)*i(-1)/k(-2)+c_i(2)*(ur_star-ur)/ur+c_i(3)*.125*q/q(-8)+c_i(4)+ec_i
genr ec_i=resid
```

This step can be bypassed if you are sure that the present equation is correct, and you know the estimation results. However, we advise it as it displays the situation and helps you control the process.

  - Transfer the coefficients into the vector elements, making any desired change in the values. Here we want to set the tension coefficient to 0.5.

```
p_i(1)=c_i(1)
p_i(2)=0.50
p_i(3)=c_i(3)
```

- Replace all “c_ “coefficients by “p_” parameters, except for the constant term (this will allow EViews to estimate the equation, a process which require at least one coefficient).[111]
- Estimate the equation using the normal name.
- Store the new residual value.

```
equation eq_i1.ls i/k(-1)=p_i(1)*i(-1)/k(-2)+p_i(2)*(ur_star-ur)/ur+p_i(3)*.125*q/q(-8)+c_i(4)+ec_i
genr i_ec=resid
_fra_1.merge _eq_i1
```

Of course, we can modify as many coefficients as we want, but not the constant term if we want the residuals to add up to zero, and EViews to recognize the equation as econometric.

To apply this technique, we can consider two general strategies.

- Start with no changes, and transfer in succession each equation which requires it.
- Copy at the beginning all equation-building elements into the forecasting program, with no value changes (all the “c_“ elements are transferred into equivalent “p_”s). The specifications have changed, but not the numerical properties.
- The requested paragraphs will be changed in turn later, as needed. This calls for an initial investment, but makes the following changes easier and less error prone.

As you can guess, we favor the second option. But one should consider the share of equations he contemplates changing[112].

Note: when you are testing new values, you should keep the old one as a comment, which will allow you to backtrack if the situation deteriorates, and also to recall the level of your change.

8.5.5 UPDATING THE MODEL SPECIFICATIONS

The above tests, or other considerations like a desire to introduce a new theoretical feature, can lead the modeler to update the model. This can represent a very simple change, like replacing the unemployment rate by its logarithm in the wage equation, or a very complex one, like changing the whole production function.

This process can lead to the introduction of new variables, or completely new behavioral equations.

Once the new set has been established, two options are available:

- If the model specifications are produced by a single program, one can simply edit it. Of course, one should keep a version of the old program, in case mistakes are made or the new model proves unacceptable, for any reason.
- One can keep the previous model creation program (or programs) and create an additional updating program applying the changes.

[111] We have no simple solution if the equation has no constant term (as for the change in inventories in our example).

[112] Which is difficult to forecast.

As you can expect, we favor the first option, which is both safer (recreating the model is completely safe forward) and clearer (the specifications of the model can be observed by sight, especially if it is not too large).

You can also (if you choose that option[113]):

- Drop any kind of equation, using :

model_name.drop *variable_name* (for identities)

model_name.droplink *equation_name* (for estimated equations)

For instance you can specify:

```
_fra_1.drop GDP    '   GDP is made exogenous

_fra_1.droplink eq_i    '   eq_i, the equation for investment is dropped, and I becomes exogenous
```

- Replace the formula using:

model_name.replace *variable_name* (for identities)

model_name.replacelink *old_equation_name new_equation_name* (for estimated equations)

```
_fra_1.replace GDP = ……

_fra_1.replacelink eq_i eq_i_new
```

- Finally (the most important) one can replace one variable name by another in all model occurrences.:

[113] Remember that in our opinion, the most efficient method for editing a model is to destroy the previous version and create a new one from scratch.

```
model_name.replacevar old_variable_name  new_variable_name
```

For instance:

```
_fra_1.replacevar GDP Q    '   replaces GDP by Q in the whole model (but not in estimated relations)
```

There are at least three interesting applications.

- You receive a model from some source and you want to change the names to ones you are familiar to.
- The names follow a language and you want to change it. For instance moving from English to French you want to replace FD by DF ("demande finale").
- You obtained a model from the World Bank using its complex names (like "ny_gdp_mktp_kn" for GDP) and want to change it to plain "Q".

The good point of this feature is that you can switch between versions, provided you follow the right order. So you can work with modelers using World Bank names, and exchange the same model, each partner using his own names.

You have to be careful, however. If in your model QV stands for GDP at current prices, and Q at current prices, QV should be processed first. And if you have specified QV_RES as a residual, it should be treated at the very first.

This means that this method applied essentially to models with names long enough, to avoid mix-ups.

It should be clear however that to follow the process, using programs is mandatory.

8.5.6 MODEL PROTECTION

One can protect a model to avoid modifying or deleting by accident.

This requires access to the Protect / Unprotect item in the Proc menu of a given model.

In case of protection, a password must be specified to update it.

Apparently this feature is only available by menu.

8.6 SOME RECENT USEFUL FEATURES TO BETTER MANAGE SIMULATIONS

Since Eviews 9, features have been added which do not modify the computing process, but make its management easier. We will present them using the model we have just described.

8.6.1 THE GRAPH OF DEPENDENCIES

You can now asks for a graph showing the links between model elements.

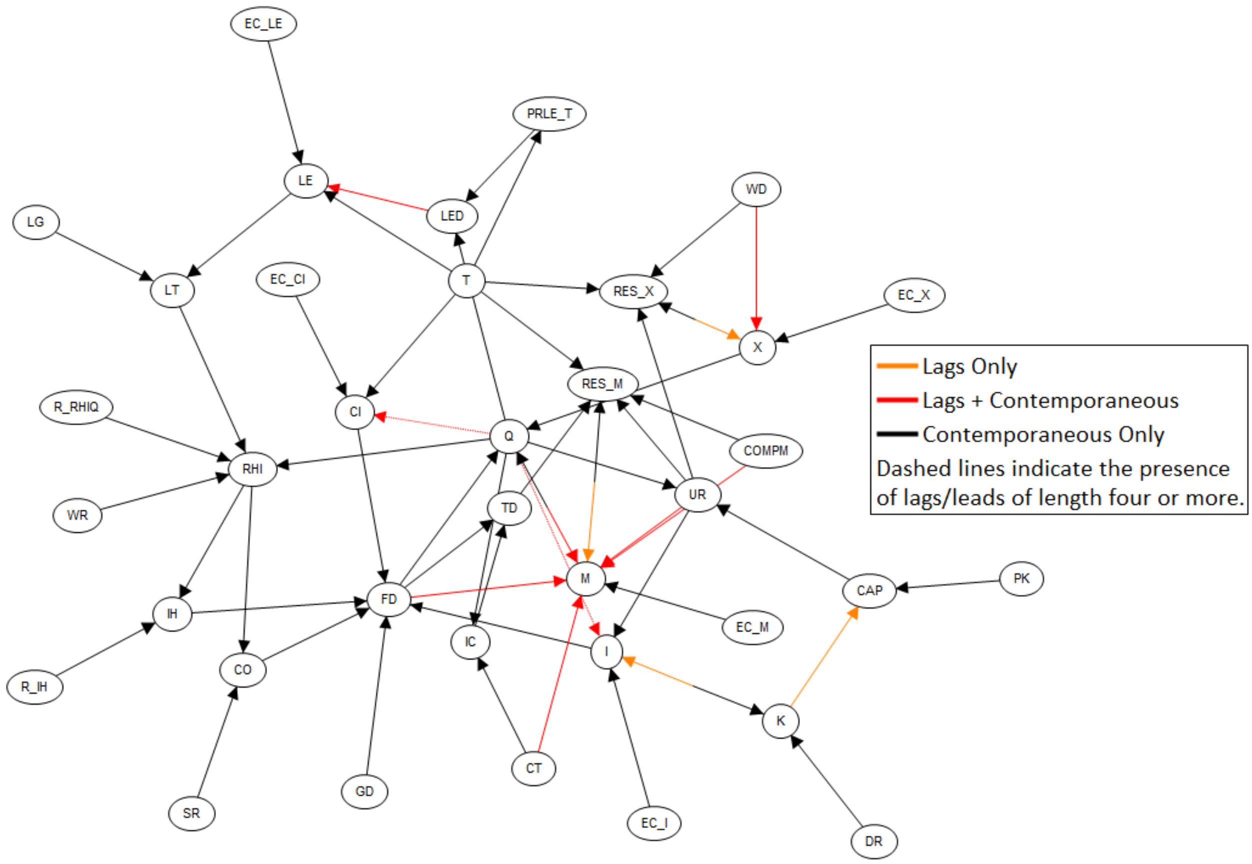
This obtained by

Model>View>Dependency Graph

Arrows allow to observe the nature of links (simultaneous, lagged or both) according to a color.

This feature is just right for this size of model. It would have no use for a very small one, and be too complex for a large one. We shall prove this later.

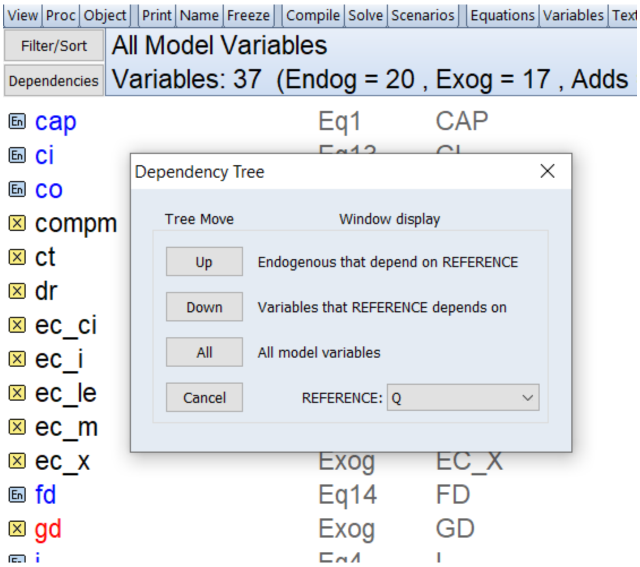
This graph evidences the density of the links (for instance the role of Q and FD). The only drawback is that it does not take into account the economic implications. We think the graph we have designed is more helpful in presenting the model structure (but it is longer to produce).



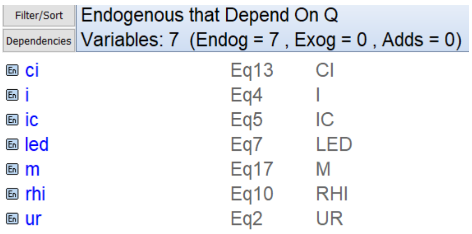
8.6.2 INTERPRETING THE LIST OF VARIABLES

For a given variable, EViews allows to identify the elements it depends on, and influences (only endogenous of course). The causality path can then be followed further for one of these elements.

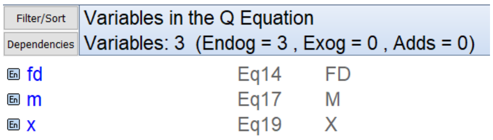
Let us apply this feature to Q (GDP)



These variables depend on Q:

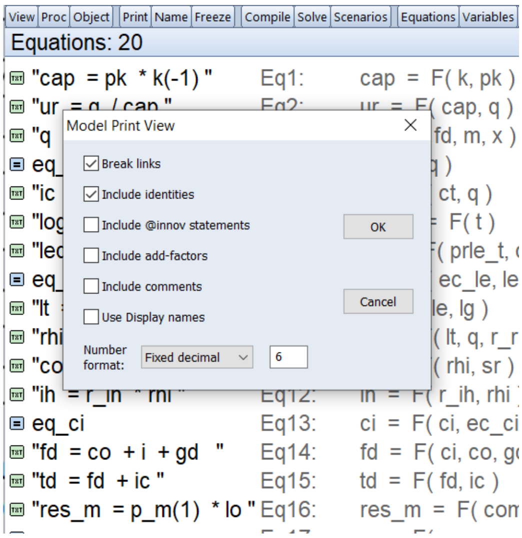


These variables determine Q:



8.6.3 THE "PRINT"DISPLAY

To get the text of the model in the best editable form, one should call for the Print View" item in View menu of the model. You should "include identities" and display coefficients as "Fixed decimal".



The estimated coefficients will be replaced by their numerical values.

Later you can add equation numbers. For that, we propose is the following sequence, a little long but not error prone.

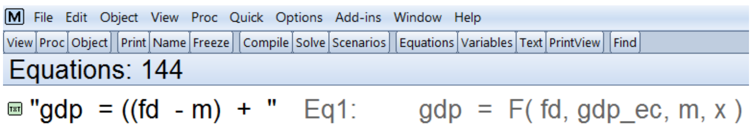
- You copy the text to Excel (or equivalent).
- Most of the equations have the endogenous at the left. For the others (like Dlogs, maybe 5% of the total) you type the name of the variable before the expression.
- You insert a column with numbers starting at 1.
- You sort the sheet according to the formulas.
- You delete the blank lines at the end.
- You insert two columns of "[" before the numbers and "]" after.
- You sort the page according to the (now odd) numbers.
- You delete this first column of numbers and insert a new one between the brackets.
- You delete (by hand) the names you have inserted in the "5%" formulas.
- You select the full zone and "Copy special" it to Word (or equivalent) using the "a" option (the last one).
- You replace the tabulations (hidden but accessible as "^t") by a null element or a blank . By the way, the item "insert a jump to the next line" is accessible as "^p". This can allow you to transform a column of items into a line, separated by blanks or commas.

Et voilà!

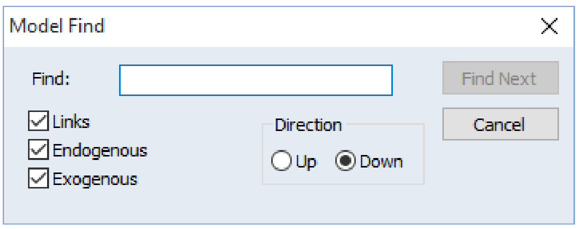
8.6.4 EQUATION SEARCH

One can examine the set of equations, looking for the presence of a given variable (or aa particular equation).

This is done under the "Equations" view using the "find" button at the end of the line.



One can limit the search to endogenous / exogenous, or links.



8.7 RATIONAL EXPECTATIONS

EViews calls them « model consistent expectations » and their rationale is a little different from the usual one.

8.7.1 THEFRAMEWORK

In a rational expectations framework, we suppose that some of the agents (at least) are aware of the future assumptions, and are able to evaluate their influence on some (at least) of the model variables. This does not need the knowledge of the actual model equations, just the mathematical « application » from the assumptions to the endogenous.

As they will use this knowledge in their present decisions, some of the present variables will depend on future values, either of assumptions or endogenous elements depending on these assumptions.

To take into account rational expectations one does not need to believe in them. Interpreting the differences in economic behavior (and their consequences for the equilibrium) between forward and backward looking agents is quite interesting if only from a theoretical point of view. The following example will shed some light on this point.

8.7.2 CONSEQUENCES FOR MODEL SIMULATIONS

In this context, we can no longer compute the solution for each period separately, moving from the past to the future. The solution for a given (future period) will depend on values for which the solution has not been yet obtained.

This introduces two problems:

- Finding a way to take into account future values, belonging to the forecasting period.
- Finding a way to end the simulation, the last simulations taking into account values outside the forecasted period. This calls for the definition of « terminal conditions ».

8.7.2.1 Inside the forecasting period

The problem is purely technical. Two strategies are considered at present.

- The Fair – Taylor algorithm: it solves the model for each period in succession (using the Gauss-Seidel, Newton or Broyden algorithm each time), then goes back to the first one and iterates (in a Gauss – Seidel way) until convergence of the whole system. This is the technique used by EViews.
- The Laffargue (1990) algorithm: basically, it adds a time dimension to the model equations (duplicating them as many times as there are periods) and solves the associated model as a whole, using the Newton algorithm with possible improvements (using the fact that the matrix is band diagonal, the number of non-zero matrixes on a given line depending on the number of backward and forward lags).

Obviously, applying the last algorithm was only made possible by the increase in the computing power of computers, and more importantly of the memory capacity, as a very large set of values has to be managed at the same time.

8.7.2.2 Outside the forecasting period

EViews provides the usual options for terminal conditions, which are controlled by

```
t=arg
```

in the SOLVEOPT statement

arg can take the values:

```
"u" (user supplied – actuals, default)
"l" (constant level)
"d" (constant difference)
"g" (constant growth rate)
```

8.7.3 TECHNICAL ELEMENTS

EViews will detect automatically the presence of future endogenous variables, and will switch to forward looking algorithm (actually to Fair-Taylor). One can use the same statements as usual, or change the terminal conditions from default values, if deemed useful.

8.7.4 OUR EXAMPLE

In our example, we shall limit the case to the investment equation.

In the present model, we use:

```
I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.125*Q/Q(-8)+C_I(4)
```

Dependent Variable: I/K(-1)
Method: Least Squares (Gauss-Newton / Marquard
Date: 03/25/20   Time: 17:42
Sample: 1980Q1 2004Q4
Included observations: 100
I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*(UR_STAR-UR)/
+C_I(4)+EC_I

	Coefficient	Std. Erro
C_I(1)	0.944504	0.01330
C_I(2)	-0.002560	0.00089
C_I(3)	0.042822	0.01358
C_I(4)	-0.004488	0.00167

R-squared	0.985646	Mean depe
Adjusted R-squared	0.985198	S.D. deper
S.E. of regression	0.000260	Akaike info
Sum squared resid	6.51E-06	Schwarz c
Log likelihood	685.4474	Hannan-Q
F-statistic	2197.406	Durbin-Wa
Prob(F-statistic)	0.000000	

Now we shall consider four options:

The basic case (full backward looking)

```
I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.25*(Q-Q(-4))/Q(-4)+C_I(4)
```

A mixed backward – forward looking case

```
I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.25*(Q(2)-Q(-2))/Q(-2)+C_I(4)
```

A full forward looking case

```
I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.25*(Q(4)-Q)/Q+C_I(4)
```

Another full forward looking case, extending the knowledge to a longer horizon

```
I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.05*(Q(20)-Q)/Q+C_I(4)
```

The global influence of Q is the same in each case, only the lags differ.

8.7.5 THE TEST

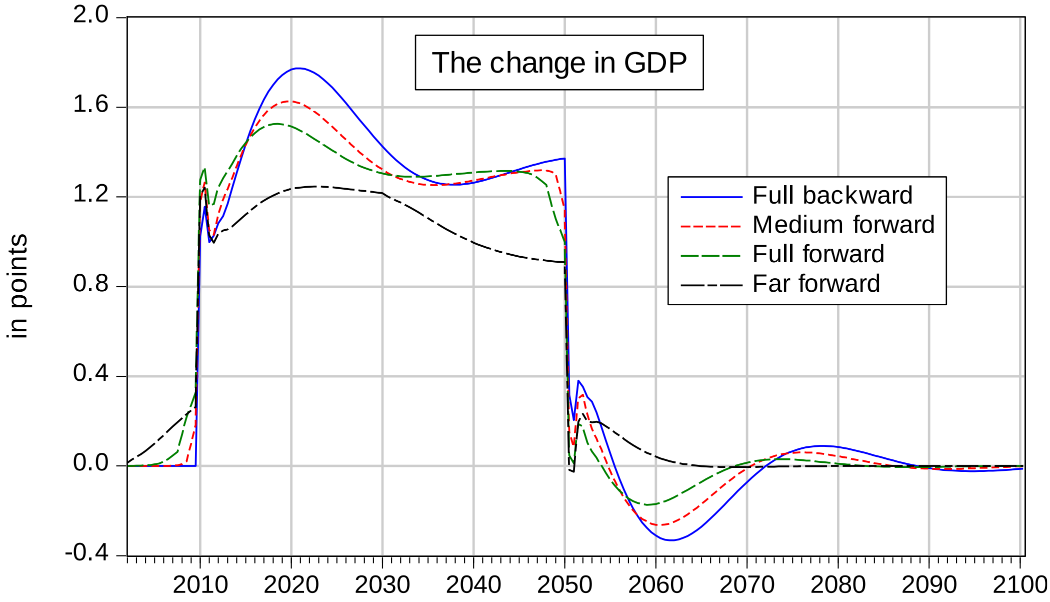
We shall consider that the State increases its demand by 1% of GDP (as in the previous shocks) but only from 2010 to 2040 (considering perhaps that extending the decision would lead to an unacceptable debt level). The length of this period might seem too high, but it will be necessary for the analysis[114].

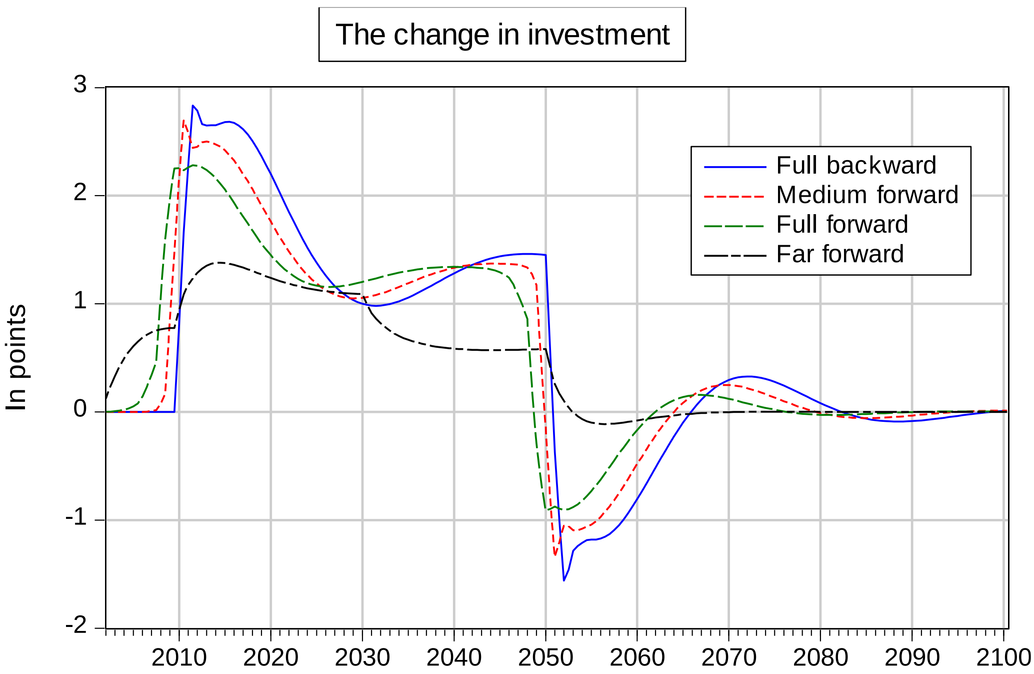
Of course, one could also apply the technique to other elements: employment and change in inventories in particular. One should be able to do it easily.

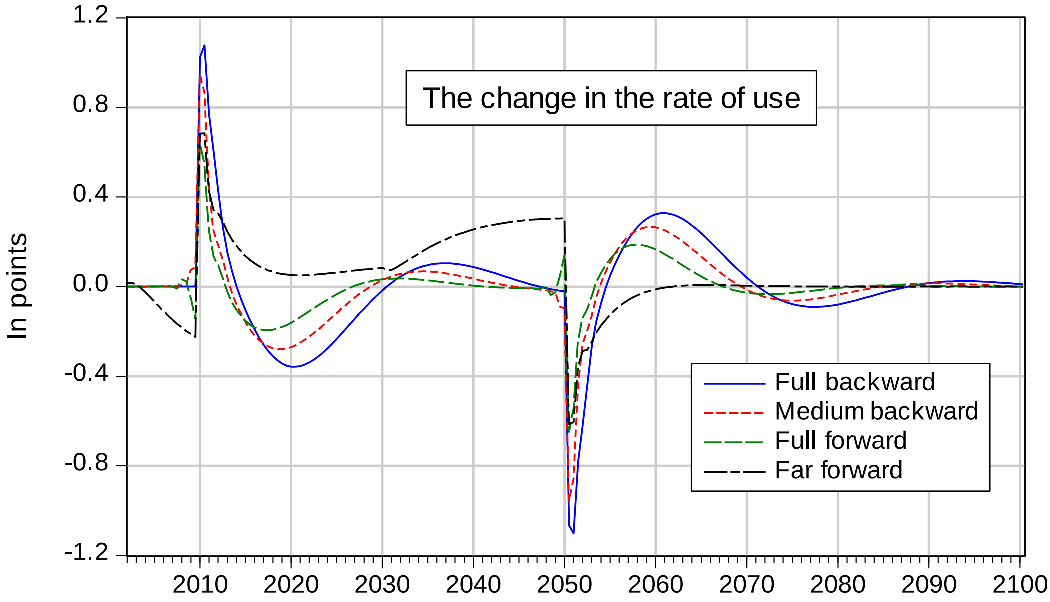
8.7.6 THE RESULTS

We shall concentrate on a few graphs. All changes are measured in percentage deviations from the baseline.

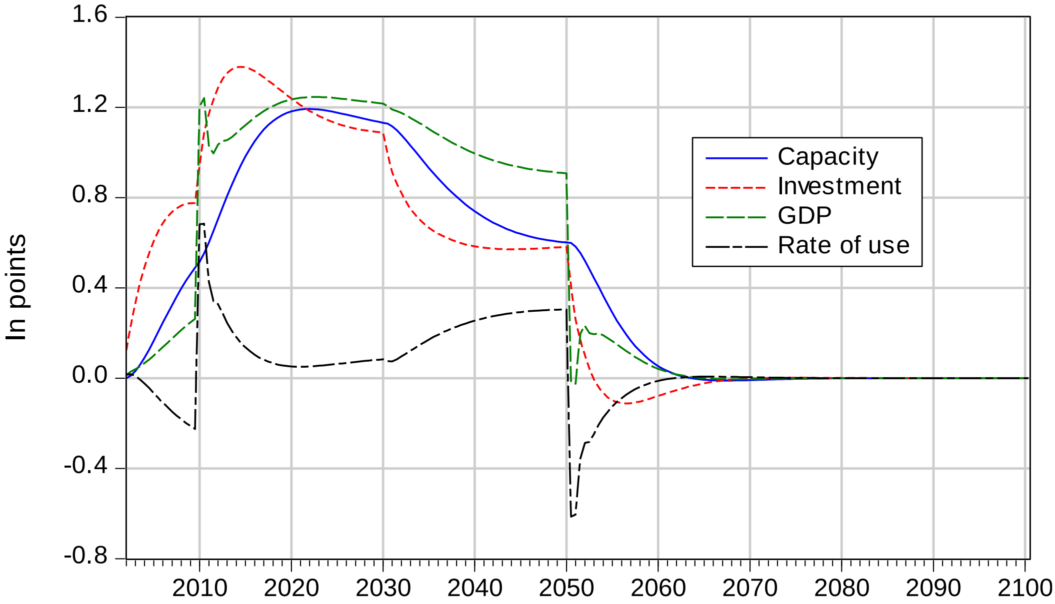
[114] In fact, this decision will lead to an increase in the debt primary of the State by 40 percentage points of GDP (excluding the additional interest). The level of debt is certainly unbearable, reducing the interest of the study, particularly concerning the reaction of other agents.


The change in GDP
in points
Full backward
Medium forward
Full forward
Far forward
2.0
1.6
1.2
0.8
0.4
0.0
-0.4
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100


The change in investment
In points
Full backward
Medium forward
Full forward
Far forward
3
2
1
0
-1
-2
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100



The 20 years ahead case



We can see very clearly generally logical elements:

- The more the investment decision is based on future GDP, the more investment adapts in advance.
- This is also true if the horizon increases. In this case, the weight of the periods associated with a change in decisions goes up.
- This produces a smoother adaptation to a change in Government decisions, in both events. The intensity of the cycles decreases, and they converge faster.
- With a 40 year horizon (case 4), the output gap does not change too much. Its higher values appear before the changes, when the investment decisions (and the resulting capital changes) are made according to the future, not the present, conditions. The negative overshooting of capacities observed in the last period is not present, and the convergence of the rate of use is monotonous.

- The changes appear sooner than the shock than the forward horizon in the equation itself. The early response of forward looking firms is taken into account even earlier by other firms, introducing a cascading effect (rather limited, however).

8.8 STOCHASTIC SIMULATIONS

For scientific purposes but also the analysis of a given model, the uncertainty of the model in forecasting may be directly measured. This error can come from several sources:

8.8.1 PURELY STATISTICAL ERRORS

This means

- The presence of a random residual term in behavioral equations
- The subsequent uncertainty on coefficients, even with an exact specification.

When a model is produced, some of the equations are estimated using an econometric technique, from simple ordinary least squares to more sophisticated techniques like cointegration. This process provides:

- An estimate of the standard error of the estimation.
- Estimates on the standard errors of the estimated coefficients.

The precision of these two elements depends on sample size, in a different way.

- A larger sample should produce better estimates of coefficients, and reduce their error.
- But for the estimation residual, once the sample is large enough for estimation to be applied, enlarging it provides essentially a better knowledge of its distribution (in practice its standard error). This is due to the fact that even with a large sample, introducing new explanations beyond five or six variables leads very quickly to collinearity problems, and the imagination of the modeler will dry up anyway.

8.8.2 FORMULATION AND ASSUMPTION ERRORS

Actually, these two errors are only half of the types facing model forecasts. One must also consider

- The fact that the estimated formula itself could be wrong. Even if a formula passes all tests, one can always find another formula (sometimes with very different properties) which will show the same statistical quality, or better. And even if the formula applied to the past, the agents could very well modify their behavior later.
- The error on the assumptions used for the forecast, for which the model should not be made responsible.

8.8.3 CONSIDERING THE ERRORS

In practice only the last first types can be measured without too much difficulty, using statistics (discarding the last two): the estimation of coefficients gives under some assumptions an estimation of the distribution of estimation residuals, as well as of the distribution of coefficients themselves. The moments of forecasted variables then can be computed, either:

- By drawing at random a sample of residuals, undertaking the associated projections, and observing the statistical characteristics of results (a « Monte Carlo » technique).
- By determining, analytically or numerically, the linear transformation from residuals to the model solution, and by applying this transformation directly to the law of residuals, to obtain the law of forecasted variables.

- More simply and under clearly less restrictive assumptions, by drawing randomly a sample from the sequence of observed residuals (with or without reintroducing the elements selected into the pool). This technique is "bootstrapping".

The same type of method applies to the uncertainty due to coefficients:

- By repeating enough times the following sequence of operations: drawing a vector of residuals on the estimation period (using the estimated law or the “bootstrap" technique), simulating the model on this period to obtain a new sample of coherent variables (in the model terms), re-estimating coefficients based on this sample, and forecasting the re-estimated model. One will then measure the moments of the sample of projections.
- By using the distribution of coefficients to generate a sample of models, which when forecasted will provide a sample of variable values.
- By inferring analytically the distribution of the projected variable from the distribution of coefficients, and using the matrix transformation (Jacobian) to link the variables to the coefficients.

Three main error types should be considered:

- The bias: in the case of a model with non - linear properties (relative to the endogenous, both instantaneous and lagged), the mean of the solution will not be identical to the deterministic solution (obtained with a null residual): this introduces a bias.

To improve the measurement of the bias, one can use antithetic residuals, by repeating the simulation using the same residuals with the opposite sign. This eliminates the first order term, leaving only the second order one, the actual source of the bias. The negative correlation between the results reduces the error on the sum.

The difference comes of course from the non-linearities of the equations, and any economic model presents non-linearities. Obvious cases are the presence of variables at current prices, product of a variable at constant prices by a deflator, or variables computed by applying a growth rate to their past value.

If one is only interested in this accuracy, a single experiment (with a large number of replications) will tell of the stochastic improvement. If it is small, one can stick with deterministic (and cheaper) simulations. This test can be made as soon as the model becomes operational, and applied to all subsequent uses.

One must be aware that coefficient uncertainty can introduce non linearities even if the model is linear relative to coefficients (actually it is not linear to the set: variables + coefficients put together).

Let us take the simplest model (Q=GDP, C= private demand, g=public demand):

$$\boldsymbol{Q_t = C_t + g_t}$$
$$\boldsymbol{C_t = a \cdot Q_t}$$

It can be written as:

$$\boldsymbol{Q_t = g_t/(1 - a)}$$

which is not linear relative to a.

- the standard error: this time we consider the distribution of solutions around their average: this criterion will allow for assessing the reliability of results (starting possibly from the mean of stochastic simulations) by evaluating a confidence interval, or range of possible values.
- The distribution itself: what we want is a graph of the probability distribution of the random solutions. Obviously producing a consistent graph calls for more simulations than the previous cases. One interesting message is its symmetric character Actually EViews takes into account dissymmetry, even without a graph, by producing a confidence interval eliminating the highest and lowest 2.5% of the values.

The analysis of the uncertainty is not necessarily limited to values of variables. It also can be measured on multipliers (or derivatives of variables relative to the exogenous), therefore on the efficiency of economic policies, or on eigenvalues of the structural form, thus on the probability of model convergence (see the following paragraph).

Finally, one could consider the covariance between equations. Cointegration should eliminate the problem, as it individualizes a group of variables or concepts which evolve independently from the rest of the model.

If one considers also the error on coefficients, the process is a little more complex. Again, the sets of cointegrating equations are independent, but they always use more than one coefficient (one has to consider the dynamic equation and the long-term relationship). In this case, it is necessary to take into account the correlation between coefficients, and the drawing will call for a multivariate normal law, with a non-diagonal covariance matrix.

The two last forms of error, on the functional form and the assumptions, are more difficult to assess. However:

- If we have recent information on some variables, the main aggregates (supply-demand balance, inflation) and especially on the model assumptions (external environment, public policy), we can simulate the model over these periods and observe if the evolution of the main aggregates is correct.

A rough evaluation of the functional error can be obtained by reducing the estimation sample, and simulating the model over the remaining periods (as in an ex-post forecast). Then the size of the residuals can be compared with the theoretical standard error obtained from the previous method (which supposed the use of the true equations). If the error is much higher than the standard error (for instance systematically more than twice its value) one can surmise that the estimated functions are unable to forecast the future, and the functional form must be wrong (or the agent has changed his behavior, which is in fact the same). This technique is similar to the Chow forecast test.

This technique does not face the critique presented above, as we do not consider getting back to the choice of model.

- As to the error on assumptions, it is outside the responsibility of the model (only of the forecast producers). But an estimate of it can be obtained by applying an ARMA process to the external assumptions (not the instruments), and observing the reliability of the associated simulation. The error will then supposed to be the part of the assumption which cannot be forecasted.

8.8.4 THE INTEREST OF THE TECHNIQUE

Stochastic simulations have a cost, and an advantage. The cost is the increased length of computations.

To get an accurate measure of the distribution, a minimum of 500 replications is considered necessary, 1000 or even 10000 being a better figure. This is clearly the main reason for the delay in the use of this technique: 50 years ago, solving 1000 times even a small model could take several hours on a mainframe computer. Now the price can be afforded in most cases: on an average microcomputer, the simulation over 20 periods of the 705 equations model MacSim-CAN takes less than one second. Multiplied by 1000, we get about 15 minutes, a reasonable figure which might look large however to contemporaneous modelers, especially if simulations are part of an iterative process looking for the best forecast.

8.8.5 BACK TO OUR EXAMPLE

We shall solve the _fra_1 model over 100 periods, starting from 2005Q1.

The only necessary change concerns the **solveopt** statement, where « d=d » is followed by « d=a »:

In addition, the **solveopt** option « r=*integer* » controls the number of replications (default: r=1000), and « b=*number* » controls the confidence interval (default: b=.95)

```
_fra_1.solveopt(n=t m=1000,c=1e-6,o=g,d=d,s=a)

_fra_1.solveopt(n=t m=1000,c=1e-6,o=g,d=d,s=a,r=500,b=.98)
```

Actually, the « a » option is not the only one associated with stochastic simulations. It is the one which gives the most information, including the confidence intervals, while «m» gives only the mean, « s » only the standard deviation and « b » the means and the standard deviation.

In our case, the solution will create for each variable in the model:

- A series for the mean, adding « m » to the name of the simulated variable. For Q and a current suffix of F, the name of this variable will be « Q_FM ».
- A series for the standard error, using « S », such as Q_FS
- Series for the higher and lower bounds, using « H » and « L »., such as GDP_FH and GDP_FL

8.8.6 THE RESULTS

We shall perform a stochastic simulation on our small model, under the same conditions as the deterministic forecast, which we will actually conduct again, to measure the error it introduces.

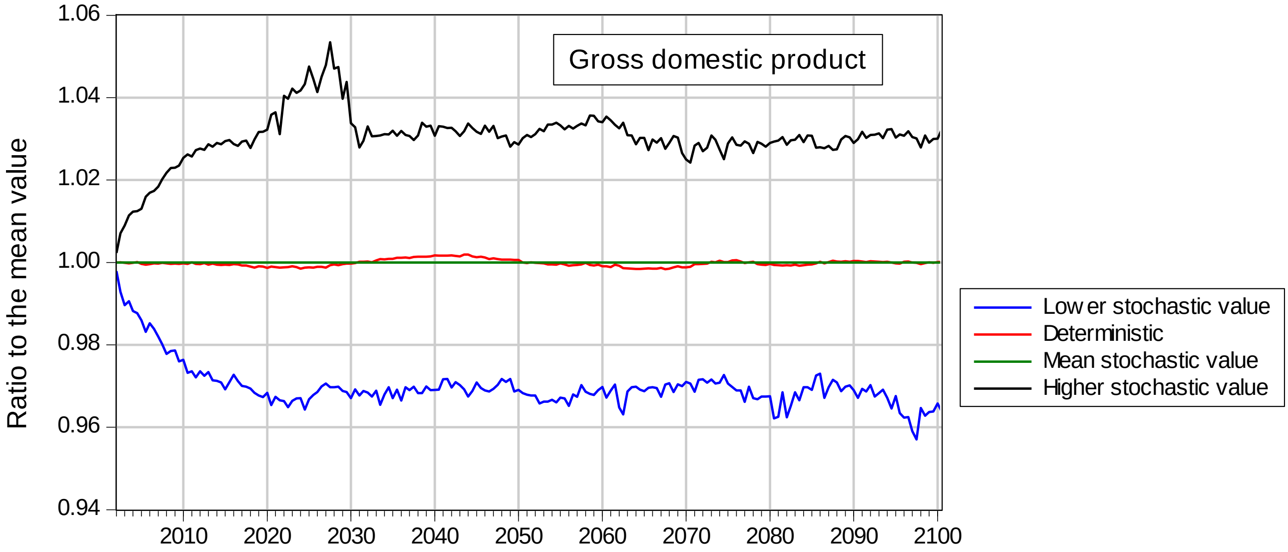
As the model and the application are not changed from the previous edition, we will be using the same graphs, more explicit in our opinion.

The program is very similar to the previous one, except for the line:

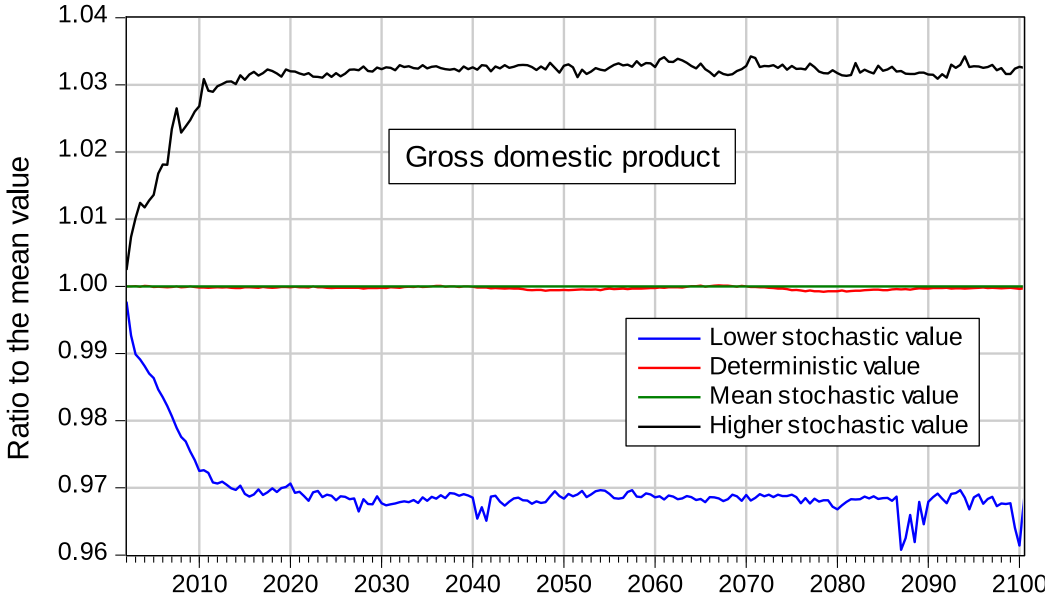
```
_fra_1.solveopt(n=t m=1000,c=1e-6,o=g,d=d,s=a)
```

The following graphs * show:

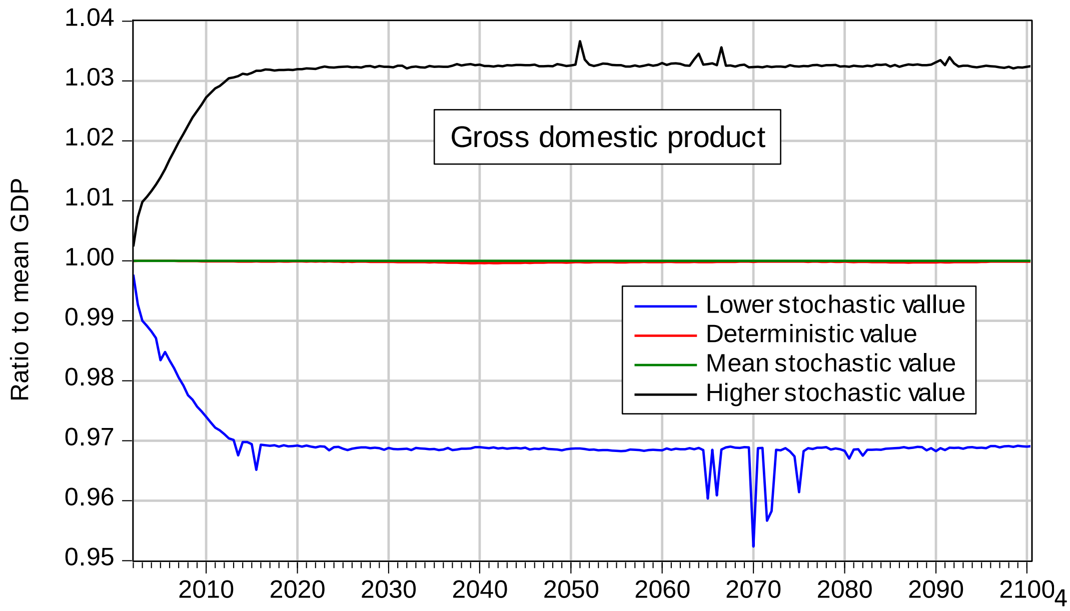
- For GDP, the evolution of its higher bound, lower bound and deterministic value, as a ratio to the average value.
- The ratio of the standard error to the average value, for a set of variables. Two graphs are actually provided, one with 100 replications, the other with 1000.

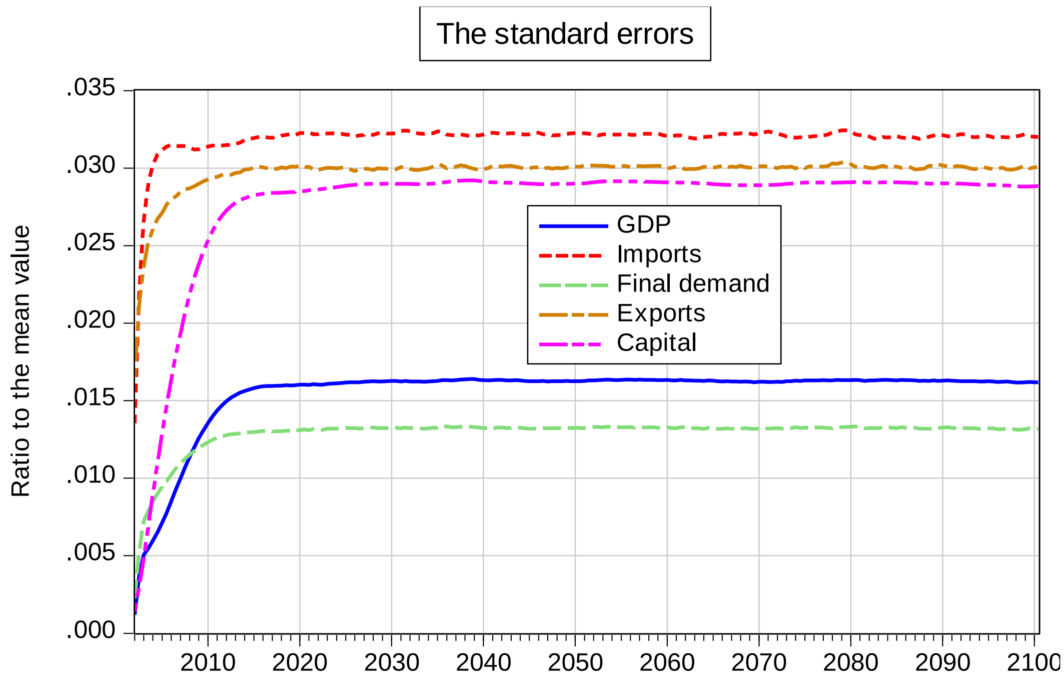


Gross domestic product
Ratio to the mean value
1.06
1.04
1.02
1.00
0.98
0.96
0.94
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100
Lower stochastic value
Deterministic
Mean stochastic value
Higher stochastic value



Gross domestic product
Ratio to the mean value
1.04
1.03
1.02
1.01
1.00
0.99
0.98
0.97
0.96
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100
Lower stochastic value
Deterministic value
Mean stochastic value
Higher stochastic value





In the above graph, one will find from top to bottom: Imports, exports, capital, GDP, final demand.

In addition, the following table gives the average relative difference from the mean, compared with the error obtained from the ex-post simulation on the past.

Variable	Stochastic error	Ex-post error
CAP	2.85	1.69
CO	1.31	1.43
FD	1.38	1.36
I	4.36	2.85
K	2.85	2.70
LE	1.68	1.68
LED	1.59	1.36
LT	1.28	1.29
M	3.23	2.81
Q	1.62	1.36
RHI	1.31	1.30
UR	1.35	1.00
X	2.99	1.77

One can observe:

- That the error is relatively small.
- That the error increases at first for all variables, as the sources increase.
- That it converges in the medium-term, faster for the trade elements (which use a cointegrating framework). This is consistent with stationarity properties.
- That the evolution is much smoother for capital (its values are highly correlated).
- That trade elements show the highest variability (and are correlated). This correlation does not come from Q.
- Comparing the ex post error, we can observe that the stochastic error is generally larger, in particular for the capacity variables. This can be due to the fact that the sum of the residuals is no longer zero, which helped the solution to get back to the baseline in the end.

8.9 GOING FURTHER: STUDYING MODEL PROPERTIES

In addition to the use of direct simulations, the mathematical functions provided by EViews allow to apply more complex techniques.

8.9.1 EIGENVALUE ANALYSIS

This technique can help to understand the dynamics of model convergence or divergence, and interpret them in economic terms.

Its application calls for the model to be linearized around a reference solution (now Δ stands for the difference between two simulations):

$$\boldsymbol{\Delta y_t = A_t \cdot \Delta y_{t-1} + B_t \cdot \Delta x_t}$$

If A is supposed to be constant with time, a shock Δx undertaken at time t (only) on the vector of the exogenous x will lead at period t+k to a change in y:

$$\boldsymbol{\Delta y_{t+k}} = \boldsymbol{A^k} \cdot \boldsymbol{B} \cdot \boldsymbol{\Delta x}_{\boldsymbol{t}}$$

and long-term consequences will depend therefore on the evolution of powers of A^k, when k tends to infinite. Three cases can occur:

A^k decreases to 0.

A^k increases to infinite (or at least one of its terms does).

A^k stabilizes at a finite, non-null limit.

Each of these three cases associates to the measure of the spectral radius of A (larger modulus of its eigenvalues): this radius will have to be respectively lower, higher or equal to unity. This will lead us to reason in eigenvalues, by stating:

$$A = V^{-1} \cdot \Lambda \cdot V$$

(Λ diagonal, containing eigenvalues of A, and V base of associated eigenvectors).

One obtains then:

$$V \cdot \Delta y_{t+k} = V \cdot A^k \cdot \Delta y_t = \Lambda^k \cdot V \cdot \Delta y_t = \Lambda^k \cdot V \cdot B_t \cdot \Delta x_t$$

Therefore, for a given shock on x, its dynamics can be studied by decomposing first period effects $B_t \cdot \Delta x_t$ on the eigenvector base of A, each element of this decomposition evolving then with time proportionally to the corresponding eigenvalue. Evolutions can be:

- Divergent, convergent or stationary, according to the modulus of the eigenvalue (higher, lower or equal to one).
- Monotonous or cyclic, according to the real or complex nature of the eigenvalue.

Actually, the stability of A over time is much more assured if one considers relative variations:

$$\boldsymbol{\Delta y_t / y_t} = \boldsymbol{A_t} \cdot \boldsymbol{\Delta y_{t-1} / y_{t-1}} + \boldsymbol{B_t} \cdot \boldsymbol{\Delta x_t / x_t}$$

This property has two explanations

- Most estimated equations are formulated in logarithms.
- By dividing the variation by the value, the result should be stable with time.

Indeed, this modification is equivalent to replacing each variable by its logarithm in model equations.

This does not change fundamentally the above considerations, except that convergence and non-convergence will be associated now to relative variations: an eigenvalue larger than one will mean that it is the relative difference to the variable to the base value which will tend to infinite.

As most variables grow with time, this criterion for convergence will be less restrictive than the first. But the following question must be considered: must one associate to divergence a growing absolute variation or a relative one? For instance, let us suppose households benefit from a tax exemption, at two different periods in time. To compare the two, do they consider levels at constant prices, or the ratios to their present income? The truth is certainly between the two.

8.9.1.1 First example: a very simplified model

As an example, let use an extreme simplification of our usual model.

$$(1)\ I_t = a \cdot (Q_t - Q_{t-1}) + b$$

Firms invest to adapt their productive capacities to the evolution of demand.

$$(2)\ CO_t = c \cdot Q_t + d \cdot Q_{t-1} + e$$

Households consume a share of the present and previous production levels (introducing a lag makes the example more interesting).

$$(3)\ Q_t = I_t + CO_t + g_t$$

Production adapts immediately to demand, composed of investment, consumption, and State demand.

Reasoning in variations, one gets:

$$\begin{pmatrix}\Delta \mathrm{I}\\ \Delta \mathrm{CO}\\ \Delta \mathrm{Q}\end{pmatrix}_{\mathrm{t}}=\begin{pmatrix}0&0&\mathrm{a}\\0&0&0\\1&1&0\end{pmatrix}\begin{pmatrix}\Delta \mathrm{I}\\ \Delta \mathrm{CO}\\ \Delta \mathrm{Q}\end{pmatrix}_{\mathrm{t}}+\begin{pmatrix}0&0&-\mathrm{a}\\0&0&\mathrm{c}\\0&0&0\end{pmatrix}\begin{pmatrix}\Delta \mathrm{I}\\ \Delta \mathrm{CO}\\ \Delta \mathrm{Q}\end{pmatrix}_{\mathrm{t-1}}+\begin{pmatrix}0\\0\\1\end{pmatrix}\Delta \mathrm{g}_{\mathrm{t}}$$

or

$$\begin{pmatrix}\Delta I\\ \Delta CO\\ \Delta Q\end{pmatrix}_{t}=1/(1-a)\begin{pmatrix}0&0&(c-1)\\0&0&c(c-a)\\0&0&(c-a)\end{pmatrix}\begin{pmatrix}\Delta I\\ \Delta CO\\ \Delta Q\end{pmatrix}_{t-1}+1/(1-a)\begin{pmatrix}a\\0\\1\end{pmatrix}\Delta g_t$$

This process contains a single lagged, variable, and a single eigenvalue:

$$\boldsymbol{\lambda=(d-a)/(1-a-c)}$$

and the multiplier at period k of a unitary change in State demand at period 0 will be:

$$\boldsymbol{1/(1-a-c)\cdot\lambda^k}$$

At this level of simplification a formal interpretation is still possible. The eigenvalue adds two lagged dynamic effects:

- The sensitivity of consumption to past households income: d.
- The sensitivity of investments to the previous level of production, representative of capacity levels acquired earlier: - a.

These two effects are amplified by the multiplier mechanism, due to the present influence of investment and consumption on present production: 1 / (1 – a -c).

One can therefore logically suppose:

- That a, c and d are positive.
- That 0 < c + d < 1: households consume a part of their additional revenue, itself lower than additional production.

Our purpose is not to theorize on these formulations. However, we can identify some special cases:

- d + c = 1 (unitary impact of Q on consumption); then the eigenvalue is 1, and while the dynamic process does not converge, nor does it diverge: the increase in consumption maintains the level of production in the next period, which stabilizes investment, which has no further impact on growth Q. A demand shock at period 1 will maintain its effects indefinitely.

In other words, the multiplying effect: 1/(1-a-c) is compensated in the next period by the reducing effect: d-a or (1-a-c) – (1- d-c) lower than (1-a-c) if d+c<1.

Or: with c+d<1, the multiplying effect will spread to the two periods with an intensity lower than the initial shock. The process will converge.

- c = 0 (consumption independent of Q); then the eigenvalue is - a /(1 - a), and the dynamics converge if a < 1 / 2. Otherwise, getting back to the initial level of investment represents more than half the previous increase of production, and the multiplier more than doubles this change, which leads to an alternate diverging process.

- a = 0; then the eigenvalue is c, leading to a general convergence if $|c| \in]0,1[$

More generally, when 0 < a < 1 and 0 < c +d < 1 the process will converge if:

$$c + d > 2a - 1$$

Which is a less restrictive condition than a > 1 / 2 (if c=0), because the positive effect of the consumption can then compensate the alternate divergence presented higher.

We can see that even for a very simple example the formal interpretation can present some complexity. For more realistic cases, one will have to resort to numerical techniques, interpreting results in terms of economic mechanisms by associating them to variables of the model.

For this one can (following a methodology introduced by Deleau and Malgrange (1978)) eliminate for each variable in turn its lagged influence (by deleting the associated column and line of A, then by re-computing the eigenvalues). In the most favorable case, a single eigenvalue will disappear, creating a natural association to the eliminated variable. In other cases research will be more difficult, and might need the simultaneous elimination of groups of variables, which one will associate to a group of eigenvalues of the same dimension.

Our example is too simple to present this method, as it contains only a single lagged influence (and therefore a single eigenvalue). But the analysis we have just made gives a basic illustration of the method. To extend the dynamic characteristics of the above model, one could have added an equation for capital and productive capacity.

8.9.1.2 Second example: the role of the Taylor rule

We shall now explain, on a very crude example, the role of the real element of the Taylor rule in our type of model.

Let us remind that the Taylor rule represents the behavior of a central bank which uses the interest rate to target inflation (and also reduce its volatility). See TAYLOR JB. (1993).

It will increase the rate if:

- Inflation is higher than the target (2% for the European Central Bank).
- Tensions appear on productive capacities (the “output gap”) signaling potential inflation.

This action should reduce demand (investment and consumption), activity and therefore inflation.

The original formula is:

$$\boldsymbol{IR_t = tx(p_t) + 0.5 \cdot (tx(p_t) - tx^*(p_t)) + 0.5 \cdot gap}$$

With

- IR the interest rate
- P the price index (deflator of consumption?)
- tx*(p) the inflation target
- gap the output gap indicator, normally the relative difference between the actual and potential values of output.

Let us consider the following model, measured in real terms, a simplification of our example without external trade

(1) $i_t/k_t = 0.65 i_{t-1}/k_{t-1} + 0.10[(q_t - q_{t-1})/q_{t-1} + (ur_t - ur*)/ur*] + d + a(ut_t - ur*)$

(2) $c_t = r_t q_t (1 - b(ur_t - ur*))$

(3) $ur_t = q_t/(k_t pk_t)$

(4) $k_t = k_{t-1}(1 - dd_t) + i_{t-1}$

(5) $q_t = c_t + i_t + g_t$

with the endogenous variables:

c household consumption
i productive investment
k capital
q GDP
ur the rate of use of productive capacity

the exogenous variables:

dd the depreciation rate
g government consumption
pk the productivity of capital
r the ratio of consumption to GDP
ur* the target rate of use

For the Taylor rule, we have only considered the role of the output gap, which we have represented (this can be questioned) by the rate of use of capacities.

So we shall use the parameters:

a the sensitivity of investment to the rate of use
b the sensitivity of consumption to the rate of use

If we compute the dynamics of this model by differentiating its equations numerically,

$$\Delta y_t / y_t = A \Delta y_t / y_t + B \Delta y_{t-1} / y_{t-1} + C \Delta x_t / x_t$$

where A, B and C are relative Jacobians supposed constant with time,

we get

$$\Delta y_t / y_t = (I - A)^{-1} (B \Delta y_{t-1} / y_{t-1} + C \, \Delta x_t / x_t)$$

To describe the dynamic process, we must now compute the eigenvalues of $(I - A)^{-1} B$, a 6x6 matrix.

But of course, independently from the values of a and b, we should get only two non-zero eigenvalues, as there are only two equations with lagged elements (and three such elements).

Starting with a=b=0, we get two complex eigenvalues, with

Modulus: .977 period: 18.8

Now if we put the value of a to 0.03, consistent with our Taylor rule (coefficients of -.06 in the investment equation, and .5 in the Taylor rule itself) we get:

Modulus: .872 period: 20.6

The period not much affected, but the convergence is much faster. After 10 periods an initial shock is reduced by 75% (instead of 21%).

Considering b (the role of the interest rate in consumption) we get only a small increase in the period, and in the speed of convergence.

In conclusion: the dynamics are almost fully explained by the real elements of capital formation, including the real interest rate.

The elements allowing to product the above study can be very simply implemented using EViews.

8.9.1.3 Steady state growth paths

Let us go further in determining the conditions for the existence of a path of model solutions, such that the whole set of variables grows at a constant rate (a « steady state growth path »). The application of these conditions to formulations of the model will lead, by editing some equations, to the production of a long-term model, with potentially very different causalities compared to the original model. One will study also the stability of this path, by introducing an initial gap through a one-time shock, and observing if the long-term dynamics leads back to the path or away from it (again, reasoning in absolute or relative terms).

These techniques can be clarified through our simple model. Let us recall its formulations (in EViews format):

```
[1] cap = pk * k(-1)

[2] log(prle_t) = 9.8398+ 0.009382* (t - 2002) + 0.02278* (t - t1) * (t<t1) + 0.01540* (t - t2) * (t<t2)

[3] DLOG(x) = 0.4940*DLOG(wd) - 0.09142*res_x( - 1) + 0.01960+ ec_x

[4] ur = q / cap

[5] i/k( - 1) = 0.944*I( - 1)/K( - 2) - 0.00256*(ur-ur_star)/ur + 0.0428*.125*q/q( - 8) - 0.00448+ ec_i

[6] led = q / prle_t

[7] DLOG(le) = 0.2954*DLOG(led) + 0.1990*LOG(led( - 1)/le( - 1)) + 0.0004748* + 0.01759**((T = 1968.25) - (T = 1968.50)) + ec_le

[8] ic = ct * q

[9] lt = le + lg

[10] rhi = wr * lt + r_rhiq * q

[11] co = rhi * (1 - sr)

[12] ih = r_ih * rhi

[13] ci/q( - 1) = -0.02680*(T = 1968.25) + 0.6128*ci( - 1)/q( - 2) - 0.0021490 + 0.2193*@PCH(q) + 0.1056*@PCH(q( - 1)) + 0.05492*@PCH(q( - 2)) + 0.03918*@PCH(q( - 3)) + 0.03026*@PCH(q( - 4)) + ec_ci

[14] fd = co + i + gd + ci + ih

[15] td = fd + ic

[16] DLOG(m) = 1.676*DLOG(td) + 0.2133*DLOG(ur) - 0.1057*DLOG(compm) - 0.1028*res_m( - 1) - 0.2757+ EC_M
```

[17] q + m = fd + x

[18] res_x = p_x(1) * log(x / wd) + 0.9815 * log(ur) + 0.001441 * (@trend(60:1) * (t<=2004) + @elem(@trend(60:1) , "2004q4") * (t>2004))

[19] res_m = log(m / (fd + ct * q)) ‘ 1.517 * log(ur) +-0.552 * log(compm) + 0.00579 * (@trend(60:1) * (t<=2004) + @elem(@trend(60:1) , "2004q4") * (t>2004))

[20] k = k(-1) * (1 - dr) +i

If we call

- txq the growth rate of quantities
- txn the growth rate of populations.

We get:

[1] cap = pk * k / (1+txq)

[2] log(prle_t) = 9.8398+ 0.009382* (t – 2002)

[3] log(1+txq) = 0.4940*log(1+txq) - 0.09142*res_x+ 0.01960+ ec_x

[4] ur = q / cap

[5] i/k*(1+txq) = 0.944* i/k*(1+txq) - 0.00256*(ur-ur_star)/ur + 0.0428*log(1+txq)- 0.00448+ ec_i

[6] led = q / prle_t

[7] log(1+txn) = 0.2954*log(1+txn) + 0.1990*log(led /le) + ec_le

[8] ic = ct * q

[9] lt = le + lg

[10] rhi = wr * lt + r_rhiq * q

[11] co = rhi * (1 - sr)

[12] ih = r_ih * rhi

[13] ci/q*(1+txq) = 0.6128* ci/q*(1+txq) - 0.0021490 + 0.2193*txq + 0.1056*txq + 0.05492*txq + 0.03918*txq +0.03026*txq + ec_ci

[14] fd = co + i + gd + ci + ih

[15] td = fd + ic

[16] log(1+txq) = 1.676*log(1+txq) - 0.1028*res_m- 0.2757+ EC_M

[17] q + m = fd + x

[18] res_x = p_x(1) * log(x / wd) + 0.9815 * log(ur)

[19] res_m = log(m / (fd + ct * q)) ‘ 1.517 * log(ur) +-0.552 * log(compm)

[20] k = k / (1+txq) * (1 - dr) +i

First we can observe that in each equation, provided that all the explanatory elements grow at the theoretical rate, the explained element will also. In particular, all additive terms should grow, taking into account the variables which enter them, at this rate.

This means that there exists a consistent solution in which all elements grow at the theoretical rate. As the model has only one solution, this is the solution it should reach, provided the error correction mechanisms lead in the right direction (from the coefficients, we can suppose they do).

We can also build the long-term model, which follow a logic somewhat different from the dynamic one, as disequilibria have been stabilized.

- From (20) we get I/K as the rate which allows to adapt capital to both growth and depreciation:

i/k = 1 - 1/ (1+txq) * (1 - dr)

- From (5) we get UR depending on the speed of adaptation. Even if adaptation was immediate (dynamic homogeneity) UR would still depend on txq, as a higher expected growth calls for an effort in creating additional capital, for use at the next period:

i/k*(1+txq) = 0.944* i/k*(1+txq) - 0.00256*(ur-ur_star)/ur + 0.0428*log(1+txq)- 0.00448+ ec_i

- From (18) we get RES_X.
- From (16) we get RES_M.
- From (9) we get LED/LE.

If adaptation was immediate the three would not depend on txq or txn. This would also be the case for UR if the coefficients complied exactly with the error correction framework.

Here the gap will increase with the growth rate, and decrease with the speed of adaptation.

- From RES_X, UR and (18) we get X/wd and X:
- From RES_M, UR and (19) we get M/TD
- From UR, (1) and (4) we get Q/K;
- from (5) we get I/Q
- From (13) we get CI/Q
- From (6) and (7) we get LE/Q
- From (9) and (10) we get RHI as a term proportional to Q, plus wr*lg
- From (11) and (12) we get CO and IH as proportional to RHI, thus to Q and wr*lg

- From (17) we get FD as proportional to Q plus a term proportional to wr*lg, plus gd
- This gives M/Q
- This gives Q from (2)
- And all the other elements.

Finally, all the terms in the supply – demand equilibrium are defined as proportional to Q, except for:

- Household consumption and investment which are partially proportional to government employment.
- Government demand which is exogenous[115].

The sum of these elements makes demand partially exogenous.

- Exports which are proportional to world demand.

We can see that the logic of the model causalities is quite different from the dynamic model. In particular,

- The rate of use depends on the growth rate of the economy and the depreciation rate.
- The gap between targets and actual values too (with UR as an eventual intermediary).
- Knowing this gap we can define all the real elements as a share of GDP, except for exports and an exogenous element of demand. Let us call it dx.

The supply – demand equilibrium

$Q + M = FD + X$

becomes:

$Q + a . (b . Q + dx) = b . Q + dx + c . wd$

$Q = ((1 - a) . dx + wd) /(1 – (1 – a) . b)$

With a and b constant elements depending on exogenous assumptions and the growth rate txq (and also txn).

This means that if

- gd was defined as a constant share of demand
- lg was defined as a constant share of LE (or of LT = LE + lg)

(which seem logical and even necessary assumptions),

[115] Of course, the assumptions will give them values proportional to Q.

the exogenous world demand level would define all elements in the local economy (as world inflation will drive the local one with a more complex model).

8.9.2 THE CASE OF ERROR CORRECTION MODELS: A SIMPLE EXAMPLE

Let us now show how an error correction formulation makes the interpretation of model dynamics both easier and more meaningful.

Let us introduce the following model

(1) $CO_t = a \cdot Q_t$

(2) $Q_t = CO_t + I_t + g_t$

(3) $K_t = K_{t-1} \cdot (1 - dr_t) + I_t$

(4) $\Delta K_t = \alpha \cdot b \cdot \Delta Q_t + \beta \cdot (b \cdot Q_{t-1} - K_{t-1})$

We have defined an error correction model, where the target consists in making K proportional to Q (this equates to a constant rate of use of capacities if one supposes constant the productivity of capital).

Let us consider the dynamics of this model, along the lines of the eigenvalue analysis presented earlier.

By eliminating CO and I, we can reduce the system into:

(1) $Q_t = a \cdot Q_t + K_t - K_{t-1} \cdot (1 - dr_t) + g_t$

(2) $K_t = \alpha \cdot b \cdot Q_t + (\beta - \alpha) \cdot b \cdot Q_{t-1} + (1 - \beta) \cdot K_{t-1}$

$$\begin{pmatrix} \Delta Q_t \\ \Delta K_t \end{pmatrix} = \begin{pmatrix} a & 1 \\ \alpha b & 0 \end{pmatrix} \begin{pmatrix} \Delta Q_t \\ \Delta K_t \end{pmatrix} + \begin{pmatrix} 0 & -(1\text{-d}) \\ b(\beta - \alpha) & 1\text{-}\beta \end{pmatrix} \begin{pmatrix} \Delta Q_{t\text{-}1} \\ \Delta K_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} g$$

We can apply to this system the same technique as before, with normally two eigenvalues. The process might be a little too complex for this simple formulation. We shall simplify it by supposing that the speed of adaptation is the same for a new gap, due to a change in the target, and a pre-existing gap:

$\alpha=\beta$

The system becomes

$$\begin{pmatrix} \Delta Q_t \\ \Delta \mathrm{KP}_t \end{pmatrix} = \begin{pmatrix} 1-a & -1 \\ 1-\alpha b & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -(1\text{-d}) \\ 0 & 1\text{-}\alpha \end{pmatrix} \begin{pmatrix} \Delta Q_{\text{t-1}} \\ \Delta K_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} g_t$$

with one eigenvalue:

$$\lambda = 1 - \alpha + \alpha\,(d - \alpha) / (1 - a - b\,\alpha)$$

Let us interpret this element.

If d = α, the eigenvalue is (1 - α). This means an initial gap between the variable and its target will close with a reason 1 - α. The reason for this is clear. If we consider the equation system, the capital will depreciate at the speed d, which is exactly the rate at which the firms « want » it to decrease if a gap has appeared in previous periods, due for instance to a shock on g. The ex-ante variation will be the same as the desired one, and no additional mechanism will appear.

On the contrary, if d is higher than α for instance, the natural decrease in capital will be higher than the desired one. Firms will have to invest the difference, leading to an increase in capital made larger by the presence of the multiplier.

But for the process to diverge, the difference between d and α must be large, or the multiplier effect very important (which would lead to problems with any model).

It might look strange that an ex-ante decrease in capital (through a higher d) could lead to an ex-post increase; This is because the behavior of firms as to capital is set by (2), taking into account the variation of production but not the additional depreciation rate, which is compensated automatically through higher investment (thus production, and investment again).

Let us now consider the long-term model.

It is now clear that all variables: CO, Q, I, g, CAP grow at the same rate on the steady state growth path. This rate (let us call it q) is set by « g ».

The long-term model is

(1) $CO = a \cdot Q$

(2) $Q = CO + I + g$

(3) $K = K \cdot (1 - dr)/(1 + q) + I$

$$(4)\ K \cdot (q+\beta)/(1+q) = (\alpha \cdot q + \beta) \cdot b \cdot Q/(1+q)$$

or

$$(1)\ CO = a \cdot Q$$

$$(2)\ Q = CO + I + g$$

$$(3)\ I/K = q + dr$$

$$(4)\ K/Q = b \cdot (\alpha \cdot q + \beta)/(q + \beta)$$

We observe that none of the causal explanations of the original model has changed. The ratio of production to capital is fixed, actually based on the target in the original formulation. And investment maintains this constraint, by leading to an implicit evolution of productive capacity equal to the growth rate (taking into account depreciation).

But one also can see that the target is reached only in special conditions. This is a standard characteristic for an error correction model.

Equation (4) shows that the cases are:

$\alpha = 1$ (dynamic homogeneity, gaps are closed immediately)

$q = 0$ (the target does not move)

This is true for any starting values, including those which meet K /Q =b: if q or α are not zero, a gap will appear, which will never close.

We shall stop here, as our purpose was only to illustrate the process, and show that error correction models make the dynamics more interesting, as well as easier to study.

8.10 MISCELLANEOUS MODELLING FUNCTIONS

Versions 12 and 13 allow new features making the modeling process easier.

8.10.1 MODEL DEPENDENCIES

For a given model, it determines the elements it Influences or depends on.

The syntax is

Model_name.@depends(string)

Model_name.@upends(string).

For example, if the model "mod" contains the trade equations:

BAL=XV-MV

RX=XV/MV

Mod.@depends("XV") will give the string "BAL RX"

Mod.@upends("RX") will give the string "XV MV"

This feature can be useful to check that the requested variables are available, or

8.10.2 DEBUGGING TOOLS FOR EVIEWS PROGRAMS ("PROGRAM DEBUGGING").

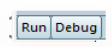
Until version 13, the main debugging possibilities were:

- Introducing "STOP" statements in the program.
- Using the mouse to select part of a displayed program and running it using "Run Selected".

The new "Debug" option makes the process easier.

First one should set the breakpoints in the program.

click on the "Debug" item (right of "Run").



Let us use the "pfor_b.prg" program for running the PIC_B model on the future, and concentrate on the lines 57 To 77.

```
57 for %1 prle_t
58 genr {%1}={%1}(-1)*(1+txq)/(1+txn)
59 next
60
61 c_prle(2)=4*log((1+txq)/(1+txn))
62
63 wfsave pic_p
64
65 smpl 2005Q1 2100Q4
66 pic_b.scenario "Scenario 1"
67 pic_b.append assign @all _p
68 pic_b.override
69 pic_b.solveopt(n=t m=1000,o=g,d=d)
70 solve  pic_b
71
72 for !i=1 to g_vendo.@count
73 %2=g_vendo.@seriesname(!i)
74 genr tc_{%2}=100*log(({%2}_p+({%2}_p=0))/({%2}_p(-1)+({%2}_p(-1)=0)))
75 genr tc0_{%2}=100*log(({%2}+({%2}=0))/({%2}(-1)+({%2}(-1)=0)))
76 genr dtc_{%2}=tc_{%2}-tc0_{%2}
77 next
78
79 genr gd_v=gd+.01*q_p*(t>=2007)
```

Clicking on “79” we introduce a breakpoint (signaled by the red dot).

```
79 genr gd_v=gd+.01*q_p*(t>=2007)
```

When we run the program under the “Debug” option it will stop at this point and give access to three items: Watch, Breakpoints and Callstack.

```
79 genr gd_v=gd+.01*q_p*(t>=2007)
```

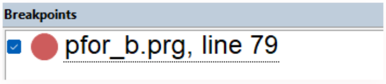
Watch | Breakpoints | Callstack

“Watch” will give the value of the current parameters.

Watch

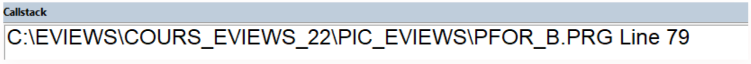
Name	Value
%1	prle_t
%2	X
!i	19

“Breakpoint” will give the list of current breakpoints (only one at present).



"Callstack" will give the list of currenly opened programs.

Of course, you can also acceess all the cuerrent elements and display their value.



8.10.3 PROGRAM DEPENDENCY TRACKING ("PROGRAM DEPENDENCY LOGGING").

For instance, the program pdon_a.prg uses the file fra_xls and creates the file pic_q.wf1.

```
FILE: c:\eviews\cours_eviews_22\pic_eviews\fra_q.xls
(c:\eviews\cours_eviews_22\pic_eviews\pdon_a.prg: 34)
FILE: c:\eviews\cours_eviews_22\pic_eviews\pic_q.wf1
(c:\eviews\cours_eviews_22\pic_eviews\pdon_a.prg: 187)
```

CHAPTER 9: USING MODELS

In chapter 2 we began to present modelling applications. We are going now to be more precise, concentrating on practical aspects.

9.1 OPERATIONAL DIAGNOSES

The first application we quoted, and the most natural apparently, is to request from the model a realistic diagnosis of economic problems. This diagnosis can lean on scenarios or shocks.

The originator of this type of study can be:

- A governmental organization, such as the local Ministry of Finance, which wants to get information on the future evolution of the economy it manages, or the efficiency of the policy decisions it contemplates.
- An international institution, such as the IMF or the OECD, or a Non-Governmental institute, which wants to assess the future of the world economy.
- An academic research unit or individual, which wants to study the impact of policy decisions or structural changes, according to a given theoretical model.
- A private firm dealing in economic services, which provides forecasts and studies to subscribers, either as a publication or a dedicated study.

9.1.1 SCENARIOS AND THEIR DIFFERENT TYPES

Two types of scenario may be distinguished.

- **Trend forecasts**, which will use the most probable assumptions (in the case of a unique scenario), or a set of possible assumptions, covering the scope of foreseeable evolutions.

Assumptions will often be developed by experts from outside the area of model building. Let us take as an example: the forecasts of MESANGE, a French macroeconomic model considering policy decisions as exogenous, managed by the French Ministry of Finance, in two of its Directorates: at INSEE, the National Institute for Statistics and Economic Studies at the General Direction of Treasury and Economic Policy (DGTPE).

Variables describing foreign environment are valued through simulations of a set of tools:

- The MZE (Modèle Zone Euro or EuroZone model) model of the CEPII (a French economic institution specialized in international studies), managed in the same institutions.
- The NIGEM world model, managed and leased by the British NIESR (National Institute for Economics and Social Research).
- Results from economic forecasts of OECD, the UN and the IMF.
- Numerical evaluations made by experts in international economics and trade, without using any kind of model.
- The observation of international developments, like the Asian financial crisis, and the assessment of potential evolutions.

The government decisions are specified following discussions with the DGTPE and especially the people which produce the numerical evaluation of next year's budget.

As for assumptions on the evolution of the population and its structure, they will use the results of studies produced by INSEE or Ministry of Labor demographists.

This type of scenario will give indications of the most probable future evolution, or on the range of foreseeable evolutions, to the best knowledge of the available experts.

Production of the final forecast will require a large number of model simulations, accompanied by meetings and discussions. It lasts usually several weeks (let us say 2 to 4). The improved speed of computation helps, but not too much, as it is mostly used to improve the quality of the results.

- **Normative forecasts**, in which results are asked to meet a certain number of conditions. These constraints are set before the development of assumptions.

This type of scenario describes a context in which efforts are made, relative to the foreseeable evolution of the economy, to reach some targets (such as meeting the criteria for joining the European Monetary Union, usually called the Maastricht criteria). It will essentially give information on the extent of expected difficulties, and efforts required, associated to a possible path.

In general the client of this type of study will also be the State, or a public institution. In cooperation with the model managers, it will set the constraints, and the policy elements used to meet them[116]. But non-governmental institutions and researchers can also be interested in this type of issue, perhaps more than the production of a forecast as it involves more economic reasoning.

If precise goals are set (such as a budget deficit of 3 GDP points), one can work either by a sequence of approximations on the normal model, or by modifying model specifications by an explicit reversal of causalities. This last technique introduces two problems:

[116] However in some cases forecasters will also include in the potential set of instruments some assumptions which are not manageable by the State, such as an exogenous change on the behavior of agents, or structural parameters such as the productivity of factors. This is not unacceptable, provided the option is clearly displayed in the associated publication. The initial question was then: how should the behavior of agents, or structural parameters, have to change in order to reach a specific objective?

- A formal one: for the model to preserve a unique solution, the number of endogenous elements which are now given a set value must be equal to the number of instruments that will allow to reach them. For example, if one wants to balance both the commercial balance and the State budget, it will be necessary to free two decision variables and two only, to avoid either impossibility or an indetermination[117].
- A technical one: often it will be too complex to produce an identified model which takes into account this reversal explicitly, and if the solving algorithm requests such a formulation, it will be necessary to revert to the other technique, determining the solution by a sequence of approximations. Let us suppose for instance that we want to balance foreign trade by increasing subsidies on investment. To transform the model, one will have to identify a chain of causalities leading from subsidies to foreign trade (which can be quite long) and revert each one in turn, through the associated equation. To this obviously difficult if not unfeasible technique, one can prefer a numerical, iterative one: computing outside the model an approximation of the necessary subsidy amount, using it to solve the standard model, then correcting the change from the observation of the remaining error. This sequence of approximations should decrease to an acceptable value.

We will propose soon an alternate method, which allows solving the model for given values of a number of endogenous variables, provided the same number of exogenous variables are released. This method does not work all the time, but if it does the solution is quasi-immediate, and can be applied under EViews without accessing the model code. The program itself is provided, and can be adapted in a few minutes to any existing model managed under EViews.

But in any case, results cannot be taken for more than what they are, the economic equilibrium associated to assumptions deviating from normal evolution, whose probability of occurrence cannot be judged. To earn the right to use this method, one has to display clearly and probably justify the options he has chosen.

Under those restrictions, one is not limited in the elements he can use. For instance one can modify the estimation residual in the investment equation to measure by how much private firms should increase it (all things being equal) for the trade balance to get back to equilibrium in the long run. The less controlled the instruments, the less the simulation can be considered as a policy result. But this does not make the results (and the observation of causal sequences) less interesting from a theoretical point of view.

9.1.2 MANAGING ACTUAL FORECASTS: MANAGING THE RESIDUALS

On the estimation period, the sensible option for residuals was of course to set them to the difference between the actual values and the one given by the estimated formula. Now we have to decide on values, starting with the first forecast period.

Two main (and extreme) options are available.

- Setting them to zero (in principle, the value with the highest probability).
- Keeping the last value in the estimation period (the most recent information we have).

In other words: should we rely on the recent past, or summarize all the past information we have, giving equal weight to each element, independently from its age?

Each choice can be dangerous:

---

[117] Except if the number of instruments is larger than targets, but limits are set on their variations, as well as a priority order. Then each instrument will be used in turn, switching to a new one if the constraint has been reached. In practice we will be left with as many unknowns as constraints. One can also create a synthetic instrument by applying predefined weights to a larger set.

- The first will refuse to take into account recent elements, like a feature which appeared only in the last periods, and could not be evidenced as too recent (otherwise it should appear in the actual formula). Also, it will produce shocks on the initial growth rates, which will be difficult to explain. For instance if the last residual on household consumption was 2%, a structural decrease of 1% (taking only into account the formula and its variables, and excluding the residual) will be transformed into a growth of 1%.
- The second will violate the zero mean assumption, and suppose that a possibly high shock will be maintained for the whole period of the forecast.

In our opinion, no option clearly dominates under all circumstances. One should first consider the profile of the last residuals.

First, their size: obviously if the last residual is close to zero, the problem is minor.

This means of course that one will favor estimations with a low last residual. From a statistical point of view, the standard error should not depend on the period, but the modeler will concentrate on the last values, the closest to the computations he will have to manage.

If the residual is large (whatever its sign), one should look at the immediately preceding values. If they are too, there is clearly a problem with the equation, and something should be done to reduce it. This is also true if they are smaller but growing: a trend (or a variable with a trend) might be missing.

Maybe the best option is to decide that the last residual is indeed going to last for a while, but with an intensity decreasing to zero.

But I any case one should first try to understand the reason for the value. This might make things clear as to the way to manage it.

9.1.3 MANAGING ACTUAL FORECASTS: THE TARGETING OF SIMULATIONS

In the normative case, one has to manipulate the exogenous in order to reach a certain solution. But this happens also in trend forecasts. One cannot expect the model to give spontaneously an image of the future consistent with what we expect, especially in the short run where we already have some (partial) information about what is happening.

A targeting procedure has to be applied, which may be automated, by the application of specific techniques.

9.1.3.1 on the past

We might want the base simulation to reproduce exactly the known past. This operation is not futile: it is advisable to start the analytic shocks from the historical value, eliminating second-order errors.

If all values are known, the technique is simple: we apply to behavioral equations the estimation residual. Each individual equation will provide historical values, and the global absence of error makes historical values the solution of the system.

9.1.3.2 in forecasts

In forecasts this method will seek to give some variables values belonging to a certain interval, through the choice of assumptions and of estimation residuals.

In the short-term the goal will be to adapt results to already known values (or known intervals) of some elements of the model. At the end of the year, we know already global growth and some of its components, global inflation, unemployment and the trade balance, and the forecast should provide their values, if only for show. Reproducing wrongly the already known past casts strong doubts on the forecasting quality of the model.

In the medium-term one will seek to keep the general evolution within reasonable limits, in particular meeting some general conditions. Thus one will not accept too large a gap between the country described by the model and the rest of the world, as to inflation and growth.

But the fragmentary character of the information will prohibit the use of the method given in the previous paragraph. The model builder will work in general by iterations, sometimes relaxing some constraints if they prove difficult to meet. Or he can try the specific algorithm we are proposing.

9.1.4 DIFFERENT TYPES OF SHOCKS

We are going now to categorize the different types of shocks.

9.1.4.1 Analytic and complex shocks

We have already presented analytical shocks, used to interpret and validate model properties, by comparison with economic theory and properties of the other models. This set of tests also allows the interpretation of model responses, making the later interpretation of more complex shocks easier. Actually, operational modelling teams always produce a “shocks catalogue "[118] giving the responses of the model to changes in the main assumptions. Given the numerical quasi-linearity of models, one will get a realistic enough approximation of a complex set of changes by combining linearly the results of these individual shocks, using a spreadsheet for instance or more simply an EViews program (combining in one page the contents of several “single shock” pages with the same structure).

Going even further, a few months spent working with a model will give its user the ability to anticipate (at least roughly) its reactions to given changes, and this without any simulation. He will already have a basic “shocks catalogue” available in his mind.

But if the intensity of the shock changes with time (for instance to take into account a gradual loosening of quotas in a trade agreement) we face a larger problem. As we have just stated, most macroeconomic models are roughly linear at a given period, but they their Jacobian is not stable over time, mostly because of the growing role of external trade. Of course, the slower the change, the larger the problem. But in the above case, the expected change in the structure of exports and clients should call for a specific forecast anyway.

One will also use this type of shock for operational studies

- If requested diagnoses are associated to simple decisions.
- If one seeks to decompose the effects of more complex policies.

But to get the exact consequences of complex shocks, one will have to apply them simultaneously to the base simulation, in a true operational context.

9.1.4.2 External and policy shocks

All these shocks fall also into three categories:

- External environment shocks, which seek to measure the consequences of non-controllable events. For a single country model, the question asked is then: *how will a change in the situation in the world affect the economy of my country?*

Two subcategories can be considered.

- Shocks on foreign assumptions, associated for single country models with the situation in other countries, or elements defined at the world level. A typical shock would represent a depression in Asia, or an increase in the price of oil, followed possibly by a decrease in world demand and world inflation. Technically, these elements should not be , as the modelled country or set of countries have some impact on the economy of the rest of the world. But the cost of modelling this feedback is generally much too high (it could involve the production and management of a world model) compared to the limited improvement of model properties. Of course, the approximation (and the associated error) increases with the size of the modelled country. If it is reasonable to consider the world outside Denmark as exogenous to this country's events (we have taken voluntarily a country outside the Euro Zone), this is not so true for Germany (a large country which shares a currency with a significant part of the world) and even less for the whole Euro Zone itself.
- Shocks on local uncontrolled variables. Examples include technical progress, population, or climate. Again, one can argue that each of these elements is not completely exogenous: strong economic growth accelerates technical progress, the birth rate changes with wealth and health of households, growth brings pollution and climate changes. But the associated mechanisms are difficult to establish, formally complex and limited in their short and medium-term impact. Therefore the cost of taking them into consideration is generally much higher than the benefits, except for very specific fields.

The only exception concerns long-term studies. But although it is true that we have solved our model over a very long period, the goal was only to observe its properties from a technical point of view, and not to produce an actual forecast. A true long-term model would require a rather different framework, with new formulations relying more on a formal economic reasoning than on actual estimations, as the sample is generally too small to establish statistically long-term relationships.

- Economic policy shocks, where interest lies in the consequences of choices on institutional instruments, decision variables if one considers that the State (or its impersonator) is the client of the study. The question is now: what will happen if I take the following decisions? And when these shocks are given an objective (for example: to reduce unemployment by a given margin without eroding too much the other criteria), they will call for an iterative targeting process analogous to that of normative scenarios.

These elements are not exogenous to the country modeled, as the decisions taken by the Government are largely based on the local economic situation. But the producer of the study wants to be allowed to manage the decisions, and not to see the model do it in his place. Moreover, the State has few goals (unemployment, inflation, budget and foreign deficits, growth) and a lot of potential instruments. Evidencing statistically the link between decisions on the instruments and the present state of the goals is very difficult if not impossible. The few studies on that subject do not give in general reliable answers.

- The two approaches may also be combined however, if we look for the decisions which allow to face (or take advantage of) a given evolution of external conditions. One could for instance determine how the State can use his own tools to fight losses in trade balance coming from a given depression in world demand. Or an oil producing country can decide on the use of the additional revenue coming from an increase in the price of a barrel. This type of exercise is parent again to normative scenarios.

It will be necessary, in all these cases, to care for the presence of formal connections between exogenous variables. For example, one cannot suppose an increase of the price of oil [119]without taking into account its inflationary effects

[119] Otherwise, the modeled country will be unduly subject to losses in competitiveness coming from the increase of its own inflation.

on the production price in non-oil-producing countries, or the depression it will cause in world economy. Similarly it seems difficult to increase the number of civil servants without considering an increase in State consumption (heating, lighting, and various equipment) and investment (offices, hospitals, schools) according to the type of positions created.

A very important issue is the impact of public spending (investment, consumption or hiring of civil servants, separately or together) on local supply. Of course, they modify demand (creating jobs in the public sector increases household revenue and reduces unemployment). But they should also have a positive effect on potential production, for many reasons. To give just two examples:

- Creating or improving roads, or the lighting of streets, increases the productivity of the transportation sector, and of the whole economy (agricultural goods can be delivered fresh over longer distances, attrition from accidents will diminish…). Some completely new productions can become feasible (access to water can allow to produce new kinds of vegetables).
- Buying educational books, or hiring more teachers, will improve students' skills and their productivity when they join the workforce.

Unfortunately, the influence of these elements is very difficult to quantify. It is left to the discretion of the modeler, who more often than not keeps it null. A more reasonable option would certainly be to suppose that the effect on global productive capacity would be the same as actual productive investment, but with a longer delay in its impact. This would mean that the main difference in behavior between the public and private sectors would be the lower concern for quick returns, the more widespread effect, and the fact that the Government has the power and resources to make a decision that others agents cant individually, even if it would be optimal for all of them.

- Finally, one might want to modify structural parameters, such as the productivity of factors, or the depreciation rate. Even if it does not come from an actual decision, an increase in factor productivity can come from a new educational policy, or research on the productive process. It can be interesting to observe its consequences on the whole economic equilibrium.

  Falling in the same category will be the residuals on some behavioral equations: the ratio of household consumption to revenue can be modified ex-ante, as well as the sharing of value added between firms and workers. According to the formulation, the error-correcting mechanisms and multipliers can lead to increased or decreased ex-post changes[120].

Of course, these shocks fall in the first category concerning their feasibility, as external uncontrolled elements.

9.1.5 THE FORECAST: TECHNICAL ASPECTS

It would appear that once a coherent model has been developed, with its global precision and economic properties validated by the previous tests, the only efforts remaining for producing a correct forecast would focus on the definition of assumptions. It is unfortunately not so, and a good deal of know-how will be necessary for the model builder to produce acceptable forecasts, bearing especially on the following points:

- How to judge the likely character of the forecast.
- How to integrate evolutions of the endogenous elements, known but not yet precisely quantified (concerning especially the first forecasted periods). This goes from partially known values (as the recent monthly evolution of unemployment for a quarterly simulation) to vaguer elements, such as the anticipations provided by non-modelling experts.

[120] The price-wage loop will increase the deflationary effect of a wage hike, but the impact of additional demand on imports will be reduced by capacity adaptation.

- Once anomalies concerning the near future are corrected, how these corrections should be taken into account for next periods. For example, if one decides that estimates of exports have to be decreased in the first period compared to the value given by the equation, should this gap be maintained in the medium-term?
- How to interpret results, beyond the simple observation of figures.

In addition, some specific difficulties can appear when generating assumptions:

- In general, they need the intervention of external expertise, as the macro-economic model builder cannot be fully competent in the whole set of areas concerned. In some cases, it even will be necessary to call for another model to obtain coherent assumptions (for example a world model will provide assumptions on international environment).
- External information, even quantified, is often presented neither in the units nor in the categories used by the model (like trade agreements or the results of wage negotiations).
- The transformation of some qualitative information (climate, strikes, tensions, expectations.....) into numerical values is even more complex. The present coronavirus outbreak is a huge example of this issue. Even if some consequences can be foreseen, their intensity and the full set is not conceivable, even when (if?) the pandemic itself disappears.

9.1.6 CHANGING MODEL SPECIFICATIONS

9.1.6.1 The reasons

Normally, the modeler should make its forecasts on the same model he has estimated. To proceed differently, there should be some acceptable reason.

One can think of four cases:

1. It is known that a behavior has changed, in a given way: for instance, the Central Bank has decided to change its method for inflation targeting, or it has just become independent and can decide on the interest rate. Or a new government has just decided to stabilize the social security deficit, putting the burden on the contributions of firms and workers, in equal shares.
2. A behavior was present from the first, but we had not enough data, or the variations of the explanatory element were too small to support it. For instance, the country faced no supply problem, but now capacities are growing slower than demand and disequilibrium has to be considered[121].
3. A behavior appeared late in the sample period, and there were not enough periods to support it. For instance, a new law introduced part-time work, with a strong impact on the employment process.
4. It is known that the economy of the future will follow different behaviors compared to the past.

In the first case, the behavior is supposed to be known. The associated equation should be introduced, for instance a new Taylor rule.

[121] It was present from the first but had no opportunity to play a role.

The second case is more difficult to solve. It applies for instance to a period for which inflation was flat, and its impact therefore statistically insignificant, while in the future we can expect it to fluctuate much more. Or local inflation was similar to that of partner countries (stabilizing competitiveness) while differences are starting to appear.

In this case there is no other way than taking values either from economic theory, measurement on other countries and the observation of resulting properties.

The last two cases can be generally associated with:

- Developing countries
- Countries in transition, in principle from socialist to market economy (some countries can remain in between like China or Vietnam).

Identifying transition countries

In transition countries, the behavior of all agents has changed as market economy mechanisms have appeared. But although the technical transition could be quite fast (like German reunification) the transition process has been generally much slower.

One can separate two cases:

- The Central and Eastern Europe countries, with the addition of Russia and maybe the former USSR republics.
- Asia: mostly China and Vietnam, possibly Laos and Cambodia.

The difference of course is that the later are still under Communist regimes, and that they can be considered as developing countries, which is also probably the case of the less advanced elements of the first set, like Bulgaria, Romania and Ukraine[122].

Actually in our opinion, if we consider the present situation, the type of regime is not the most important, as for the most part China and Vietnam behave economically like market economies (perhaps more than present Russia). The only issue is the date at which the transition can be considered as advanced enough for the economy to be modelled in this way. Officially, the date was 1978 for China, and 1986 for Vietnam. But the process was quite gradual, and one should consider delaying the application of a structural model by at least 10 years.

For the Eastern bloc, the transition started at the beginning of the nineties, and can be considered to have proceeded much faster, in particular because the economy of these countries had been following western standards (in framework and level) 50 years before . This means modelling could start in 1994-1995, with additional elements:

- The statistical system took some time in changing methods, in particular from the Social Accounting Matrix system to western-type National Accounts (like the UN defined SCN95, now SCN08).
- Now almost all of these countries provide quarterly accounts. This multiplies the number of observations by four, which is crucial in this case (although the problem is decreasing with time). Unfortunately, the gain in information is much lower.
- Some specific events took place during the period, like the partition of Czechoslovakia in 1993, or joining the European Union (even the EuroZone).

Concerning the data

Now several situations can appear:

[122] Although the appreciation is quite dependent on criteria (look at China for instance).

- We now have enough data to estimate equations. This is for instance the case of Poland, where data is available from 1995, using the same methodology, and with behaviors which have converged gradually to market economy.

But we have also to consider the last two cases:

- In case 4, the equation has to be entirely calibrated, with values taken either from economic theory, measurement on other countries and the observation of resulting properties.
- In case 3 also, except that some additional information can be obtained from estimations. But it should not be used alone, even if all statistics look good (unless we get extremely good statistics on a sample not too small).
- In any case, if the influence of an element (such as profitability on capital or unemployment on the wage rate) is considered to be growing in the present period, the associated coefficient should also show a growth in forecasts, probably reduced with time to converge to a constant. This process (initial level of the increase, speed of convergence) is the responsibility of the modeler, helped by the observation of model properties.

9.1.6.2 The technique

We can use exactly the same technique as for testing model properties over the future, except that the role of changes in parameter values and formulations will increase.

9.1.7 OPTIMAL CONTROL

This technique consists in computing the values of instruments which gets model variables as close as possible to a given objective function, under constraint of the model. Compared to the reversal of causalities, the model user will set here a unique target, normally unreachable, formalized in most cases by a quadratic function of one or several variables. Then he will apply a maximization algorithm under constraints (equations of the model and bounds on the evolution of instruments). This will have to be done outside the modelling package (while spreadsheets provide a simplified version of this technique).

This technique has been quite popular in the fifties and sixties in western countries like France (managed by the Planning Agency) and remained so in the socialist countries for a longer time[123]. Now they have been more or less abandoned, and the remaining applications limited to economic research.

9.2 TEACHING WITH MODELS

The associated models will allow educational or scientific applications. Diagnoses can be observed for themselves, as the quantified illustration of some theories. But their small size makes them efficient for the application of specific validation techniques, useful especially if they have been designed to reproduce the properties of a larger operational model.

The simplicity of these models gives them the natural role of a learning instrument. According to the public and the goal, different tools may be used.

- Extremely simplified representations, based on formal examples and fixed parameters, clarify theoretical frameworks (Solow model, IS-LM model, Mundell-Fleming Fair model...). By formalizing them on a numerical example and solving them under various assumptions, one can better appreciate their properties (the

[123] Not surprisingly, Russia has produced the best theoreticians in this field (like Pontryagin) and in the connected field of matrix theory (like Gantmacher)

simplicity of these models often allows a graphical display of their equations as dynamic curves). Courses in macroeconomics are now often accompanied by computer programs running these representations.

- More applied models, reproducing as faithfully as possible the properties of larger models used for operational diagnoses. Rather than illustrating an economic theory, the goal here is to put together a limited number of elementary mechanisms, evidencing the main interactions governing the economic equilibrium. This will generally be done through shocks, concentrating on the impact of the different economic policy instruments and the orders of magnitude of their effects.

In this last case, even if the precision of these models remains limited, producing acceptable diagnoses requires to apply them to real cases. Hence they generally will be based on real data, and often estimated econometrically. And to make the interpretation easier, one will generally process current issues: for example measuring consequences of a decrease in work duration, for the French economy of 1998.

9.3 PRESENTATION OF RESULTS

9.3.1 GENERAL ISSUES

The role played in the efficiency of the simulation process by the quality of results presentation is often underestimated. To researchers, it represents often a minor task, to which they are not very well prepared. Nevertheless, this element is often a necessary condition to capture the attention of the audience and insure the success of a study. While a mediocre but well-presented work will sometimes have some success, a work of quality but badly presented will generally fail in getting any audience, by discouraging its potential listeners or readers, including the referees[124].

A good presentation has to be:

- Error free (this is more difficult than it seems). A single visible error will cast doubt on the whole set of results.
- Explicit (the meaning of figures presented has to be evident, in particular their classification).
- Not too long.
- Pleasing to the eye (not too dense), but not too flashy (one must not think the author is trying to hide problems, or has diverted time from his research). PowerPoint slides should not be packed.
- Clearly synthesizing in the right order the main teachings, and only them. In a PowerPoint presentation, one must not be diverted from listening to the presenter.
- Allowing fast access to a particular element of information.
- Adapted both to the nature of the work and to its public (for different audiences, different presentations can be produced for the same work[125]).

Once these principles have been followed, one should not spend to much time sophisticated features, the danger being criticized for trying to hide the low quality of the results behind the gaudiness of the presentation.

The two main elements of any presentation are tables and graphs, and one must keep in mind that these two types of presentation, while complementary, must be conceived simultaneously.

9.3.2 TABLES

[124] But the truth of this argument depends also heavily on the celebrity of the author.

[125] This is made easier by presentation packages.

There are two main kinds of graphs, depending to the public:

- Basic tables, designed for the model builder and his working partners, use simple formats to summarize the whole set of model information. They allow a quick interpretation of the main characteristics of simulations and shocks. For a macro economic model, the table will generally contain the evolution of the main components of the supply-demand equilibrium (at constant prices) as well as the evolution of prices, employment, external and budgetary balances, and some ratios.

The definition of elements presented will be limited, sometimes even to a name representative of the concept. It will be simpler to implement if adequate procedures have been created by the user inside the model-building software.

For more detailed models, one will not have to consult the whole set of results, except to find reasons for the abnormal evolution of some aggregated elements.

- More elaborate tables, designed for presentation to a more or less initiated public (the general public, the clients of the study, partners that have collaborated directly to it, other researchers, non-specialists of the subject). Even in the case of a working paper, a certain level of quality is needed. This is made easier by the most sophisticated model-building packages, and their production can be used for some types of publications not requiring too high a level. The quality of EViews tables has increased dramatically with the last versions, allowing to choose fonts, symbols and borders. In our opinion, it is no longer necessary to transfer information to presentation packages, except in very specific cases.

In practice, one observes that, even for if the model is completely assimilated, the production of relatively elaborate tables makes work easier, and allows to locate, not only the properties, but also the problems. It is a good investment to produce them as soon as possible, starting with the development phase. Once their framework has been defined, these tables can be reused indefinitely at no additional cost.

9.3.3 GRAPHS

The advantage of graphs is to clarify series evolutions by a more telling presentation. The disadvantage is the absence of precise figures. Graphs therefore should be used:

- Along with tables, to complement the detailed results by more synthetic and easily accessible information.
- Alone, if the message to transmit represents a simple synthesis.

The main graph type uses:

- Points, connected or not by segments, or more rarely by curves (differentiated by a color, a type of line, and often a symbol).
- Histograms, each observation being associated to a set of vertical bars (juxtaposed or piled) each measuring the size of a variable.
- Pies[126], associating to a single observation a circle whose slices correspond to the share of each variable in the total.
- Two dimensional bubble graphs (since EViews 10), using a third dimension for the diameter of the bubble.

This list is not restrictive, as the imagination can go far in search of visual synthesis. For example, bars in histograms are often replaced by appropriate drawings (persons, factories, products...).

[126] In French, "Camembert", a regionTapez une équation ici.al cheese with a round shape.

Concerning graphs (and presentations as a whole) the modeling packages have made considerable progress in recent years. Some offer sophisticated graphic features available through a large set of user-friendly functions, and EViews is not the last in that regard. One will be able for example, immediately after simulation, to display the evolution of historical and simulated variables or of selected expressions. This display mode will be used in the model development process, allowing a better interpretation of intermediate results, and revealing problems.

Since EViews 8, you can:

- Decide on the time scope of the graph using a slide line, available at the bottom.
- Add custom arrows.
- Export the graph to a PDF file.

And since EVIews 9:

You can select (as a rectangle) a rectangular part of a graph, which can be useful to evidence a feature of the series, for a given subsample, without changing the graph itself.

9.3.3.1 Graph of model solutions

For graphs, model solutions represent a specific case.

The most interesting feature is the MAKEGRAPH command, specially adapted to model solutions. The syntax is:

model_name.makegraph(options) *graph_name model_series_names*

If no option is specified only the current scenario solution is displayed.

The main options are:

a : include actuals

c : compare active to baseline

d : include deviations from baseline (as an additional graph in the same frame)

n : do not include active scenario.

The results can be presented raw or transformed, and for stochastic simulations a confidence interval can be displayed.

Once the graph is created, the usual modifiers can be applied (legend, type of line, colors….).

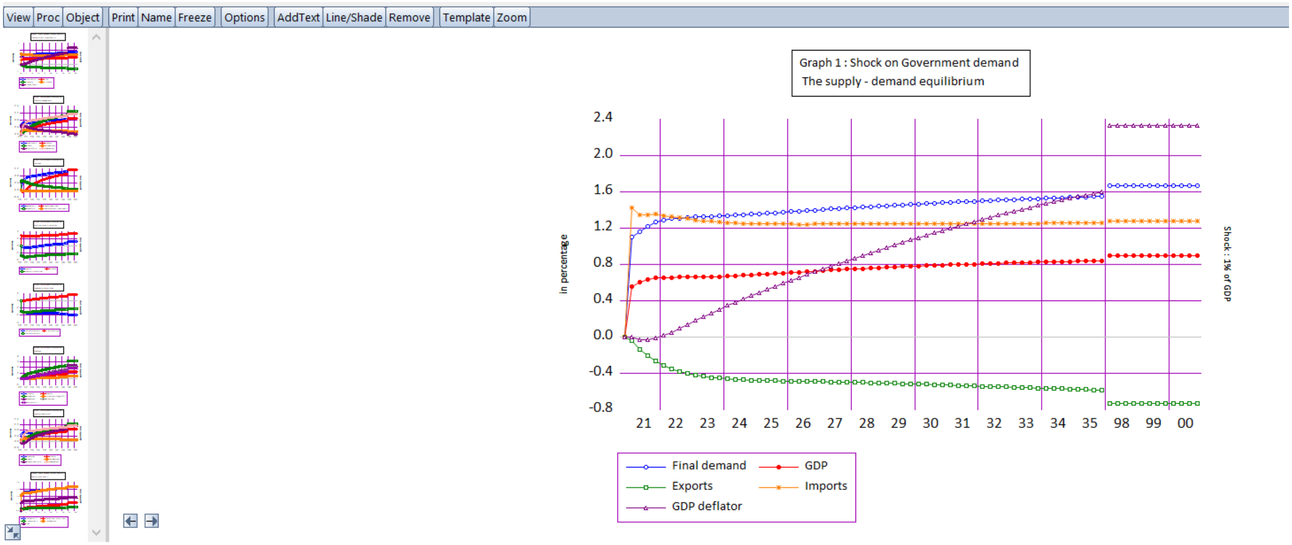
One has to note that the graph must not preexist. If it does it must be deleted first (using the “noerr” modifier to avoid a possible error message).

delete(norerr) *name_of the graph*

9.3.3.2 Multi-graph slideshow

Being allowed to move from graph to graph in a set is very useful if one wants to display the evolution of a group of important variables in a forecast, or different aspects of the model response to a shock: demand and supply, deflators, external trade….. This feature might lead to using in presentations an EViews session in parallel to PowerPoint slides.

This done through Zoom +/-



9.3.3.3 An example of model graph

We will now present an example of graph corresponding actually to the results of a shock, which we will present later.

We are not going to present the full syntax, just give some basic elements. Editing the statements to adapt to your case should be rather easy.

The variables represent the relative changes in final demand, GDP,, exports, imports and added value deflator following a 1% increase in Government demand (considered as the most meaningful shock available.

First we delete any previous version. Using a preexisting name will cause an error.

```
delete(noerr) gr1
```

Then we chose the type of graph. We chose lines.

```
graph gr1.line p_fd pg_gdpm pg_x pg_m pg_pq
```

Then we decide to place horizontal and vertical grids with given width and color (black). The three figures represent the amount of the three basic colors : red, green, blue.

```
_gr_g1.axis(left) grid font(11)
_gr_g1.axis(bottom) grid  font(11)
_gr_g1.options gridwidth(0.375) gridcolor(160,0,181)
We add the title at top(t), units at left(l) and size of shock at right( r)
_gr_g1.addtext(t, font=10,x) Graph 1 : Shock on Government demand \r The supply - demand equilibrium
_gr_g1.addtext(l,font=8) in percentage
```

```
_gr_g1.addtext(r,font=7) Shock : 1% of GDP
```

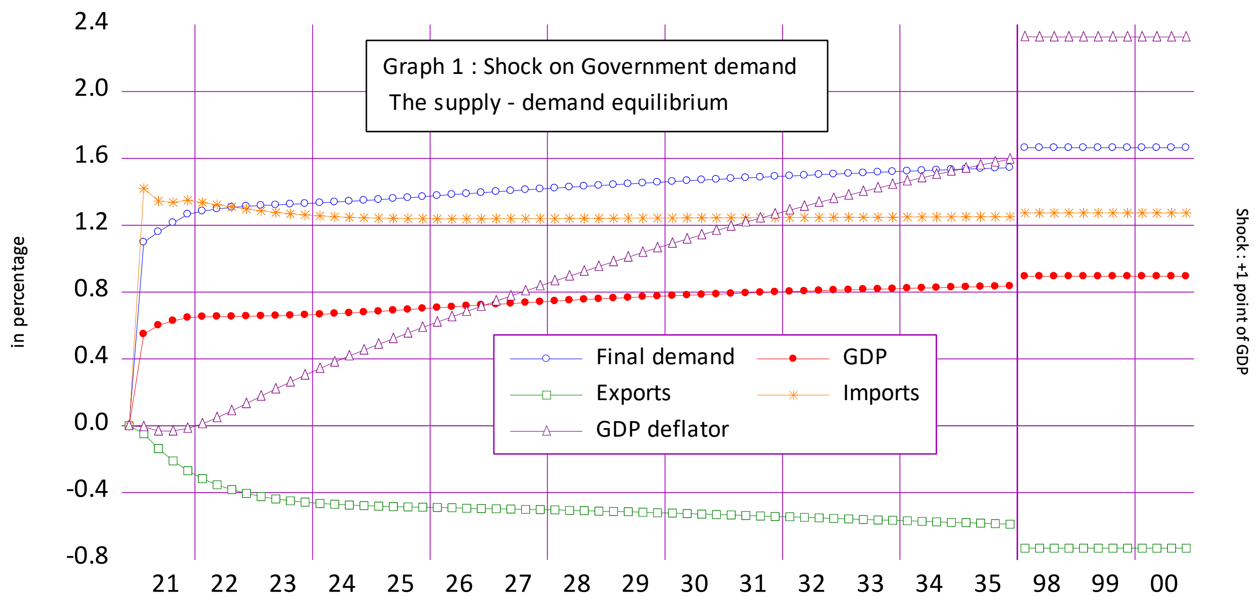
The legend will use two columns, with a black frame, positioned at the bottom on the left.

```
_gr_g1.legend columns(2)
_gr_g1.legend framecolor(160,0,181)
_gr_g1.legend position(botleft) font(10)
```

Now we define for each of the five lines the symbol, the color, the width, the pattern and the text.

```
_gr_g1.setelem(1) symbol(1) lcolor(blue) lwidth(0.25) lpat(1) legend(Final demand)
_gr_g1.setelem(2) symbol(2) lcolor(red) lwidth(0.25) lpat(2) legend(GDP)
_gr_g1.setelem(3) symbol(8)lcolor(green) lwidth(0.25) lpat(3) legend(Exports)
_gr_g1.setelem(4) symbol(4) lcolor(orange) lwidth(0.25) lpat(4) legend(Imports)
_gr_g1.setelem(5) symbol(10) lcolor(purple) lwidth(0.25) lpat(5) legend(GDP deflator)
```

We get the following result, after moving the legend and title to adapt their position to the rest of the graph, the best we could.



CHAPTER 10: APPLYING THE ABOVE PRINCIPLES TO OPERATIONAL CASES

In the preceding chapters, we have addressed all the technical aspects of models construction and applications, illustrating them by a simple example. Even if the techniques proposed have been quite general, we have not addressed the more economic features of operational models. This is what we are going to do now.

Having mastered the techniques for producing a very small model, the professional will face two kinds of difficulties:

- The structure of the new model has to be more consistent with economic theory (the one he has in mind). The detail of the specifications must meet his requirements, in terms of assumptions, mechanisms and results.
- The methodology for producing and managing the model call for slightly different techniques, adapted to a problem of higher complexity.

In the following pages, we shall address both these questions. As we are not giving a course in macroeconomics[127], we shall only describe the elements which in our opinion are requested for the production of an operational model, the choices one has to make, and give some directions as to the most promising options. We leave to the reader the choice of actual decisions and the introduction of additional features he considers adapted to his problem.

So please do not criticize our explanations for being too basic[128].They should be considered as a starting point.

If this is your feeling:

- You are a good enough economist to proceed by yourself
- Or you can use any of the more advanced textbooks in the list we provide.

Again, let us stress that you are not reading a book on sophisticated econometrics or economics. Our goal is to show how the two fields can be merged into a working model, concentrating on the technical aspects of the task.

In any case, the elements we will describe should be interesting, if interpreted as a description of the main options used by mainstream operational structural models (maybe a little on the old fashioned side).

For this reason, we do not cite a source in which the mechanisms we describe are presented. Rather, we give a list of books on macroeconomics, or full model presentations, in which most of them will be found in higher (and probably better) detail.

We will be more directive on the EViews aspects, of course.

10.1 THE ACCOUNTING FRAMEWORK

Before we start dealing with macroeconomic theory, we must describe the framework and the concepts we shall use. We have already started, of course, but in operational cases we must be much more precise and consistent.

---

[127] Neither did we give a course in econometrics.

[128] Some people will certainly decide they are wrong. This is another problem, but most of the features we present have been used in several models (this is not a proof of their validity of course).

As for econometrics, we do not have the ambition of giving a course in national accounting. In some cases, we will use simplifications which will make the presentation not entirely rigorous.

At the end of the book one will find a list of associated publications, from simple courses to official manuals stating the official concepts used at present by local or international organizations.

In our opinion, the best reference in that field is still:

LEQUILLER F, BLADES D, (2006), Understanding National Accounts, OECD Publishing, downloadable at: https://www.oecd-ilibrary.org/economics/understanding-national-accounts_9789264027657-en

For data and methodology, one can refer to:

http://www.oecd.org/sdd/na/

To define the framework of a macroeconomic model, one must define:

- The agents involved
- The operations they share
- The products associated with these operations.

According to the problem the model is designed to solve, these elements will be more or less detailed.

We shall start with a single good, single country model, then extend the classification. The specifics of the building process will also be dealt with separately.

In a special chapter, we shall address Stock-Flow Consistent modelling, a field which is much more rigorous in defining agents and operations, requiring the production of fully balanced accounting tables.

10.1.1 THE AGENTS: A FIRST DEFINITION

The operations described in a model will be managed by agents, characterized by their role. We shall separate:

- The firms, which buy and produce goods and market services, distribute revenue.
- The households, who receive revenue and consume part of it.
- The financial institutions, which manage financial transactions: lending, borrowing, insuring.
- The administrations, which gather taxes from the above agents, and address social needs through redistribution and production of associated elements.
- The rest of the world, representing the foreign agents trading with the first four.

10.1.2 THE OPERATIONS

They are divided into:

- Operations on goods and services: a good (material, such as a bottle of wine or a DVD player) or a service (immaterial, such as transportation from one place to another or a music lesson) is traded between two agents, generally in exchange for money. But if can be bartered (a kingdom for a horse) or given for free, most of the time by some administration (free school or free medicine). In the last case the transaction is considered as “non-market”.
- Transfers from one agent to the other, which can come in exchange for something, separately from the above transaction. For instance, wages are paid to households by firms, but they do not buy the actual good,

just the contribution to its production. Alternatively, pensions are financed by contributions, but not necessarily paid by the beneficiaries, and in any case with some delay. And income taxes are not linked with any service (although they will be used in part for the good of the taxpayer, at least in principle).

Things are not always so simple: an independent taxi driver will sell a service (and use the money as revenue). If he belongs to a company he will earn wages (a transfer). The company will sell the service and transfer part of the revenue to him.

10.1.3 THE INTEGRATED ECONOMIC ACCOUNTS

Once these elements have been defined, they can be presented in a table, allowing one line per operation, and one column per agent. Actually, an additional column should be introduced for the first “goods and services” part: for each operation, on the left for total expenditures, on the right for total revenue.

This is particularly useful for transfers: in general, wages will be an expenditure for firms and a revenue for workers. But all agents pay wages, including households: this technique allows to separate the two types: paid and received.

- In addition, the table will describe the steps in the economic process: the operations will be distributed in sequential sub-tables, each presenting a balance as the last line.

We shall find in succession:

- The “production” account, describing all operations linked to production, but also imports and exports. It gives value added. This part will call for an additional column, as lines are not balanced (exports represents an expenditure for the Rest of the world, but the revenue is not distributed among agents).
- The “primary distribution of income” account, introducing the immaterial expenses: wages, direct taxes, subsidies (-). It gives the operating surplus.
- The “secondary distribution of income” account, further distributing revenue: dividends, revenue of property, and also social transfers from the administration. The balance is disposable income.
- The “use of disposable income” account, describing essentially household consumption. The balance is savings.
- The “capital” account, describing the use of savings to acquire goods. It includes various forms of investment and the change in inventories. Its balance is the financial capacity.
- The “financial” account, which explains how the financial capacity is used(if positive) or met (if negative). It describes money, loans, credits and equities. Its balance is zero.

10.1.4 SECTORS, BRANCHES AND PRODUCTS,

Now we can separate the “firms” or rather “non-financial firms” into categories, according to its type of production. Three notions appear here:

- The sector: it categorizes a firm according to its main activity (car making for the Peugeot car manufacturer).
- The branch: it separates the firm into each of its activities (spare parts production and distribution for the Peugeot car manufacturer).
- The product: it represents the good or service actually produced by the firm. It is normally equivalent to the branch, but not in all countries: the difference can come from “fatal products” coming automatically from a process designed to create another good (such as hydrogen coming from oil refining).

As we shall see later, this decomposition calls for the definition of intermediate consumption, a two-dimensional variable associated to the consumption of one good to produce another. This can be true also for investment (in one product by one branch) or inventories[129].

10.1.5 AGENTS SUBDIVISIONS

National accounts can further subdivide the agents. For instance one can separate:

- Households into individual (a family) and collective (a convent).
- Financial institutions into the Central Bank, other banks, and insurance companies.
- Firms into State owned, individual firms and companies, local or foreign owned.
- The administration into central government, local agencies and specific agencies, such as the social security agency.
- The rest of the world into countries and zones (even for a single country model). It will allow to make the model assumptions more explicit.

Of course, other more statistical classifications can be used, like the revenue level of the household or the occupation of its head.

10.1.6 A MULTI COUNTRY MODEL

We shall develop later the specifics of a multi-country model. Let us just say for now:

- Of course, the system must be duplicated for each of the countries described. Specific elements can be introduced in some cases (like individualizing oil for producing countries).
- Consistency must be enforced between the exports of each country and the imports of others. This can be done either by identifying individual trade flows, or by creating a specific export demand variable for each country, based on the global demand of each of its clients.
- Finally, the categories can be countries, associations (like the European Union) or related zones (like Sub-Saharan Africa).

10.2 A SINGLE COUNTRY, SINGLE PRODUCT MODEL

We shall start with the economic developments needed by our model to reach operational status.

10.2.1 THE ECONOMIC ASPECTS

The very small model we have built as an example is of course too limited to be used for operational studies. It presents nevertheless:

- A link between production and revenue.
- External trade, depending on available productive capacity.
- Production defined as local demand plus net exports.
- A simple production function.

---

[129] Even though in this case some elements are identically null, such as investment in the energy product or inventories in financial services.

- A description of the output gap, its consequences and the way it is closed.

What do we need in addition, at the least?

- A price system, and its links with the real sector (in both directions).
- A better description of the behavior of firms and households.
- A financial sector.
- A full description of the State budget and its instruments.

This will also call for a redefinition of the scope of assumptions.

10.2.1.1 The productive process

This part of the model (one speaks often of "blocks") will not define production, but rather potential production (or productive capacity), as a function of available factors.

Why not actual production itself? There are two ways to consider production:

- Actual local production, contributing along with foreign exporters to the satisfaction of demand (both local and foreign), in a share depending on relative prices and available capacities).
- Potential production, given by the production function, taking into account technology and the level of factors (capital and labor), themselves chosen by firms according to their relative costs, expected demand, and profits conditions.

We want our model to follow the most logical causal sequence, which is:

1. Firms form expectations on demand.
2. They decide to try adapting their capacity.
3. They decide on the optimal level of factors allowing this capacity.
4. Technical difficulties, and concerns in adapting the factors too fast will create inertia in the definition of actual factors.
5. Actual factors will define actual capacity.
6. A gap appears with target capacity.
7. Firms will try to close this gap by increasing factor levels, correcting the decision in 4.

Alternate version:

2. Firms decide on the production target taking into account profitability.

The comparison between actual and potential production will play an important role in some behaviors.

This is the sequence that the model will describe, actual production being obtained late in the process, once demand is known (as in the small model).

This capacity for production will be measured:

- For employment, in man - years or man - quarters according to model periodicity
- For capital, at constant prices, in the currency of the country.

The function can also include:

- Energy consumption
- Intermediate goods (like raw materials).

Actually, capacities are generally defined in terms of value added, a more reliable notion as we have explained earlier. This means the two last elements are not considered, or rather their level will come automatically from value added itself.

The first issue concerns the logical link between capacity and factors. We have already seen:

- Complementary factors. For a given capacity, there is a single optimal process requiring a fixed combination of labor and capital. Starting from an optimal combination, adding to the process a quantity of one factor does not increase capacity, or allow using less of the other factor. This capacity is obviously optimal regardless of the relative costs. Actually labor productivity has generally some flexibility, and capital is the truly constraining factor, as temporary and limited increases in labor productivity can be achieved (for instance by increasing the number of hours worked).

This is the simplest option, in its formulation, estimation and understanding of properties. Operational models use generally more sophisticated frameworks:

- Cobb-Douglas. The elasticity of substitution is unitary. This means that if the ratio of the cost of labor to capital changes by 1%, the optimal ratio of capital to labor will change by 1% too, for a given capacity requirement. So the sum of the elasticities of capacity to labor and capital is one.
- CES (Constant elasticity of substitution). Now the elasticity can take any fixed value (with the right sign).

Of course, the CES option covers both others (with fixed elasticities of 0 and 1 respectively).

The framework calls also for:

- A definition of the relative cost.

The relative cost of labor and capital is not just measured by the ratio of the wage rate to the investment deflator. One has to take also into account:

  - Social contributions of firms: they contribute to the cost of labor.
  - The interest rate: while capital is bought immediately[130], labor can be bought (rented) when the time comes (slavery has been abolished for some time now). So a firm which has money can save it, and one which has not does not have to borrow. If capital itself can be rented, the distinction disappears.
  - The depreciation rate: capital wears out, while when a worker "wears out" through old age or sickness, he will leave and can be replaced by a new worker at no cost except training (pensions have already been saved as a share of wages).
  - The future evolution of wages: if wages are currently growing faster than inflation, firms can expect labor to become less competitive in the future. The gain from having output transferred to fast developing countries becomes lower as they reduce the gap with developed ones. This applies in particular to present Asian developing countries like Vietnam or Cambodia.

- The possible changes in technology.

The issue here is to decide if the technology decided at investment time (which defines the roles of labor and capital) can change later.

[130] Actually, some forms of capital (like buildings, computers or patents) can be rented or leased.

Basically, the options are:

- A single available technology (Clay-Clay).
- A technology chosen at installation time, with no possible later change (Putty-Clay). This means basically that the "complementary factors" option applies to factors once they are installed.
- A technology with a permanent possibility of change (Putty - Putty). The same substitution option applies to factors at any period.

The reader fill find easily examples for the three options.

10.2.1.1.1 A specific problem: the statistical determination of productive capacity

To determine the actual value of capacity, we have several options, depending on the available information

- In some countries (such as France), a survey asks a sample of firms by how much they could increase their production using the present factors (hiring more people if necessary). This gives the firm's capacity. Then a synthetic measure of capacity will be obtained using the same weights as for computing actual production, , and the rate of use as a ratio of the global values.

Then we shall use the capacity series to estimate its equation. For this, we can specify the actual behavior of firms, and optimize their profits under a capacity constraint using the formula we want to estimate. This applies when the factors are substitutable (otherwise the optimum solution is set from the start, and does not depend on relative costs). Taking the derivative of the function according to both labor and capital will give a set of equations with common coefficients, which one can estimate as a system. This method takes into account fully and explicitly the role of the relative costs.

- If we know only the level of factors (capital is sometimes missing in the country's statistics), we can specify the production function, and estimate its parameters over the actual values of production. We can suppose that the estimated formula gives normal production, and the residual is the output gap. Again, the ratio of actual to "normal" production gives the rate of use, but this time to a constant factor (the average rate of use).

We can also (a better solution in our opinion) apply the first method, using actual production instead of capacity. Again, the estimated capacity (reconstructed by applying the production function to the estimated factors, considered as optimal) will give a normal level of production, and the difference to actual production the output gap. However, we have to consider a delay before this level is reached, through the same inertia we have described earlier.

- If we do not have this information, we can always smooth production, and use the result as a "normal production" level (at a normal rate of use of capacities). For this, applying to actual data a Hodrick-Prescott filter is the most usual technique. If we suppose the "normal" rate of use of capacities constant over time, we get capacity at an unknown multiplicative factor.

This technique does not require a choice of production function, or the availability of a series for capital (which is often absent or unreliable). Neither does it provides it, which will be a problem for model specification and interactions.

10.2.1.2 The change in inventories

We see no specific reason to modify the framework used by the small model. More sophisticated formulations could use:

- A full error-correction framework, provided we knew the level of inventories.

- An influence of demand: if it goes up suddenly, some of it can be met by relying more on inventories. This element will be difficult to introduce, as it calls for a negative influence, while value added has a positive one, and both elements are positively correlated. This means the over estimation of one coefficient can be compensated by over estimating the second.
- An influence of prices: the more expensive the inventories, the shorter the time they will be stored.

Measuring the change in inventories at current prices introduces a specific problem, as one must also consider the change in the price of previously stored inventories, especially for a quarterly model.

10.2.1.3 Unemployment

This is a new concept, which would have fitted easily into the small model (increasing its size a little[131])

We shall consider that the variations of employment do not transfer fully to unemployment. Job creation will attract to the labor market previously inactive persons, who shall take some of the jobs offered: the work force (employed + unemployed) will increase.

For instance, creating a small firm in the Vietnamese mountains will allow housewives to combine employment with domestic work.[132] Or employees of a closing down factory will not necessarily remain in the labor market if their qualification is not required elsewhere.

But the level of unemployment should also influence its dynamics. If it is high, the incentive to join the work force will be lower. Favorable employment prospects will lead young people living with their parents to start their working life. On the contrary, a depressed labor market will persuade aged workers to retire earlier (and they will be incited to). And some of the unemployed will stop looking for a job, and leave the work force.

Also, the higher the unemployment level, the higher the quality of the best unemployed. Observing the situation, the average unemployed people will lower their expectations for getting a job, leading them to leave the work force.

On the contrary, at a low level of unemployment, the unemployed will feel that they stand a good chance over their competitors, most of them being either inefficient or not really looking for employment.

This obviously corresponds to an error correction framework, leading to a target rate of unemployment (and also of participation of potential workers to the labor force, as we shall see).

10.2.1.4 The price system

The role of prices in a model is of course essential. But it is not so simple to introduce, even for a minimal model like the one we presented above. In this case, several deflators have to be introduced simultaneously, associated with the elements in the supply - demand equilibrium:

- GDP
- Final demand
- Exports

---

[131] We have tested it, and it works. You can try too. And you can also link the real wage rate, making it endogenous.

[132] Which is not considered as employment (maybe because it is not paid, and does not affect GDP, even if paid housework does).

- Imports.

And in addition:

- Wages (possibly including social security contributions)
- Possibly, deflators for each element in the decomposition of demand: consumption, investment, government demand....
- The price of foreign currency (the exchange rate)
- The prices of lending and borrowing (the interest rates)

Moreover, trade prices have to be defined including and excluding taxes. This distinction applies to external trade (for defining competitiveness and trade balance) and local demand (for defining final and intermediate consumptions). And the price of labor (the wage rate) must take into account social contributions.

Not all these elements have to be estimated. Behaviors should be associated with:

- GDP (firms decide on the price at which they sell, once they take into account the cost of input).
- Exports (local exporters do the same).
- Imports (now we consider foreign exporters)133.
- Wages (the result of a negotiation between workers and firm managers).

The demand price is obviously missing. Its role will be to balance the supply-demand equilibrium at current prices:

$$Pfd \cdot FD + Px \cdot X = Pq \cdot Q + Pm \cdot M$$

Or

$$Pfd\,.FD = (Pq \cdot Q - Px \cdot X) + Pm \cdot M$$

In which demand at current prices appears as the sum of two demands : supplied by local firms (production – exports) and foreign firms (imports).

- In the system, the deflators will depend on each other. For the time being, we will only give indications. A more detailed reasoning will come with actual estimations.

  - The GDP deflator depends on the wage rate, or rather the wage cost per unit produced.

If the wage cost goes up, firms will have to increase prices to keep their margins. If the slice of the pie decreases, the size of the pie can compensate.

---

[133] Remember we are building a single country model. The description of trade will be different with several connected countries.

They do not have to do it immediately, and they have also to consider their competitiveness on the local and foreign markets (for exporting firms).

And if the firm maximizes its margin, it will be more accurate to consider the global cost, including amortization of capital.

- The wage rate depends on the consumption price, but maybe also on the value added price.

If prices go up, workers ask for a raise in wages to sustain their purchasing power. But again, firms are more likely to accept raises if they were able to increase their own price.

- Trade prices depend on the cost supported by the exporter, and on the price set by its competitors. This means they have to maintain their margins and their competitiveness at the same time.

This behavior is obviously based on production prices, the price at which they sell. This means the cost of intermediate consumptions has to be taken into account. For instance, a country having access to cheap oil will be able to export at lower prices, even at the same cost in value added (and the same margins). But this introduces a problem, as until now the single product feature allowed us to neglect intermediate consumption, a variable difficult to manage as its value depends on the classification and the length of the production process.

The behavior also has to apply to the same currency. If the export price uses the currency of the exporter, the price of its competitors measured in foreign currency has to be corrected by the exchange rate.

- The price of demand depends on the price at which local producers and foreign exporters sell on the local market.

This uses the identity above.

Another important issue concerns the behavior of firms selling on both on local and foreign markets.

Two behaviors can be considered:

- The firms define both selling prices separately. Local firms start by defining a price for selling on the local market, using the above behavior. Then their export price will average this price and that of competitors.
- The firms define first a global selling price, allowing to reach a global margin rate, then they chose a combination of the two prices which meets this target. This means that a decrease in the export price (possibly designed to stay competitive facing foreign deflation) will have to be compensated by an increase in the local selling price.
- Of course, this behavior is completely justified for firms which sell only on one market, like producers of Nike shoes in Vietnam or private teachers in France.

The choice will have a strong impact on the price system. Actually the second option will increase the intensity of the price-wage loop: if local costs go up, firms refuse to apply completely these costs to exports (as they do not want to lose their competitiveness), and maintaining global margins calls for a larger increase in local selling prices (which does not happen if targets are defined separately).

- This equilibrium is subject to external influences, either endogenous or exogenous.

  - Endogenous

    - If labor productivity goes up, firms need fewer workers and can pay them more. They can also lower their prices.

* If output is too low compared to capacities, firms can first lower prices to sell more (later they can adapt their capacities).
* If unemployment goes down, workers can increase their demands without the risk of firm managers to look elsewhere.

- Exogenous.

The concerned variables are the indirect tax rates.

One will generally consider:

* VAT.
* The other indirect tax rates, such as the tax on tobacco, gas, alcohol.
* Tariffs.

And also

* The rate of social security contributions by firms.

If indirect[134] tax rates (such as VAT, tax on gas, cigarettes, social contributions paid by firms...) go up then firms should adapt their price if they want to keep their margins.

Two elements:

- First, it is quite important to separate these taxes in a model, for the usual reason: their base is different, and their impact on the economy also.

* VAT applies only to demand. But the most important feature is that it does not apply to exports (exporters can deduct it before they sell abroad), and they apply to imports. VAT on foreign cars is the same as on local ones, and applies to the total value. And when the car producer looks for electronic equipment, increasing VAT on this good will not change its decision on its origin as it can deduct VAT anyway[135].
* On the contrary the other indirect taxes apply only to local productions, even though the impact of this difference is not so high, as imported goods are often taxed at the moment they are sold. For instance, the tax on tobacco applies also to imported cigarettes, and the tax on alcohol to imported whisky.
* Concerning local tariffs, they are not deductible in the general case. This means that a change in their rate will affect directly the competitiveness of imported goods, unless the importer decides to compensate the effect by adapting its margins.
* As to tariffs applied to local products by foreign countries, they affect directly the competitiveness of exports. This means they have to be considered, even though their statistical value is not directly obtained from the national accounts of the country.

- Second, formalizing the role of taxes relies obviously on the rates, the variables decided by the state (or foreign states for tariffs on local exports). These rates will affect deflators, and allow computing the amount of the tax once the base is known.

---

[134] These taxes are called indirect because they are not paid directly by the ultimate payer, contrarily to income tax, corporate tax....

[135] It might not apply it fully, however, if it fears a drop in its sales.

It should be clear that the right way to formalize these taxes is to set the rate as an exogenous decision variable, and not to estimate the amount as some modelers might be tempted to do. This allows handling the decision easily, both in forecasts and shock analysis. And on the past, the technique is quite simple: the tax amount and the base are known, and this allows computing the rate, used as an exogenous ratio. The associated identity (tax = rate x base) will hold exactly true. We shall see later how to handle these rates on the future.

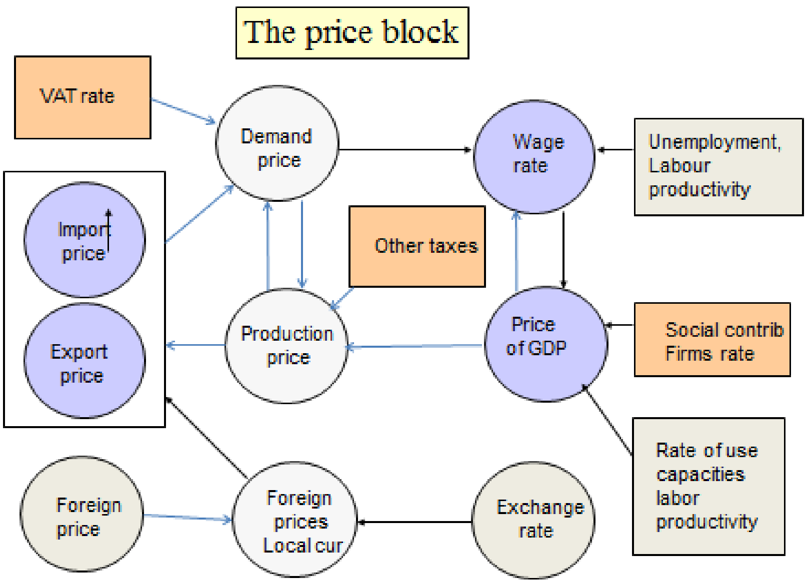
Obviously, the rate obtained will be different from the legal one (generally lower!).

This technique is consistent with the general approach: make the government decisions exogenous, but identify first what is the true decision variable.

- Prices can also influence real elements

  - The selling price of local producers determines the quantities they will sell. This is also true of exporters, through the comparison between their export price and the price on the market on which they sell
  - The relative costs of labor and capital influence the choice of the factors in the productive process.
  - More generally, ratios of prices affect the ratios of elements (or the shares in a total). For a given global consumption level, reducing the price of one good will increase its share.
  - A higher inflation reduces the purchasing power of previous savings, calling for a larger reconstitution effort.
  - And of course prices enter the definition of variables at current prices, when they are separated into volume and deflator (elements in the trade balance, wages...). For the elements in a sum, a different evolution of deflators will change the shares at current prices. And some ratios are computed at current prices, with a different deflator on the numerator and denominator.

All this is described by the following graph.



10.2.1.5 The account of firms

We have already dealt with the supply side, defining the adaptation of production factors: employment and capital, to target capacity, as well as the decision on prices, based generally on the short-term maximization of profits.

This means that most of the remaining equations will be definitions, describing the firms account without calling for any theoretical elements.

There are three exceptions:

- The tax on profits, which should be again computed by applying a rate to a base. This is more complex than usual, however, as:

  - Computing profits in a model is quite complex, and not all models are able to do it. Sometimes it is necessary to use a proxy, making the apparent rate more difficult to interpret.
  - The timetable for the tax calls for a dynamic equation, as the tax is not generally paid in the same period as the associated profits (but there can be a provision to pay immediately). So a formula describing the mechanism must be established.
  - The tax on negative profits is not negative, but null, introducing a bias on the apparent rate.

- The income tax, which can be subtracted from wages, or paid later, with a possible advance system.

- Dividends paid by firms, which can be estimated or constructed through an identity (using a rate in the same manner as taxes). Again, one must decide on the dynamics, as dividends follow the profits. Also, the beneficiary of dividends has to be identified (sharing must be done between the five usual agents).

Of course, the complexity of formulations (and even the identification of elements such as dividends) depends on the role of the model, whether it is used by researchers trying to answer global theoretical issues, or by policy advisers addressing in detail the evolution of the next State budget.

10.2.1.6 The behavior of households

On the contrary, we have not yet addressed the decisions of households.

Basically:

- Households obtain revenue from several sources, the main ones being:

  - Wages
  - The revenue of individual workers
  - Remittances
  - Social benefits of various kinds
  - Interests from loans in local or foreign currency
  - Interests from bills and bonds
  - Dividends from equities
  - Rent transfers with other households (a service)[136]

- They use this revenue:

  - To pay the income tax
  - To consume various goods and services
  - To save, and in particular in housing, but also in deposits, bonds, stocks and goods (such as art).

To be considered operational, even a single product model must use some detail, as the economic processes through which these revenues are obtained, and the consequences of spending decisions, are quite different from each other.

---

[136] Actually, it would be strange to consider that if a household buys the apartment it was renting, the service disappears and GDP decreases. For that reason, housing owners are considered by National Accounts as paying to themselves a fictitious rent.

Another principle of modelling: favor the detail which allows separating behaviors.

Basically

- For revenue:

  - Wages paid by firms should be the product of an average wage rate (coming from the price block) by the number of workers (from the production block).
  - The number of civil servants will generally be exogenous, but not the wage rate, which can be different from firms'.
  - Wages paid by households (mainly for housekeeping) can be identified or not, according to the type of model.
  - Remittances will depend on the wage rate in foreign countries, corrected by the exchange rate.[137]
  - Social benefits are generally separated in five types: sickness, family subsidies, unemployment benefits, invalidity from working accidents, pensions. It is clear that:

∗ Each of these elements depends on inflation, but at different degrees.
∗ Most of them depend on population, and often a given type of population. For instance, the number of children, the number of people having reached retirement age, or of unemployed.
∗ All of them depend on economic activity, again in a variable way. For instance, unemployment benefits decrease with GDP, working accidents increase, and pensions should increase (in principle) with the revenue from the contributions which finance them.
∗ They also depend on a decision made by the State (for instance the purchasing power is maintained).

This means an operational model should try to separate these items, to take into account their differences in behavior.

In this way the model will show naturally:

∗ The change in benefits with the number of beneficiaries.
∗ The change in benefits with the government policy.

  - The interests will be described globally, in a subsequent paragraph (and in more detail in the chapter on SFC models). Let us only stress that for households the interest rates (lending and borrowing) can be deviate from market values through state intervention. In France, a limited number of savings benefits from a higher guaranteed rate, and borrowing to buy housing can be done at a lower rate (0% in some cases).
  - Dividends will be treated later with the firms' account.
  - As to the revenue from housing (rents), its role in a model is limited, as it mostly represents a transfer from households to other households. For owners of real estate, it is even a transfer within the same household. There are reasons to consider it, however: it can be subject to taxation, and it enters GDP.

It is not essential to consider marginal elements, such as lottery winnings, inheritance, donations, fines...

  - Finally, one can formalize the transfers from abroad (or to abroad). For developing countries remittances can represent a sizable share of household revenue (more than one third of GDP for Tajikistan). For a single country model, they should be exogenous, perhaps even in current terms (a notable exception to the general principle, with particular consequences).

- For expenditures:

[137] Modelling the number of expatriates is an interesting issue.

- The income tax should be computed as a rate applied to revenue before tax, obtaining the historical values of the apparent rate by dividing the amount by the base. The model will then get the tax by applying the exogenous rate to the base. This base poses the usual timing problem: the tax can be paid after the revenue is obtained (with a provision mechanism).

Also, applying an average rate to all households can be acceptable for forecasts (which allow this rate to change with time), but less for the shocks addressed to a category of households at one extremity of the spectrum: in a traditional macroeconomic model, a decrease in the tax on large incomes or an increase in benefits for the poor, of the same ex ante size, will have the same ex post consequences[138]. To eliminate this error, an ad hoc correction has to be made on the savings rate itself.

This problem appears in most models, coming from the fact is that the tools to solve it are not available. National Accounts separate firms using the goods they produce, but not households according to any categorization, including the level of revenue. Some surveys address the problem, and their teachings could be used to create specific data. This means some solution might be found, but without doubt at a high cost. Actually, the same problem arises if one wants to separate firms not according to sectors, but to size, considering that small firms act differently from large ones.

- Once the disposable income is known, the first task is to separate it into consumption and savings, considered as whole in most models (for multi-product models the situation will be more complex).

The most common technique is to compute consumption first, as a ratio to revenue, then savings as a residual. We shall develop this with estimations.

Consumption is generally determined at constant prices (which means in purchasing power). The usual determinants are:

- The level of revenue (measured also in purchasing power). The higher the revenue, the higher the consumption level, but the lower the share of consumption (the poor do not save, and remember that buying a house is considered as savings).
- The recent evolution of revenue. Households take some time in adapting their behavior to an increase (or decrease) in revenue. And a sudden hike (especially if it is destined to be permanent, like a promotion) can lead them to invest in housing, which can actually decrease consumption for a while.
- Inflation (the “real holdings” or Pigou effect). Present savings contain a large share of monetary elements (deposits, bonds with fixed rates...). Current inflation reduces their purchasing power, which has to be complemented by additional savings. The effort is proportional to the inflation level.
- The unemployment rate. For employed workers, an increase in the chance of losing their jobs (measured more by the change in the rate than its value[139]) leads them to save a larger share of their present revenue, if they want to optimize their utility across time.
- The (short-term) interest rate: in general, people prefer satisfying a given need now than later. But this has a cost, the interest they have to pay. The lower the rate, the more they will indulge in immediate consumption.

This is particularly true for durable goods: if a household wants to watch TV on a set using the latest technology (and thinks that after its purchase, in its whole life it will have enough resources to afford this kind of set) the only reason for not buying one right now (and increasing its satisfaction for a period) is the actualized cost, which is lowered with a decrease in interest rates. What it has to consider is not the cost of the good, but the cost of making its acquisition earlier.

---

[138] Of course, the impact on consumption will be higher if the increase concerns the poor.

[139] Although the actual rate plays also a role: a higher value implies a higher turnover, and a high risk of participating in the turnover.

If the good is perfectly durable, and can be sold back at its original value at constant prices, things are just as if it was renting the good. If the interest rate is divided by 2, the "price" of the increase in satisfaction is divided by 2.

For non-durable goods, the situation is different. The household has already optimized its consumption over time. If the interest rate changes, it might be tempted to consume earlier, but if the marginal utility of the good is decreasing fast, the pattern of consumption will not be much affected. A person dreaming of visiting the pyramids, and saving for that purpose, might make the trip earlier but will not do it again.

What matters is the real rate:

- They allow comparing goods at constant prices.
- If households assume their revenue will grow with inflation, they will optimize in real terms.

Once consumption is determined, savings are computed as a residual, and generally as a global element. This option can be discussed, as different kinds of savings can be assumed to follow different behaviors.

In particular, housing investment is negatively affected by interest rates (a specific rate, but one can assume it follows the global rate) while financial savings are positively so. Buying a house calls for obtaining a given good and asking another agent to provide the collateral, in return for interests. Buying a bond means lending collateral to another agent to use it as a spending tool (maybe to buy a durable good), in return for interests but this time in the other direction.

We will develop this feature when we address SFC models.

10.2.1.7 External trade

In a single country model, the rest of the world is exogenous.

This means that we consider only influences from the world to the country, and not the other way around.

Of course, this is not really exact, even for the smallest of countries (or in that regard for a country region, a town or an individual): by increasing your consumption and so local production, you create a fraction of a job, a small amount of household revenue, and again more consumption.

What we consider is that the influence is too small to have a sizable effect, and that the cost of producing (and running) a model describing it is too high compared to the gain in the accuracy of results. This is essentially true for lower or medium sized countries like Latvia or Bolivia, much less so for larger countries like France, and quite untrue for the USA or the European Union considered as a whole. For instance, when we used the MacSim world model for a shock analysis, the French Keynesian multiplier for 2000 was 1.3 if we ran the full model, but only 1.1 if we ran the French model by itself. The iterative feedbacks of German imports from France, coming from the increase of German exports, will have the largest share in the difference. Considering the evolution of world trade, the present difference should be even wider.

This means that the exchanges of the country have to be considered from the point of the country:

- Exports are the share of production of goods and services which is sold by the country to the rest of the world.
- Imports are the share of local demand for goods and services which is not produced in the country but bought from the rest of the world.

Both elements will be computed using the currency of the country. However, using elements at constant prices will mean using the exchange rate of the base year, so the currency issue is not relevant, introducing only a scaling by a constant factor[140].

However, the trade elements having the same nature, their logical determinants will be the same. The main difference will come only from the relative size of the two markets (buyer and seller) in the trading operation: the single country's importance (or GDP) will always be much lower than that of the rest of the world, although this is less obvious again if we model the USA or the European Union as a whole.

These elements will be:

- Demand: for a country to sell a given good to a partner country, demand for this good must be present, part of this demand must be addressed to the world market, and the quality of local products must appeal to the importing country. For instance, French exports of wine will depend on the world demand for wine, and the natural preference of importing countries for foreign wine (starting with their status as wine producers).and French wine in particular. And at present Vietnam will not export so much rice and coffee to markets asking for high quality products. However, one has to consider evolutions of this criterion, as the limits on land availability and technical progress is changing the nature of supply.

To define world demand, one can:

  - Use a global index, provided for instance by an international institution like the OECD or the World Bank.
  - Weigh roughly the growth of GDP or better demand in the main trading partners.
  - Use the precise weighting of partners and the evolution of their imports, to build an accurate index.

The best but rather complicated option (which we have actually implemented) is to establish a recursive prologue, in which foreign demand to the country is computed by weighting the global demand from each partner by the share of this partner in exports. The data requirements are higher, but a combination of local trade statistics and international data (from the World Bank WDIs?) should provide the information.

The same technique can be used to compute competitors' deflators, using the same weightings (and different ones for imports).

This framework allows to study the consequences of detailed shocks (like a decrease in Chinese growth for a Vietnamese model).

Defining demand introduces a problem.

For imports, we have already seen that including intermediate consumption in the supply-demand equilibrium (thus considering production on one side and total local demand on the other) is quite a problem for models, as the level of intermediate consumption depends on the number of steps in the production process. The single product feature has until now eliminated the need for considering intermediate consumption. But imports do contain intermediate goods, whether they represent energy (oil, gas, even electricity) or primary goods (from untreated wood to electronic components). And these intermediate goods are necessary to export.

A simple solution is to consider the ratio of intermediary consumption to value added. Looking at the figures, we can indeed observe that the "technical coefficients", the number of units needed to produce a unit of value added or GDP, is rather constant.

[140] This is only true if we consider a single rest of the world, or we measure it in a single currency. More on this later.

So we have just to consider a composite demand, as the sum of final demand itself, and intermediate consumption as a function of GDP (or rather value added, as intermediate consumption excludes VAT).

- Price competitiveness: to decide whether to buy a good from a local or foreign producer, a country will compare the local price with the foreign exporters' price. And to choose among potential sellers, the importing country will consider their relative price at a given quality (remember that the deflators consider goods at the same quality level, an increase in quality improving the value at constant prices instead of the deflator).

We generally observe that the relative price is less of an issue when the buyer contemplates buying local or foreign goods, than when he has to choose between foreign sellers. This follows economic logic: local goods are supposedly designed for the local market, and some goods are not or hardly substitutable (local bus tickets, meals in restaurants or newspapers).

This means in our case that the sensitivity of exports to price competitiveness should be higher than for imports. Exports depend on world demand to the world market, and once a country has decided to import, the price will play a more important role than in the import decision itself.

Finally, of course measuring competitiveness must use deflators defined in the same currency. It can be any currency, as applying the same exchange rate to both elements of the ratio will not change its value. In the case of exports, this means that measuring their deflator in local currency calls for a foreign price measured in the same units. As the exchange rate is identified, this foreign price will be endogenous, as the product of two assumptions: the foreign price in foreign currency, and the correcting exchange rate (a deflator). It is perhaps more logical (and equivalent in practice) to consider both prices in foreign currency, the local one being corrected by the symmetric exchange rate.

To compute precisely this element, one can use the same technique as for foreign demand: weighting foreign production prices by the share in the country's imports and exports (giving two separate indexes).

- The available capacities

The third element is the potential to supply additional demand, which means the presence of available productive capacities. The relevant variable is naturally the rate of use of capacities, independent from the size of the economy.

The choice of this option is not so straightforward, however. One could argue that as long as the rate is lower than one, additional demand can be satisfied. We have already shown that this is not true: demand concerns a wide range of products, and one cannot generally be substituted for another, in particular in the short-term. And some products may see their capacity completely saturated.

Let us explain the process again, this time in an import-oriented way.

Actually, the average rate of use is based on a distribution of rate values, from zero (hopefully a few cases) to one (probably a sizable number). When global demand increases, it addresses a range of firms, in variable intensity. Some of these demands will be addressed to firms already unable to provide more, and some others will make them reach that level. The proportion of firms working at full capacity will grow. Of course, in some cases, another available product can represent a substitute. But the simplest option is to import the same product, as the missing product should be available somewhere in the world (maybe at a higher price, but this should be treated by the price competitiveness).

The "missing" demand increases with the share of firms which cannot increase their production, having reached capacity.

Of course, this phenomenon applies essentially in the short-term, as firms will react by investing, which will increase capacity and close the output gap with time. But this process can be slow, even if full adaptation should be obtained in the very long run.

But if we follow the above reasoning, we observe:

- That the larger the country, the lower the probability that a given absolute, but also **relative** increase in demand will face local supply problems. This increase in demand will be more diversified, and the available capacities will be more diversified too[141]. At the world level, the issue disappears.
- That in our case, the rest of the world should not face any supply problem, which means that for both our country's imports and exports, only the local rate of use should be taken into account.

  Of course, this is not true for a multi-country model. This issue will be addressed later.

- And a last condition can appear for the exporting country. If the world requires a given good, the characteristics of that good produced in the country must also be adapted to the particular demand, which can be quite specific. For instance, facing an increase in the foreign demand for cars, a country might be able to supply them at a competitive price, but the type of cars they propose might be unsuitable. This might concern the size, the metrics used, the sophistication, the powering energy, the security features... Of course, one cannot consider going in such a detail in a model, even if the data was available (which it is not).

Unfortunately, finding an element describing this feature is less straightforward than above, especially for a single product model. The simplest idea is to use the age of capital, assuming that a younger productive process will be better adapted to present demand[142]. For instance, a recently built car factory might follow market trends in producing smaller cars, or more energy efficient ones. The age of capital can be derived simply from the chronology of investment and depreciation, if we consider that this depreciation applies equally to all generations of present capital, or that capital disappears brutally a given number of years after its implementation. Another assumption leads to more complex, but manageable, formulas.

- A last element (linked to the previous one) is the share of foreign investment in the productive capacity. The product from foreign firms can be better adapted to foreign demand (and they might have been designed for the foreign market in the first place). This feature is implemented in some of or models

10.2.1.8 The Budget

In operational models, describing fully and consistently the Government budget is an absolute requirement.

This is true even if the model is not going to be used by Government advisers, but rather by experts in applied economics. A crucial goal of these researchers is to assess the consequences for the economy of Government decisions, external events, or structural changes, considering the most probable impact or the range of possibilities, possibly under different model formulations (like different options on the interest rate). The approach might be more or less applied (the advisers might try to produce an image of the next budget, to be presented to the Parliament, and the scientists will try to see how the adoption by the Central Bank of a Taylor rule will stabilize the economy), but the tool required is quite similar.

As we have stated above, the best way of defining the associated equations is to build identities, computing an endogenous revenue or expenditure as the product of an endogenous base by an exogenous rate. The equations will hold true over the past, and the modeler will be responsible for (and allowed to) establishing future assumptions on the rate. Of course, he does not have to keep this rate constant, and can rely on an estimated trend as a starting base. But the final decision will be his.

---

[141] This would not happen if the additional demand was in a specific good.

[142] Especially foreign demand as its role is increasing with time, so its influence on the nature of investment will be higher in later periods.

This technique answers to the following objection: if we consider VAT, even with constant legal rates, the apparent rate will change (grow) with the affluence of households, able to increase the share of highly taxed products in their consumption. One solution is to establish a trend, used as a base value, and to deviate from this trend as a policy decision.

If these principles are followed, it will be possible to produce a table showing the evolution of all budget elements, in current terms and in GDP points, both features being obviously required for presentations.

Another important principle of modelling: if you cannot choose between the possible presentations for a given concept (value at constant prices, at current prices, growth rate, and ratio to another variable) just look at how this concept is presented in economic publications (focusing on the ones designed for the general public). Or wait until you will have to use the figures in your own presentations, then measure your reaction and that of the public.

10.2.1.9 Financial and monetary elements (basic version)

This feature will be much improved in the chapter on Stock-Flow Consistent models. For now, we give only some pre-SFC elements, for modelers with limited interest in financial features.

In any model, this represents the most variable and controversial part. The first models had little or no financial equations. Even at this stage, the financial block can be limited to the definition of a few rates, and their impact on the real sector (these rates can even be exogenous, generally in real terms). On the contrary, this block can be so developed that the purpose of defining a real sector can be considered as a way to complete the links between financial elements, for instance describing the creation and origin of additional lending if a decrease in interest rates draws investment upward.

In our opinion, even a real side oriented model should include:

- A base interest rate set by the Central Bank of the country.
- A short and a long-term rates in the currency of the country.
- An average rate on current net borrowings.
- A rate on the present debt, being computed from the chronology of past rates, perhaps as an autoregressive function.
- One or several foreign rates, applied to borrowings in foreign currency both in the country and in the rest of the world.

- The net interests paid by all (five) agents, considering two currencies for the interests paid to the Rest of the World.

From this basic option, developments can consider:

- Identifying the debt of agents (or their financial holdings).
- Separating it into currencies (local, US Dollars, maybe Euros for non EMU countries).
- Separating it into short-term and long-term.

- In addition, one or more forms of money supply can be formalized.

Most of these equations should be established as identities, based on available data or assumptions. Exceptions can concern:

- The Central Bank rate, following perhaps a Taylor rule, but not necessarily. Actually, the same model should allow several options (using a separating parameter).
- The short-term and long-term rates could include a risk premium, depending for instance on the current budget deficit or its most recent values.

- The spread between long and short-term could depend on growth expectations (truer if they are partly or totally rational) and the health of the local economy.

We will develop all these aspects later.

10.2.2 THE EVIEWS PROGRAMS

We shall now present EViews programs producing the data and the framework for a model of the type we have just described.

Note: all the programs presented here are available on the EViews site, or on request from the author (who shall provide an access through Google Drive)...

Just as in the previous example, we shall start by stating completely the identities, but limiting the definition of the behavioral equations to a declaration of the type of function we intend to estimate (this will be done later).

As we have stated before, producing the data and stating the equations can be done in any order, but creating the groups and predicting the residual check must wait for both tasks to be completed.

However, we shall separate the program which creates the data, and the one which creates the equations (and produces the residual check).

The reasons for this shall be made clear later, but we can already state the main issue: during the model building process, both tasks will evolve, but separately most of the time. Changing the model specifications will be much more frequent than changing the data set, and we want to avoid running a task without reason.

It seemed to us that our presentation would be clearer if we started with model production. It might be also the more natural sequence for the model builder: first considering the scope of data available (without studying individual series in detail[143]) then writing down the equations in sequence, observing more precisely if each series is present (directly or through a transformation), then creating the data building program.

10.2.2.1 General elements on running programs

In the general case, you just have to access the program and run it.

But you can also run part of a program, by selecting it with the mouse (in the usual Windows way), clicking on the right button, and choosing “Run Selected”.

This is generally more efficient than the previous method of copying the selected part into a blank program, and running it. However, the new method does not allow editing, useful when one wants to run a selected AND modified set.

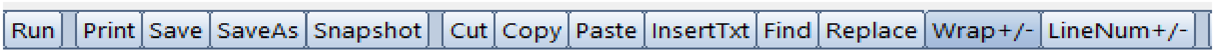
Symmetrically one can exclude temporarily from execution part of a program, by “commenting it out”. To do this, one should select the relevant part, click on the right button, and choose “Comment Selection”. To reactivate the statements, one should select them again and use “Uncomment Selection”.

This can be a little dangerous, especially if you (like myself) have the reflex of saving the program before each execution. To avoid destroying the original, one can save first the modified program under another name[144].

Finally, one can ask a column of numbers to be displayed left of the program lines, switching the “LineNum” option to “+”.

[143] For instance one can start from the list of variables presented at the end of the book.

[144] Only once of course.



This is particularly efficient if you use the “Go To Line” statement[145], but also to locate the error position when the error message provides the information.

Unmatched parenthesis in "_GR_G1.LEGEND COLUMNS(2" on line 10.

The only drawback is that the statement has to be displayed in a single line, which requires activating the “Wrap+” option. This means long statements will be cut to the screen size.

10.2.2.2 Producing the model

Let us now see how the economic elements we have described can be combined into an EViews program, along the lines of the simpler case.

In this program, we shall expand the principles of the method we have outlined earlier, to create the framework of a small but operational model. As in the example, we shall stop before any estimation. What we should get is a set of equations in which:

- Identities are fully defined.
- Behavioral equations are defined as identities presenting, using specific notations, the explained variable and its explanatory elements.

The following text will essentially contain the EViews statements, and comments present in the file itself (a quote mark appears at the start of the line). Additional comments will appear without quote mark, and use a different font. They are generally associated to features requested longer explanations, which would have taken too much place in the program.

‘ We start with the usual definition of the directory

```
cd "c:\program files\eviews11\book"
```

‘ We state that any test results will be exported to a Rich Text Format file named _mod_1.rtf

‘   output(r) _mod_1 would create a basic text file.

```
output(r,o) mod_1
```

‘ We close the input and output file

[145] However, you must be careful to update the numbers when the program changes.

‘ in case a version is already open

```
close data_1
close mod_1
open data_1
wfsave mod_1
```

‘ We define the sample period as the maximum available

‘ restrictions will appear with actual estimations

```
smpl 1962Q1 2017Q4
```

' We now need to create a new (blank) model
‘
‘ Trick
‘ We give the model a name starting with « _ »
‘ « _ » is the first element in any alphabetical ordering
‘ As the workfile elements are displayed in alphabetical order
‘ this means it will appear first in the display
‘ (of course among elements using the same trick).
‘ This is quite useful for frequently accessed elements
‘ In case the model already exists, we delete it (the “noerr” option allows failure if the model does not exist).

```
delete(noerr) _mod_1
```

‘ So this means we start afresh

If we do not do this, the previous equations will be retained, even if the new equations redefine the same variables (even using identical formulations). This means variables would be defined twice, which EViews will obviously refuse, producing an erroneous model. This happens because identities are « appended » as text to the model specification.

We will not present them, as we consider much more efficient to state the full model in the same program. We have already presented the reasons for this method: clarity of the model text, better error identification, easier transfers, and easier management of programs.

This is not true for estimated equations, as we shall see later.

```
model _mod_1
```

‘ We define the f scalar, for the production of the behavioral « identities »

```
scalar f
```

We start with the production block

```
‘----------------------------------------------
‘ The production block
‘----------------------------------------------
```

‘ The rate of use of capacities is the ratio of actual GDP to potential GDP

```
_mod_1.append UR=Q/CAP
```

‘ GDP balances the supply – demand equilibrium

```
_mod_1.append GDP+M=FD+X
```

‘ To get value added we substract VAT at constant prices, the product of the base year VAT rate by final demand excluding VAT, at constant prices

Identifying value added is necessary to compute firm’s margins and deflators excluding VAT.

In a model at constant prices, the VAT rate would be more or less constant, and VAT proportional to GDP. But here the use of VAT at current prices, and a deflator excluding VAT, calls for excluding VAT from a denominator measured at constant prices. We shall use value added Q in the production function (even though GDP would provide a similar explanation). .

```
_mod_1.append Q=GDPM-r_vat0*FD/(1+r_vat0)
```

‘ Investment depends on GDP, the rate of use, profitability, previous capital and the long-term real interest rate. If we consider capital – labor substitution, it will depend on the relative cost of factors.

We see that we can already introduce lags (for the previous level of capital) and elements of formulations (for the real interest rate).

```
_mod_1.append I=f*(Q+UR+RPROF+K(-1)+(IRL-100*@pchy(PC))+RELC)
```

We have chosen to measure capital at the end of the period.
It is the sum of the non-discarded previous capital
and the investment implemented during the period.

Actually what we consider here is the investment purchased during the period. There is no guarantee that it can be used readily for production.

```
_mod_1.append K=K(-1)*(1-rdep)+I
```

‘ The change in inventories will depend on GDP

```
_mod_1.append IC=f*(Q)
```

‘ Employment by firms too, and also possibly on the relative cost of factors.

```
_mod_1.append LF=f*(Q+RELC)
```

‘ Productivity of labor is the ratio of value added to firms’ employment

```
_mod_1.append PL=Q/LF
```

‘ Productive capacity depends on firms’ employment and the initial level of capital)

```
_mod_1.append CAP=f*(LE+K(-1))
```

```
‘   Total employment introduces civil servants
‘    actually we should also identify households’ employees
‘    such as maids and janitors
```

```
_mod_1.append LT=LF+lg
```

```
‘    The actual work force depends on employment and the potential work force,
‘     in practice the population in age of working.
```

```
_mod_1.append POPAC=f*(LT+pop65)
```

```
‘     The unemployed are the jobless inside the work force
```

```
_mod_1.append UN=POPAC-LT
```

```
‘     The unemployment rate
```

```
_mod_1.append UNR=UN/POPAC
```

```
‘--------------------------------------------------------
‘   The price block
‘--------------------------------------------------------
```

```
‘   The unitary wage cost is the wage
‘   (including contributions by firms)
‘   necessary to produce one unit of value added
```

```
_mod_1.append UWC=WR*(1+r_scf)/PL
```

```
‘    The value added deflator will depend on UWC
```

‘ and the rate of use

```
_mod_1.append PQ=f*(UWC+UR)
```

‘ The production price aggregates the value added deflator
‘ and the price of intermediary consumption
‘ weighted by its influence :
‘ the demand price excluding VAT
‘ It will be needed for defining the trade prices.

```
_mod_1.append PP=(PQ+tc*PFDXT)/(1+tc)
```

‘ The deflator of final demand is the ratio of demand
‘ at current and constant prices.
‘ It balances the supply – demand equilibrium at current prices
‘ All other elements are or will be computed elsewhere

```
_mod_1.append PFD=(GDPMVAL+MVAL-XVAL)/(GDPM+M-X)
```

‘ We compute also the deflator excluding VAT
‘ by inverting the relation with PFD:

```
_mod_1.append PFDXT=PFD*(1+r_vat0)/(1+r_vat)
```

‘The deflators for demand elements depend on the global one.
‘The relation can be estimated, or we can apply an exogenous ratio

```
_mod_1.append PC=f*(PFD)
_mod_1.append PI=f*(PFD)
_mod_1.append PIG=f*(PFD)
```

‘The wage rate depends on deflators:
‘PC from the point of view of workers
‘PQ from the point of views of firm managers
‘And also on labor productivity and unemployment

```
_mod_1.append WR=f*(PC+LP+UNR+PQ)
```

‘The trade deflators in local currency depend on the local and foreign production prices
‘which have to be converted in local currency through the exchange rate

```
_mod_1.append PX=f*(PP+ppx+ER)
_mod_1.append PM=f*(PP+ppx+ER)
```

‘ The exchange rate can be exogenous
‘ or depend on the inflation differential (PPP assumption)

```
_mod_1.append ER=f*(PP+ppx+ERX)
```

‘ The short-term interest rate can be exogenous
‘ in nominal or real terms
‘ or depend on inflation and the output gap (Taylor formula)

```
_mod_1.append IRS=f*(IRSX+(IRSR+100*@pchy(PC))+ (150*@pchy(PC)+50*(UR-urd)/urd))
```

‘ The long-term interest rate depends on a lag structure of short-term rates
‘ possibly with a spread

```
_mod_1.append IRL=f*(IRS+spread)
```

‘ The rate on new borrowings is an average of short and long-term rates

```
_mod_1.append IR=f*(IRS+IRL)
```

‘ The rate on previous borrowings depends on the previous rate
‘ and the current rate
‘ according to the reimbursement speed

```
_mod_1.append IRM=f*(IRM(-1)+IR)
```

```
'---------------------------------------------------------
The households block
'---------------------------------------------------------
```

```
'    The firms' wages
```

```
_mod_1.append WF=WR*LF
```

```
'    The civil servants wages
'    could be different on average from firms'
```

```
_mod_1.append WG=WR*lg
```

```
'    Total wages
```

```
_mod_1.append W=WF+WG
'
```

```
'    Social benefits are exogenous in purchasing power per head
'    They could be separated into risks
```

```
_mod_1.append SOCB=socbr*PC*popt
```

```
'    Additional revenue can be linked to value added
```

```
_mod_1.append REVQ=r_revq*QVAL
```

```
'    or be exogenous in purchasing power
```

```
_mod_1.append REVX=r_revx*PFD
```

' More complex formulations can be used
' Social security contributions use an exogenous rate

```
_mod_1.append SCW=r_scw*W
```

' Household income is defined as a sum of its elements

```
_mod_1.append HI=W-SCW+REVQ+REVX+SOCB
```

' Income tax applies to the revenue of the previous year
' again, this can be made more complex
' or use the present value of revenue
' according to the country's rules

```
_mod_1.append ICT=r_ict*HI(-1)
```

' Now we compute disposable income

```
_mod_1.append HDI=HI-ICT
```

' also in purchasing power

```
_mod_1.append HRDI=HDI/PC
```

' Household consumption will depend on:
' * real disposable income
' * inflation
' * unemployment

‘   * the short-term interest rate

```
_mod_1.append HCO=f*(HRDI+PC+UNR+(IRS-100*@pchy(PC)))
```

‘   * a more explicit relation could be used

‘   _mod_1.append HCO=f*(HRDI+HCO(-1)/HRDI(-1)+PC+d(UNR)+(IRS-100*@pch(PC)))

‘   or even:

‘   _mod_1.append Dlog(HCO) =f*(Dlog(HRDI)+Log(HCO(-1)/HRDI(-1)+Dlog(PC)+d(UNR)
‘   +(IRS-100*@pch(PC)))

‘ This will not change the diagnosis on model structure
‘   But will state the type of relationship one wants to estimate
‘   As a personal reminder or a presentation to partners

```
‘----------------------------------------------------------------
'    The firms block
‘----------------------------------------------------------------
```

‘   Value added at current terms

```
_mod_1.append QVAL=PQ*Q
```

‘   The value added tax

```
_mod_1.append VAT=r_vat*PFDXT*FD/(1+r_vat0)
```

‘   Gross Domestic Product

```
_mod_1.append GDPMVAL=QVAL+VAT
```

‘   And its deflator

```
_mod_1.append PGDPM=GDPMVAL/GDPM
```

' Subsidies are proportional to value added

```
_mod_1.append SUBS=r_subs*QVAL
```

' Margins are computed as:
' value added plus subsidies minus « other indirect taxes »
' minus wages including firms' social security contributions

```
_mod_1.append MARG=QVAL*(1+r_subs-r_oit)-WR*LF*(1+r_scf)
```

' The margins rate

```
_mod_1.append RMARG=MARG/QVAL
```

' The tax on profits is based on past profits

' A more complex rule can be applied

```
_mod_1.append IFP=(PROF(-1)+IFP(-1))*r_ifp
```

' Profits exclude:
' * household revenue from production
' * the tax on profits
' * net interests paid

```
_mod_1.append PROF=MARG-REVQ-IFP-NIF
```

' The profits rate applies to capital
' at the cost of renewal

```
_mod_1.append RPROF=PROF/(PI*K(-1))
```

‘ We can also compute the ratio of margins to capital
‘ a more stable and reliable concept

```
_mod_1.append RPROB=MARG/(PI*K(-1))
```

‘ The balance of firms excludes spending on
‘ * productive investment
‘ * the change in inventories
‘ as margins includes yet unsold value added

```
_mod_1.append FCAPF=PROF-PI*I-PFD*IC
```

‘ Net interests paid depend on
‘ * their past value
‘ * the rate on past debts
‘ * the rate on new debts
‘ * the balance

```
_mod_1.append NIF=f*(NIF(-1)+IRM+IR+FCAPF)
```

‘ --
' The external trade block
‘ --
‘

The import price including tariffs

```
_mod_1.append PMT=PM*(1+r_tar)/(1+r_tar0)
```

‘ Import price competitiveness compares
‘ * the local production price
‘ * the import price including tariffs

```
_mod_1.append COMPM=PMT/PP
```

‘ Final demand at constant prices is the sum of its components
‘ including a residual proportional to GDP
‘ (a reasonable assumption)
‘ Housing investment could be identified

```
_mod_1.append FD=COH+I+IC+CG+IG+fdxr*Q
```

‘ Imports depend on

‘ * Final demand and intermediate demand
‘ (proportional to value added)
‘ * The rate of use of capacities
‘ * Price competitiveness

```
_mod_1.append M=f*(FD+Q+UR+COMPM)
```

‘ Export price competitiveness compares
‘ * the foreign production price
‘ * the local export price including tariffs
‘ The exchange rate corrects the currency difference

```
_mod_1.append COMPX=PX*(1+r_tarx)/(1+r_tarx0)/(PPX*ER)
```

‘ Exports depend on
‘ * World demand (both final and intermediate)
‘ * The rate of use of capacities
‘ * Price competitiveness

```
_mod_1.append X=f*(WD+UR+COMPX)
```

‘ Trade flows are computed in current terms

```
_mod_1.append MVAL=PM*M
_mod_1.append XVAL=PX*X
```

```
'       The export import ratios are computed:
'       * at current prices
'       * at constant prices
'       * for the deflators
```

```
_mod_1.append RCVAL=XVAL/MVAL
_mod_1.append RCVOL=X/M
_mod_1.append TTRAD=PX/PM
```

```
'     The trade balance
```

```
_mod_1.append TRB=XVAL-MVAL
```

```
'       Net interests paid to the Rest of the world
'       are separated depend on currency.
'      They depend on:
'       * their past value
'       * the rate on past debts
'       * the rate on new debts
'            the international rate is used for interests
'            in foreign currency
'       * the balance
'       * the exchange rate for the debt in foreign currency
```

```
_mod_1.append NIXL=f*(NIXL(-1)+IRM+IR+TRB)
_mod_1.append NIXX=f*(NIXL(-1)+IRMX+IRX+TRB+ER)
_mod_1.append NIX=NIXL+NIXX
```

```
'    The financing capacity
```

```
_mod_1.append FCAPX=TRB-NIX
```

```
'-----------------------------------------------------
'     The State budget block
'-----------------------------------------------------
```

```
'     Most of its elements have been computed already
```

‘ as transfers between the State and another agent
‘ Social contributions paid by firms

```
_mod_1.append SCF=r_scf*Wf
```

‘ Other indirect taxes

```
_mod_1.append OIT=r_oit*QVAL
```

‘ Tariffs

```
_mod_1.append TAR=r_tar*MVAL
```

‘ Social contributions paid by the State

```
_mod_1.append SCG=R_SCG*WG
```

‘ Total revenue

```
_mod_1.append REVG=SCF+SCG+SCW+OIT+IFP+ICT+VAT+TAR+r_revg*QVAL
```

‘ Government investment at current prices

```
_mod_1.append IGV=IG*PIG
```

‘ Government consumption at current prices

```
_mod_1.append CGV=CG*PFD
```

‘ Government demand at current prices

```
_mod_1.append FDGV=CGV+IGV
```

‘ Net interests paid by the State depend on
‘ * their past value
‘ * the rate on past debts
‘ * the rate on new debts

‘ * the balance

```
_mod_1.append NIG=NIG(-1)*IRM/IRM(-1)-IR/100*FCAPG
```

‘ Total expenditures

```
_mod_1.append EXPG=FDGV+WG+SUBS+SOCB+NIG+SCG+r_expg*QVAL
```

‘ Government balance

```
_mod_1.append FCAPG=REVG-EXPG
```

‘ Government balance in GDP points

```
_mod_1.append FCAPGP=100*FCAPG/GDPMVAL
```

‘ Total GDP
‘ The sum of market and non-market GDP
‘ equated to the total wage cost of the State

```
_mod_1.append GDPVAL = GDPMVAL+WG+SCG
```

```
'--------------------------------------------------------------
'      End of model specifications
'--------------------------------------------------------------
```

10.2.2.3 Producing the data: an OECD example

Now we will present a program creating the data for the above model. We have used a very simple case, in which the model builder has access to the OECD "Economic Perspectives", a file containing about 5000 quarterly series describing the world economy. Each OECD country is described individually using the same concepts (with a few exceptions), and other important countries (like China) or zones (like Latin America) using less detail. However, the definitions vary slightly across countries, and although the same set of series is always technically present, some of them do not contain any value, or very few. For instance, the notion of savings is not always the same, and capital stock is not always available.

As we shall see, our base contains all the data we need for France, with one or two minor exceptions.

Obviously, this program is quite case dependent, and a user starting from another base will use a rather different one. A program with the same purpose is available for the World Bank World Development indicators, with a lot more countries and series, but an annual periodicity.

However:

- The user might actually start from this OECD base, a quite popular one.
- Some of the tasks (creating the workfile, producing a trend...) will appear in all cases.
- And actually the program will not be so different. Almost of the variables accessed in the OECD base will be available in any base of this type, and one will just have just to replace the "OECD" names, changing some of the concepts if needed.

The only additional problems for the prospective user of the program will come from:

- Unavailable data which has to be estimated, or guessed.
- Additional data required by the introduction of new elements in the model.

```
'=================================
'     An example of data transfer
'=================================

'    This program will start from the original French data
'     provided by OECD Economic perspectives and named fra_*,
'     the prefix used by OECD to identify French statistics
'     We decide on the directory
'     This is not generally necessary
'     except if one works on several projects
'     or maintains several directories for the same project
'     It guarantees trial versions do not destroy official ones

cd  " d:\eviews\_pic_2020\pic_eviewsf"

'    It will create the data for our model
```

'   with the prefix f_*
'--
'   This technique can be used with any source
'   where the original series use the same prefix
'   the results will be created with any different prefix
'   In the best case, if the set available is the same (or larger) than the OECD set
'   one has just to replace the OECD names in the following statements
'--
'   we close the original file fra_1
'   containing the sole fra_* data
'   and some global OECD series named OECD_*
'   we close also the file which will receive the French data
'   in case it is already open
'   having two versions of the same file open in memory is quite dangerous...

```
close fra_1
close data_1
```

'   We open the original file (presently closed)
'   and save it under the name data_1 for the French data

```
open fra_1
save data_1
```

'   Now the file should contain only original data
'   called fra_*
'   we delete any existing f_* series
'   just in case, this should not happen...

```
delete f_*
```

'   We have to make an assumption on the sharing of indirect taxes

```
'    into VAT and other indirect taxes
'    as OECD provides only a global variable
'    p_oit = assumption on the share of oit in indirect taxes
'

scalar f_p_oit=0.2

'    we create a time trend
'    with the value of the year for the first quarter
'    to which we add 0.25 for each following quarter of the year
'     1994    : 1994Q1
'     1994.25: 1994Q2
'     1994.50: 1994Q3
'     1994.75: 1994Q4
'     1995.00: 1995Q1 ....
'   This will be quite useful to:
'     create yearly time trends
'     replace actual dummy variables by expressions using logical conditions
'     much easier to manage

smpl 1962Q1 1962Q1
genr f_t=1962
smpl 1962Q2 2017Q4
genr f_t=f_t(-1)+0.25

'    Now we start with the supply - demand equilibrium

smpl 1962Q1 2004Q4
genr f_gdpval=fra_gdp

‘  Gross domestic product at current prices

genr f_gdp=fra_GDPV

‘  Gross domestic product at constant prices

'-------------------------------------------------------------
```

```
'    Supply
'------------------------------------------------------------

'   The model separates market GDP

genr f_gdpmval=fra_gdp-fra_cgw

‘ =  gross domestic product at current prices – Government wage consumption at current prices

genr f_pcog=fra_cg/fra_cgv

‘ = Government total consumption at current prices / = Government total consumption at constant prices

genr f_gdpm=fra_GDPV-fra_CGw/f_pcog
genr f_pgdpm=f_gdpmval/f_gdpm

'      Trade at current and constant prices, deflators

genr f_m=fra_MGSV
genr f_x=fra_XGSV
genr f_pm=fra_PMGS
genr f_px=fra_PXGS
genr f_xval=f_px*f_x
genr f_mval=f_pm*f_m

'      The deflator of final demand (including VAT and OIT)

genr f_pfd=(f_GDPmVAL+f_MVAL-f_XVAL)/(f_GDPm+f_M-f_X)
genr f_fd=fra_TDDV-fra_CGw/f_pcog

'      Identifying the indirect taxes This separation is important as VAT applies to final demand

genr f_oit= f_p_oit* fra_TIND
genr f_vat = (1-f_p_oit)* fra_TIND
genr f_r_vat =f_vat/(f_fd*f_pfd-f_vat)
smpl 2014q1 2014q4
```

```
scalar f_r_vat0=@mean(f_r_vat)
scalar f_r_oit0=@mean(f_r_oit
smpl @all
genr f_r_oit=f_oit/(f_gdpmval-f_vat-f_oit)
genr f_pfdxt=f_pfd*(1+f_r_vat0)/(1+f_r_vat)

'      Value added excluding VAT (but including OIT)

genr f_qval=f_gdpmval-f_vat
genr f_q = f_gdpm-f_r_vat0*f_fd/(1+f_r_vat0)
genr f_pq=f_qval/f_q

'      Capital, rate of use, capacity

genr f_k=fra_KTPV
genr f_ur=fra_GDPV/fra_GDPVTR
genr f_cap=f_q/f_ur
genr f_pk=f_cap/f_k(-1)
genr f_urd=1

'-------------------------------------------------------------
'    Demand
'-------------------------------------------------------------

'      Intermediate consumption
'      Not identified in the OECD data base
‘      We specify the ratio as 1 (a common value)

genr f_tc=1
genr f_ic=f_tc*f_q
genr f_td=f_fd+f_ic

'      Production price

genr f_pp=(f_qval+f_ic*f_pfdxt)/f_td
```

```
'     Elements of demand and their deflators
'     Household consumption

genr f_coh=fra_CPV
genr f_pcoh=fra_cp/fra_cpv

'     Investment, depreciation rate

genr f_i=fra_IBV
genr f_pi=fra_ib/fra_ib
genr f_rdep = (f_k(-1) + f_i - f_k) / f_k(-1)

'      Housing investment

genr f_hih=fra_IHV

'      Government investment (excluding State firms)

genr f_ig=fra_IGV
genr f_igv=fra_IG
genr f_pig=f_igv/f_ig

'      Government consumption (excluding State firms)

genr f_cogv=fra_cg-fra_cgw
genr f_cog=(fra_cg-fra_cgw)/f_pcog

'      Government demand

genr f_fdgv=f_cogv+f_igv
genr f_fdg=f_cog+f_ig

'      Change in inventories

genr f_ci=fra_iskv
```

```
'      Individual demand deflator ratios

genr f_r_pi=f_pi/f_pfd
genr f_r_pcoh=f_pcoh/f_pfd
genr f_r_pig=f_pig/f_pfd
genr f_r_pcog=f_pcog/f_pfd

'     Wage rate

genr f_wr=fra_WAGE/(fra_ET-fra_ES)
genr f_r_scf =  fra_WSSS/fra_WAGE -1

'       Checking that the decomposition of final demand is correct

genr f_fdxr=(f_fd-f_coh-f_i-f_hih-f_ic-f_fdg)/f_q

'--------------------------------------------------------------
'    employment, unemployment, population
'--------------------------------------------------------------
'      Employment

genr f_lt=fra_ET
genr f_lg=fra_EG
genr f_lf=fra_ETB

'     Labour productivity

genr f_pl=f_q / f_lf

'     Wages and wage cost

genr f_w=f_wr*f_lt
genr f_wf=f_lf*f_wr
genr f_uwc=f_wr*(1+f_r_scf)/f_pl

'      Unemployment
```

```
genr f_un=fra_UN
genr f_unr=fra_UN/(fra_ET+fra_UN)

'       Population

genr f_pop=fra_POPT
genr f_pop65=fra_POPT
genr f_popac=f_lt+f_un

'----------------------------------------------------------------------------
'   Financial elements
'----------------------------------------------------------------------------

'      Exchange rate

genr f_er=1/(fra_EXCHEB/@elem(fra_EXCHEB,"1995"))
genr f_erx=f_er

'     Nominal interest rates

genr f_irs=fra_IRS
genr f_irl=fra_IRL
genr f_ir = fra_irwyp
genr f_irm=f_ir
genr f_irsx=fra_irsaf
genr f_irlx=fra_irlaf
genr f_irmx=fra_irfor
genr f_irx=fra_irfor

‘     Real interest rates
‘     Here we assume the formulas used for the short-term
‘         and the average interest rate on past debts
‘         their logic will be explained later

genr f_irr=f_ir-100*@pchy(f_pcoh)
```

```
genr f_irl_ec=0
genr f_irsr=f_irs-100*@pchy(f_pcoh)
genr f_irst=f_irs-(150*@pchy(f_PCoh)+50*(f_UR-f_urd)/f_urd)
scalar p_f_irm=0.8
genr f_irm_er=f_IRM-(p_f_irm*f_IRM(-1)+(1-p_f_irm)*f_IR)

'     relative cost (for substitution between labor and capital)

genr f_relc=f_wr*(1+f_r_scf)/f_pi/(f_ir/100-@pchy(f_pcoh)-4*log(1-f_rdep))
genr f_spread=3

'--------------------------------------------------
'     Households
'--------------------------------------------------

'      Global elements

genr f_hi =fra_YRH-fra_TRPH
genr f_ict=fra_tyh
genr f_r_ict = fra_TYH / f_hi(-1)
genr f_hdi=f_hi-fra_tyh
genr f_hrdi=f_hdi/f_pcoh
genr f_sr=1-f_coh/(f_hrdi)

'     Individual elements

genr f_socb=fra_trrh
genr f_scw= fra_TRPH  - f_r_scf*f_w
genr f_r_scw=f_scw/f_w
genr f_rpro=f_hi-(f_w-f_scw+f_socb)
genr f_revx=(1-0.5)*f_rpro
genr f_revq=0.5*f_rpro
genr f_r_revx=f_revx/f_pfd
genr f_r_revq=f_revq/f_qval
genr f_wg=f_wr*f_lg
genr f_hdir=f_hdi/f_pcoh
```

```
'-----------------------------------------------------
'     external trade
'-----------------------------------------------------

'       Tariffs

genr f_tar=0
genr f_r_tar=f_tar/f_mval
genr f_r_tarx=0
scalar f_r_tarx0=@elem(f_r_tarx,"1995")
scalar f_r_tar0=@elem(f_r_tar,"1995")

'       Foreign elements

genr f_wd = fra_XMVMKT
genr f_ppx=OECD_PGDP

'      Competitiveness

genr f_PMT=f_PM*(1+f_r_tar)/(1+f_r_tar0)
genr f_compm=f_pmt/f_pp
genr f_COMPX=f_PX*(1+f_r_tarx)/(1+f_r_tarx0)/(f_PPX*f_ER)

'       Ratios and balances

genr f_rcvol=f_x/f_m
genr f_rcval=f_xval/f_mval
genr f_ttrad=f_px/f_pm
genr f_trb=f_xval-f_mval
genr f_FCAPX=fra_FBGSV

'      Interests paid
‘      In the OECD data set, the interests paid to the RoW are not identified
‘      We apply the formulas used by the model
‘      They will be explained later
```

```
smpl 1990Q1 2004Q4
scalar p_nix=0.5
genr f_nixl=-f_ir*f_trb/40*p_nix
genr f_nixx=-f_irx*f_trb/40*(1-p_nix)
smpl 1990Q2 2004Q4
genr f_nixl=(f_nixl(-1)*f_irm/f_irm(-1)-f_ir*f_trb/400*p_nix)/(1-f_ir/400)
genr f_nixx=(f_nixx(-1)*f_irmx/f_irmx(-1)*f_er/f_er(-1)-f_irx*f_trb/400*(1-p_nix))/(1-f_irx/400)
genr f_nix=f_nixl+f_nixx
genr f_fcapx=f_trb-f_nix
smpl 1962Q1 2004Q4

'--------------------------------------------------------
'    Government budget
'--------------------------------------------------------

'       Revenue

genr f_scf=f_r_scf*f_wf
genr f_r_scg=  f_r_scf
genr f_scg=f_wg*f_r_scf
genr f_ifp=fra_tyb
genr f_r_oit=f_oit/(f_qval)
genr f_revg=f_ict+f_oit+f_vat+f_scf+f_scg+f_tar+f_scw+f_ifp
genr f_r_revg=0
genr f_recg=0
genr f_r_recg=0

'       Expenditures

genr f_socbr = fra_TRRH /f_pcoh/f_pop
genr f_subs=fra_TSUB
genr f_r_subs = fra_TSUB / (f_qval)

'       Interest rates and balance
```

```
genr f_nig=fra_gnintp
genr f_fcapg=-FRA_NLG
genr f_nig_er=(f_NIG-(f_NIG(-1)*f_IRM/f_IRM(-1)-f_IR/400*f_FCAPG))/f_qval
genr f_expg=f_revg-f_fcapg
genr f_r_expg=(f_expg-(f_fdgv+f_wg+f_scg+f_nig+f_socb+f_subs))/f_qval

'       Balance in GDP points

genr f_fcapgp=100*f_fcapg/f_gdpval

'--------------------------------------------------------------
'      Firms account
'--------------------------------------------------------------

'       Margins

genr f_marg=f_qval*(1+f_r_subs-f_r_oit)-f_wf*(1+f_r_scf)
genr f_rmarg =f_marg / f_qval

'      Interests paid, profits and balance

genr f_prof=fra_PROF
smpl 1978Q1 2004Q4
genr f_fcapf=-fra_NLB
smpl 1978Q1 1978Q1
genr f_nif=f_ir*f_fcapf/11
smpl 1978Q2 2004Q4
genr f_nif=f_nif(-1)*f_irm/f_irm(-1)-f_ir/400*f_fcapf
smpl 1962Q1 2004Q4
genr f_rprob = f_marg/(f_pfd*f_k(-1))
genr f_prof1=f_marg-f_revq-f_ifp-f_nif
genr f_prof_er=(f_prof-f_prof1)/f_qval
genr f_r_IFP=f_ifp/(f_PROF(-1)+f_IFP(-1))
genr f_rprof=f_prof/(f_pfd*f_k(-1))
genr f_fcapf_er=(f_fcapf-(f_prof-f_pi*f_i-f_pfd*f_ci))/f_qval
save data_1
```

10.2.2.4 Creating the model groups and checking the data – equations consistency

These two tasks can be started only when the model has been defined, and the data created.

Creating the groups needs not only the model, but also the data, as EViews groups can only be built from (existing) series.

This program can be run after the two tasks, in the same session. This is why we did not introduce the usual statements specifying the directory and opening the workfile.

The program is model independent. It:

- Creates a list of endogenous and exogenous variables.
- Solves the model equations separately from each other twice, for different values of the scalar « f ».
- Separates the endogenous into identity and behavioral using the fact that the equations for the latter contain the scalar « f »
- Computes the level and relative differences between historical values and the results of one of the simulations above.

```
' ==========================================
'   creating the groups
'==========================================
'    We create groups for the endogenous and exogenous
'    The statement « makegroup »
'    * applies to the model « _mod_1 »
'    * creates a group of endogenous (@endog) or exogenous (@exog)
'    * with a name _g_vendo or _g_vexog
'     The modifiers specify the elements in the group
'    « n » means we do not want the suffixed elements
'     (suffixed by the current suffix, actually « _c »)
'    « a » means we want the actual names (no suffix)
'    This overrides strange EViews default options

_mod_1.makegroup(a,n) _g_vendo @endog

_mod_1.makegroup(a,n) _g_vexog @exog

'  Now we want to separate identity and behavioral variables
'  We shall use a trick
'   The future behavioral equations use the scalar "f"
'   We solve the model twice for different values of "f"
'   using different suffixes for the solution
'    We need to "update" the model
```

```
'    for the parameter change to be taken into account

smpl 1995Q1 20017Q4
f=1
_mod_1.append assign @all _d
solve(d=f) _mod_1
f=2
_mod_1.update
_mod_1.append assign @all _c
solve(d=f) _mod_1

'    We create two empty groups

group _g_vbeha
group _g_viden

'   Then we consider in turn each endogenous variable
'   _g_vendo.@count is the number of elements in _g_vendo
'   _g_vendo.@seriesname(!i) is the name of the ith element in the group
'   We consider the variable built from this name and the two suffixes "_c" and "_d"
'   If the values are different then the variable is behavioral
'   We test the condition on a given period
'   (but not the base year as there is a chance the right hand side gives zero
'   in which case the method does not work)
'   If it is true we add the element to the list of behavioral elements
'   otherwise to the list of identity elements
'   « endif » close the condition
'   « next » closes the loop
'    the brackets « {»   and  « }» delimitate the parameter
'    they are dropped after the replacement

for !i=1 to _g_vendo.@count
%1=_g_vendo.@seriesname(!i)
scalar  {%1}_eq=(@elem({%1}_c,"2014Q1")<>@elem({%1}_d,"2014Q1"))
if  {%1}_eq=1 then
_g_vbeha.add {%1}
```

```
else
_g_viden.add {%1}
endif
next

' ==========================================
'   computing the residuals
'==========================================
'   Now we check the residuals
'   We define the suffix as « _c »

_mod_1.append assign @all _c

'   We reduce the sample to a period
'   for which each equation can be computed
'   taking into account the presence of lags

smpl 1986Q2 2017Q4

'     But only for identities
'     However, we were able to compute the « behavioral »
'     over the given period
'     which means that we will be able to estimate
'    The dc_{%1} are the absolute differences
'    The pc_{%1} the relative ones
'     To avoid dividing by zero, we use a trick
'     We add to the variable a boolean one
'     Testing if the variable is zero
'     If it is true we get 0 divided by 1 = 0
'     If not the computation is not affected
'
'     This means we have just modified the outcome of divisions.
'     so that dividing zero by zero now gives zero

for !i=1 to _g_viden.@count
%1=_g_viden.@seriesname(!i)
```

```
genr dc_{%1}={%1}-{%1}_c
genr pc_{%1}=100*dc_{%1}/({%1}+({%1}=0))
next
```

10.2.2.5 Estimating the equations: examples

We shall now apply the above principles to the estimation of equations. Basically, this means replacing in the previous program the identities declaring the modeler's intentions by actual equations.

We shall only present the new program sequences, associated with this process. The full model building program is available as an annex, and includes a summary of the comments we are going to make. It is also available as a .prg file on the model site.

Of course, the results we shall obtain are specific to the French case. They should only be considered as a (working) example, and will probably prove much less directly useful than the previous programs.

One will observe that we are not always following a pure and clean methodology. Our opinion is that we behave better than most. And anyway rather than a perfect element, what we are proposing here an example of the general methodology used by model builders, including deviations from the "politically correct" path.

For instance, we shall use cointegration only once, and build most error correction models in one step. You can guess (rightly) that we could not find a formulation fitting the required conditions, either because one of the elements was stationary, cointegration was not present, or the coefficients in the cointegrating equation were not significant or satisfactory.

We shall deal with the equations in sequence, considering two cases for the production function: complementary factors and Cobb-Douglas. The choice will affect other elements, as estimating a different capacity will affect the rate of use, an explanatory variable.

10.2.2.5.1 The complementary factors case

Let us first deal with the complementary factors case. The framework is very similar to the one we have seen earlier. Investment and employment are estimated separately.

10.2.2.5.2 Estimating investment

As in the small model, capacity will be defined by capital.

As to investment, it is given by:

$$I_t = K_t - K_{t-1} \cdot (1 - dr_t)$$

The investment equation will be:

$$I_t/K_{t-1} = a \cdot \frac{I_{t-1}}{K}_{t-2} + b \cdot \frac{Q_t}{Q_{t-1}} - c.(UR * - UR_t)/UR_t) + d$$

or in EViews terms:

```
equation _eq_i.ls(p) i/k(-1)=c_i(1)*i(-1)/k(-2)+c_i(2)*(q-q(-1))/q(-1)-c_i(3)*(@mean(ur)-ur)/ur+c_i(4) +i_ec
```

In our equation, accumulation rate will come from:

- Its previous value.
- The change of value added over the last period, representing the firms' expectations for future growth.
- The target change in the rate of use.

The results are rather satisfactory even though the lagged influence looks quite high, which can lead to problems in the period following a shock..

10.2.2.5.2.1 ESTIMATING EMPLOYMENT (COMPLEMENTARY FACTORS CASE)

Again, we shall use the same framework as in the small model. We start by defining the labor productivity trend. The same breakpoint tests applied to labor productivity leads to the two following breakpoint periods: 1972Q4 and 1992Q4.

Chow Breakpoint Test: 1972Q4 1992Q4
Null Hypothesis: No breaks at specified breakpoints
Varying regressors: All equation variables
Equation Sample: 1963Q1 2005Q4

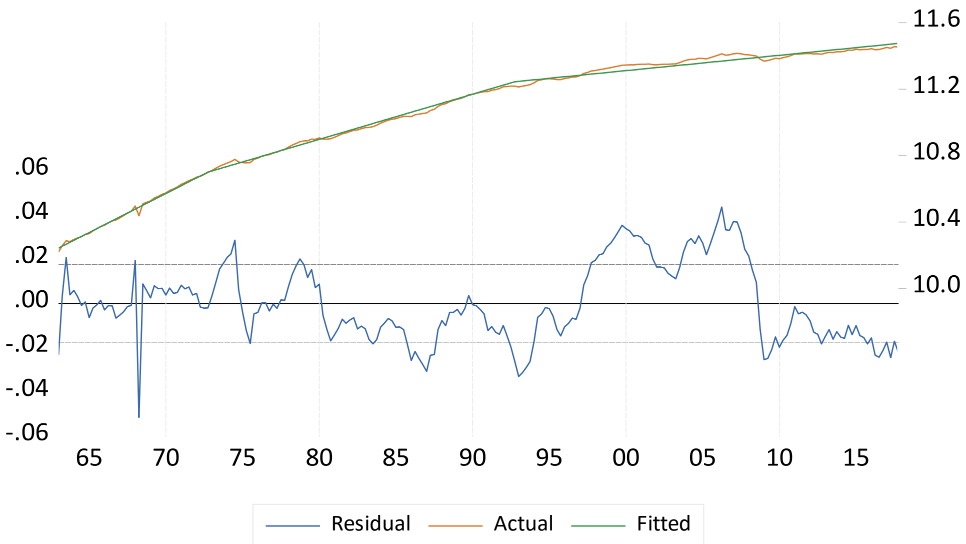
F-statistic	1491.695	Prob. F(4,166)	0.0000
Log likelihood ratio	620.8195	Prob. Chi-Square(4)	0.0000
Wald Statistic	5966.779	Prob. Chi-Square(4)	0.0000

Taking into account these breaks, estimating the trend in labor productivity gives:

Dependent Variable: LOG(Q/LF)
Method: Least Squares
Date: 04/02/20 Time: 17:58
Sample (adjusted): 1963Q1 2017Q4
Included observations: 220 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.46786	0.003141	3651.269	0.0000
T-2017	0.009231	0.000195	47.42078	0.0000
(T-1972.75)*(T<=1972.75)	0.019739	0.000826	23.89066	0.0000
(T-1992.75)*(T<=1992.75)	0.017956	0.000378	47.51469	0.0000

R-squared	0.997510	Mean dependent var	11.05890
Adjusted R-squared	0.997475	S.D. dependent var	0.348982
S.E. of regression	0.017536	Akaike info criterion	-5.231159
Sum squared resid	0.066419	Schwarz criterion	-5.169457
Log likelihood	579.4275	Hannan-Quinn criter.	-5.206242
F-statistic	28840.75	Durbin-Watson stat	0.232734
Prob(F-statistic)	0.000000		




```
smpl 1962Q1 2017Q4
coef(10) c_plt

_eq_plt.ls(p)LOG(Q/LF)= c_plt(1) + c_plt(2)*(t-2018)*(t<=2018)
 +4*log((1+txq)/(1+txn))*(t-2018)*(t>=2018)+c_plt(3)*(t-1972.75)*(t<=1972.75)
+c_plt(4)*((T<=1992.75)*(T-1992.75))+c_plt(5)*(t-2008.25)*(t<=2008.25)+c_plt(6)*(t>=2008.50)
  _eq_plt.resids(p)
 close _eq_plt
 genr log(lpt)=log(q/lf)-resid
 genr lfd=q/plt
```

To the two breaks which take place in 1972 and 1992 (last quarters in each case), we needed to add a third break in 2008Q2, and a dummy variable starting in the third quarter. All these additions are validated by the relevant tests.

As in the previous case, we observe that actual labor productivity is stationary around the trend.

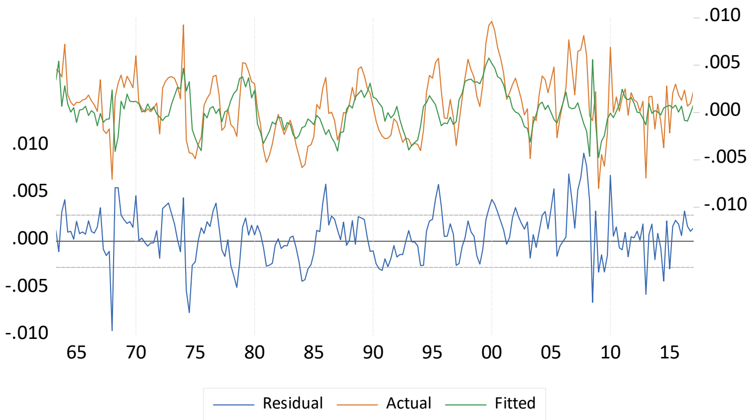
This gives us target employment, from which we build the error-correction model:

```
coef(10) c_lf
smpl 1962Q1 2017Q1
genr lf_ec=0
equation _eq_lf.ls(p) dlog(lf)=c_lf(1)*dlog(lfd)
   +c_lf(2)*log(lfd(-1)/lf(-1))+c_lf(3)*(t=1968.25)+c_lf(4)*(t=1968.50)+c_lf(5)*(t=1968)+lf_ec
_eq_lf.resids(p)
close _eq_lf
genr lf_ec=resid
```

Dependent Variable: DLOG(LF)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/02/20 Time: 22:37
Sample (adjusted): 1963Q2 2017Q1
Included observations: 216 after adjustments
DLOG(LF)=C_LF(1)*DLOG(LFD)+C_LF(2)*LOG(LFD(-1)/LF(-1))+C_LF(3)
*((T=1968.25)-(T=1968.50))

	Coefficient	Std. Error	t-Statistic	Prob.
C_LF(1)	0.211368	0.027205	7.769419	0.0000
C_LF(2)	0.151266	0.018871	8.015851	0.0000
C_LF(3)	0.007869	0.002699	2.915794	0.0039

R-squared	0.306256	Mean dependent var	0.000853
Adjusted R-squared	0.299742	S.D. dependent var	0.003294
S.E. of regression	0.002757	Akaike info criterion	-8.935719
Sum squared resid	0.001619	Schwarz criterion	-8.888840
Log likelihood	968.0577	Hannan-Quinn criter.	-8.916780
Durbin-Watson stat	1.056421		



As for the small model, we had to introduce a dummy variable, with opposite signs, for the middle quarters of 1968.

During the whole second quarter, a so-called "student revolution" actually blocked the economy, in particular transportation. GDP went down sharply, but employment did not really follow, as managers guessed (rightly) that this situation was transitory. This calls for a dummy for that quarter (with a positive coefficient). In the next quarter, things got back to normal, calling for a high growth in GDP, but employment was there to meet it.

One will also observe that we ended the estimation period in the first quarter of 2017. The evolution of employment was difficult to explain for the second and fourth quarters. And rather than introducing dummy variables (a difficult task at the end of the sample), we have elected to leave the management of residuals to the forecast step.

10.2.2.5.3 The Cobb-Douglas case

In the framework of a policy model, the complementary factors option is quite questionable and limited. But it is also quite simple to implement, and leads to easier interpretation of properties. Moving to a Cobb-Douglas specification might be simple at first sight. For instance we could consider modifying slightly the complementary factors framework, defining employment and capital separately based on a capacity target, and simply establishing the capacity using the Cobb-Douglas function.

But then we lose the most interesting property of this framework: taking into account the (endogenous) sensitivity of labor and capital factors to the ratio of their relative costs.

This means that a shock producing a 1% increase in target capacity would lead in the long run to the same relative increase in both factors, even if the same shock has modified the ratio of labor and capital costs.

We think this is too simplistic, and our formulas will take into account this effect. This will show for instance that a decrease in the social security contributions of firms will especially favor employment, and that the effect of a demand shock on unemployment and wages will favor capital.

The Cobb-Douglas assumption supposes a unitary elasticity of the share of factors to the relative cost.

However, using this assumption calls for developing a much more complex framework. Let us consider its elements in turn.

10.2.2.5.3.1 THE FRAMEWORK

We shall now describe the framework in detail.

10.2.2.5.3.1.1 MARGINS MAXIMIZATION

In this framework, firms will try to maximize their margins

(1) $pq_t \cdot Q_t - w\,cos\,t_t \cdot LE_t - k\,cos\,t_t \cdot K_{t-1}$

under constraint of the production function:

$$(2)\ Log(CAP_t) = \alpha \cdot Log(LE_t) + (1-\alpha) \cdot Log(K_{t-1}) + \beta \cdot t + \gamma$$

which leads us to maximizing

$$(3) pq_t \cdot Q_t - w\,cos\,t_t \cdot LE_t - k\,cos\,t_t \cdot K_{t-1} - \lambda \cdot (CAP_t - exp(\beta \cdot t + \gamma) \cdot LE_t^{\alpha} K_{t-1}^{1-\alpha})$$

relative to both LE_t and K_{t-1}.

Derivation of (3) gives:

$$-w\,cos\,t_t - \lambda \cdot \alpha \cdot exp(\beta \cdot t + \gamma) \cdot LE_t^{\alpha-1} K_{t-1}^{1-\alpha} = 0$$

Or

$$(4)\ -w\,cos\,t_t - \lambda \cdot \alpha \cdot exp(\beta \cdot t + \gamma) \cdot (K_{t-1}/LE_t)^{1-\alpha} = 0$$

and equivalently

$$(5)\ -k\,cos\,t_t - \lambda \cdot (1-\alpha) \cdot exp(\beta \cdot t + \gamma) \cdot (K_{t-1}/LE_t)^{\alpha} = 0$$

Dividing (4) by (5) on both sides gives:

$$(6) w\,cos\,t_t / k\,cos\,t_t = \alpha/(1-\alpha) \cdot (K_{t-1}/LE_t)$$

which shows indeed the unitary elasticity of the ratio of factors to the ratio of costs.

From (2) and (6) we get:

$$(7)\quad Log(LE_t/CAP_t) = (1-\alpha)\cdot Log(LE_t/K_{t-1}) - \beta\cdot t - \gamma$$
$$= (1-\alpha)\cdot Log(\alpha/(1-\alpha)k\cos t_t\,/w\cos t) - \beta\cdot t - \gamma$$

$$(8)\quad Log(K_t/CAP_t) = -\alpha\cdot Log(LE_t/K_{t-1}) - \beta\cdot t - \gamma$$
$$= -\alpha\cdot Log(\alpha/(1-\alpha)k\cos t_t\,/w\cos t) - \beta\cdot t - \gamma$$

But to apply this framework to a full model, we have to take into account several elements.

10.2.2.5.3.1.2 TARGETS AND ACTUAL VALUES

The above presentation applies to targets: knowing the relative costs, firms will estimate a target level of capacity, then the target levels of factors which allow reaching this capacity.

This means that we have to define this capacity, then compute the target level for both factors, then define the process leading from target values to actual values.

10.2.2.5.3.1.3 CAPACITY

For defining capacity, we shall suppose that firms have a target rate of use, constant over the period. This means that target capacity is proportional to production (or rather value added). In the estimation of equations (7) and (8), we will replace the ratio of factors to capacity by the ratio of factors to production, without loss of generality (the target rate of use will be absorbed by the constant term).

10.2.2.5.3.1.4 THE TIME FACTOR

In the above framework, we have used only instantaneous elements. But we have to consider the nature of our variables.

- Capital and capacity are measured at a given point in time
- Employment, in our definition, is an average level across one period (one quarter).

We shall suppose that target capital, employment (and implicit target capacity) are given by the system, using as target the actual level of production, but that actual capacity for the period is given by actual employment and the initial level of capital. As capital is measured as end-of-period, we shall use the lagged value.

10.2.2.5.3.1.5 THE INERTIA OF FACTORS

We shall also assess that optimal decisions are not implemented immediately. As explained before, the reasons are both technical (the length of the investing process) and psychological (risk aversion). So firms go only part of the way to the target, starting from the previous decision level. We shall try to estimate the inertia factor, allowing different values for labor and capital.

The first factor should be less inert, as the penalty for errors is lower (we have to consider annual wages compared to full cost of capital), and managing their consequences is also easier (laying down workers is easier than selling back unneeded equipment).

10.2.2.5.3.1.6 THE RELATIVE COST

Of course, it should compare the price of capital (actually investment[146]) to the price of labor (the wage rate). Actually, things are a little more complex.

- The wage rate should include social contributions.
- Once purchased, capital can be used as long as it is not destroyed or obsolete[147], whereas employment is bought for a single period[148].
- The price of capital should take into account the fact that it has to be purchased at once, whereas the alternate factor, labor, is paid for at the moment it is used, or even later. This delay should call for the introduction of the interest rate.
- The price of labor is measured in time spent. This means it is expected to increase in real terms with labor productivity.
- Capital depreciates over time. For workers, the efficiency generally increases then decreases, depending on the product. But firms can always replace older workers by new ones, at minimal cost (retirement financing is generally included in the wage cost).

We shall actually compare the yearly wage cost with the price of investment, equivalent to the price of capital at the cost of renewal. One has to consider that the increase in the efficiency of capital is included in the variable at constant prices, not in the deflator (this is called the « quality effect »). To spread the cost of capital over its period of use, we shall divide its deflator by an estimated factor, which should associate more or less with the number of periods of its productive life.

$$relc = (wf \cdot (1 + r\_scf)/pi) \cdot 1/((ir - \dot{pc}) - 4 \cdot Log(1 - rdep))$$

Some explanations are needed for the last term. The increase in the cost of capital for a given period is the sum of the interest rate and the annual depreciation:

$$z \;= (ir - \dot{pc}) + 4 \cdot Log(1 - rdep)$$

---

[146] Remember that the evolution of capital quality changes its value at constant prices, not its deflator, measured at a given quality (or efficiency). This means that the present deflator applies also to past values of capital, as the efficiency of a quantity is independent from the time at which it has been purchased. Applied to a given capital level, it will give the value of investment necessary to replace this capital, with the same efficiency.

[147] Maybe with a decreasing efficiency.

[148] This was different in the Roman Empire, when you could buy slaves. Then labor and capital had closer properties.

This means the ratio between the cost of labor and capital will give to the first an advantage of:

$$1/(1+z^k)$$

Or over the full period of

$$1/(1+z) + 1/(1+z)^2 + 1/(1+z)^3 + \ldots = 1/z$$

10.2.2.5.3.2 THE ESTIMATION

Let us see how we can apply this technique in our model, under EViews. We have seen that we have to estimate two equations, with common coefficients. This will be done through a system, a feature which we will have now the opportunity to present.

First let us compute the relative cost, as explained above:

```
smpl 1962Q1 2017Q4
genr relc=wr*(1+r_scf)/pi/(ir/100-@pchy(pc)-4*log(1-rdep))
genr k_ec=0
genr lf_ec=0
```

Now we have to estimate a system of two equations with common coefficients.

- As usual we create two vectors: coefficients and parameters.
- We destroy any preexisting system of the same name.
- We initialize the values with a guess.
- (this is important as convergence of this system is a little tricky under EViews)

```
coef(10) c_cd
vector(10) p_cd
delete cd
smpl 1962Q1 2017Q4
system cd
```

```
c_cd(4)=0.1
c_cd(1)=100
c_cd(3)=0.02 ' the yearly growth of total factor productivity: guessed at 2%
c_cd(2)=0.65 '  the share of employment in the process, guessed at 0.65[149]
p_cd(8)=0.0
p_cd(9)=0.0
cd.append log(k/q(-1))=-c_cd(3)*(t-2017)-c_cd(7)+c_cd(2)*log(relc)
cd.append log(lf/q(-1))=-c_cd(3)*(t-2017)-c_cd(1)+(c_cd(2)-1)*log(relc)
```

Two methods are available; Full Information Maximum Likelihood (FIML), and Seemingly Unrelated Regression (see Zellner (1962)). We test both but we chose the latter.

System: __CD
Estimation Method: Seemingly Unrelated Regression
Date: 04/03/20 Time: 11:59
Sample: 1980Q4 2017Q4
Included observations: 149
Total system (balanced) observations 298
Linear estimation after one-step weighting matrix

	Coefficient	Std. Error	t-Statistic	Prob.
C_CD(3)	0.003820	0.000672	5.683930	0.0000
C_CD(7)	4.539931	0.214654	21.14995	0.0000
C_CD(2)	0.462223	0.017133	26.97902	0.0000
C_CD(1)	4.962403	0.214680	23.11530	0.0000
Determinant residual covariance	4.62E-05			

Equation: LOG(K/Q)=-C_CD(3)*(T-2017)-C_CD(7)+C_CD(2)*LOG(RELC)
Observations: 149

R-squared	-4.729373	Mean dependent var	1.005121
Adjusted R-squared	-4.807857	S.D. dependent var	0.051824
S.E. of regression	0.124893	Sum squared resid	2.277333
Durbin-Watson stat	0.066958		

Equation: LOG(LF/Q)=-C_CD(3)*(T-2017)-C_CD(1)+(C_CD(2)-1)
*LOG(RELC)
Observations: 149

R-squared	0.260933	Mean dependent var	-11.26714
Adjusted R-squared	0.250809	S.D. dependent var	0.158847
S.E. of regression	0.137491	Sum squared resid	2.759949
Durbin-Watson stat	0.074370		

The estimation shows a 0.46 coefficient for labor (rather low compared to usual values) and a significant positive trend in global factor productivity.

We generate the estimated ("desired") series as target values.

[149] A figure also close to the share of the wage cost in the total.

```
smpl 1963Q4 2017Q4

genr log(kd/q(-1))=-c_cd(3)*(t-2017)-c_cd(7)+c_cd(2)*log(relc)

genr log(lfd/q(-1))=-c_cd(3)*(t-2017)-c_cd(1)+(c_cd(2)-1)*log(relc)
```

We estimate the actual results as weighted averages of the estimated result and the previous actual value, with an additional trend.

Actually we have fixed the inertia factors, at levels which give to the model reasonable properties.[150]

```
smpl  1970Q4 2017Q4

coef(10) c_k

vector(10) p_k

genr k_ec=0

coef(10) c_lf

vector(10) p_lf

genr lf_ec=0

p_lf(1)=0.20

p_k(1)=1.00

equation _eq_k.ls(p) log(k/q(-1))=p_k(1)*log(kd/q)+(1-p_k(1))*log(k(-1)/q(-1))+c_k(2)

 +c_k(3)*(t-2017)*(t<=2017)+k_ec

genr k_ec=resid

equation _eq_lf.ls(p) log(lf/q(-1))=p_lf(1)*log(lfd/q)+(1-p_lf(1))*log(lf(-1)/q(-1))+c_lf(2)

 +c_lf(3)*(t-2017)*(t<=2017)+lf_ec

genr lf_ec=resid
```

[150] OK, we do a little cheating here, please forgive us.

Finally we generate capacity applying the estimated formula to actual factor values, and we compute the rate of use of capacities.

```
genr log(cap)=c_cd(7)*(1-c_cd(2))+c_cd(1)*c_cd(2)

+c_cd(3)*(t-2017)*(t<=2017)+c_cd(2)*4*log((1+txq)/(1+txn))*(t-2017)*(t>2017)

+c_cd(2)*log(lf)+(1-c_cd(2))*log(k(-1))

genr ur=q/cap
```

We have to observe that this rate of use is different from the one we got with the complementary factors function. This means that the equations in which it enters will have to be estimated again: value added deflator, exports and imports.

10.2.2.5.4 The change in inventories

We shall use the same explanation as the simple model, simplifying the formula by using the present growth rate of value added.

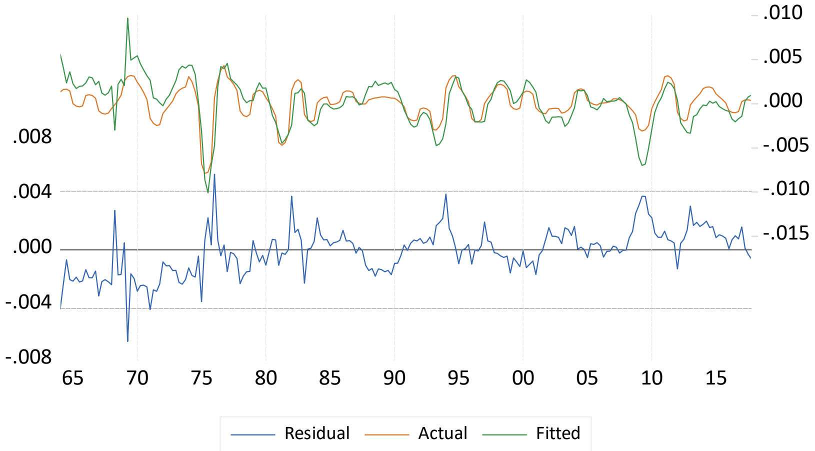
The unusual profile of the dependent variable is due to the use of an original yearly series.

```
coef(10) c_ci

smpl 1962Q1 2017Q4

genr ci_ec=0

smpl 1963Q1 2017Q4


equation _eq_ic.ls(p) ic/q(-1)=c_ic(1)*c_ic(2)*@pchy(q)+c_ic(2)+(1-c_ic(1))*ic(-1)/q(-2)+c_ic(3)+ic_ec

_eq_ci.resids(p)

close _eq_ci

genr ci_ec=resid
```

Dependent Variable: IC/Q(-1)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/03/20 Time: 16:18
Sample (adjusted): 1964Q1 2017Q4
Included observations: 216 after adjustments
Convergence achieved after 5 iterations
Coefficient covariance computed using outer product of gradients
IC/Q(-1)=C_IC(1)*C_IC(2)*@PCHY(Q)+C_IC(2)+(1-C_IC(1))*IC(-1)/Q(-2)
+C_IC(3)+IC_EC

	Coefficient	Std. Error	t-Statistic	Prob.
C_IC(1)	0.281439	0.041693	6.750247	0.0000
C_IC(2)	0.065153	0.011786	5.528009	0.0000
C_IC(3)	-0.065597	0.011851	-5.535271	0.0000

R-squared	0.706556	Mean dependent var	0.000118
Adjusted R-squared	0.703801	S.D. dependent var	0.001890
S.E. of regression	0.001029	Akaike info criterion	-10.90758
Sum squared resid	0.000225	Schwarz criterion	-10.86070
Log likelihood	1181.019	Hannan-Quinn criter.	-10.88864
F-statistic	256.4312	Durbin-Watson stat	1.105253
Prob(F-statistic)	0.000000		



10.2.2.5.5 Unemployment

This is a new equation compared to the small model. Actually it seemed clearer to us to model the work force (employment + unemployment). As it depends on employment, the quality of estimation will be exactly the same (only the R-squared will change). The coefficient of the work force POPAC will be higher by 1 compared to a formulation using unemployment UN.

Otherwise, the equation follows fully the framework defined above.

The short-term sensitivities are reasonable, and the long-term ones too, with a slightly lower value. We had to decide on the rate of convergence to the long-term relation.

It is logical to expect that a permanent improvement of the labor market will attract more and more job seekers.

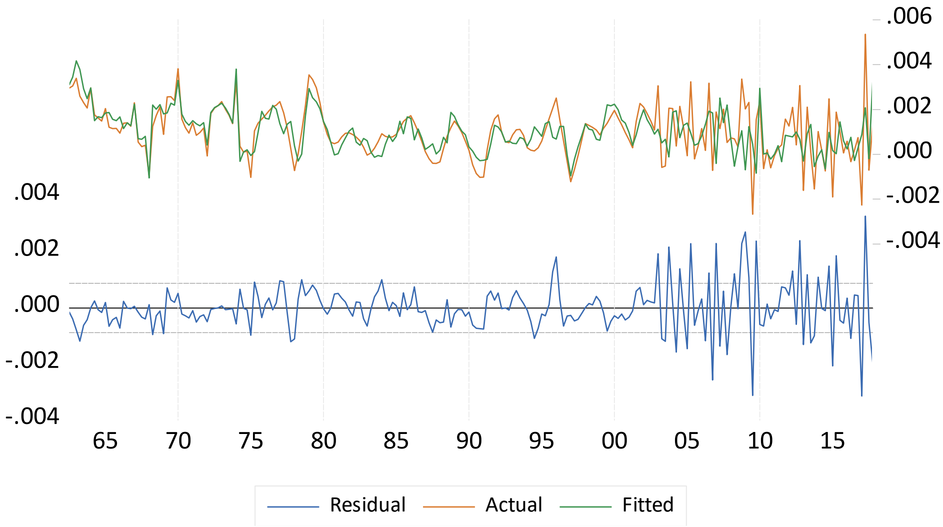
However, the erratic evolution of residuals in the later period is rather disquieting.

```
coef(10) c_popac

 p_popac(3)=0.50

 smpl 1962Q1 2017Q4

 genr popac_ec=0

 d(popac)/pop65(-1)=c_popac(1)*d(lt)/pop65(-1)+c_popac(2)*d(pop65)/pop65(-1)

 -p_popac(3)*(popac(-1)/pop65(-1)-c_popac(4)*lt(-1)/pop65(-1)-c_popac(5))+[ar(1)=c_popac(6)])

 +popac_ec

 _eq_popac.resids(p)

 genr popac_ec=resid

 close _eq_popac
```

Dependent Variable: D(POPAC)/POP65(-1)
Method: ARMA Conditional Least Squares (Marquardt - EViews legacy)
Date: 04/03/20 Time: 16:55
Sample (adjusted): 1962Q3 2017Q4
Included observations: 222 after adjustments
Convergence achieved after 8 iterations
D(POPAC)/POP65(-1)=C_POPAC(1)*D(LT)/POP65(-1)+C_POPAC(2)
D(POP65)/POP65(-1)-0.5(POPAC(-1)/POP65(-1)-C_POPAC(4)*LT(
-1)/POP65(-1)-C_POPAC(5))+[AR(1)=C_POPAC(6)]

	Coefficient	Std. Error	t-Statistic	Prob.
C_POPAC(1)	0.608286	0.048134	12.63745	0.0000
C_POPAC(2)	0.183219	0.086401	2.120565	0.0351
C_POPAC(4)	0.446176	0.093014	4.796890	0.0000
C_POPAC(5)	0.357420	0.051223	6.977662	0.0000
C_POPAC(6)	0.977937	0.014378	68.01539	0.0000

R-squared	0.423523	Mean dependent var	0.001048
Adjusted R-squared	0.412896	S.D. dependent var	0.001146
S.E. of regression	0.000878	Akaike info criterion	-11.21481
Sum squared resid	0.000167	Schwarz criterion	-11.13817
Log likelihood	1249.844	Hannan-Quinn criter.	-11.18387
F-statistic	39.85605	Durbin-Watson stat	2.427417
Prob(F-statistic)	0.000000		
Inverted AR Roots	.98		



10.2.2.5.6 The value added deflator

The formula will change according to the type of production function: the measure of the rate of use will be different, and in the Cobb Douglas case the cost will include amortization of capital.

In both cases the equation follows an error correction format: a long-term relationship between the rate of use and the margins rate (or rather the share of the cost in value added), and a dynamic equation freeing the elasticity of the deflator to the cost.

In the Cobb-Douglas case, we have to consider the total cost: wage and capital.

```
genr cost=(wr*lf*(1+r_scf)+0.05*pi*k(-1))/q
```

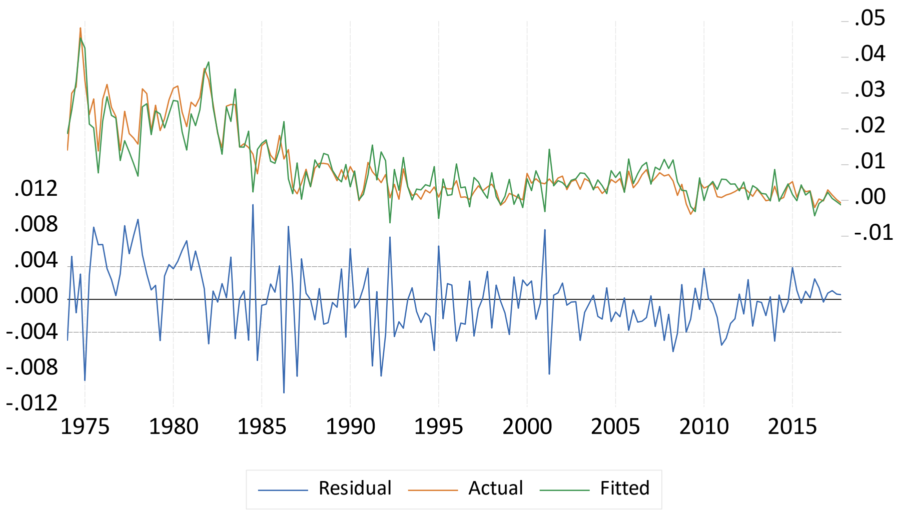
We distribute the cost of capital over 20 periods (5 years). Reducing this value would not change much the properties.

The coefficients are rather different from the previous case: stronger dynamic effect, slower correction.

Dependent Variable: DLOG(PQ)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/03/20 Time: 17:57
Sample: 1974Q1 2017Q4
Included observations: 176
Convergence achieved after 5 iterations
Coefficient covariance computed using outer product of gradients
DLOG(PQ)=C_PQ(1)*DLOG(COST)+C_PQ(2)*DLOG(UR)+C_PQ(3)
*(LOG(PQ(-1)/COST(-1))-C_PQ(2)*LOG(UR(-1)))+C_PQ(5)+PQ_EC

	Coefficient	Std. Error	t-Statistic	Prob.
C_PQ(1)	0.652817	0.023028	28.34928	0.0000
C_PQ(2)	0.365519	0.061221	5.970488	0.0000
C_PQ(3)	-0.018863	0.006386	-2.953605	0.0036
C_PQ(5)	0.026904	0.008276	3.250874	0.0014

R-squared	0.869624	Mean dependent var	0.009297
Adjusted R-squared	0.867350	S.D. dependent var	0.010069
S.E. of regression	0.003667	Akaike info criterion	-8.356375
Sum squared resid	0.002313	Schwarz criterion	-8.284318
Log likelihood	739.3610	Hannan-Quinn criter.	-8.327149
F-statistic	382.4193	Durbin-Watson stat	1.976983
Prob(F-statistic)	0.000000		



An alternate version uses cointegration.

We test the cointegration between the rate of use and the margins (as the ratio between the value added deflator and the cost). It reflects the fact that facing a low rate of use firms will reduce their margins to increase demand.

It works.

Date: 04/03/20 Time: 17:23
Sample (adjusted): 1972Q2 2017Q4
Included observations: 183 after adjustments
Trend assumption: Linear deterministic trend (restricted)
Series: LOG(PQ/COST) LOG(UR)
Lags interval (in first differences): 1 to 4

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.100608	29.59230	25.87211	0.0164
At most 1	0.054149	10.18772	12.51798	0.1189

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.100608	19.40458	19.38704	0.0497
At most 1	0.054149	10.18772	12.51798	0.1189

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b'*S11*b=I):

LOG(PQ/COST)	LOG(UR)	@TREND(62Q2)
19.03573	-32.15614	-0.003083
-0.295312	66.21976	-0.001362

Unrestricted Adjustment Coefficients (alpha):

D(LOG(PQ/...	-0.001096	-0.001306
D(LOG(UR))	0.000840	-0.000750

1 Cointegrating Equation(s): Log likelihood 1422.333

Normalized cointegrating coefficients (standard error in parentheses)

LOG(PQ/COST)	LOG(UR)	@TREND(62Q2)
1.000000	-1.689252	-0.000162
	(0.79135)	(0.00025)

Adjustment coefficients (standard error in parentheses)

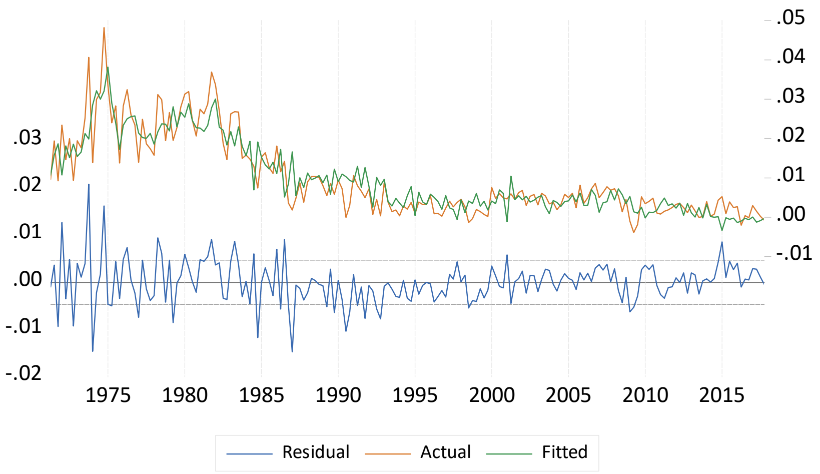
D(LOG(PQ/...	-0.020863
	(0.00940)
D(LOG(UR))	0.015996
	(0.00592)

The dynamic equation works also, with a low correcting coefficient.

Dependent Variable: DLOG(PQ)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/03/20 Time: 17:38
Sample (adjusted): 1971Q2 2017Q4
Included observations: 187 after adjustments
DLOG(PQ)=C_PQ(1)*DLOG(COST)+C_PQ(2)*DLOG(UR)+C_PQ(3)
RES(-1)+C_PQ(5)+C_PQ(6)(T-2017)*(T<=2017)

	Coefficient	Std. Error	t-Statistic	Prob.
C_PQ(1)	0.462592	0.036840	12.55669	0.0000
C_PQ(2)	0.227275	0.080422	2.826041	0.0052
C_PQ(3)	-0.022939	0.006275	-3.655477	0.0003
C_PQ(5)	0.028252	0.008094	3.490630	0.0006
C_PQ(6)	-0.000304	3.77E-05	-8.067195	0.0000

R-squared	0.796039	Mean dependent var	0.009865
Adjusted R-squared	0.791557	S.D. dependent var	0.010248
S.E. of regression	0.004679	Akaike info criterion	-7.865245
Sum squared resid	0.003984	Schwarz criterion	-7.778851
Log likelihood	740.4004	Hannan-Quinn criter.	-7.830238
F-statistic	177.5822	Durbin-Watson stat	2.087218
Prob(F-statistic)	0.000000		



10.2.2.5.7 The wage rate

The wage rate estimation is the only occasion which allows us to use cointegration, and thus to conform to the accepted principles of econometrics.

The estimation process is exactly the same as for the small model.

We suppose that the wage cost is indexed in the long run: for 50% on the value added deflator (the firms want to stabilize the share of wages in value added) and for 50% on the consumption deflator (workers want to their purchasing power to follow the gains in productivity). The weighting could be adapted through a parameter.

So first we test that the stationarity of _luwc = log(uwc/(0.5*pq+0.5*pc))

```
genr _luwc=log(uwc)-0.5*log(pq)-(1-0.5)*log(pcoh)

uroot(p) _luwc
```

Then that of the unemployment rate itself.

```
uroot(p) unr
```

Both fail.

Then we test the cointegration of the two elements, which works.

```
coint(b,p) _luwc unr
```

Date: 04/03/20 Time: 18:08
Sample (adjusted): 1973Q4 2017Q4
Included observations: 177 after adjustments
Trend assumption: No deterministic trend (restricted constant)
Series: _LUWC UNR
Lags interval (in first differences): 1 to 6

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.100436	23.92640	20.26184	0.0149
At most 1	0.028906	5.191742	9.164546	0.2631

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.100436	18.73466	15.89210	0.0174
At most 1	0.028906	5.191742	9.164546	0.2631

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b'*S11*b=I):

_LUWC	UNR	C
-12.04745	-55.71775	8.626833
15.27940	-4.114727	-4.265647

Unrestricted Adjustment Coefficients (alpha):

D(_LUWC)	0.002981	-0.000523
D(UNR)	0.000222	0.000250

1 Cointegrating Equation(s): Log likelihood 1461.761

Normalized cointegrating coefficients (standard error in parentheses)

_LUWC	UNR	C
1.000000	4.624858	-0.716071
	(0.96386)	(0.08071)

Adjustment coefficients (standard error in parentheses)

D(_LUWC)	-0.035914
	(0.00888)
D(UNR)	-0.002676
	(0.00152)

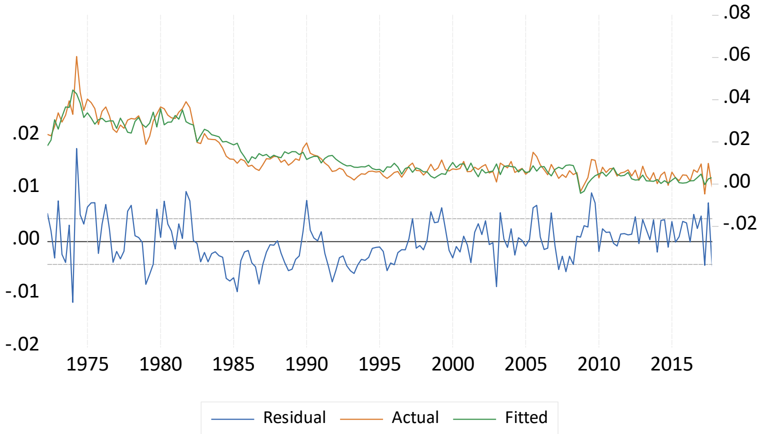
The coefficient for UNR in the cointegrating equation is significant with the right sign. It remains to be seen if this equation gives good properties to the full model.

For the dynamic equation, we had to set the indexation as globally unitary (dynamic homogeneity) with a lag structure. All the elements are quite significant, but the error correction is quite slow (a result we found in other circumstances for the same equation).

Dependent Variable: DLOG(WR)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/04/20   Time: 20:24
Sample (adjusted): 1972Q2 2017Q4
Included observations: 183 after adjustments
DLOG(WR)=C_WR(1)*DLOG(PC)+C_WR(2)*DLOG(LPROD)+C_WR(4)
    RES_WR(-1)+C_WR(5)(T-2018)*(T-2018)+C_WR(6)

	Coefficient	Std. Error	t-Statistic	Prob.
C_WR(1)	0.598216	0.055015	10.87363	0.0000
C_WR(2)	0.151471	0.062906	2.407877	0.0171
C_WR(4)	-0.014572	0.003991	-3.650994	0.0003
C_WR(5)	7.51E-06	9.67E-07	7.764355	0.0000
C_WR(6)	-0.001007	0.000909	-1.108496	0.2691

R-squared	0.865434	Mean dependent var	0.013277
Adjusted R-squared	0.862410	S.D. dependent var	0.011659
S.E. of regression	0.004325	Akaike info criterion	-8.021948
Sum squared resid	0.003329	Schwarz criterion	-7.934257
Log likelihood	739.0082	Hannan-Quinn criter.	-7.986402
F-statistic	286.1925	Durbin-Watson stat	1.269681
Prob(F-statistic)	0.000000		



10.2.2.5.8 The trade deflators

For this equation we will not use an error correction format, rather a homogenous indexation. This means we shall not separate the short and long-term sensitivities.

As usual for this type of equation:

- Exporters show a higher attention to their costs than to the price of their competitors. This is particular true for the exporters to France, as the role of French prices is not completely significant.

This will dampen the dynamics of the price-wage loop. The impact of trade on the price of demand can be explained in the following way.

- Imports are a share of global demand. They are bought at the import price. The higher its sensitivity to foreign costs, the higher the difference to the local production price, and the higher the reducing impact of imports on the global demand price.
- If local producers decided on their selling price on the local and foreign markets independently (a possible behavior that we did not consider), the sensitivity of the demand price to local costs would clearly be less than one.
- But in our framework, the production price is decided globally, and the lower sensitivity of the export price has to be balanced by a higher sensitivity of the price at which they sell on the local market. The higher the impact of local costs on the export price, the lower the necessary compensation.

One can see that in the transition from production to demand price, the higher the role of the production cost in the price set by the exporter, the higher the first (negative) effect and the lower the second (positive) one.[151]

In the extreme, if all exporters take only into account their costs, the import price will not be affected, and as the export price will change just as the global production price, no compensation will be needed. The damping effect will be maximal.

- The additional trend is negative and quite significant. It probably represents a structural but permanent shift in traded goods to the ones which present the lowest price.

To reach a steady state in the long run, these trends will have to be suppressed after a while. Here we did it immediately, but true forecasts should call for a gradual decrease.

```
 coef(10) c_px
 smpl 1962Q1 2017Q4
 genr px_ec=0
 smpl 1962Q1 2017Q4
 equation _eq_px.ls(p) log(px)=c_px(4)*log(pp)+(1-c_px(4))*log(ppx*er)
+c_px(6)+c_px(5)*(t-2017)*(t<=2017)+[ar(1)=c_px(7)]+px_ec
_eq_px.resids(p)
close _eq_px
```

```
coef(10) c_pm
smpl 1962Q1 2017Q4
genr pm_ec=0
smpl 1962Q1 2017Q4
equation _eq_pm.ls(p) log(pm)=c_pm(4)*log(pp)+(1-c_pm(4))*log(ppx*er)
  +c_pm(6)+c_pm(5)*(t-2017)*(t<=2017)+[ar(1)=c_pm(7)]+pm_ec
_eq_pm.resids(p)
close _eq_pm
```

[151] This could be formalized easily, but we hope the message is already clear.

Dependent Variable: DLOG(PX)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/03/20 Time: 19:40
Sample (adjusted): 1975Q2 2017Q4
Included observations: 171 after adjustments
Convergence achieved after 3 iterations
Coefficient covariance computed using outer product of gradients
DLOG(PX)=0.7*DLOG(PP)+C_PX(2)*DLOG(PPX*ER)+C_PX(3)
*(LOG(PX(-1))-C_PX(4)*LOG(PP(-1))-(1-C_PX(4))*LOG(PPX(-1)*ER(
-1)))+C_PX(6)+C_PX(5)*(T-2017)*(T<=2017)

	Coefficient	Std. Error	t-Statistic	Prob.
C_PX(2)	0.203108	0.038859	5.226755	0.0000
C_PX(3)	-0.089478	0.030611	-2.923041	0.0040
C_PX(4)	0.635238	0.126448	5.023700	0.0000
C_PX(6)	-0.006800	0.003914	-1.737285	0.0842
C_PX(5)	-0.000839	0.000399	-2.106586	0.0367

R-squared	0.195857	Mean dependent var	0.005023
Adjusted R-squared	0.176480	S.D. dependent var	0.022168
S.E. of regression	0.020117	Akaike info criterion	-4.945705
Sum squared resid	0.067179	Schwarz criterion	-4.853843
Log likelihood	427.8577	Hannan-Quinn criter.	-4.908431
F-statistic	10.10771	Durbin-Watson stat	1.816121
Prob(F-statistic)	0.000000		

Dependent Variable: DLOG(PM)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/03/20 Time: 19:43
Sample (adjusted): 1975Q2 2017Q4
Included observations: 171 after adjustments
Convergence achieved after 4 iterations
Coefficient covariance computed using outer product of gradients
DLOG(PM)=C_PM(1)*DLOG(PP)+C_PM(2)*DLOG(PPX*ER)+C_PM(3)
*(LOG(PM(-1))-C_PM(4)*LOG(PP(-1))-(1-C_PM(4))*LOG(PPX(-1)
ER(-1)))+C_PM(6)+C_PM(5)(T-2017)*(T<=2017)

	Coefficient	Std. Error	t-Statistic	Prob.
C_PM(1)	0.726492	0.187994	3.864435	0.0002
C_PM(2)	0.266741	0.030791	8.663057	0.0000
C_PM(3)	-0.078071	0.037872	-2.061446	0.0408
C_PM(4)	0.625968	0.113616	5.509510	0.0000
C_PM(6)	-0.005655	0.002856	-1.980227	0.0493
C_PM(5)	-0.000673	0.000315	-2.136540	0.0341

R-squared	0.465638	Mean dependent var	0.005416
Adjusted R-squared	0.449445	S.D. dependent var	0.020259
S.E. of regression	0.015032	Akaike info criterion	-5.522834
Sum squared resid	0.037283	Schwarz criterion	-5.412601
Log likelihood	478.2023	Hannan-Quinn criter.	-5.478106
F-statistic	28.75589	Durbin-Watson stat	1.562780
Prob(F-statistic)	0.000000		

ls dlog(px)=c_px(1)*dlog(pp)+c_px(2)*dlog(ppx*er)+c_px(3)*(log(px(-1))-c_px(4)*log(pp(-1))-(1-c_px(4))*log(ppx(-1)*er(-1)))+c_px(6)+c_px(5)*(t-2017)*(t<=2017)

ls dlog(pm)=c_pm(1)*dlog(pp)+c_pm(2)*dlog(ppx*er)+c_pm(3)*(log(pm(-1))-c_pm(4)*log(pp(-1))-(1-c_pm(4))*log(ppx(-1)*er(-1)))+c_pm(6)+c_pm(5)*(t-2017)*(t<=2017)

10.2.2.5.9 Household consumption

Our equation follows as usual an error correction specification (estimated in one step!) following almost completely the framework presented earlier.

The change in consumption depends on:

- The change in real income (over the last year).
- The change in unemployment.
- Inflation (the « real holdings » effect).
- An error correction term.
- A negative time trend, representing the increase in the wealth of households and their saving potential (in particular their accession to housing ownership).

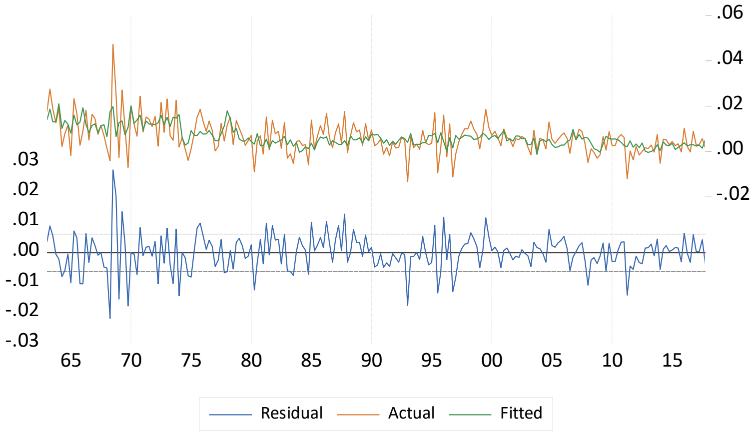
The only influence we could not evidence is that of the real short-term interest rate.

```
 coef(10) c_coh
smpl 1962Q1 2022Q4
genr coh_ec=0
  smpl 1962Q1 2017Q4
  equation _eq_coh.ls(p) dlog(coh)=c_coh(1)*.25*log(hrdi/hrdi(-4))+c_coh(2)*dlog(unr)
  +c_coh(3)*log(pc/pc(-4))+c_coh(5)*dlog(coh(-1))+c_coh(6)+c_coh(7)*log(coh(-1)/hrdi(-1))
  +c_coh(8)*(t-2017)*(t<=2017)+coh_ec
  genr coh_ec=resid
```

Dependent Variable: DLOG(COH)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/03/20 Time: 12:39
Sample (adjusted): 1963Q1 2017Q4
Included observations: 220 after adjustments
DLOG(COH)=C_COH(1)*.25*LOG(HRDI/HRDI(-4))+C_COH(2)
*DLOG(UNR)+C_COH(3)*LOG(PC/PC(-4))+C_COH(5)*DLOG(COH(
-1))+C_COH(6)+C_COH(7)*LOG(COH(-1)/HRDI(-1))+C_COH(8)*(T
-2017)*(T<=2017)+COH_EC

	Coefficient	Std. Error	t-Statistic	Prob.
C_COH(1)	0.417154	0.128560	3.244825	0.0014
C_COH(2)	-0.048968	0.011362	-4.309732	0.0000
C_COH(3)	-0.035504	0.017628	-2.014029	0.0453
C_COH(5)	-0.213361	0.066355	-3.215459	0.0015
C_COH(6)	-0.304138	0.098582	-3.085123	0.0023
C_COH(7)	-0.064356	0.020801	-3.093930	0.0022
C_COH(8)	-0.000153	4.79E-05	-3.195567	0.0016

R-squared	0.342567	Mean dependent var	0.006449
Adjusted R-squared	0.324048	S.D. dependent var	0.007509
S.E. of regression	0.006173	Akaike info criterion	-7.305809
Sum squared resid	0.008118	Schwarz criterion	-7.197830
Log likelihood	810.6390	Hannan-Quinn criter.	-7.262204
F-statistic	18.49791	Durbin-Watson stat	2.091141
Prob(F-statistic)	0.000000		



10.2.2.5.10 Exports

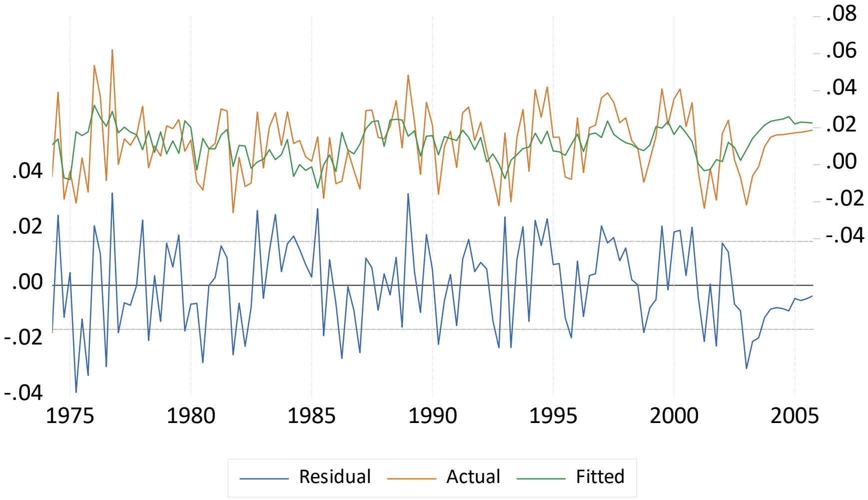
For exports we shall use again an error correction framework, estimated in one pass.

- The substitution effect appears through the average of the rate of use over the last two quarters (the dynamic and long-term effects are not otherwise separated). Price competitiveness uses the same technique.
- A significant trend had to be introduced and will be stopped in the future. It is possible that the measure of world demand is biased.

Dependent Variable: DLOG(X)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/03/20 Time: 20:48
Sample (adjusted): 1974Q2 2005Q4
Included observations: 127 after adjustments
Convergence achieved after 4 iterations
Coefficient covariance computed using outer product of gradients
DLOG(X)=C_X(1)*DLOG(WD)+C_X(2)*LOG(X(-1)/WD(-1))+C_X(5)
+C_X(6)*T+[C_X(3)=AR(1)]

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.421293	0.084847	4.965310	0.0000
C_X(2)	-0.158407	0.041881	-3.782334	0.0002
C_X(5)	2.257588	0.689654	3.273506	0.0014
C_X(6)	-0.001118	0.000343	-3.256103	0.0015

R-squared	0.247388	Mean dependent var	0.011818
Adjusted R-squared	0.229031	S.D. dependent var	0.018045
S.E. of regression	0.015844	Akaike info criterion	-5.421050
Sum squared resid	0.030877	Schwarz criterion	-5.331469
Log likelihood	348.2367	Hannan-Quinn criter.	-5.384654
F-statistic	13.47692	Durbin-Watson stat	2.002881
Prob(F-statistic)	0.000000		



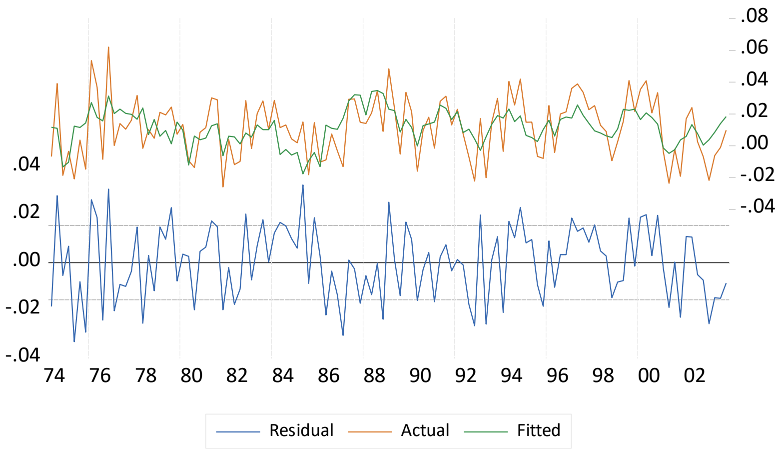

```
genr x_ec=0
smpl 1962Q1 2003Q4
equation _eq_x.ls(p) dlog(x)=c_x(1)*dlog(wd)+c_x(2)*log(x(-1)/wd(-1))
+c_x(3)*0.5*(log(ur)+log(ur(-1)))+c_x(4)*0.5*(log(compx(-1))+log(compx))+c_x(5)
+c_x(6)*(t-2005)*(t<=2005)+x_ec
_eq_x.resids(p)
 close _eq_x
```

```
genr x_ec=resid
```

Dependent Variable: DLOG(X)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 04/03/20 Time: 20:51
Sample (adjusted): 1974Q2 2003Q4
Included observations: 119 after adjustments
DLOG(X)=C_X(1)*DLOG(WD)+C_X(2)*LOG(X(-1)/WD(-1))+C_X(3)*0.5
*(LOG(UR)+LOG(UR(-1)))+C_X(4)*0.5*(LOG(COMPX(-1))
+LOG(COMPX))+C_X(5)+C_X(6)*(T-2005)*(T<=2005)

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.415130	0.083749	4.956817	0.0000
C_X(2)	-0.201800	0.048941	-4.123297	0.0001
C_X(3)	-0.254627	0.121300	-2.099154	0.0380
C_X(4)	-0.202750	0.060155	-3.370481	0.0010
C_X(5)	-0.038971	0.018205	-2.140639	0.0345
C_X(6)	-0.007775	0.001966	-3.953931	0.0001

R-squared	0.326431	Mean dependent var	0.011474
Adjusted R-squared	0.296627	S.D. dependent var	0.018593
S.E. of regression	0.015593	Akaike info criterion	-5.434828
Sum squared resid	0.027477	Schwarz criterion	-5.294704
Log likelihood	329.3723	Hannan-Quinn criter.	-5.377928
F-statistic	10.95263	Durbin-Watson stat	2.071445
Prob(F-statistic)	0.000000		



10.2.2.5.11 Imports

For imports we use the same framework, but:

- We had to set the dynamic demand coefficient to unity.
- The rate of use is not lagged.
- Price competitiveness is measured over 6 periods.

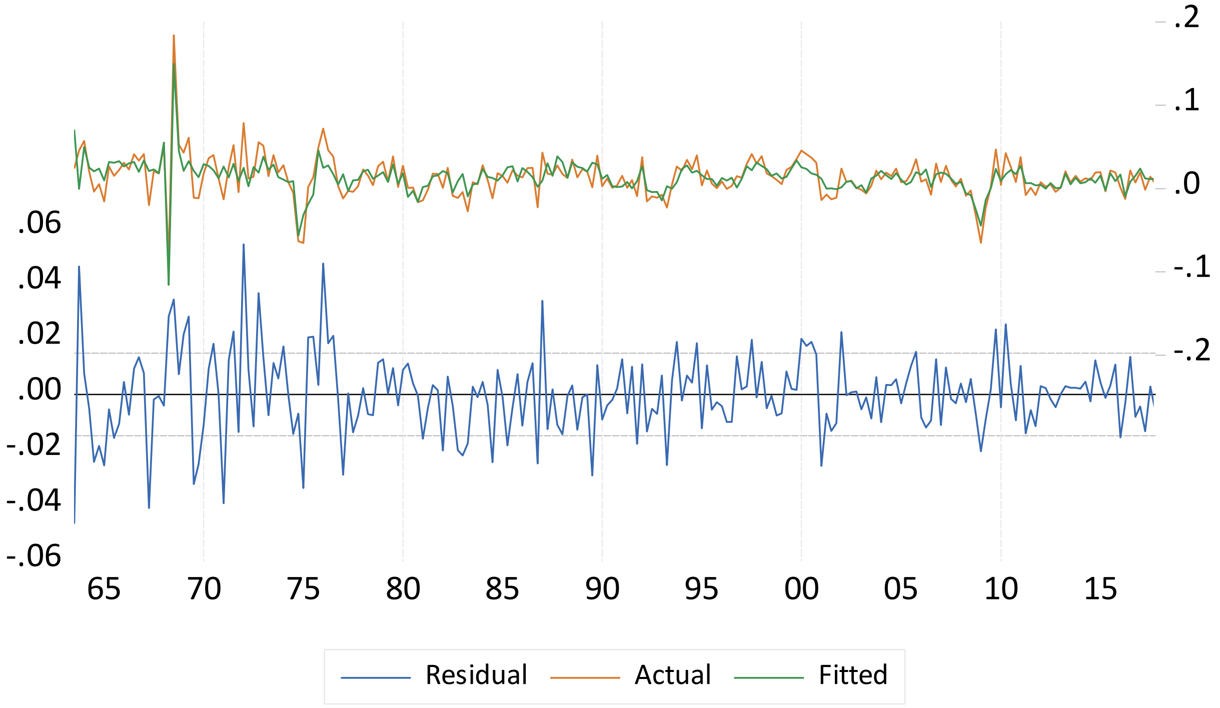
```
'    estimating imports
'show err_m
coef(10) c_m
smpl 1962Q1 2022Q4
genr m_ec=0
smpl 1962Q1 2017Q4
equation _eq_m.ls(p) dlog(m)=dlog(fd+tc*q)+c_m(2)*log(ur)+c_m(3)*dlog(compm)
   +c_m(5)+c_m(6)*(t-2017)*(t<=2017)+[ar(1)=c_m(7)]+c_m(8)*log(m(-1)/(fd(-1)+tc*q(-1)))+m_ec
genr m_ec=resid
```

Dependent Variable: DLOG(M)
Method: ARMA Conditional Least Squares (Marquardt - EViews legacy)
Date: 04/03/20   Time: 21:54
Sample (adjusted): 1963Q3 2017Q4
Included observations: 218 after adjustments
Convergence achieved after 8 iterations
DLOG(M)=C_M(1)*DLOG(FD+TC*Q)+C_M(2)*LOG(UR)+C_M(3)
    DLOG(COMPM)+C_M(5)+C_M(6)(T-2017)*(T<=2017)
    +[AR(1)=C_M(7)]+C_M(8)*LOG(M(-1)/(FD(-1)+TC*Q(-1)))

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(1)	2.154307	0.116071	18.56025	0.0000
C_M(2)	-0.020114	0.045942	-0.437819	0.6620
C_M(3)	-0.103709	0.042707	-2.428405	0.0160
C_M(5)	-0.046274	0.029940	-1.545553	0.1237
C_M(6)	0.000941	0.000504	1.865146	0.0635
C_M(8)	-0.029705	0.018188	-1.633248	0.1039
C_M(7)	0.096754	0.073690	1.312989	0.1906

R-squared	0.684084	Mean dependent var	0.013658
Adjusted R-squared	0.675101	S.D. dependent var	0.024959
S.E. of regression	0.014227	Akaike info criterion	-5.635817
Sum squared resid	0.042706	Schwarz criterion	-5.527140
Log likelihood	621.3040	Hannan-Quinn criter.	-5.591921
Durbin-Watson stat	1.961930		

Inverted AR Roots	.10



10.2.2.6 Checking again model – data consistency

Now that equations have been estimated, we should check again data-equations consistency, this time on all equations.

The identities should still hold true, and adding the estimation residual at the end of the behavioral equations should make them hold true, too.

So this process could look unnecessary. This will generally be true if all estimations have been repeated immediately before the check, which means that they have been grouped, along with the definition of identities, in the same program. This is actually the best technique, as it allows producing the full model through a single program, and not through a sequence (which needs some organization and can lead to errors).

But if the estimates somewhat dated, and if one is not sure that the data have not been modified since, errors might have appeared.

And in any case, this check is almost free (the control that all the residuals are negligible can be done with a program, with a single answer) so there is no reason not to do it.

Of course, our example gives acceptable results[152].

[152] But achieving this result took some time…

```
smpl 1995Q2 2017Q4

'     Now we compute the residuals
'      for all the endogenous
'
'      The process is the same
'
'      The dc_{%1} are the absolute differences
'      The pc_{%1} the relative ones
'      To avoid dividing by zero, we use trick 1
'      We add to the variable a boolean one
'      Testing if the variable is zero
'      If it is true we get 0 divided by 1 = 0
'      If not the computation is not affected
'
'      This means we have just modified the outcome of divisions.
'      Dividing zero by zero now gives zero

for !i=1 to _g_vendo.@count
%1=_g_viden.@seriesname(!i)
genr dc_{%1}={%1}-{%1}_c
genr pc_{%1}=100*dc_{%1}/({%1}+({%1}=0))
next
```

10.2.2.7 Solving the model on the future

We are coming now to the crucial part of model testing, observing how it performs on the field where it will be used: the future.

Globally, the technique is the same as explained earlier. The differences are minimal.

We start with the usual statements: changing the directory, creating a special expanded file called “proj_1”.

```
cd "d:\eviews\__fra_2020"



'     we set the workfile to proj_1, quarterly from 1962 to 2100
'

close proj_1.wf1
close mod_1.wf1
open mod_1.wf1
save proj_1
pagestruct(end=2100Q4) *
```

Now we are using deflators. For the long-term assumptions, we need to include their growth rate:

```
' 	we define long-term growth rates

scalar txq=exp(0.005)-1
scalar txn=exp(0.002)-1
scalar txp=exp(0.006)-1
```

Obviously, the set of exogenous variables, and of assumptions to define, is larger.

This program simulates the model over the future

```
' we expand series for projection: EXOGENOUS variables
	' constant expansion

for %1 COH_EC ERX FDXR  K_EC CI_EC IR_ER IRL_ER IRMX IRSR IRST IRSX IRX LF_EC  M_EC NIF_ER NIG_ER PK PM_EC POPAC_EC PQ_EC PX_EC R_EXPG R_ICT R_IFP R_OIT R_PCOH R_PCOG R_PI R_PIG R_REVG R_REVQ R_SCF R_SCG R_SCW R_SUBS R_TAR R_TARX R_VAT RDEP TC URD  WR_EC X_EC IRM_ER relax_q   relax_pfd fcapf_er prof_er smpl 2005Q1 2100Q4
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)
next

	' expansion with GDP growth rate

for %1  COG   IG   WD R_REVX HIH
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txq)
next

	' expansion with population growth rate

for %1 LG POP65 POP
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txn)
next

	' expansion with deflators growth rate

for %1  PPX
smpl 2004 2100 if {%1}=na
```

```
genr {%1}={%1}(-1)*(1+txp)
next

            ' expansion with special growth rate

for %1   SOCBR
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txq)/(1+txn)
next

genr t=t(-1)+0.25
```

It will be also useful (but not necessary) to create values for the endogenous:

```
smpl 2005Q1 2100Q4

' we expand series for projection: we initialize ENDOGENOUS variables
 '    constant expansion

for %1    COMPM COMPX ER   FCAPGP IR IRL IRM IRS RCVAL RCVOL RES_WR   RMARG RPROB RPROF TTRAD UNR  UR
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)
next

'      expansion with long-term GDP growth rate

for %1  CAP COH  FD HRDI I CI K  M  Q  X GDPM IC
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txq)
next

            ' expansion with population growth rate

for %1 LFD  LF  LT   POPAC UN  popt
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txn)
next

            ' expansion with deflators growth rate

for %1    PCOH PFD PFDXT PI PIG PM PMT  PP PQ PX UWC
```

```
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txp)
next
next

        ' expansion with variables at current prices  growth rate

for %1 CGV  EXPG FCAPF FCAPG FCAPX FDGV HDI HI ICT IFP IGV MARG MVAL NIF NIG NIX NIXL NIXX OIT PROF GDPMVAL QVAL  REVG REVQ REVX  SCF SCG SCW SOCB SUBS TAR TRB  VAT W WF WG  XVAL
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txp)*(1+txq)
next

        ' expansion with labour productivity growth rate

for %1  PL  PLT
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txq)/(1+txn)
next

        ' expansion with labor productivity at current prices growth rate

for %1 WR
smpl 2004 2100 if {%1}=na
genr {%1}={%1}(-1)*(1+txp)*(1+txq)/(1+txn)
next
```

But the essential issue is the guarantee that a stationary path will be obtained.

10.2.2.7.1 General elements

The same elements as above apply to this case.

But the higher complexity of the problem makes the methods proposed earlier more efficient if not necessary:

- Making shocks on the main assumptions and observing the consequences.
- Excluding some variables, or changing (increasing, decreasing or suppressing) local explanations.

10.2.2.7.2 Ensuring convergence in the long run: further elements

Let us return to the constraints the model must follow, and see how the above principles apply to a more complex case.

- First, care must be taken to avoid any remaining trend.

As we have said earlier, all elasticities should be unitary in the estimated equations, except of course for elements without dimension (ratios). The easiest method is obviously to use an error correction framework: the constraints can be set in the cointegrating equations (estimating them in one pass does not change this property, just the validity of the method….). They can be released in the VAR (the dynamic equation) at no cost to convergence, in principle.

One will observe that almost all our estimated equations, even if cointegration was applied only twice, contain a long-term expression linking ratios without dimension. The only exceptions are:

- The trade prices where no difference is made between short and long-term, but homogeneity is obtained by constraints on coefficients.
- For investment:

o In the complementary factors case, most elements represent derivatives (with no dimension). The only exceptions are the rate of use and the profits rate, both ratios which should stabilize in the long run. However, the rate of use is no longer fixed as in the simpler case.

o In the Cobb-Douglas case, the capital – output ratio is linked to a ratio of prices (labor to capital).

The only real problem lies with employment, which has a specific dimension.

o We have already treated the complementary factors case. The trend in labor productivity is estimated. In the long run, both trends in employment (populations) and value added (a variable at constant prices) are fixed, and so is the trend in labor productivity, with a value which must replace the estimated one[153].

o For the Cobb-Douglas case, three elements are taken into account to define productive capacity:

* The trend in global factor productivity, at present estimated.
* The contribution of labor: α txn
* The contribution of capital: (1- α) txq

Basically, the total growth rate of capacity should be the same as all variables at constant prices, which gives:

$$txq = a + \alpha \cdot txn + (1 - \alpha) \cdot txq$$

$$a = \alpha \cdot (txq - txn)$$

This formula is easy to interpret: capital units will have a constant productivity (remember that the value at constant prices includes the increase in quality). Labor units do not, so to achieve their share of the increase in capacity their productivity must increase as the difference between growth of quantities and growth of labor.

In this way the total growth of capacity will meet the constraint.

[153] Or one of trends must be set to agree with the estimated value.

To present the issue in a clearer way, it might be better to consider, not employment, but employment "efficient units" in which each worker is value at his potential contribution to production, just like capital. The value of these units at constant prices will grow as production, and there will be no need for an additional productivity trend.

Once these conditions are set, the model should have a long-term solution, to which the formulations (in particular error-correction behaviors) should make the model converge.

10.2.2.7.3 Making the model converge in the short run

The problems are the same as in the simple case, but

- The danger is higher as more elements and mechanisms are concerned. The probability of spurious cycles grows.

- Deflators are generally more volatile than quantities (being less "real").
- The risk of making an error on assumptions is higher, as the process of definition is more complex: trying to produce an accurate forecast on a partially known future will lead us to define some of them explicitly.
- The trends which have to be blocked are more numerous, and they might appear in more complex equations.
- Finally, it is possible that the management of residuals in the first forecast periods brings high variations, and in particular strong cycles.

10.2.2.7.4 Making the model converge in the medium run

Again, the main danger lies in the cycles, with a higher probability.

But the fact that the model has converged at least for some periods will provide us with important information, and the possibility to conduct additional tests, as we have already seen.

10.2.2.7.5 Specific forecast elements

Now that we have made the model converge over the whole period, we can test its properties through responses to shocks. Then we can move to actual forecasts.

10.2.2.8 Producing a forecast

The technique is the same as usual. We can use any algorithm.

```
' we solve the model in the future (no shock)
'    The letter for the base simulation is b

  _mod_1.append assign @all _b
   smpl 2020Q1 2100Q4
  _mod_1.scenario "Scenario 1"

'    We make sure that the full model is solved
'     and that we use the normal values of exogenous (with no suffix)
'    We use the Gauss-Seidel method

  _mod_1.override
  _mod_1.exclude
```

```
  _mod_1.solveopt(n=t m=1000,o=g,d=d)
  _mod_1.solve

'     We compute the growth rates
'     with an additional term in the denominator
'     it avoids dividing by zero (which gives an error message)
'     if the value is zero the expression gives 0 / (0 + 1) = 0
'     if not zero it doean not change the result as the condition gives zero

for !i=1 to _g_vendo.@count
%1=_g_vendo.@seriesname(!i)
genr td_{%1}=100*d({%1}_b)/({%1}_b(-1)+({%1}_b(-1)=0))
genr td0_{%1}=100*d({%1})/({%1}(-1)+({%1}(-1)=0))
genr dtd_{%1}=td_{%1}-td0_{%1}
next
```

10.2.2.9 Producing shocks on assumptions

Compared to the previous example, the process will be a little more complex. In particular, we shall consider eight shocks.

Comments are imbedded in the program.

```
'     Now we shall produce a set of shocks (in the present case 7)
```

```
  smpl 2000 2100
```

```
'    The group called shocks_v will contain the list of 7 shocked variables
```

```
group shocks_v ig erx r_vat r_tar r_tarx wd m_ec
```

```
'    The group called shocks_l will contain 7 letters associated to each shock
'     but the letters must be known as series
'     as in EViews groups can only contain series
'     We create the artificial series
'      but only for the non-existing ones of course
```

```
  for %z  g r t f y w n

  if @isobject(%z) =0 then
```

```
genr {%z}=na
endif
next
```

' We create the group

```
group shocks_l g r t f y w n
```

' Now we compute the additional change for each assumption,
' using the name of the variable
' The shock will only start in 2021
' leaving one unshocked period
‘ to check that the difference comes only from the shock
‘ The name of the assumption will combine the shocked variable with the letter
' 1 - Shock g: +1 GDP point on Government demand

```
genr ig_g=ig+.01*gdpm_b*(t>=2021)
```

' 2 - Shock r: 1% devaluation of the Euro

```
genr erx_r=erx*(1+.01*(t>=2021))
```

' 3 - Shock t: -1 point on the VAT rate

```
genr r_vat_t=r_vat-.01*(t>=2021)
```

' 4 - Shock f: -1 point on the local tariffs rate

```
genr r_tar_f=r_tar-.01*(t>=2021)
```

'   5 - Shock y: -1 point on the foreign tariffs rate

```
genr r_tarx_y=r_tarx-.01*(t>=2021)
```

'   6 - Shock w: +1% on World demand addressed to France

```
genr wd_w=wd*(1+.01*(t>=2021))
```

'   7 - Shock n: +1% on ex-ante imports

```
genr m_ec_n=m_ec+.01*(t=2021) -.01*c_m(8)*(t>2021)
```

The loop on the shocks

```
for !j=1 to shocks_v.@count
smpl 2005Q1 2100Q4
```

'   We get the name of the shocked variable and the associated letter

```
%2=shocks_v.@seriesname(!j)
%3=shocks_l.@seriesname(!j)
```

'   We set the solution suffix using the letter
'   We control no endogenous is excluded from the simulation
'   We override the variable associated to the current shock
'   We solve the model

```
_mod_1.scenario "scenario 1"

_mod_1.append assign @all _{%3}
```

```
_mod_1.exclude

_mod_1.override {%2}

_mod_1.solve(n=t m=1002 c=1e-6 o=g d=d)
```

' We compute the difference to the base simulation, in absolute and relative terms

```
for !i=1 to _g_vendo.@count
%1=_g_vendo.@seriesname(!i)
series d{%3}_{%1}={%1}_{%3}-{%1}_b
series p{%3}_{%1}=100* d{%3}_{%1}/({%1}_b+ (%1}_b=0))
series dv_{%1}=d{%3}_{%1}
series pv_{%1}=p{%3}_{%1}
next
```

' We create groups for variations for a specific list of important variables

‘ the number is restricted to make a table legible

```
group g_v{%3} P{%3}_PFD P{%3}_PM P{%3}_PX P{%3}_PQ P{%3}_WR P{%3}_FD P{%3}_gdpm
P{%3}_X P{%3}_M P{%3}_I P{%3}_COH P{%3}_UR P{%3}_CAP P{%3}_LF P{%3}_K P{%3}_UNR
D{%3}_FCAPGP P{%3}_RCVAL P{%3}_RCVOL P{%3}_TTRAD
```

' or very important ones
‘ displayed in a graph
‘ the number is further restricted to make the graph legible

```
group g_w{%3} P{%3}_FD P{%3}_gdpm P{%3}_X P{%3}_M P{%3}_PQ
```

' We create a group for variations for all variables

```
group g_v2 pV_* dV_*
```

' We store all shocks in Excel files (2025 and 2100)

```
'       using the letter associated to the shock
```

```
smpl 2020Q1 2025Q4
write(t=xls) v2_{%3}.xls g_v2
smpl 2020Q1 2100Q4
write(t=xls) w2_{%3}.xls g_v2
next
```

10.2.2.10 Applying the programs

We shall now present the results obtained from the previous programs.

10.2.2.10.1 Producing the model framework

Basically, the program creates an item called _mod_1 in the workfile. Of course, this model cannot be solved at this time. However, accessing this item gives alternately

- the list of variables,
- the list of equations,
- the original source code,
- the block structure.

The list of equations and the code repeat the model creating statements, and are of little interest. The list of variables, as stated earlier, helps locate problems such as logically endogenous variables with no equation (they appear as exogenous) or typing errors in variable names (they create also an exogenous item).

The last element is the most interesting. Let us show the result for our model.

```
Number of equations: 88
Number of independent blocks: 7
Number of simultaneous blocks: 3
Number of recursive blocks: 4
Largest simultaneous block: 52 Equations (3 feedback vars)
Block 1: 1 Recursive Equations
        ict(43)
Block 2: 52 Simultaneous Equations (3 feedback vars)
        lfd(5)      lf(8)       lt(11)      popac(14)   un(15)
        unr(16)     pl(9)       pc(22)      wr(26)      wf(50)
        wg(39)      w(40)       scw(41)     socb(38)    revx(37)
        pi(23)      cost(27)    cap(13)     ur(3)       pq(18)
        qv(47)      revq(36)    hi(42)      hdi(44)     hrdi(45)
        coh(46)     ic(10)      irl(32)     irs(31)     ir(33)
        er(30)      pfdxt(21)   pp(19)      px(28)
        compx(64)   x(65)       xv(67)      pm(29)      pmt(60)
        compm(61)   kd(4)       k(6)        i(7)        fd(62)
        m(63)       mv(66)      vat(76)     gdpmv(48)   gdpm(1)
        relc(35)    q(2)        pfd(20)
```

Block 3: 10 Recursive Equations					
	id(12)	uwc(17)	pig(24)	res_wr(25)	irm(34)
	pgdpm(49)	subs(51)	marg(52)	rmarg(53)	ifp(54)
Block 4: 3 Simultaneous Equations (1 feedback var)					
	prof(56)	fcapf(59)	nif(55)		
Block 5: 18 Recursive Equations					
	rprof(57)	rprob(58)	rcval(68)	rcvol(69)	ttrad(70)
	trb(71)	nixl(72)	nixx(73)	nix(74)	fcapx(75)
	scf(77)	oit(78)	tar(79)	scg(80)	revg(81)
	igv(82)	cogv(83)	fdgv(84)		
Block 6: 3 Simultaneous Equations (1 feedback var)					
	expg(86)	fcapg(87)	nig(85)		
Block 7: 1 Recursive Equations					
	fcapgp(88)				

What the above shows is that the 88 equations in our model can be decomposed into a sequence of seven « blocks ». Three of them are defined as recursive, two as simultaneous.

Let us define these notions, which are associated to the model solving process.

- Recursive means that the set can be arranged in such a way that no variable appears in any equation before the one in which it is computed. Obviously, this means that computations will give to each variable its exact value, and after each variable has been computed once no further computation is needed.
- Simultaneous means that each variable in the set depends on a variable computed later, either directly or through a sequence of (possibly recursive) influences. Therefore, this variable will not take its exact value immediately, except if the variable(s) responsible for the non-recursivity was given its exact value in the beginning.

Starting from any order, EViews is able to build an initial recursive set, as variables which depend only on , lagged, or variables which have been included earlier in the set. We will call this set a « prologue ».

It is also able to build a final recursive set, as variables which influence no variable in the rest of the set, or only variables which have already been included in the set. We will call this set an « epilogue ».

Once this is done, the rest of the model could be considered simultaneous. However, this set can eventually be separated again in two consecutive ones, if no variable in the second set influences any variable in the first. The first set can have an epilogue, the second one a prologue, which can be joined into an intermediate recursive block. And the process can be applied again to the two simultaneous blocks, until partition is no longer possible.

In our case, we observe:

- A small initial recursive block with ICT (Income tax) and IFP (Tax on firms' profits). They obviously depend on an exogenous tax rate and revenue from the previous period.
- A large (51 equations) block containing two interconnected loops: the Keynesian supply – demand – supply equilibrating process, and the wage - price loop. Taking any couple of variables in the set, one can link the first to the second, and the second to the first.
- A large (11 equations) recursive block, containing mostly descriptive elements (the terms of trade), items in the State budget (amount of tariffs) and variables influencing only the future (the capital stock)
- A small non-recursive 3 equations block, linking Firms profits, Firms Financial capacity, and net interests paid (depending on the capacity and entering profits These elements depend on the rest of the model, but have no influence.

- A large (18 equations) recursive block, containing mostly descriptive elements (the terms of trade), items in the State budget (number of tariffs) and variables influencing only the future (the capital stock)
- A small non-recursive 3 equations block linking Government balance (revenue – expenditures) interests paid (depending on the balance) and expenditures (including interests).
- A 1 equation final block for the balance in GDP points.

10.2.2.10.2 Producing the data

We hope the comments included in this program make it self-explanatory. Basically, the OECD data is used to create the model series in sequence, and the results are saved. Of course, elements used in any computation must have been created in a previous statement (if the sequence defined a model it would have to be recursive).

To adapt this program to another model using the same concepts, we advise to replace all references to the original "FRA_" by the specific name in the original file. This should produce a nearly correct version, with some exceptions:

- Unavailable variables which will have to be created (a common example is capital)
- Variables known using a different definition. This will change the logic of the statements. For instance OECD provides wages including and not including contributions, but not the contributions themselves, which have to be computed as a difference. If they are available, a direct transfer is possible.

10.2.2.10.3 Creating the model groups and checking data – equations

This program is also self-explanatory. The elements produced are:

- Groups for the endogenous (separated automatically into identity and behavioral) and .

As explained in the program, the fact that the behavioral equations, and those only, contain the element « f » is used to separate them from the identities.

On has just to run two « residual check » simulations with different values of "f", and identify as behavioral the variables for which the results are different.

- A set of absolute and relative differences between the historical series and the result given by the associated equation.

Let us make a comment on checking that the residual is null.

Actually, the value of the residual is seldom exactly null. Due to the limited precision of EViews, we generally get a value like 10^{-6} (percent, so zero to 8 digits precision). Once the model reaches a certain size, it becomes difficult to check by sight that the value is sufficiently small (the error can concern a limited range of periods, even a single one).

There are at least two ways to treat the problem, one immediate but limited, the other harder to implement but more efficient and exact.

- The first is to produce an EViews graph. If all residuals are negligible, we should get a set of Brownian motions, with a very low higher and lower maximum values. So we know immediately if there are no errors. But it is difficult to identify the culprits (and an EViews graph is limited in series display).
- The second is to export the relative errors to an Excel sheet, sort the file for a given year, and concentrate on the top and bottom values. The process is repeated until only negligible errors remain.

10.2.2.10.4 Estimating the equations

We have already presented its elements in detail.

10.2.2.10.5 Solving the model on the future

The solving process gives the same results as usual. If convergence is achieved (our case), no message is produced. If not, we have described earlier the techniques one can use to solve the problem.

10.2.2.10.6 Producing a forecast

In our case, the goal is not to produce an actual forecast, but a simulation giving results acceptable enough to be used as a base for the shocks which will follow.

This property can be controlled by displaying the growth rates of the main variables in the short and medium runs.

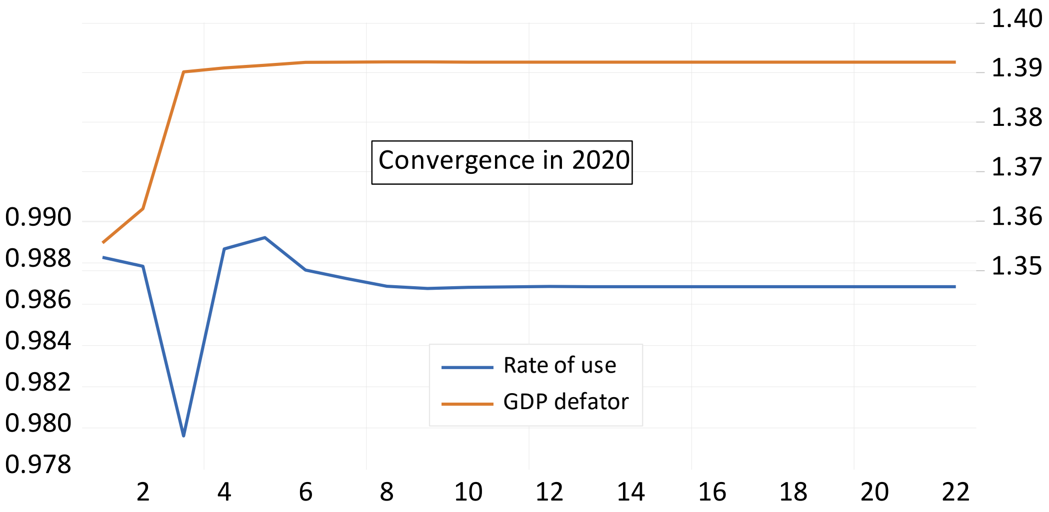
In our case, the model converges normally under Gauss-Seidel, to values not very different from history for the first periods (we have only presented the convergence of the rate of use UR and the value-added deflator PQ).

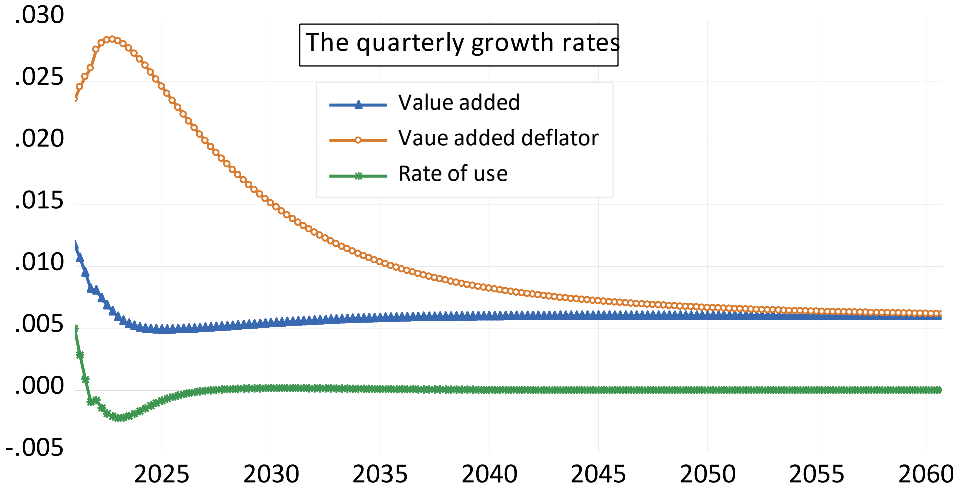
```
Model: _MOD_1
Date: 04/14/20   Time: 17:41
Sample: 2017Q1 2020Q4
Solve Options:
        Dynamic-Deterministic Simulation
        Solver: Gauss-Seidel
        Max iterations = 5000, Convergence = 1e-06
Scenario: Scenario 1
Solve begin 17:41:45
        2017Q1    Block 1 - 1 eqns     Solved (recursive block)
        2017Q1    Block 2 - 51 eqns    Convergence after 24 iterations
        2017Q1    Block 3 - 11 eqns    Solved (recursive block)
        2017Q1    Block 4 - 3 eqns     Convergence after 6 iterations
        2017Q1    Block 5 - 18 eqns    Solved (recursive block)
        2017Q1    Block 6 - 3 eqns     Convergence after 7 iterations
        2017Q1    Block 7 - 1 eqns     Solved (recursive block)
        2017Q2    Block 1 - 1 eqns     Solved (recursive block)
        2017Q2    Block 2 - 51 eqns    Convergence after 29 iterations
        2017Q2    Block 3 - 11 eqns    Solved (recursive block)
        2017Q2    Block 4 - 3 eqns     Convergence after 6 iterations
        2017Q2    Block 5 - 18 eqns    Solved (recursive block)
        2017Q2    Block 6 - 3 eqns     Convergence after 7 iterations
        2017Q2    Block 7 - 1 eqns     Solved (recursive block)
        2017Q3    Block 1 - 1 eqns     Solved (recursive block)
        2017Q3    Block 2 - 51 eqns    Convergence after 26 iterations
        2017Q3    Block 3 - 11 eqns    Solved (recursive block)
        2017Q3    Block 4 - 3 eqns     Convergence after 7 iterations
        2017Q3    Block 5 - 18 eqns    Solved (recursive block)
        2017Q3    Block 6 - 3 eqns     Convergence after 7 iterations
```

2017Q3	Block 7 - 1 eqns	Solved (recursive block)
2017Q4	Block 1 - 1 eqns	Solved (recursive block)
2017Q4	Block 2 - 51 eqns	Convergence after 28 iterations
2017Q4	Block 3 - 11 eqns	Solved (recursive block)
2017Q4	Block 4 - 3 eqns	Convergence after 7 iterations
2017Q4	Block 5 - 18 eqns	Solved (recursive block)
2017Q4	Block 6 - 3 eqns	Convergence after 8 iterations
2017Q4	Block 7 - 1 eqns	Solved (recursive block)

2020Q1	Block 1 - 1 eqns	Solved (recursive block)
2020Q1	Block 2 - 51 eqns	Convergence after 22 iterations
2020Q1	Block 3 - 11 eqns	Solved (recursive block)
2020Q1	Block 4 - 3 eqns	Convergence after 8 iterations
2020Q1	Block 5 - 18 eqns	Solved (recursive block)
2020Q1	Block 6 - 3 eqns	Convergence after 11 iterations
2020Q1	Block 7 - 1 eqns	Solved (recursive block)
2020Q2	Block 1 - 1 eqns	Solved (recursive block)
2020Q2	Block 2 - 51 eqns	Convergence after 22 iterations
2020Q2	Block 3 - 11 eqns	Solved (recursive block)
2020Q2	Block 4 - 3 eqns	Convergence after 8 iterations
2020Q2	Block 5 - 18 eqns	Solved (recursive block)
2020Q2	Block 6 - 3 eqns	Convergence after 8 iterations
2020Q2	Block 7 - 1 eqns	Solved (recursive block)
2020Q3	Block 1 - 1 eqns	Solved (recursive block)
2020Q3	Block 2 - 51 eqns	Convergence after 25 iterations
2020Q3	Block 3 - 11 eqns	Solved (recursive block)
2020Q3	Block 4 - 3 eqns	Convergence after 8 iterations
2020Q3	Block 5 - 18 eqns	Solved (recursive block)
2020Q3	Block 6 - 3 eqns	Convergence after 8 iterations
2020Q3	Block 7 - 1 eqns	Solved (recursive block)
2020Q4	Block 1 - 1 eqns	Solved (recursive block)
2020Q4	Block 2 - 51 eqns	Convergence after 22 iterations
2020Q4	Block 3 - 11 eqns	Solved (recursive block)
2020Q4	Block 4 - 3 eqns	Convergence after 8 iterations
2020Q4	Block 5 - 18 eqns	Solved (recursive block)
2020Q4	Block 6 - 3 eqns	Convergence after 8 iterations
2020Q4	Block 7 - 1 eqns	Solved (recursive block)

Solve complete 17:41:45





If we consider now the evolution with time, we can see that all growth rates stabilize after 40 years. But the quantities and the output gap do it rather quickly, as well, while the deflator takes more time. Of course, this depends a lot on the original distance to the solution.

10.2.2.10.7 Producing shocks on assumptions

Let us now see how our model answers to shocks on its assumptions. We shall be brief, and only present and comment the results for four of them: our goal is mostly to show that the present framework and the associated estimations can provide a model with consistent properties. With a little bit of luck, you should be able to do the same, or probably better. The model(s) we are presenting are far from perfect, and we did not try to make them so[154].

All shocks will start in the first quarter of 2010 and will be sustained for the whole period. To make interpretation easier, we shall limit the periods to the first fifteen years (for the short and medium-term dynamics) and the last three (for the long-term impacts).

10.2.2.10.7.1 AN INCREASE IN GOVERNMENT DEMAND

[154] Actually, we can identify a few coefficients which, with different values, would improve slightly model properties. Of course, we will not apply these changes.

Starting in the year 2021 we increase Government demand by 1% of the GDP forecasted for the same period. This will make interpretation easier than for a constant shock; in particular concerning the “Keynesian multiplier”.

The definition of this element, in practice the most frequently used to characterize a model, is quite simple. In the presence of a given shock on demand, coming for instance from Government consumption, it measures the ratio between the change in GDP produced by the shock and its initial size. If the ratio is higher than one, it means that the economic effects of the shock on the economy are higher than the negative ones. Of course, this figure has also to be interpreted.

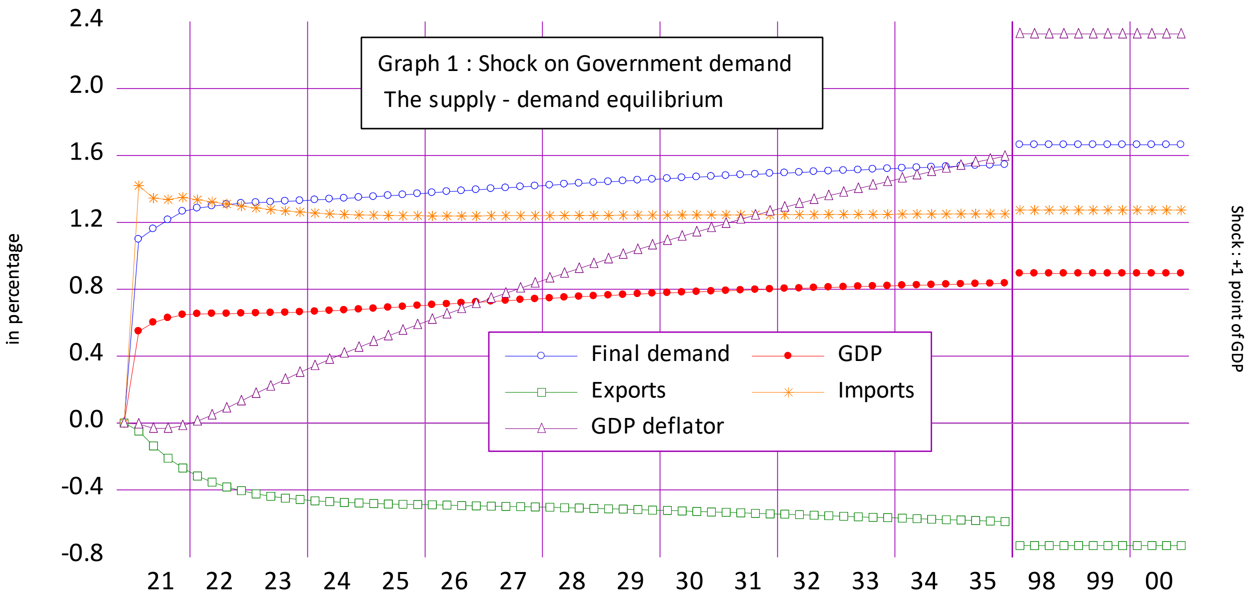
In our case, the “Keynesian multiplier” will be obtained naturally by considering the relative percent change of GDP itself. If GDP increases by more than 1%, it means that the response is higher than the shock.

We observe in Graph 1 that the multiplier grows then stabilizes to an average value of 0.9.

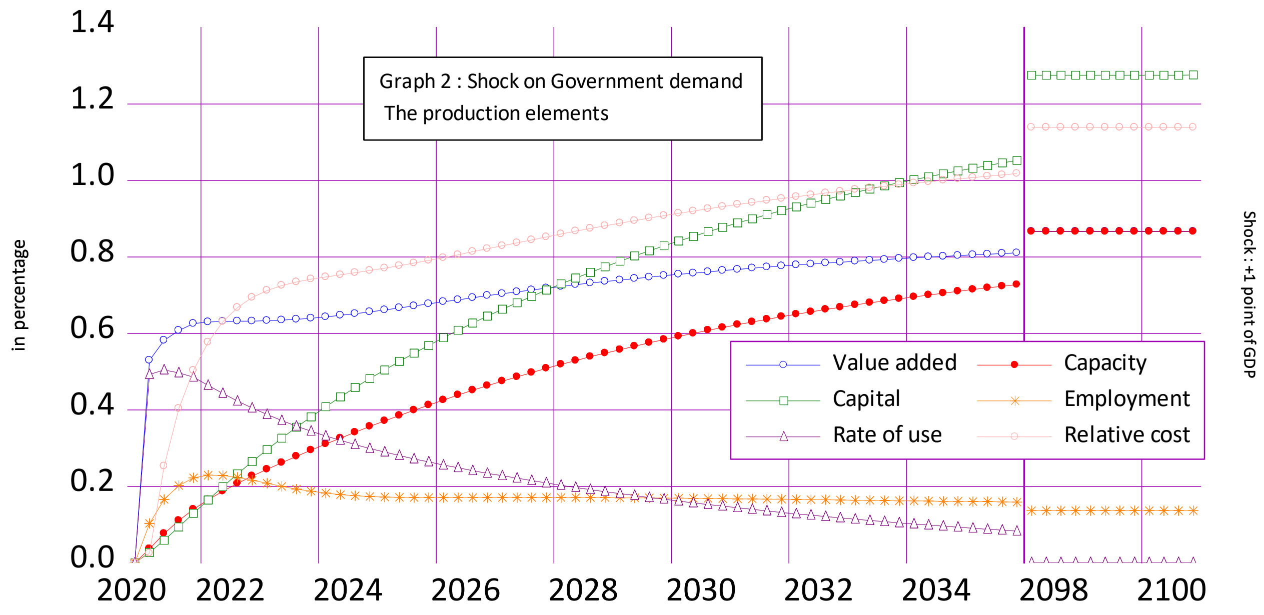
Final demand presents the same evolution, around a higher level, stabilizing at 1.6% in the long run, and imports at 1.2%. The initial decrease in exports is amplified with time, reaching 0.7%.

These elements can be easily explained, considering the model as a whole.

As local demand and GDP are of the same order of magnitude (their difference is net exports, relatively small compared to GDP), the initial (or ex-ante) shock increased demand by roughly 1%. This means that the impact of model mechanisms reduces the ex-post value of GDP growth. Let us detail this aspect.



The second graph shows that productive factors (capital and employment) grow significantly, requiring a strong effort on investment in the short run. As explained earlier, firms expect production to represent a “normal” share of capacities, defining a target on their rate of use. If production increases, a gap will appear, which firms will try to close by increasing capacities through investment and labor. However, this process is slowed by the cautiousness of firms, and the technical difficulty in implementing their decisions. This inertia applies especially to capital, for which the consequences of wrong decisions are more costly. It is only in the last period of our sample that capacities get close to the target, with a possible overshooting in the next years.



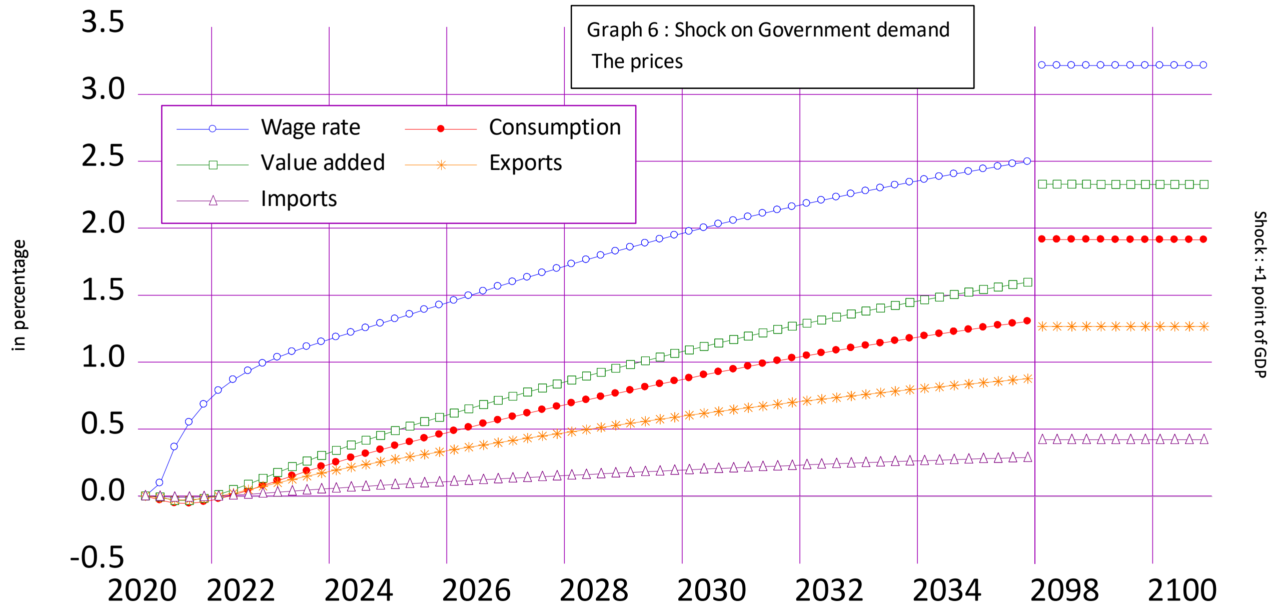
However, capital and employment evolutions diverge from capacity. In the beginning, employment grows more, as adaptation to the target is faster (the change in capital comes only from investment, itself quite inert). In the medium run, the order is reversed by the increase in the relative cost of labor, coming from the decrease in unemployment and the partial indexation on the value-added deflator, the element with the highest increase among prices. In the end the increase in employment is quite limited, and capital increases by 1.3% (representing an increase in investment of the same amount).

Consumption grows even slower. There are several reasons for this inertia.

- Employment adapts gradually to the production increase.
- The purchasing power of wages is affected initially by non-unitary indexation on prices, the gap building up with additional inflation. This loss will disappear as the price increase stabilizes.
- The short-term sensitivity of consumption to revenue is lower than one.
- Inflation reduces the purchasing power of current savings.
- A large share of household revenue is not indexed on activity.

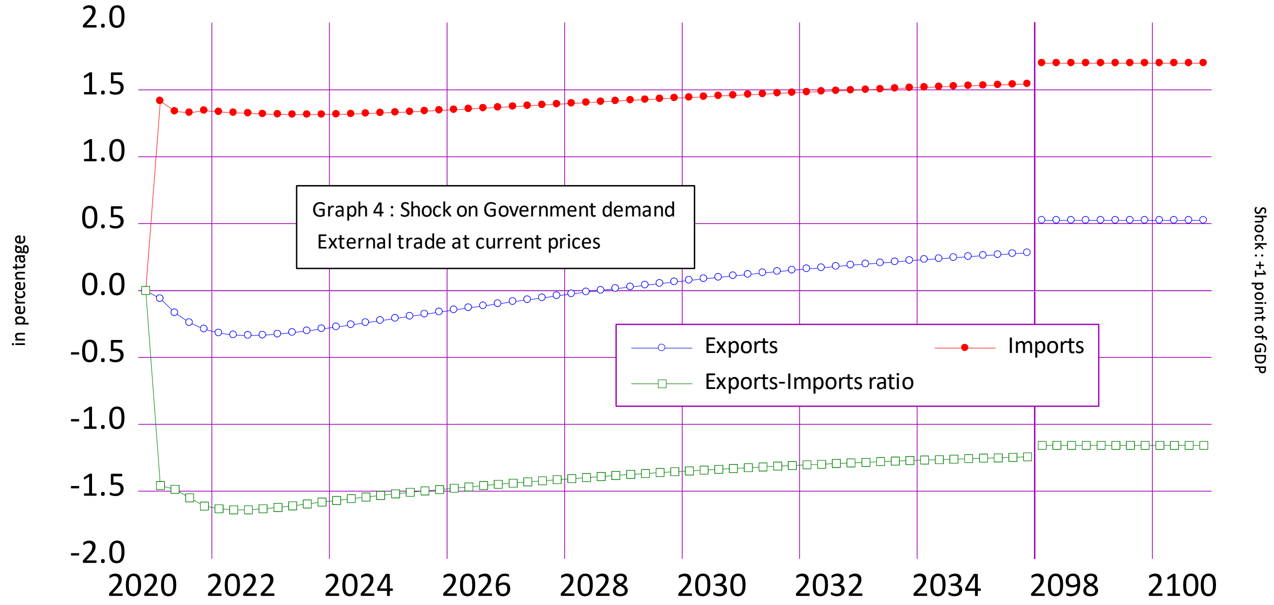
The real wage rate profits from the decrease in unemployment, but this has an adverse moderating effect on employment itself, as we have seen.

The initial decrease in unemployment is the main explanation for the evolution of **prices** (Graph 6). In the short and medium terms, the higher inflation comes from tensions on the rate of use (with higher sales perspectives, firms feel that they can apply higher prices), and on the labor market through the unemployment rate.

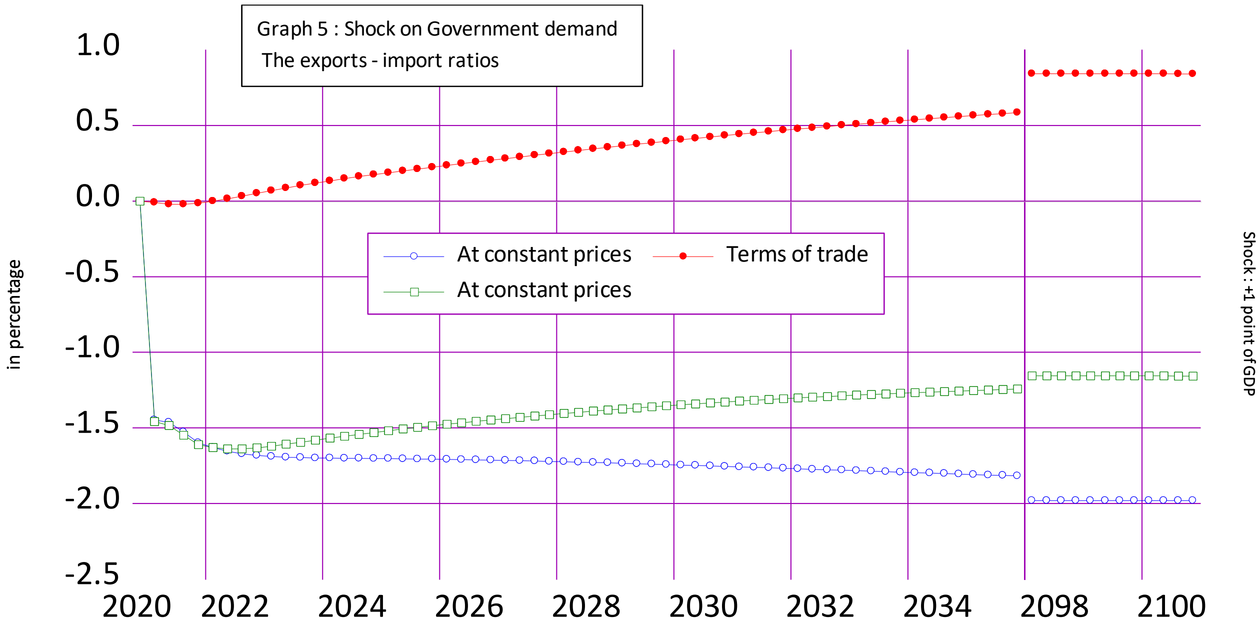


One moderating element is the short-term increase in labor productivity, due to the inertia facing an increase in production. But we have seen that the adaptation of labor to production was rather fast. [155] This element creates very temporary deflation.

Let us now consider the trade balance (Graphs 4 and 5). We know that the trade prices have a different sensitivity to local inflation (0.2 for imports, 0.6 for exports). And the short and long-term competitiveness elasticities are lower than one. But the change in local prices is not negligible, especially in the long-term.



[155] As to capital, the effect is more complex: the investment effort, destined to generate capacities for the future, is higher in relative terms than the present increase in GDP. But in the cost we only consider the share of the present period, through a correcting coefficient.



The evolution of exports is the simplest to explain: they are essentially sensitive to the loss in price competitiveness, with the highest elasticity. They follow also the rate of use of capacities, but the effect is limited by the coefficient and the evolution of the rate itself, which returns to the base value at the end of the sample. Graph 1 shows that exports decrease for the whole period, due to the high increase in inflation.

Imports are more affected: they follow local demand, with the addition of losses in market shares, coming first from the limits on available capacity (decreasing with time) then from the gradual loss in price competitiveness, which compares the local production price to rather stable import prices. Both effects roughly offset each other during the period, and the profile of imports is quite stable. The growing gap with demand comes essentially from intermediary consumption goods, a large share of which follows the decrease in exports.

Concerning the balances, the “gain” on the terms of trade is limited, and the loss terms of trade are little affected, and the loss is not so different on both nominal and real balances.

Now, as to the Government budget, the additional expenditures lead to some revenue, coming mostly from the taxes on production. But as borrowing finances this measure, the interests paid will also increase.

There are two ways to present the changes in the budget. Its variables are measured at current prices. But:

Measured in levels, each variable (including the original Government demand) increases with economic growth and inflation, which makes it difficult to interpret its evolution.

The immediate idea is to eliminate this time effect by dividing the change by a variable representative of the growth, the most neutral idea being to use GDP at current prices.

We get the following table, in which we present the changes for the first year, then for the last quarter of selected years.

	2021Q1	2021Q2	2021Q3	2021Q4	2022Q4	2023Q4	2024Q4	2025Q4	2026Q4	2030Q4	2050Q4	2100Q4
Expenditures												
Total expenditures	0.97	0.93	0.90	0.91	1.67	2.13	2.21	2.31	2.47	3.02	4.44	6.43
Demand	1.00	1.00	1.00	1.01	1.05	1.07	1.09	1.11	1.14	1.23	1.53	1.62
Wages	-0.00	0.02	0.04	0.05	0.09	0.10	0.11	0.12	0.13	0.18	0.30	0.33
Social contributions	-0.00	0.01	0.02	0.03	0.05	0.05	0.05	0.06	0.07	0.09	0.15	0.17
Social benefits	-0.00	-0.00	-0.00	-0.00	0.00	0.01	0.01	0.02	0.02	0.04	0.09	0.10
Subsidies	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.05	0.09	0.10
Interest paid	-0.05	-0.12	-0.18	-0.19	0.46	0.87	0.91	0.97	1.08	1.44	2.28	4.12
Revenue												
Total revenue	0.07	0.20	0.28	0.34	0.46	0.49	0.53	0.59	0.65	0.87	1.48	1.66
Value added tax	0.04	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.07	0.08	0.12	0.13
Other taxes on production	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.03	0.03
Tariffs	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Social contributions employers	0.01	0.05	0.08	0.10	0.14	0.14	0.15	0.17	0.18	0.24	0.40	0.44
Social contributions workers	0.01	0.06	0.09	0.11	0.15	0.15	0.17	0.18	0.20	0.27	0.45	0.50
Income tax	0.00	0.02	0.03	0.04	0.07	0.08	0.08	0.09	0.11	0.15	0.27	0.30
Tax on profits	0.00	0.02	0.01	0.01	0.00	0.01	0.01	0.02	0.02	0.03	0.06	0.08
Government balance	-0.90	-0.72	-0.62	-0.57	-1.21	-1.63	-1.67	-1.72	-1.82	-2.15	-2.96	-4.78

We observe a relative stability of the elements, with two opposite evolutions:

- An increase in revenue with wages (both through social security payments and household consumption).
- A gradual increase of interests paid, with the accumulation of debt.

The first effect wins in the short run, but the second reverses the order quickly.

Of course, it is interesting to obtain a picture of the evolution of the Government budget, measured in currency. But we have presented part of the issue

Even with a shock on real demand, the values in the table combine both a change at constant prices and an effect of inflation, and the interpretation is still difficult, in terms of gains and losses. The growth of a revenue with time might mask a decrease in real terms.

Eliminating both effects makes no sense (the multiplier disappears…), but eliminating inflation allows to reason in constant terms.

We will do this in the second table by computing the difference between elements divided by their solution deflator

$$e_t = \frac{x_t^s}{pq_t^s \cdot q_t^b} - \frac{x_t^b}{pq_t^b \cdot q_t^b}$$

eliminating inflation but keeping the change in the real value in the numerator.

We get quite a different picture as to the evolution of variations (of course, for a given period, the relative evolution remains the same). As could be expected, the figures are much reduced and more stable.

	2021Q1	2021Q2	2021Q3	2021Q4	2022Q4	2023Q4	2024Q4	2025Q4	2026Q4	2030Q4	2050Q4	2100Q4
Expenditures												
Total expenditures	0.98	0.95	0.91	0.91	1.57	1.94	1.96	2.00	2.10	2.35	2.78	4.26
Demand	1.01	1.01	1.01	1.01	1.00	1.00	1.00	0.99	0.99	0.97	0.92	0.90
Wages	0.00	0.02	0.04	0.05	0.07	0.07	0.07	0.08	0.08	0.09	0.08	0.08
Social contributions	0.00	0.01	0.02	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Social benefits	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01	-0.01	-0.02	-0.02
Subsidies	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03
Interest paid	-0.05	-0.12	-0.18	-0.19	0.44	0.82	0.84	0.88	0.97	1.24	1.73	3.24
Revenue												
Total revenue	0.07	0.22	0.29	0.34	0.39	0.37	0.37	0.39	0.40	0.44	0.44	0.42
Value added tax	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Other taxes on production	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Tariffs	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Social contributions employers	0.02	0.06	0.09	0.10	0.12	0.11	0.11	0.11	0.12	0.13	0.13	0.12
Social contributions workers	0.01	0.06	0.09	0.11	0.13	0.12	0.12	0.12	0.13	0.14	0.14	0.13
Income tax	0.00	0.02	0.03	0.04	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.05
Tax on profits	0.00	0.02	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02
Government balance	-0.90	-0.73	-0.62	-0.57	-1.18	-1.57	-1.59	-1.61	-1.69	-1.92	-2.34	-3.84

In GDP points

But the relative evolution of the ex-ante and ex-post Government balance remains the same.

For Government demand, the change keeps mostly at its original level, but not exactly. This relatively small discrepancy is of due to the deflator (GDP for the computation, final demand for the variable). If we use the final demand deflator in the above formula, we get:

	2021Q1	2021Q2	2021Q3	2021Q4	2022Q4	2023Q4	2024Q4	2025Q4	2026Q4	2030Q4	2050Q4	2100Q4
Expenditures												
Total expenditures	0.97	0.94	0.91	0.91	1.58	1.96	1.98	2.03	2.13	2.43	3.00	4.52
Demand	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Wages	0.00	0.03	0.04	0.05	0.08	0.08	0.08	0.08	0.09	0.10	0.12	0.12
Social contributions	0.00	0.01	0.02	0.03	0.04	0.04	0.04	0.04	0.04	0.05	0.06	0.06
Social benefits	0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	-0.00
Subsidies	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.04	0.04
Interest paid	-0.04	-0.11	-0.17	-0.19	0.44	0.82	0.84	0.88	0.98	1.25	1.78	3.31
Revenue												
Total revenue	0.09	0.22	0.30	0.35	0.41	0.40	0.40	0.43	0.45	0.51	0.60	0.61
Value added tax	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.06
Other taxes on production	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Tariffs	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Social contributions employers	0.02	0.06	0.09	0.11	0.12	0.11	0.11	0.12	0.13	0.14	0.17	0.17
Social contributions workers	0.02	0.06	0.09	0.11	0.13	0.13	0.13	0.13	0.14	0.16	0.19	0.19
Income tax	0.00	0.02	0.04	0.04	0.06	0.06	0.06	0.06	0.06	0.07	0.09	0.09
Tax on profits	0.00	0.02	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.03	0.03
Government balance	-0.89	-0.72	-0.61	-0.57	-1.16	-1.56	-1.58	-1.60	-1.68	-1.92	-2.39	-3.91

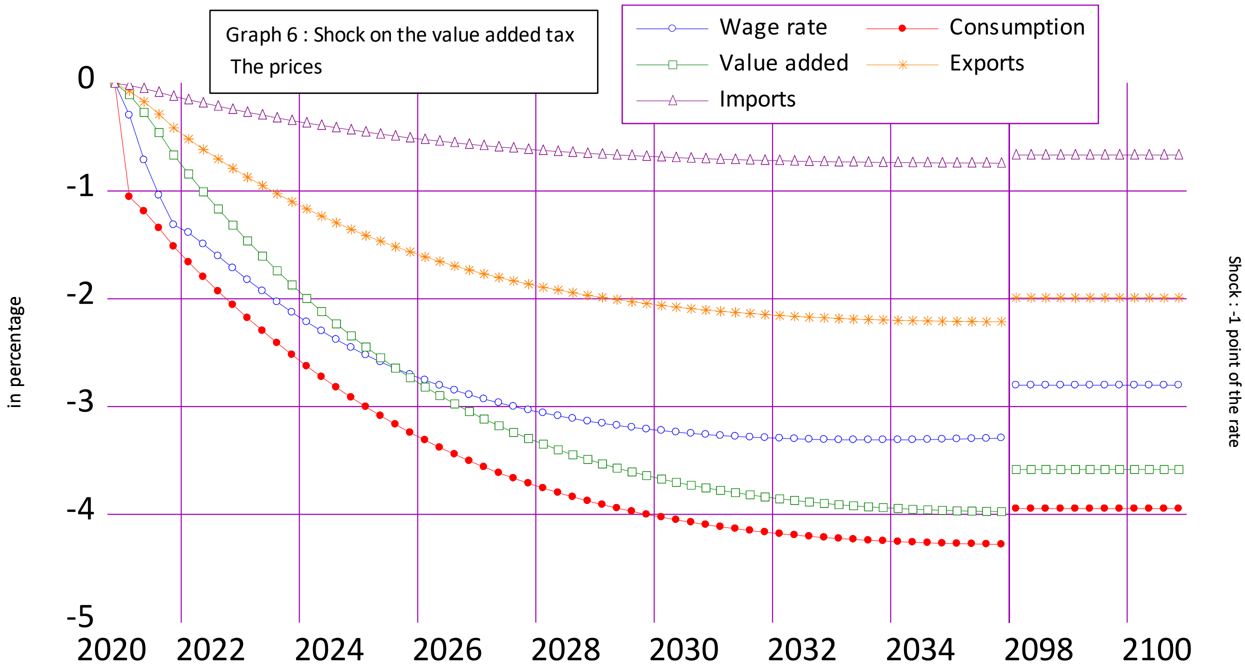
In GDP points

But of course, using different deflators in the same table is not an option, as the balances would no longer hold true.

Running the associated programs, the reader will be able to observe other elements. But he is also expected to produce his own model, and to observe how his changes (limited or extensive) modify model properties.

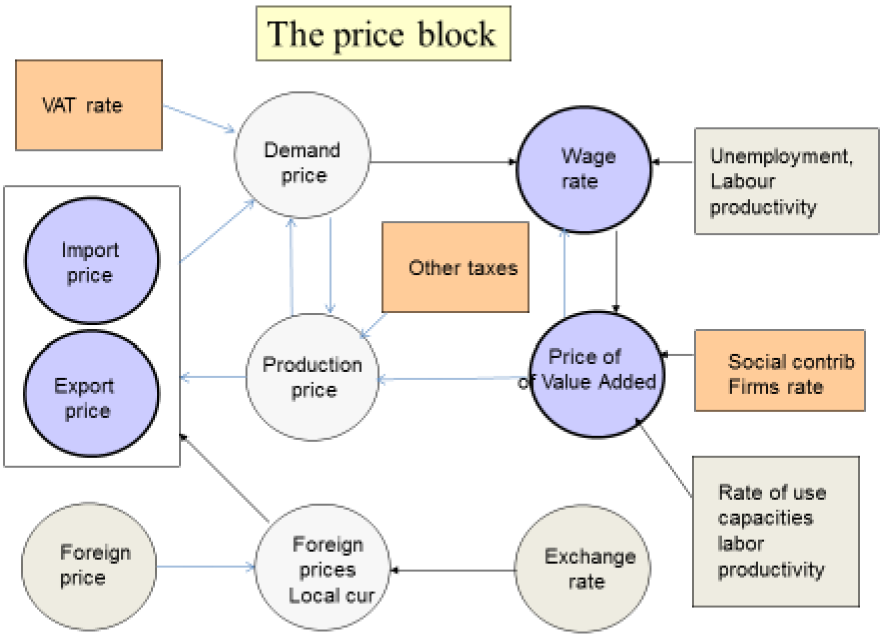
10.2.2.10.7.2 2 A SUPPLY SHOCK: DECREASING THE VAT RATE BY ONE POINT

We suppose that the Government decreases the VAT rate by 1 point.



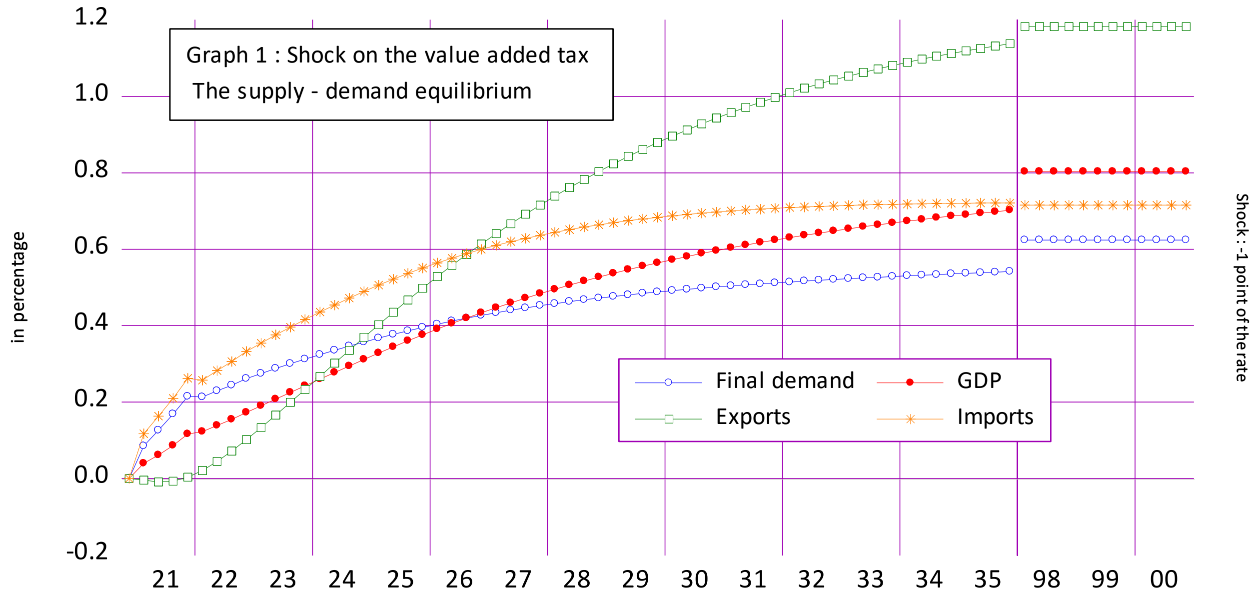
Graph 6 shows that all prices decrease significantly. Of course, the highest impact is on the consumption price (fully depending on VAT). The other prices follow, at various levels depending on their sensitivity.

The wage rate is a special case. Representing initially an average between the consumption and value-added deflators. It diverges gradually upwards, with the gains in purchasing power due to the lower unemployment.

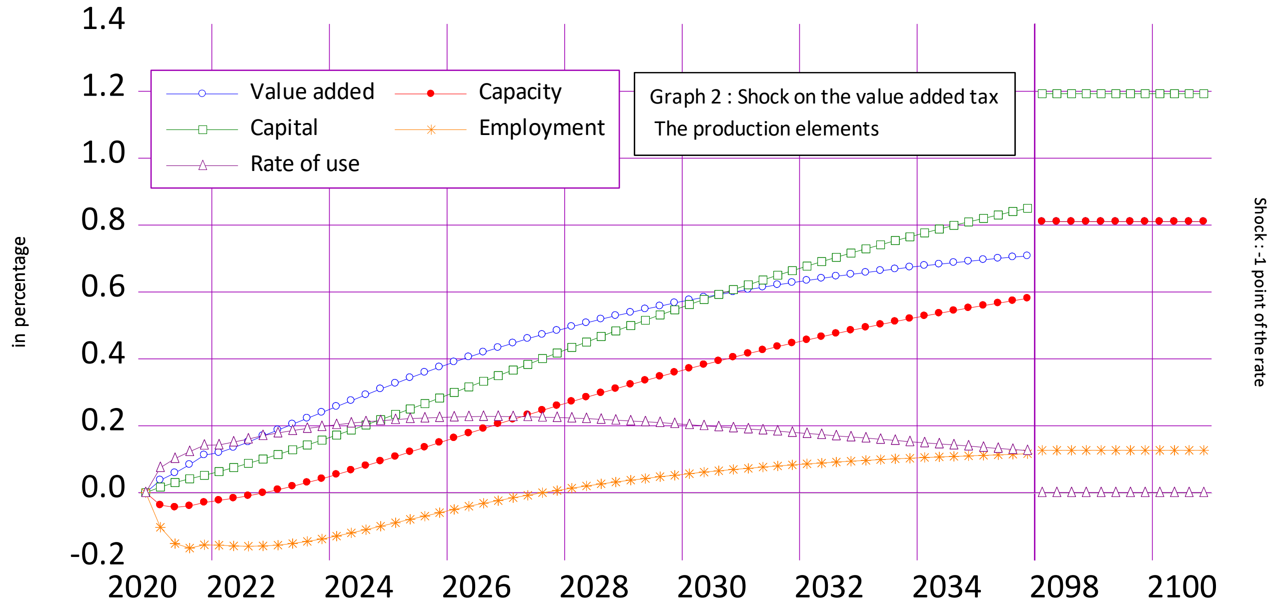


One can observe the loop between the wage rate and the value-added deflator, moderated by trade deflators (the import price is less sensitive to the GDP deflator). Compared to a complementary factors framework, the margins target includes also the capital cost, moderating the role of wages but introducing an additional loop.

The evolution of the supply-demand elements is presented in graph 1 (we have dropped inflation for clarity).

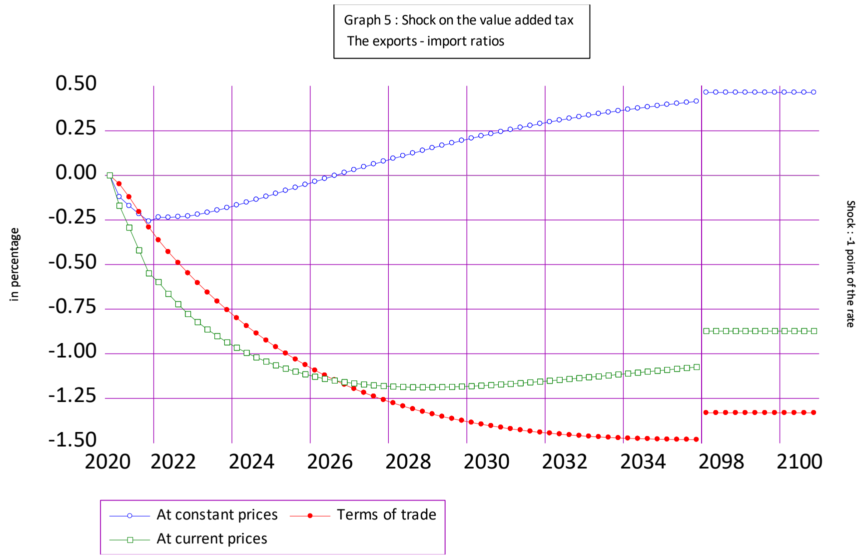


The gain in competitiveness favors exports, once capacities have started to adapt, creating growth through demand in investment and consumption goods. The subsequent increase in imports (coming from local demand but also from intermediate goods and energy requested by exports) is accentuated at first by the initial bottlenecks.



Then competitiveness, exports and GDP keep increasing. The output gap will close, but very slowly as the target is moving, with a combination of factors favoring capital due to the increase in the real wage rate, as in the previous case. As employment adapts faster than capital (which closes slowly with its target), capacity takes a little time in increasing.

The impact on the real trade balance (Graph 5) is negative at first, as the increase in GDP is followed by demand, both final and intermediate. In the medium-long-term, the gains in competitiveness on both trade elements reverses the evolution. But the loss on the current balance remains large.



As to the Government budget, its evolution is similar to the previous case, except for the interests with a growing importance of interests paid, which decrease at first with the rate itself (remember it is indexed on the evolution of CPI, which decreases for the first two years). However, the cost is lower, and the gain on GDP is higher: this instrument seems more efficient on that aspect.

10.2.2.10.7.3 A DEVALUATION (LEAVING THE EUROZONE)

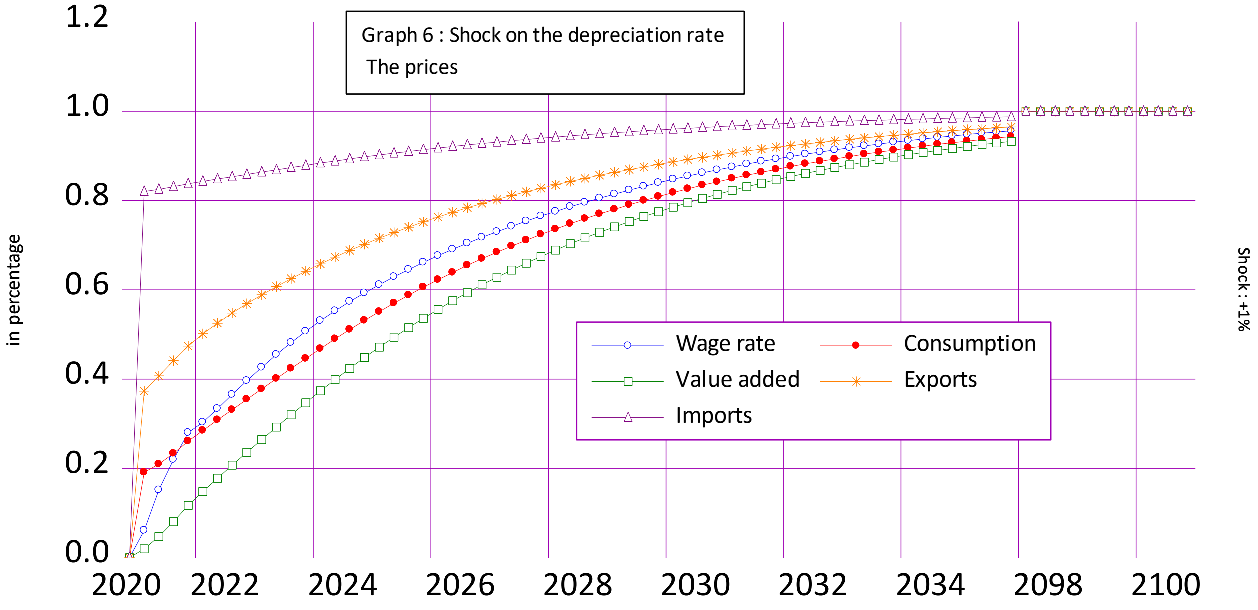
Our shock supposes that France leaves the Eurozone, and suffers a devaluation as the consequence.

We will depreciate the new currency by 1% compared to the Euro.

Of course, this shock is purely technical, as many of its consequences are not made explicit.

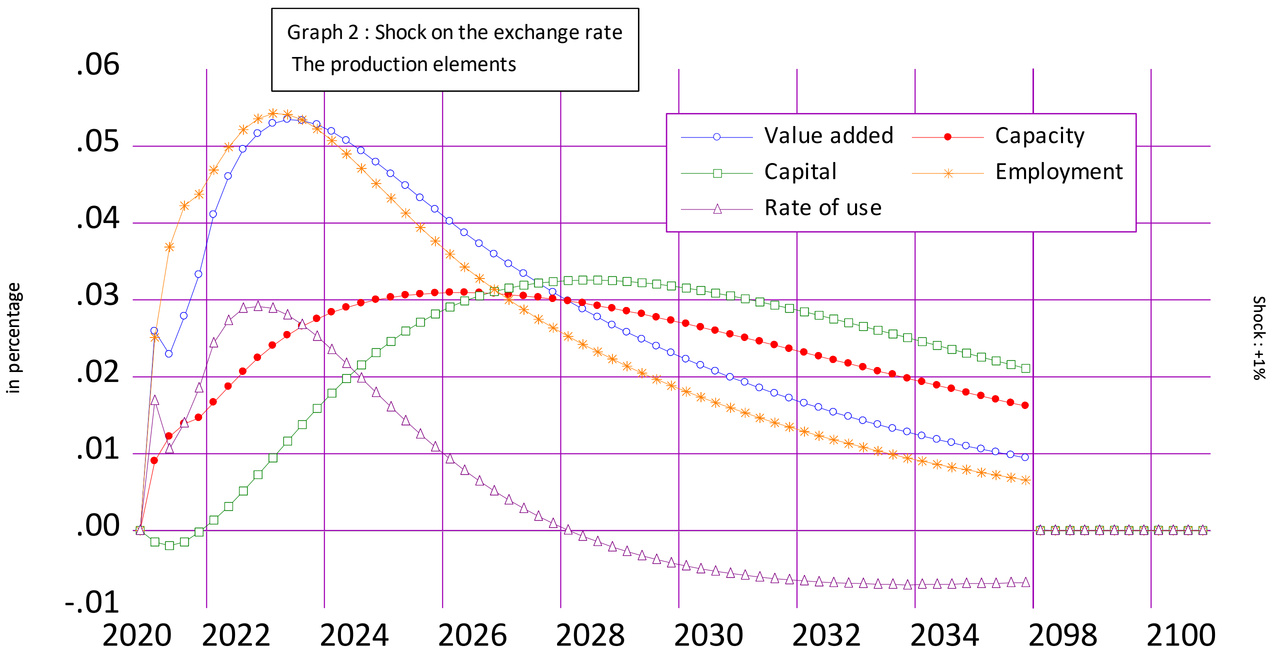
Our first graph presents prices, the second quantities.

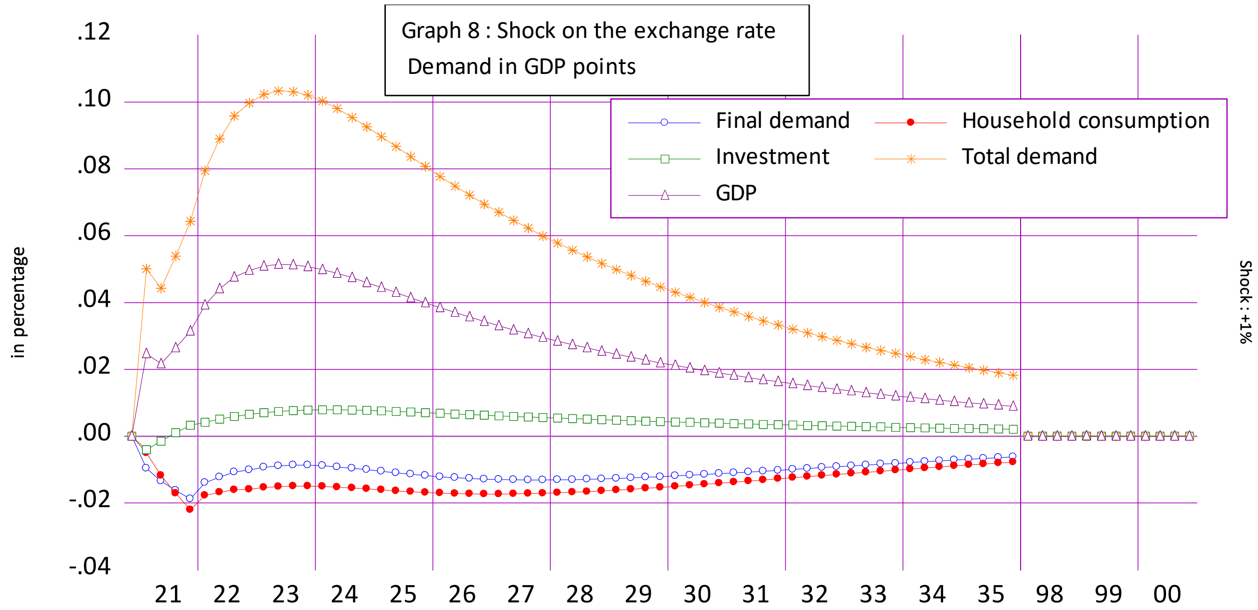
Concerning the latter, we see that in the long run they all increase by 1%, with a hierarchy consistent with the formulas. The highest growth is for the wage rate, as the shock presents a temporary increase in GDP. As it disappears, the gain disappears too.



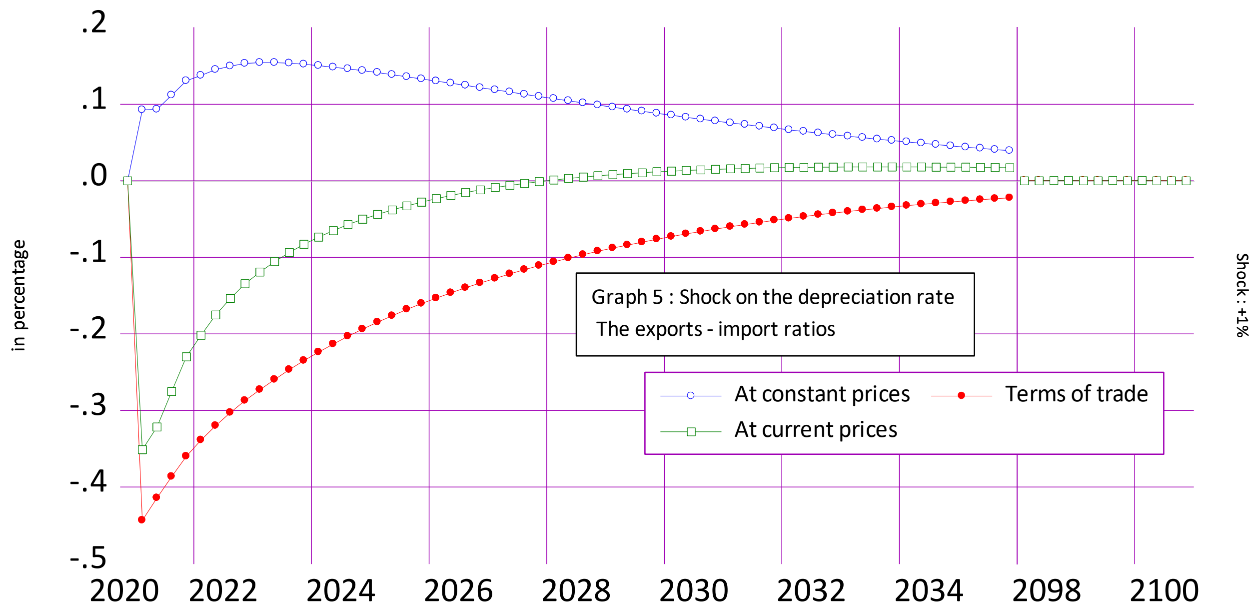
Concerning quantities, exports profit the most from the shock, improving GDP. But imports actually increase a little, the gains in competitiveness being more than balanced by the additional demand (essentially from intermediate consumption required by exports).

Graph 2 shows that the relative cost favors employment initially, as the wage rate indexation takes a few periods, while the price of investment follows the general demand deflator. The real wage rate decreases a little, and consumption follows. Final demand actually decreases.





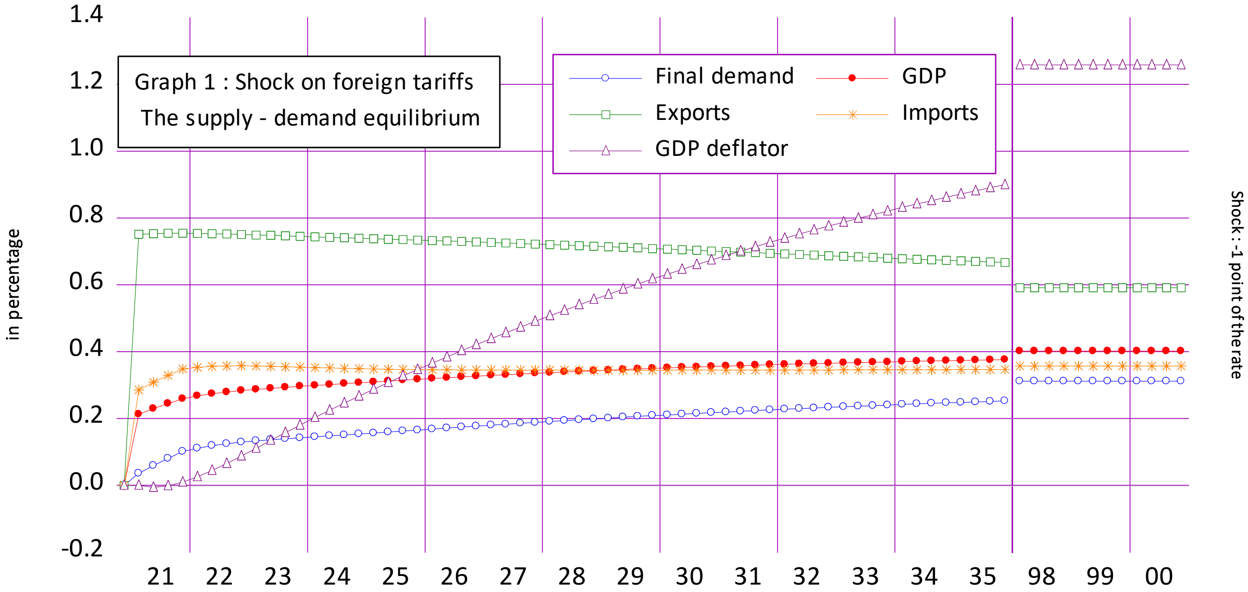
The last graph presents the evolution of trade. We see that the gains in competitiveness (associated with a loss on the terms of trade) are reduced gradually from the first, but that the gains in real trade take some time in reaching the highest level (due to the initial bottlenecks and increase in demand). A very limited gain at current prices appears only in the medium run.



10.2.2.10.7.4 A DECREASE IN THE FOREIGN TARIFFS RATE

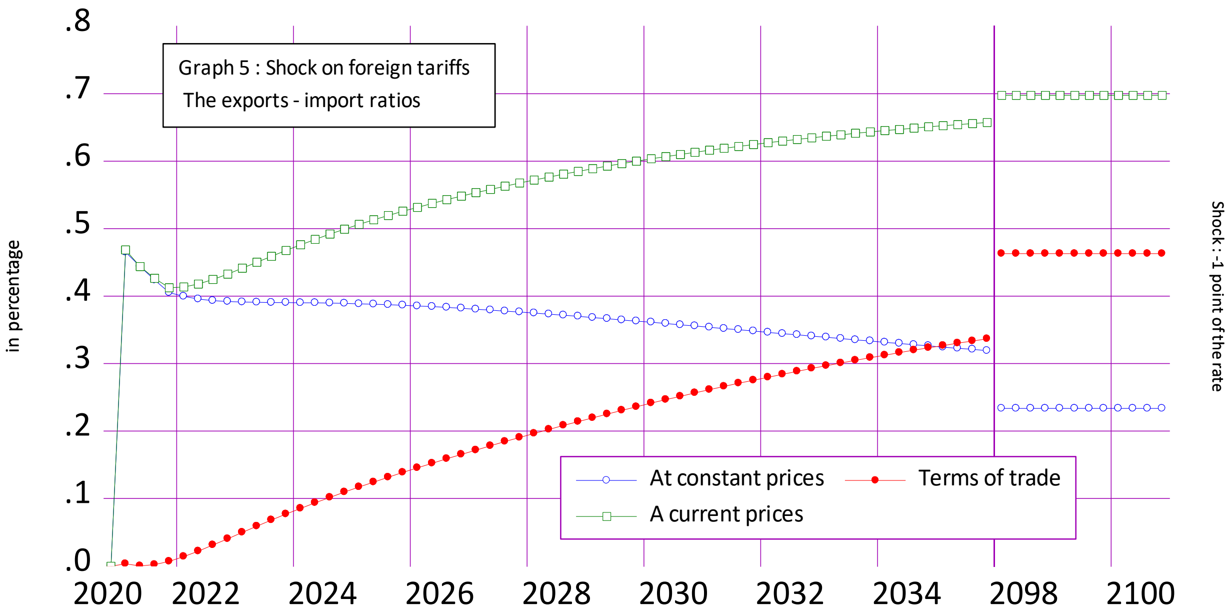
As you have certainly guessed, this shock is essentially demand-oriented:

The improved competitiveness will increase the demand addressed to France, with basically the same effects on GDP as Government demand, from the supply – demand equilibrium.



Of course, one can argue that lower import prices from France will reduce inflation in the rest of the world, which should affect French import prices. World GDP should also change (in a dubious way, positive through disinflation, negative through the higher share of France in world trade). But both these effects can be considered negligible, especially if we consider the cost of the alternative: building a reliable world model.

Of course, a model considering the whole European Union would represent a different issue.



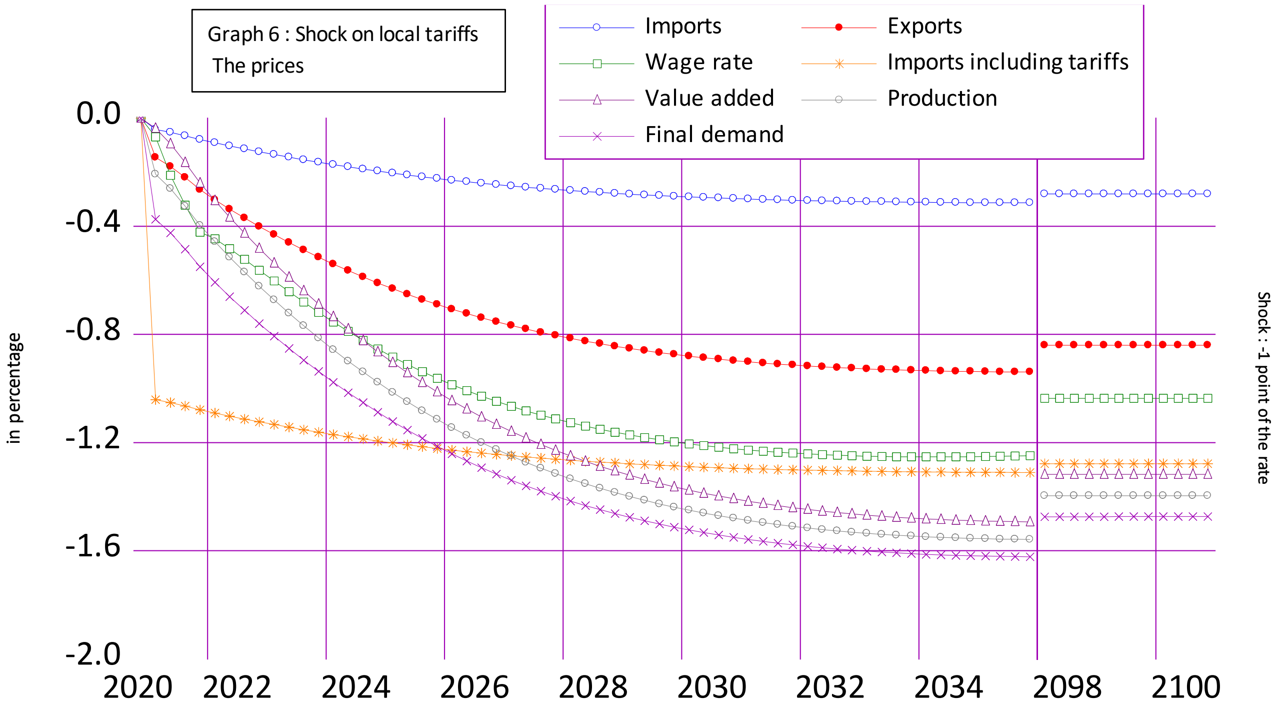
On the whole, the current gain stabilizes at constant prices after a while, when capacities have adapted to the new demand level. But the growing inflation brings a gain on the terms of trade, and the balance in current terms improves regularly.

10.2.2.10.7.5 A DECREASE IN THE LOCAL TARIFFS RATE

This shock brings traditionally the most complex mechanisms, and its consequences are quite volatile from one model to the other. Two main channels have to be considered:

- The improved imports competitiveness increases their share in local demand, reducing local output, with the traditional consequences of a demand shock, only negative.
- The lower import prices bring global disinflation, especially for demand but also for value added through lower wages (indexed on the consumer price). This disinflation helps local firms to compete with foreign producers on the foreign and local markets, reducing the initial gap on the first.
- The cost of capital (part of it imported) decreases more than the value-added price. This increases profitability and creates investment and productive capacity per se, independently from demand. Firms working at full capacity will see the constraint released[156], and local producers as a whole will gain market share on the local and foreign markets. At the same time, to improve demand for these new capacities, they will decrease their prices.

First let us consider prices, the field impacted directly:



To understand the order for the evolutions, we had to introduce additional lines, hoping this did not make the graph too crowded.

We observe in sequence a set of proportional evolutions:

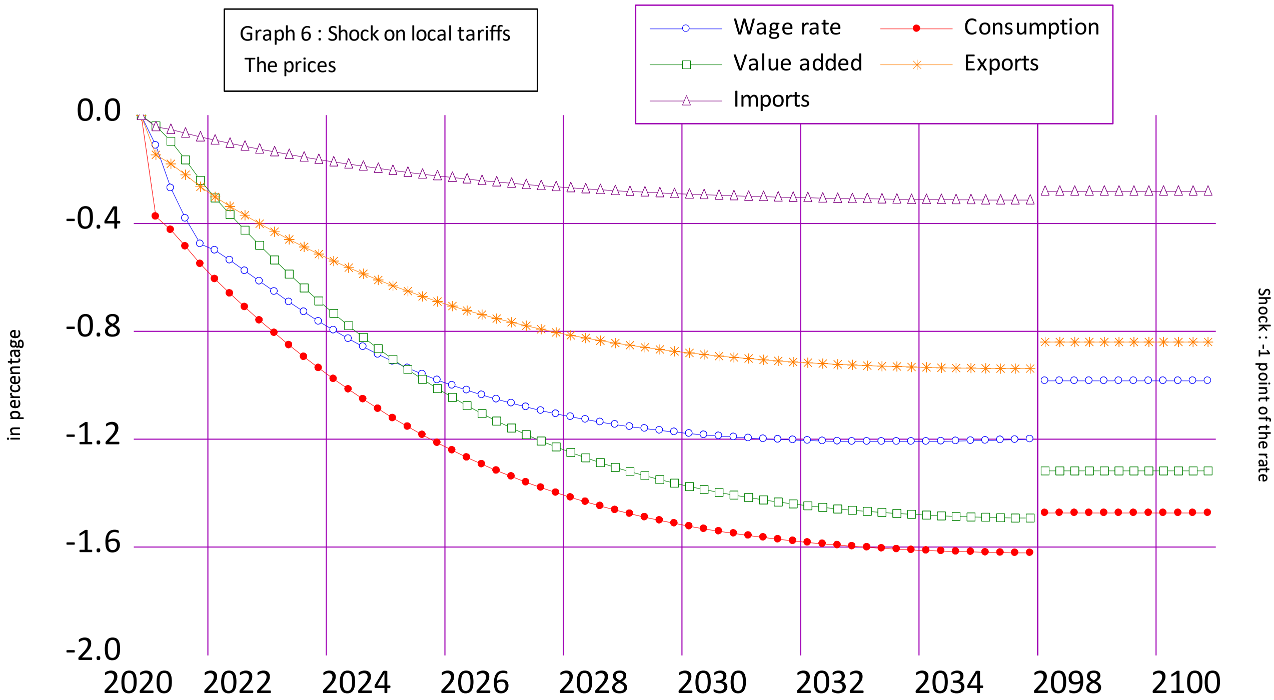
- The import price excluding tariffs, linked essentially to the world price.
- The export price, less sensitive to the world price.
- The value-added price, the most subject to deflation, as in addition to the lower consumption price wages are sensitive to the higher unemployment.

[156] And possibly reappear at a higher production level.

- The production price, a weighted average of value added and demand (intermediate consumption).
- The demand price, a weighted average of imports and local prices (with a small impact from exports).

And two special cases:

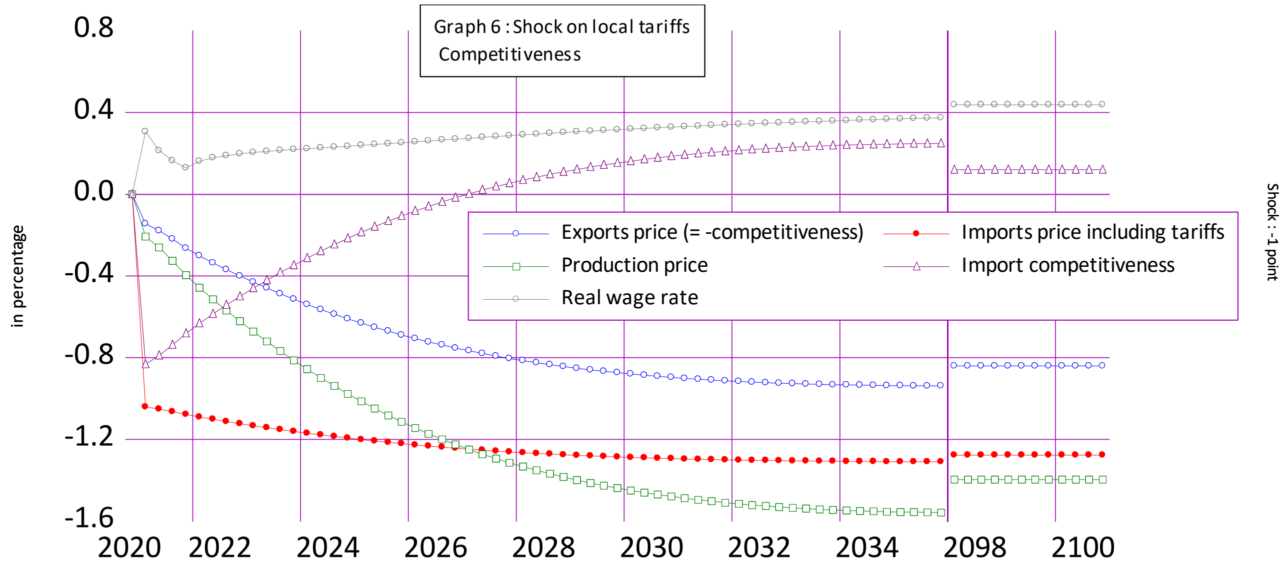
- The import price including tariffs, with a one-point difference to the base evolution.
- The wage rate, which should follow an average of value added and demand, but is also subject to inertia, and to unemployment which increases then decreases, for a reason we will explain.



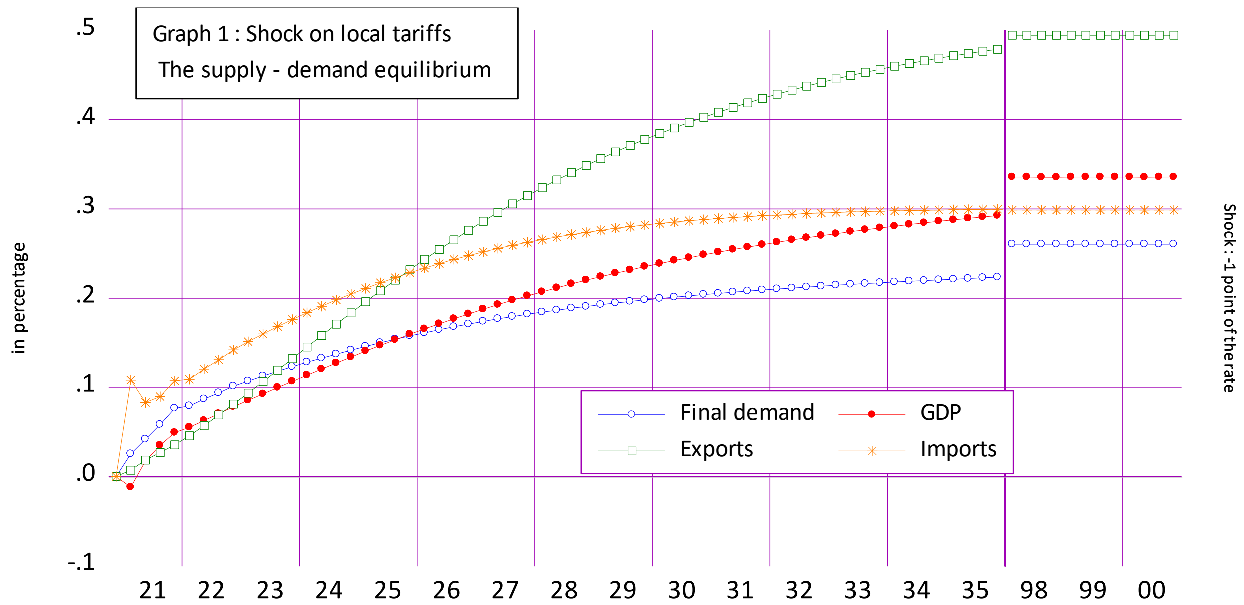
Things should be clearer on a specific graph.

- The real wage rate increases a little.
- Exports price competitiveness grows regularly then stabilizes.
- Import price competitiveness decreases and changes sign when the local production price overcomes the import price including tariffs.

So, in the long run the global outcome of the shock is a gain in competitiveness on both trade elements.

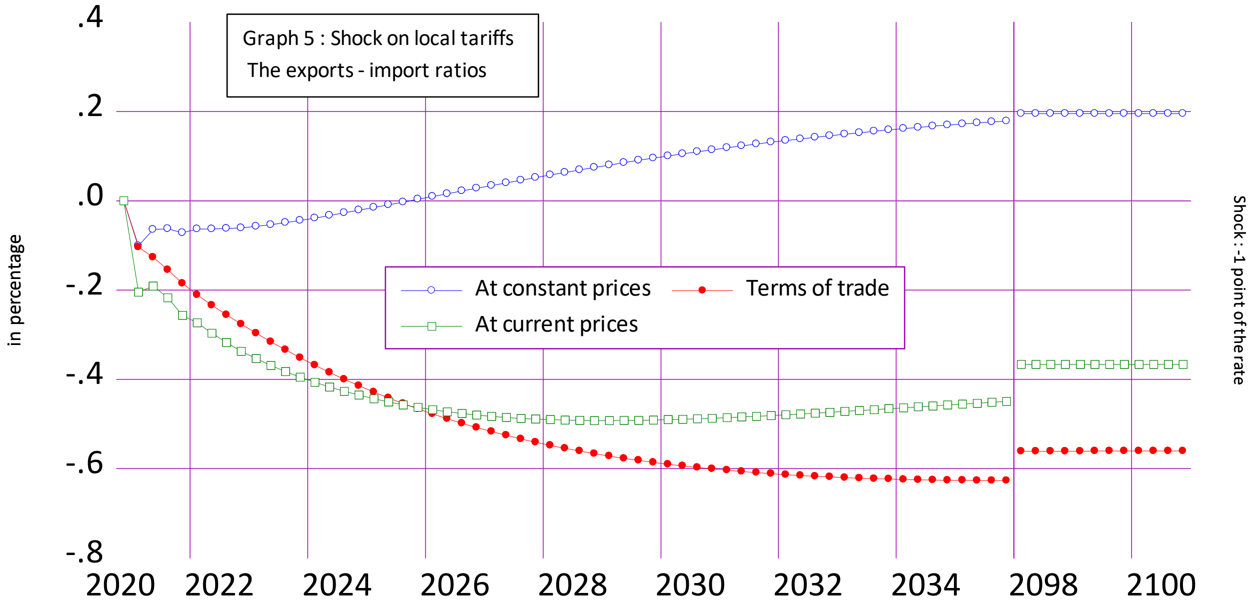


Let us now study the evolution of quantities.



In the short and medium runs, imports show the highest increase, as the only ex ante positively influenced element. But the improvement of local competitiveness limits that effect. At the same time, disinflation profits to exports, and also to demand: profitability favors investment, and the gain in purchasing power of the wage rate favors consumption and imports in turn.

After 15 years, both trade elements gain competitiveness, and the real trade balance starts improving. Demand and GDP too. The only negative evolutions are for the Government budget, but also for the trade balance: the limited gain in real terms is obtained through a large decrease in the export price, and a loss in the actual terms of trade (as seen initially on Graph 6). The local consumers do not actually buy more goods, and the Government will have to increase taxes if it wants to balance the budget.



10.3 A SINGLE COUNTRY, MULTI PRODUCT MODEL

Let us now consider separating the products of our model into several categories. We shall address in turn:

- The reasons for product decomposition
- The specific features introduced by product decomposition, and the way it can be treated.

10.3.1 THE MAIN REASONS FOR PRODUCT DECOMPOSITION

There are three main reasons for product decomposition:

- The decomposition allows to take into account differences in the values of structural elements.

This appears when structural parameters take different values from one product to the other, which implies that they will react differently to external influences (exogenous or endogenous), or exogenous assumptions concerning the product.

- The decomposition allows to evidence differences in the sensitivity to some explanatory elements.

- For some economic element or field, the links between variables follow different mechanisms, calling for different specifications which cannot be summarized by mathematical aggregation.

10.3.1.1 Structural differences

They can be identified easily within the framework of our model. The main elements are:

- Factor productivity: the quantity of labor and / or capital necessary to produce one unit of good can be different. This is particularly true for agriculture, in which labor productivity is particularly low[157]. And also, for services, where the need for capital is generally limited.
- The wage rate: the average wage obtained by workers can be quite different (linked often to labor productivity, which can offset the effect when we consider the unit cost).
- The different shares in each demand item, allowing to present different sensitivities to a given increase in consumption or in investment.
- The amount of intermediate consumption of each good needed to produce one unit of a given good.
- The sharing of production destination between the local and foreign markets (exports).
- The sharing of local demand between the local and foreign producers (imports).
- The separation between firms, households, the state and foreign agents in the local production process, with different consequences on revenue and its use.
- The taxation of goods (VAT, other indirect taxes, tariffs on local imports and exports).

All these elements will be taken into account at no cost, once you have separated the goods in the model. The only problems come from:

- Data availability, which can be partial. For instance, the separation into goods could be known, but not the full input-output table, or the account of firms.
- Model complexity, both for writing down, checking and understanding the results.

10.3.1.2 Differences in sensitivity

This can apply to all parameters in estimated equations. Of course, one should be concerned with the most important ones, and economic theory can point out for which elements significant differences can be expected. For instance, the need for inventories might be more important in the manufacturing sector, or the role of unemployment limited in defining the agricultural wage, or the sensitivity of external trade to price competitiveness can be low for services.

This does not call for specific formulas. The data should dictate the differences in properties.

10.3.1.3 Differences in logical behavior

This time the very formulation of a behavior, or its causal logic, will be different.

This applies essentially to two domains.

10.3.1.3.1 The production processes

- In the agricultural sector, one can assume that producers maximize their output. For this, they can use plantations (fields and trees), and animals, or a combination of both (like plowing, or raising trout around rice fields). They need also a minimum amount of labor, tools and intermediate consumption goods (like fodder). Increasing these levels can improve output, as well as better infrastructures (irrigation, storage and transportation facilities)[158]. But once these are defined, production follows, corrected heavily by

[157] Especially in family managed units.

[158] Remember that production includes only the goods actually bought (consider the supply-demand equilibrium) or entering inventories, which do not apply to agricultural goods (processed goods are considered industry).This means spoiled products are not considered, and improving transportation will increase agricultural production, all things being equal.

unforecasted climatic conditions. This is also mostly true of fishing. The impact of an increase in demand is limited at least in the short run (even if perhaps cattle will be slaughtered earlier, more fish will be caught, or more trees will be cut).

- In the manufacturing sector, we can assume the availability of several processes, among which the producer will chose according to the relative costs of capital and labor. Assuming a constant and unitary elasticity leads to a Cobb-Douglas formulation. We can also choose a simple complementary factors specification, as well as a more general CES function, or even a very specific set of formulas, describing for instance the production of electricity using dams, coal, oil or nuclear power, according to the level of instantaneous demand.

Actually, one of the advantages of product identification is to associate a sophisticated framework to the goods which actually deserve it. In the single good case, the elements of the alternative: using a simple function or associating a Cobb-Douglas to the whole economic spectrum are both harmful to model quality and results.

But the main advantage of this identification is to define a productive capacity, and the associated rate of use. This element will play an important role (as we have seen already) in three elements: the role of local producers in the supplying of local and foreign goods, the investment behavior, and the short-term choice by firms of the mark-up applied to the production cost.

- Finally, in the services sector, the choice of factors can depend on the relative cost (computers versus people), but the role of capacity is less clear, as in many cases producers have a considerable leverage on output, independently of installed capital and even employment: a sudden and high increase in the demand for touristic travel can generally be met by travel agencies, if customers are willing to change their plans. The quality of the service will probably decrease, but not its cost, the element by which the output is measured. Dissatisfaction of customers does not reduce the value of a purchase, except if some reimbursement is due.

10.3.1.3.2 External trade

The trade in agricultural goods, manufactured goods and services follows obviously different rules.

- For agriculture, the export price is generally set to the world price for a given quality, and a given share of the local production is proposed on the foreign market, according to the type and quality of goods, local policy, pre-established contracts.
- For manufactured goods, additionally to demand, price competitiveness is essential, as well as available capacity which can create temporary bottlenecks on the supply of specific elements.
- For services, the determinants are completely different and much less clear, but the amounts traded are less important (except for tourism which is quite difficult to model). Financial services can be formalized considering deposits and loans, rates and maturities (look at the chapter on Stock-Flow Consistent models).

10.3.1.3.3 Consequences for product decomposition

The above remarks should have made clear that the minimal decomposition should introduce three categories: primary, secondary and tertiary products.

Going further, one can consider

- Using more products (more on this later).

  - The most immediate idea concerns energy, especially for oil or gas producing countries. It is clear that the level of production can change widely according to the will of the producer, that the variable production costs are relatively small, and that the whole production should have no problem being sold, at a price highly variable but defined on the world market.

- This is also true in general of mining, which should therefore be counted with the secondary goods, in its own category, or together with energy.

- Construction can also be identified, as it follows the other products, and household revenue. The production process is relatively straightforward for a given type, and it is neither imported nor exported, although its inputs can be (like wood and steel).
- Services can be separated into Financial (managing loans) and non-Financial (transportation or meals in restaurants).

- Identifying the type of ownership. Firms can be divided in state and private, and the latter into individual firms and companies. Foreign owned firms can be identified (using FDI, repatriating profits). And this comes naturally in the above case, particularly for Stock-Flow Consistent models, for with the agents column must separate the two productions. And even if the production of individual owners is merged with firms', identifying households as an agent calls for separating the housing service.

10.3.2 INTRODUCING INTERMEDIATE CONSUMPTION

A very important development corresponds to intermediate consumption.

Any model, for any product, has to follow the equilibrium:

Production + imports = Total demand + exports

However, at the global level, subtracting total intermediate consumption on both sides give:

GDP + imports = Final demand + exports

This almost frees the model from defining intermediate consumption, a concept which is not easy to handle, as it depends on the production process itself, more precisely on its number of stages.

Its only remaining role is the fact that in trade equations, price competitiveness is based on production prices: preference for goods from a country is obviously based on the price at which they sell.

This means that the trade prices (exports and imports) must consider production prices, and that competitiveness must compare the global cost of the exporter to the price at which its competitors sell (both including intermediate consumption). This means for example that a country which has access to cheap oil or gas (like the US, Russia or Algeria) has a comparative advantage over Germany or China, both against exporters to its market and against other competitors on foreign markets[159].

[159] However, as we are using deflators, this means the trade prices for these countries will be less sensitive to the cost of energy, which represents a lower share of the total cost. They will gain competitiveness if the price goes up, but lose if it goes down (as they do not profit from the shock).

The above equation

Production + Imports = Total demand + Exports

Can be written as:

Value added + intermediate consumption by the product + imports

= Total demand + intermediate consumption of the product + exports

This calls for the identification of a matrix of intermediary consumptions, a square one with the dimension of the decomposition.

Typically, the definition of each of these elements at constant prices will suppose that producing each unit of a given good requires a given amount of each good (including itself).

At current prices, one could apply the deflator of the product used (excluding VAT, as it does not apply here) at the global demand level, or at the global intermediate consumption level. The availability of the matrix at current prices improves the situation by allowing to define an individual deflator for each cell.

Sometimes, the statisticians provide these tables only for some years, which calls for an interpolation.

10.3.3 SPECIFIC SECTORAL ISSUES

Now that we know the reasons for product decomposition, let us see how this affects model structure.

The most immediate change is of course the multiplication of categories. According to the item, the extension can apply to a different criterion, or even two criteria at the same time.

A precise list of the items belonging to each category will be proposed as an annex.

First we can consider the **products**:

- Product 1: agricultural products (such as unprocessed food)
- Product 2: manufactured goods (such as household appliances)
- Product 3: services (such as transportation or banking).

This decomposition applies naturally to traded elements, such as consumption or exports.

But the same decomposition will apply to **productive units**, according to their output. Farms, industrial plants and travel agencies will be classified in each of our three categories, in order.

And here a distinction must be made, as a given unit can produce different goods. In that case:

- The whole unit will be classified in the **branch** of its main activity.
- It will be fractioned into products, each share being allocated to the **sector** it represents.

In our model, we shall not separate branches and sectors, and we shall use alternately the two terms. But branches and products can play different roles, and appear sometimes for the same variable, introducing a double indexation.

In addition, it might look interesting to proceed further in the decomposition, in two cases:

- For agriculture (and fishing) the model would benefit from a separation into artisanal and industrial units:

  - Artisanal units use less capital (machinery) and much more labor, with a lower productivity.
  - Artisanal employment will not be very sensitive to output (family units will often include a generally inactive work force, which can be called for if needed).
  - They also use fewer intermediate goods (such as fertilizer and fodder).
  - Their revenue is entirely appropriated by households.
  - They export a lower share of their production.

This is particularly interesting in developing countries, in which the role of agriculture is still important, especially for employment (less for output). In 2015 Vietnam, it represented 41% and 15% respectively.

- The manufacturing sector includes energy and mining. Oil, gas and ore are traded through a process different from other manufacturing products:

  - The international price is fixed, and competitiveness plays no role (even for imports).
  - Capacity also plays no role in the short run: quantities exported are decided by the exporter, and imports are directly linked to demand[160].

This means we shall separate in our example:

- Agriculture and fishing production into artisanal and industrial
- Manufacturing exports and imports into energy and non-energy products. We could also include mining; as it plays an important role in the economy of some countries (Chile, Morocco, Australia…).

This is an example of the (frequent) case where decompositions can lead to further decompositions:

- it is only for agriculture that artisanal/industrial decomposition is really efficient,
- And excluding energy from manufactured trade comes from the fact that in that regard, it behaves more like the primary product. This calls for a local correction of the general decomposition, if we want the model to present reliable trade equations.
- Finally, the evolution to the integration of financial elements calls more and more to the identification of financial firms.

10.3.4 THE DATA

For detailed data, very often the information gets scarcer. The following problems can occur:

- Detail on goods and services at current prices is known only from the production side, not the demand side.

[160] But if local production goes down, imports can increase.

- For the detailed series which are available, deflators are known only at a more aggregate level, making the computation of variables at constant prices approximate.
- Investment (and capital) is not available by product, or by investing branch, or (most often) in two-dimensional detail. The same problem can appear for the change in inventories.

Fortunately, investment is made essentially in manufacturing goods and construction (plus livestock and plants for agriculture). If the model uses a global secondary product, the dimension of the products is not required[161]. In any case, many cells in the matrix will contain null values.

Intermediate consumption by product and use is available more frequently, but not always (this means the input-output matrix is not entirely known, perhaps only for some periods).

- Employment is not detailed by branch, or by categories of firm's ownership.
- The part of the account of firms giving transfers (wages, subsidies...) is not detailed by branch.
- The period for which data is known is shorter, or more difficult to collect. For instance, full tables might be produced yearly for the current year.

10.3.5 CONSEQUENCES FOR ESTIMATIONS

If the sample gets too short, we can still try to estimate. But we should not rely on results, however favorable. We should just use them as another indication on the associated behavior, along with economic theory, observations on the way the local economy works, and estimations at the global level.

But very often the above problems will not allow to estimate individual equations at the detailed level.

One can:

- Calibrate the equations using theoretical coefficients.
- Use the results estimated at the aggregate level.
- Take the coefficients from other models, describing a similar country (Thailand for Vietnam for instance, or Austria for the Czech Republic). This option presents a surprising good point: the country used will generally be more advanced (as it has a better and more developed data system) which means that its estimations are based on a period which should represent the future of our modeled country, the actual period on which economic studies will be performed.
- Estimate the equations as a system, in which some coefficients will be forced to take the same value. The size of the "sample" is multiplied by the number of categories.

And actually, this difficulty has another silver lining: partially free from econometric restrictions, we can apply the formulations which conform the most to economic theory, in specification and value of coefficients. In particular, we can apply error-correction frameworks in each formula, leading to a model which will:

- Provide a long-term solution on any future period, without any change to specifications (with a few restrictions on assumptions).
- Separate completely the long-term specification from the dynamics leading to it.
- Allow to interpret completely and easily the coefficients (using values which conform to theory).

[161] Although investment contains a share of services (architects, patents).

10.3.6 THE PRODUCTION FUNCTION

As we have stated above, we can use different solutions for each branch.

10.3.6.1 Primary product

This is a field is quite difficult to manage. Capacity can depend on:

- Capital in product 1: Land made ready for use, plantations, and cattle.
- Capital in product 2: Machinery, more or less sophisticated.
- Public capital: availability of water and electricity, road and telecommunication network.
- Services capital: transportation, storage.
- Intermediate inputs: fodder, fertilizer.
- Climatic conditions (exogenous of course).

Formulations can be more or less complex, from a simple production function (complementary or Cobb-Douglas) to a logical system including conditions and non-continuous functions. Output could also be decomposed into more categories (agriculture/fishing/forestry) or into products (rice and coffee for Vietnam, fruits for Central America…). In the last case, physical quantities could be used (tons, liters, numbers).

But basically, the short run definition of quantities produced should not depend so much on demand, but rather on potential production including climatic factors (in the long run of course, profitability and market size will affect the creation of capacities).

This has an important impact on model properties. As the supply – demand equilibrium still has to apply, we need a new balancing element, which can be imports, exports, demand or one of its components. The simplest solution is to use imports, which means that:

- Local demand has to be satisfied[162]
- Exports are determined by world demand and competitiveness (with a limited impact as agricultural export prices are mostly determined at the world level).
- The share of global demand which cannot be satisfied locally has to be imported[163].

Other frameworks can be considered, such as:

- Local demand has to be satisfied
- Imports are controlled by local agents or the State.
- What is left from local production is exported.

Or:

- Exports are controlled by the state or local agents, according to market conditions.
- Local agents are allowed to buy the non-exported share of local production.
- Imports are also controlled.

---

[162] But remember it is endogenous, so it follows global activity.

[163] But if the price of imported goods is too high, consumption will decrease and will also move to other goods.

- This gives total local demand.

It is quite possible to consider several of these frameworks simultaneously, but only if two or three categories are defined, each with its four elements which verify the particular equilibrium.

We can expect the consequences for model properties of the actual choice to be quite important.

10.3.6.2 Industrial product

One of the main purposes of the decomposition is to identify a sophisticated production function for the manufacturing sector, leaving out the other branches for which this choice is much less natural, and would have a negative impact on the quality of a global estimation.

A moderately complex option is the Cobb-Douglas function, which we presented earlier.

10.3.7 UNEMPLOYMENT

Unemployment follows the same logic as the simple model. However, one could define different sensitivities to employment according to the branch. For instance, in construction, the lower share of qualified workers could lead job creation to hire immediately productive unemployed.

10.3.8 CHANGE IN INVENTORIES

The new element is that we should develop the equation in two directions: the good in the inventory, and the branch using it. This means that in principle we will have at least 9 variables. However, services are not stored as inventories, and most agricultural products cannot be either (in particular for a yearly model).

To set the coefficients, we can consider theory, expert advice or the ratio of total inventory change to the change in value added.

10.3.9 HOUSEHOLD CONSUMPTION

The decomposition brings a new feature: consumption has to be divided into products.

The most common option is to compute global consumption according to the single product option.

Then we separate consumption into products, using an error correction framework, with a target sharing depending on the relative prices of consumption goods. After the decomposition has been applied, we need to ensure that the sum is consistent with the total (this can be done using a system or applying a correcting factor).

Depending on the product, a high level of consumption can be maintained (representing habit forming) or negatively correlated (if the good lasts several periods). For instance, a household could get used to eating meat, or wait several years to buy a new car.

The situation is simpler if we consider two products, as we only need a single parting ratio, once global consumption is known.

10.3.10 EXTERNAL TRADE

This is the second field in which product decomposition allows to introduce differences in the formulas across products.

This originality will come from:

- Different weights: the share of exports in production, and imports in demand, is different from one product to the other (and of course between imports and exports).
- Different sensitivities: estimating the role of price competitiveness can give different answers; if the equation has to be calibrated, economic theory could lead to use different values.
- Different formulations: the role of tensions on capacities can be limited to the manufacturing product.

10.3.11 WAGES

For wages, we should use the same (theoretical) framework as in the simple model.

However, the indexation process is a little more complex. We still have a choice (in particular in the long run) between an indexation on the consumption price and the value added deflator. But if the first element is measured at the global level, the second corresponds to the branch. This means we can observe the consequences of dissymetric shocks (such as an increase in subsidies to one sector), or shocks having different sectoral impacts (like an increase in the margins of exporting firms, mostly industrial).

Also, the sensitivity to unemployment can be quite different, leading to dissymetric inflationary properties.

Finally, identifying the artisanal and industrial branches of agriculture will allow considering different productivities and wage levels.

10.3.12 PRICES

10.3.12.1 The value-added deflator

With different production frameworks, the cost variable used in the long-term target will differ. And the role of the rate of use should be less intense (or even absent) outside of manufacturing.

10.3.12.2 The production price

This is a domain in which the complexity increases: the intermediate consumption of a branch is the sum of individual two-dimensional elements, each valued at its own price (which has to be defined).

10.3.12.3 The trade prices

Once production prices are known, we can move to the trade prices. To define the price of competitors, we need to identify an average foreign production price for each of our products, which will depend on the structure of traded goods, and of customers and suppliers. We have already addressed this issue.

If we only have global production prices, we can at least use them for individual equations.

10.3.12.4 The demand prices

The above deflators allow us to compute value added and trade at current prices. As we know already the values at constant prices, the demand deflators can be defined by identities.

However, one must check that, at current prices, the data still verifies the equilibrium between demand and supply. A condition difficult to enforce in practice, and even more in simulations, in which convergence can be achieved at the global level while individual elements will diverge. Here the error-correction format can help.

A small error can be accepted, and either

- Forgotten: this is dangerous as the model will never get a perfect historical fit, once estimated equations have been fed their residuals.
- Treated as a (multiplicative) residual. We have seen that it is a much better option.

10.3.12.5 The government budget

As in the simple model, we have to define a detailed government budget. The elements associated with goods and services (demand and taxes) will have to be separated. Others will not (transfers from and to households in particular).

10.3.12.6 The EViews program

We can now build a decomposed model, based on the same framework as our single product version, and using the same concepts.

We have decided (and it should be obvious) that the problem was too complex for a full model to be presented here. We shall only give a table presenting the decomposition level applied to these concepts. It can add no, one or two dimensions to the single product case.

10.3.12.1 A list of elements

We present here a list of elements, coming actually from a 3-product operational model for Vietnam. It applies the artisanal – industrial decomposition of agriculture we presented earlier.

In the “categories” column,

T stands for total,

1, 2 and 3 for the associated products (or branches, or sectors),

1a and 1i for the decomposition of agriculture into artisanal and industrial.

2e and 2m for the decomposition of product 2 into energy and the rest (mostly associated to manufacturing).

Name	Type	Description	Units	Type	categories
CAP	Endog	Productive capacity	Constant prices	Product	1,2,3
CI	Endog	Change in inventories	Constant prices	Product	T, 1,1a,1i,2,3
COG	Exog	Consumption of Government	Constant prices	Product	1,2,3
COGV	Endog	Consumption of Government	Current prices	Product	T
COH	Endog	Household consumption	Constant prices	Product	T,1,2,3
COMPM	Endog	Imports price competitiveness (import / local production price)		Product	1,2i,3
COMPX	Endog	Exports price competitiveness (export/ foreign production price)		Product	1,2e,2i,3
COST	Endog	Cost of wages and capital	Constant prices	Branch	2
CPI	Endog	Consumption deflator	Deflator base year=2014	Product	T
CRV	Endog	Export - import ratio at current prices	Ratio	Product	T
CRVOL	Endog	Export - import ratio at constant prices	Ratio	Product	T
ER	Endog	Exchange rate (value of the local currency)	Deflator base year=2014		T
EXPG	Endog	Expenditures of Government	Current prices		T
FBAL	Endog	Financing Capacity of Firms	Current prices	Product	1i,2,3
FD	Endog	Final local demand	Constant prices	Product	T,1,2,3
FDGV	Endog	Final demand of Government at current prices	Current prices	Total	T
FDI	Endog	Foreign direct investment at constant prices	Current prices	Product	1,2,3
FDV	Endog	Final local demand	Current prices	Product	T,1,2,3
FOBAL	Endog	National commercial balance	Current prices	Product	T

GBAL	Endog	Financing Capacity of Government	Current prices	Product	T
GDPM	Endog	Market Gross Domestic Product	Constant prices	Product	T,1,2,3
GDPMV	Endog	Market Gross Domestic Product at current prices	Current prices	Product	T,1,2,3
GDPV	Endog	Total Gross Domestic Product at current prices	Current prices	Product	T
HDI	Endog	Households disposable income	Current prices	Product	T
HI	Endog	Household revenue	Current prices	Product	T
HRDI	Endog	Household disposable income in purchasing power	Constant prices	Product	T
IC	Endog	Intermediate consumption at constant prices	Constant prices	Product x Branch	(1,2,3)*(1,2,3)
ICV	Endog	Intermediate consumption at current prices	Current prices	Product x Branch	(1,2,3)*(1,2,3)
IG	Exog	Government investment at constant prices	Constant prices	Product	1,2,3
IGV	Endog	Government investment at current prices	Current prices	Total	T,1,2,3
IHH	Endog	Housing households investment	Constant prices	Total	T,2,3
INCT	Endog	Income tax	Current prices	Total	T
IP	Endog	Productive investment	Constant prices	Total	(1,2,3)*(T,1a,1i,1,2,3),T *T
IPD	Endog	Target productive investment	Constant prices	Product by Branch	2*2
IR	Endog	Current interest rate on new loans	Points		T
IRL	Endog	Long-term interest rate on new loans	Points		T
IRM	Endog	Average interest rate on current debts	Points		T
IRS	Endog	Short-term interest rate on new loans	Points		T
K	Endog	Productive capital	Constant prices	Product x Branch	(1,2,3)*(T,1a,1i,1,2,3),T

					*T
KD	Endog	Target productive capital	Constant prices	Product x Branch	2*2
KV	Endog	Productive capital at current prices	Current prices	Branch	1a,1i,1,2,3
L	Endog	Total employment	Thsds persons	Branch	T,1a,1i,1,2,3
L_G	Exog	Employment	Thsds persons		
L_M	Endog	Employment	Thsds persons		
LD	Endog	Target Firms employment	Thousand persons	Branch	1a,1i,2,3
LF	Endog	Firms employment	Thousand persons	Branch	1a,1i,2,3
LH	Exog	Household productive employment	Thousand persons	Branch	1a,1i,2,3
LP	Endog	Labour productivity	Thousand persons	Branch	T,1,2,3
LPTA	Exog	Labour productivity trend	Thousand persons	Branch	1a,1i,3
LS	Endog	Wage earners	Thousand persons	Total	1,1a,1i,2,3
LS_F	Endog	Wage earners firms	Thousand persons	Total	T,1a,1i,2,3
LS_H	Endog	Wage earners households	Thousand persons	Total	T,1a,1i,2,3
M	Endog	Imports	Constant prices	Total	T,1,2e,2i,3
MARG	Endog	Margins of firms	Current prices	Branch	1,1a,1i,2,3
MV	Endog	Imports at current prices	Current prices	Total	T,1,2e,2i,3
NIF	Endog	Net interests paid by Firms	Current prices	Branch	1i,2,3
NIG	Endog	Net interests paid by Government	Current prices		
NIX	Endog	Net interests paid to the Rest of the World	Current prices		

OIT	Endog	Other indirect taxes	Current prices	Product	T,1,1a,1i,1,2,3
PCOG	Endog	Government consumption deflator	Deflator base year=2014	Product	1,2,3
PENS	Endog	Pensions	Current prices		
PENSR	Exog	Pensions rate	Constant prices		
PEX	Endog	Export deflator	Deflator base year=2014	Product	T,1,2,2e,2i,3
PEXT	Endog	Export deflator including tariffs	Deflator base year=2014, USD	Product	1,2e,2i,3
PFD	Endog	Local demand deflator	Deflator base year=2014	Product	T,1,2,3
PFDG	Endog	Government demand deflator	Deflator base year=2014	Product	T
PFDXT	Endog	Local demand deflator excluding VAT	Deflator base year=2014	Product	1,2,3
PGDPM	Endog	Market GDP deflator	Deflator base year=2014	Product	T
PIC	Endog	Price of intermediate consumption	Deflator base year=2014	Product x Branch	(1,2,3)*(1,2,3)
PIG	Endog	Government investment deflator	Deflator base year=2014	Product	2,3
PIHH	Endog	Household housing investment deflator	Deflator base year=2014	Product	T
PIM	Endog	Import deflator	Deflator base year=2014	Product	T,1,2e,2i,3
PIMT	Endog	Import deflator including tariffs	Deflator base year=2014	Product	1,2e,2i,3
PIMU	Exog	Energy import price in US Dollars	Deflator base year=2014	Product	2e
PIP	Endog	Productive investment deflator	Deflator base year=2014	Product	T,1,2,3
POP	Exog	Population	Thousand persons		
POPAG	Exog	Population in age of working	Thousand persons		
PP	Endog	Production deflator	Deflator base year=2014	Product	1,2,3

PQ	Endog	Value added deflator	Deflator base year=2014	Product	1,2,3
PROF	Endog	Profits of Firms	Current prices	Branch	1i,2,3
PT	Endog	Tax on profits	Current prices	Branch	T,1i,2,3
PWM	Exog	World price imports weighting	Deflator base year=2014, USD	Product	2i,3
PWX	Endog	World price exports weighting	Deflator base year=2014	Product	T,1,2e,2i,3
Q	Endog	Value added	Constant prices	Product	T,1,1a,1i,2,3
QA	Endog	Target value added	Constant prices	Product	2
QV	Endog	Value added at current prices	Current prices	Product	T,1,1a,1i,2,3
R_COGV	Exog	Ratio of non-wage Govt consumption to GDP			
R_CPI	Exog	Ratio of CPI to Final demand deflator			
R_EXPG	Exog	Ratio of non-identified Administrations demand to GDP			
R_IHH	Exog	Ratio of Housing investment to household revenue		Branch	2,3
R_INCT	Exog	Income tax rate			
R_IP	Exog	Ratio of productive investment to GDP		Product x Branch	3*(1,1a,1i,2,3)
R_LFA	Exog	Share of Firms employment in total		Branch	1a,1i,2,3
R_OIT	Endog	Rate of Other Indirect taxes (than VAT)		Branch	1,1a,1i,2,3
R_PCOG	Exog	Ratio of Govt consumption to final demand deflator		Branch	1,2,3
R_PFDG	Exog	Ratio of Govt demand to final demand deflator			
R_PIC	Exog	Ratio of Interm. consumption to Demand (excluding taxes) deflator		Product x Branch	(1,2,3)*(1,2,3)
R_PIG	Exog	Ratio of Govt investment to Final Demand deflator		Product	2,3

R_PIHH	Exog	Ratio of Housing investment to final demand deflator			
R_PIP	Exog	Ratio of Productive investment to final demand deflator		Product	1,2,3
R_PTI	Exog	Rate of the tax on profits		Branch	1i,2,3
R_REVG	Exog	Ratio of other Government revenue to GDP			
R_REVQ	Exog	Ratio to production of other Household revenue from production		Branch	1i,2,3
R_REVX	Exog	Non wage non production Household revenue, in purchasing power			
R_SCF	Endog	Rate of social contributions by Firms		Branch	1,1a,1i,2,3
R_SCF_G	Exog	Rate of social contributions by Government			
R_SCF_H	Exog	Rate of social contributions by Households			
R_SCW_G	Exog	Rate of social contributions by wage earners		Government	
R_SCWA	Exog	Rate of social contributions by wage earners		Branch	1a,1i,2,3
R_SUBS	Exog	Rate of subsidies (to value added at current prices)		Branch	1,1a,1i,2,3
R_TARIFF	Exog	Rate of Local tariffs		Branch	1,2,2e,2i,3
R_TARX	Exog	Rate of Foreign tariffs		Branch	1,2e,2i,3
R_UN	Endog	Unemployment rate			
R_VAT	Exog	VAT rate		Product	1,2,3
R_WG	Exog	Ratio of Government to market wage rate			
R1_A	Exog	Share of artisanal in Agriculture			
R1_I	Exog	Share of industry in Agriculture			
RDEP	Exog	Depreciation rate of capital		Product x Branch	(1,2,3)*(1a,1i,2,3)

RELC	Endog	Labour - Capital relative cost		Branch	2
REVG	Endog	Revenue of Government			Eq364
REVQ	Endog	Revenue of Households from production, other than wages.		Branch	T,1i,2,3
REVX	Endog	Non wage non production Household revenue	Constant prices		
RFDI	Exog	Share of FDI in investment		Product	1,2,3
RMARG	Endog	Margins ratio		Branch	1,1a,1i,2,3
RPROB	Endog	Margins ratio including capital amortization		Branch	1i,2,3
RPRO	Endog	Profits ratio		Branch	1i,2,3
RS_F	Exog	Share of wage earners in Firms		Branch	1a,1i,2,3
RS_H	Exog	Share of wage earners in Households		Branch	1a,1i,2,3
SCF	Endog	Social contributions by Firms	Current prices	Branch	T,1,1a,1i,2,3
SCF_G	Endog	Social contributions by Government	Current prices		
SCF_H	Endog	Social contributions by Households	Current prices		
SCW	Endog	Social contributions by Wage earners	Current prices		
SOCB	Endog	Social benefits	Current prices		
SOCBR	Exog	Social benefits in purchasing power	Constant prices		
SUBS	Endog	Subsidies to firms	Current prices	Branch	T,1,1a,1i,2,3
TARIFF	Endog	Local tariffs	Current prices	Product	T,1,2,2e,2i,3
TC	Exog	Technical coefficient		Product x Branch	(1,2,3)*(1,2,3)
TRBAL	Endog	Trade balance	Current prices	Product	T

TTRAD	Endog	Terms of trade (export - import price ratio)	Thousand persons		
UN	Endog	Unemployment	Thousand persons		
UN_R	Exog	Unemployment	Thousand persons		
UN_U	Endog	Unemployment			
UR	Endog	Rate of use of capacities		Branch	1,2,3
VAT	Endog	Value Added Tax	Current prices	Total	T,1,2,3
W	Endog	Market wage rate	At constant prices	Branch	1,1a,1i,2,3
W_G	Endog	Government wage rate	Thousands at constant prices		
WAGE	Endog	Total wages	Current prices	Branch	T,1,1a,1i,2,3
WAGE_F	Endog	Wages paid by Firms	Current prices	Total	T,2,3
WAGE_G	Endog	Wages paid by Government	Current prices		T
WAGE_H	Endog	Wages paid by Households	Current prices	Total	T,2,3
WCOST	Endog	Unitary wage cost	Thousands at constant prices	Branch	1,2,3
WD	Exog	World demand at constant prices	Index at constant prices	Product	1,2e,2i,3
WF	Endog	Wage rate paid by Firms	Thousands at current prices	Branch	2,3
WH	Endog	Wage rate paid by Households	Thousands at constant prices	Branch	2,3
X	Endog	Exports	Constant prices	Total	T,1,2,2e,2i,3
XV	Endog	Exports at current prices	Current prices	Total	T,1,2,2e,2i,3

10.4 A MULTI COUNTRY, SINGLE PRODUCT MODEL

We shall now address the issue of building, maintaining and using a model describing the economy of more than one country. We shall also consider the similar issue of a single country but multi-region model.

10.4.1 FIRST ISSUE: THE MODELS

To produce a multi – country model, we first need single country ones. Two issues have to be considered:

- The degree of similitude of the country models
- The description of the Rest of the World

10.4.1.1 The single country models

We shall suppose the country models follow the lines used above. What matters now is the differences allowed in the specifications, from one country to the other. Several cases can be considered:

A - The models are allowed to be completely different. This even applies the concepts used, such as the product decomposition. The only condition is that, once identified the elements linking the models (basically the trade variables at constant and current prices) they can be converted in such a way that they can be transferred as input to the other models. But for instance, a single product and multi product models can coexist. The transfer process will apply aggregation or disaggregation as required. In this case the names of the variables will probably be different too.

B - The models use the same concepts but the economic framework is different. For instance, a model can use a Cobb-Douglas production function, the other a CES.

C - The models use the same equations, but the coefficients are different. Some exceptions can appear, for instance an additional lagged variable, but the logic remains the same. In general, these differences should be due to estimation, which means that a common data set must be available.

D - The models use the same equations and the same coefficients. This means that either the formulas have been estimated as a panel (by polling all the countries data together) or calibrated.

The choice of option depends on the case:

- Option A can be considered if the models already exist, which simplifies the process. The case is even stronger if the models have to keep on living as independent versions.

This is the case for instance for the LINK project, which links into a multi-country version, under the management of the United Nations' Department of Economic and Social Affairs, and the University of Toronto, models maintained as independent forecasting tools by country institutes.

Another example is the grafting of existing a national model in an international environment, also existing. For example, the MODUX model managed by the Statistical Institute of Luxemburg (STATEC) has been integrated into the MacSim environment (see below) in order to enrich its projections and especially its impact analyzes. Such a possibility is provided by the managers of MacSim (which basically means the author of this book).

The drawback is of course that the models can follow different economic properties, and the differences in mechanisms, leading to different sensitivities to shocks, are not necessarily justified by reliable country specifics (if they are, then option A is clearly optimal). As stated earlier, econometrics based on the same data can justify formulations with quite different properties, even by pure chance (two economists with the same economic philosophy can reach different formulations through the meanders of successive estimations).

- Option B is option A with no initial models. Then the models will be easier to manage, but the above problems remain.

- Option C is even simpler to manage and its properties easier to interpret. The only drawback is the constraints on model formulations, which forbids introducing country-specific formulations. But of course, some minimal differences can be accepted, such as a longer lagged structure, the long run indexation of wages on different deflators.
- Option D limits even more country specifics. What only remains is the effect of the size of the country, the structure of its partners (imports and exports) and structural parameters like labor and capital productivity, the weight of the public sector or the tax structure, or the sharing of production revenue between the firms and the workers.

The only justification is just this: that the only difference in properties comes from undisputable elements, the consequences of which they describe faithfully.

From the above elements we can derive the following suggestion: if the models do not preexist, option C is clearly preferable, with some elements from option B if they are really justified, in particular by the observation of country characteristics.

For instance, for an oil producing country the trade in energy elements could be individualized. For the interest rate, identifying the policy of the central bank should lead to the introduction of the associated rule. And maybe if there is sufficient proof that factor substitution follows a zero or unitary elasticity, the associated production function (complementary factors or Cobb-Douglas) should be used.

10.4.2 SECOND ISSUE: INTER-COUNTRY TRADE

Once the single country models have been defined, they have to be merged into a single entity.

The main issue is to make consistent the imports and exports of each country, at constant and current prices.

The obvious option is to associate behaviors with the actual decision processes:

- Global imports are decided by the importing countries.
- Export prices are decided by the exporters, and import prices are computed by weighting identities, taking into account the relative exchange rate. The decision can depend on the targeted market.

Once imports of a country are decided, they must be shared between the countries which export to it. There are two options:

- Identifying the individual imports by their source and converting them into individual exports from the supplying country. This is obviously the most logical solution. It has the essential advantage of making exports and imports consistent at the global level.
- However, some models use a more devious technique: they compute global potential exports of a country by weighting by constant shares the imports of its partners.

Of course, this option gives to total exports and exports different values, and some correction must be applied. The main advantage of this method is to allow estimation of exports based on actual global values.

Also, it avoids identifying a large number of trade flows (growing as the square of the number of countries).

We favor strongly the first option.

- It identifies more variables, but the associated information is important. The interpretation of the consequences of shocks (on domestic or trade related variables) is much clearer and more informative.
- It does not call for corrections, giving automatically consistent results.
- It does not require more information: the weights used by the other method can be used to compute the trade flows.
- But the most important in our sense is that it locates the decision process where it belongs: exports are not decided by exporters, but by importing countries which, once they have decided to import, chose among the potential suppliers, essentially by considering relative prices.

Of course, it does not allow global estimation. The individual trade flow equations will have to be either calibrated (perhaps using values from global estimations) or estimated using panel techniques.

From now on, we shall concentrate on this technique.

10.4.3 A CONSISTENT METHOD: MACSIM

We shall present now a consistent method for defining in sequence the above elements. It is used by the MacSim system (2001, 2012).

We shall start, not with imports but with the price system, as prices depend only on local elements, while imports depend on price competitiveness.

10.4.3.1 The price system

We start by establishing a coherent system for trade prices. For this we shall start from the export price formulations (estimated or set). For clarity we shall use calibrated equations, but the method applies to estimated ones also.

We shall suppose that the traditional equation applies to each client country. But this time the exports destination will be identified.

We shall consider I as the exporting country, j as the destination.

To get individual export prices, we correct by the exchange rate:

$$Log(pex_{i,j}) = a \cdot Log(pp_i) + (1-a) \cdot (Log(pp_j ch_i / ch_j) + e \cdot t + f$$

Where i is the exporter and j the client, and ch_i is an index, representing the evolution (from the base year) of the value of a common currency (in practice the US dollar) compared to the currency of country i. If the currency of i depreciates, the index grows.

$$pim_{j,i} = pex_{i,j} \, . ch_j / ch_i$$

and we get the global import price through an average:

$$pim_j = \sum_i m_{j,i} \, pim_{j,i} \Big/ \sum_i m_{j,i}$$

(We shall see later how we compute the trade flows at constant prices).

10.4.3.2 The foreign rates of use

In the single country models, we have seen the fundamental role of the rate of use of capacities for the local country. But we supposed the foreign capacities to be infinite.

Now we can consider, to define the imports of a country, the rate of use for providing countries. If the rate of use increases in Germany, it will lose market shares in France, relative to other exporters but also to French manufacturers.

To get the average, we shall use the same method as above (without currency corrections).

$$urx_j = \sum_i m_{j,i} \, ur_i / \sum_i m_{j,i}$$

10.4.3.3 The global imports

We can now determine the global imports of country i, by modifying slightly the equation from our single country model, to consider the capacity of exporters:

$$Log(m_i/td_i) = b \cdot [Log(ur_i) - c \cdot Log(urx_i)] + d \cdot Log(pm_i/pp_i) + e$$

To simplify the explanation, we have presented only the long-term part. The dynamic equation should contain the same change.

This means that, as in the single country models, a general decrease in the available capacity of exporters will reduce their exports, through a substitution effect. The coefficient c, lower than one, takes into account the larger size of this set of countries.

10.4.3.4 The trade flows

Finally, we have to separate imports into individual exports. Once again, we shall take into account relative competitiveness, and fluctuations in available capacities, relative to the above average. Actually, rather than estimating an equation (we do not have the associated data) we shall use:

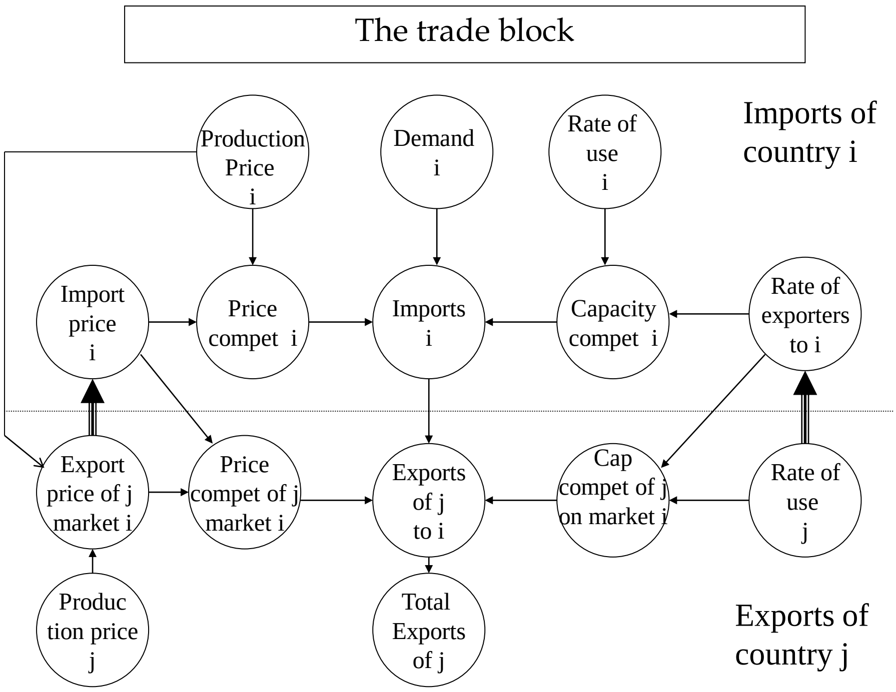
$$m_{i,j} = m_i . b_{i,j} \left[1 - \alpha_i (pex_j ch_i / ch_j - pim_i) - \beta_i (ut_j - utx_i) \right]$$

which means that (as for the single country models) exporters to one country will increase a « natural » share with competitiveness and available capacity, this time relative to their competitors. More precisely, the A and B matrixes (one derived from the other) represent the sharing of bilateral trade associated with identical rates of use and competitiveness for all countries (not necessarily the most natural assumption, nor the present situation). Scalars α and β are fixed, and will be set at a higher value than for the import equation (between 0.5 and 1.0). This should represent the higher versatility of the choice of its providers by a country, once the decision has been made not to buy local products.

One will observe that this technique guarantees the identity of the sum of individual exports with its global value, without any correction. The weighted sum of deviations from averages, computed using the same weightings, is of course zero.

Of course, the coefficients can be different from one market to another, but not within one market.

The system can be summarized with this graph:



In addition to the above, we can now introduce accounting equations:

Exports from country i to j:

$$x_{i,j} = m_{j,i} \cdot ch0_j / ch0_i$$

Where $usd0_i$ represents the base year value of the currency of country i, in US Dollars.

Total exports are computed as a sum

$$x_i = \sum x_{i,j}$$

The average export price of country i:

$$pex_i = \sum pex_{i,j} \cdot x_{i,j} / x_i$$

10.4.3.5 Third issue: common behaviors

But we have also to consider behaviors common to a set of countries, which take into account global variables. The most obvious case is the presence of a global formal agreement, such as the use of a common currency by the European Monetary Union.

This option is very easy to implement: one just has to create a new "country" using aggregated concepts, considering only the relevant variables.

The identification of the countries in the union can be done by equations considering all countries, and applying a Boolean variable (1 = yes, 0 = no) to each concept.

10.4.3.6 The Rest of the World

If the scope of the countries described is large enough, we cannot expect that shocks applied to a large subset or the whole model will have no feedback through the economy of the remaining countries. This is true in our MacSim system, which models most of Europe (the Euro zone + the United Kingdom) and also the US and Japan. Moreover, with the technique we are proposing (computing the shares of each country in each country's imports) endogenizing the Rest of the World allows its exports to conform to the global framework, avoiding any unbalance.

But of course, we are not going to estimate a ROW model: we do not have the data, and the heterogeneity is too high.

What we propose is:

- To define the trade elements as usual: imports are shared according to the above framework, and exports participate to the above competition.
- In addition, to create a set of multipliers, defining the feedback of the Rest of the World exports on its GDP and imports, and linking its imports to its final demand.

It should be quantified to give a response similar to the single country models, with the exception of infinite (or quasi-infinite) capacity and high sluggishness of prices.

This can be (and actually has been) done quite simply.

10.4.3.7 Estimating with common values, using the SYSTEM feature

If we use option C above, we do not have to give to all coefficients the same value. Of course, we have to estimate the constant term, if we want the mean residual to be zero for each country (and EViews not to refuse a system without any coefficient to estimate).

We should probably estimate also the possible trends, which should be country specific.

Let us show an example of application, based on our MacSim model. We will consider the consumption equation, using the same explanations as in the single country model. As our purpose is not studying the economic implications, our comments will be minimal.

We want the change in consumption to depend on:

- The change in real income (over the last year).
- The change in unemployment.
- Inflation (the « real holdings » or Pigou effect).
- An error correction term.
- A supposedly negative time trend, representing the increase in the wealth of households and their saving potential (in particular their accession to housing ownership).

Our system will consider 14 countries: Belgium, Germany, France, Great Britain, Italy, Japan, Luxemburg; Netherlands, USA, Spain, Finland, Greece, Ireland, Portugal (the MacSim system aggregates the last four, but we can separate the estimations). We will give to each country the same weight, without considering their size.

Using different coefficients for each equation, we will define the following system:

```
for %1 BE DE FR GB  IT JP LX NL  US ES FI  GC IR PT

genr {%1}_coh_ec=0

delete c_{%1}_coh

coef(10) c_{%1}_coh

__coh.append dlog({%1}_coh)=c_{%1}_coh(1)*.25*log({%1}_hrdi/{%1}_hrdi(-4))
+c_{%1}_coh(2)*d({%1}_unr)+c_{%1_coh(3)*log({%1}_pc/{%1}_pc(-4))
+c_(%1}_coh(4)*log({%1}_coh(-1)/{%1}_hrdi(-1))+c_{%1}_coh(5)
+c_{%1}_coh(6)*({%1}_t-2008)*({%1}_t<=2008)+{%1}_coh_ec
next
```

_coh.sur(p)

And get the following results: a table (some programming is involved).

	Dymamic	Unempl	Inflation	Error Cor.	Constant	Trend	See/DW
Belgium	0.243	-0.517	-0.049	-0.018	-0.001	-0.000	0.004
	3.491	*-3.776*	*-2.349*	*-1.609*	*-0.904*	*-4.022*	*0.808*
Germany	1.479	0.158	-0.315	-0.347	-0.027	0.000	0.007
	4.576	*0.269*	*-2.496*	*-2.400*	*-2.449*	*0.989*	*2.336*
France	0.550	-0.188	-0.006	-0.039	-0.002	-0.000	0.006
	3.383	*-0.613*	*-0.159*	*-1.067*	*-0.460*	*-0.306*	*2.554*
Great Brtain	0.299	-0.595	-0.092	-0.047	0.002	-0.000	0.010
	2.070	*-1.531*	*-3.006*	*-1.430*	*0.823*	*-2.309*	*2.286*
Italy	0.043	-0.398	-0.021	-0.019	-0.002	-0.000	0.007
	0.495	*-2.037*	*-1.410*	*-2.756*	*-0.961*	*-2.759*	*1.464*
Japan	0.391	-1.056	-0.107	-0.081	-0.004	-0.000	0.011
	2.689	*-1.351*	*-3.162*	*-2.104*	*-1.401*	*-0.143*	*2.580*
Luxemburg	0.781	0.073	-0.258	-0.118	-0.038	-0.000	0.008
	3.250	*0.067*	*-2.529*	*-1.768*	*-1.502*	*-1.143*	*1.555*
Netherlands	0.213	-0.826	0.064	-0.102	0.001	0.000	0.010
	1.247	*-1.924*	*1.422*	*-3.105*	*0.435*	*0.901*	*2.392*
USA	0.185	-0.507	-0.124	-0.035	0.005	-0.000	0.006
	1.523	*-3.040*	*-3.917*	*-1.188*	*2.435*	*-1.294*	*1.951*
Spain	0.220	-0.463	-0.018	0.006	0.008	0.000	0.006
	1.770	*-2.949*	*-0.718*	*0.427*	*2.889*	*0.108*	*1.725*
Finland	0.238	-0.573	-0.033	-0.024	0.004	-0.000	0.011
	2.274	*-2.984*	*-0.996*	*-0.708*	*1.736*	*-1.208*	*2.172*
Greece	0.274	0.252	0.058	-0.007	0.011	0.000	0.019
	1.415	*0.617*	*0.731*	*-1.094*	*1.123*	*0.603*	*2.399*
Ireland	0.794	0.912	-0.032	-0.022	0.001	-0.000	0.010
	4.256	*4.061*	*-0.683*	*-1.810*	*0.110*	*-0.957*	*2.039*
Portugal	0.350	0.284	-0.471	0.000	0.010	-0.001	0.007
	1.243	0.947	-2.449	0.008	1.170	-2.687	2.797

The estimations:

Equation: DLOG(BE_COH)=C_BE_COH(1)*.25*LOG(BE_HRDI/BE_HRDI(-4))+C_BE_COH(2)*D(BE_UNR)+C_BE_COH(3)*LOG(BE_PC/BE_PC(-4))+C_BE_COH(4)*LOG(BE_COH(-1)/BE_HRDI(-1))+C_BE_COH(5)+C_BE_COH(6)*(BE_T-2008)*(BE_T<=2008)+BE_COH_EC

Observations: 136

R-squared	0.432945	Mean dependent var	0.005667
Adjusted R-squared	0.411135	S.D. dependent var	0.005789
S.E. of regression	0.004443	Sum squared resid	0.002566
Durbin-Watson stat	0.831542		

Equation: DLOG(DE_COH)=C_DE_COH(1)*.25*LOG(DE_HRDI/DE_HRDI(-4))+C_DE_COH(2)*D(DE_UNR)+C_DE_COH(3)*LOG(DE_PC/DE_PC(-4))+C_DE_COH(4)*LOG(DE_COH(-1)/DE_HRDI(-1))+C_DE_COH(5)+C_DE_COH(6)*(DE_T-2008)*(DE_T<=2008)+DE_COH_EC

Observations: 52

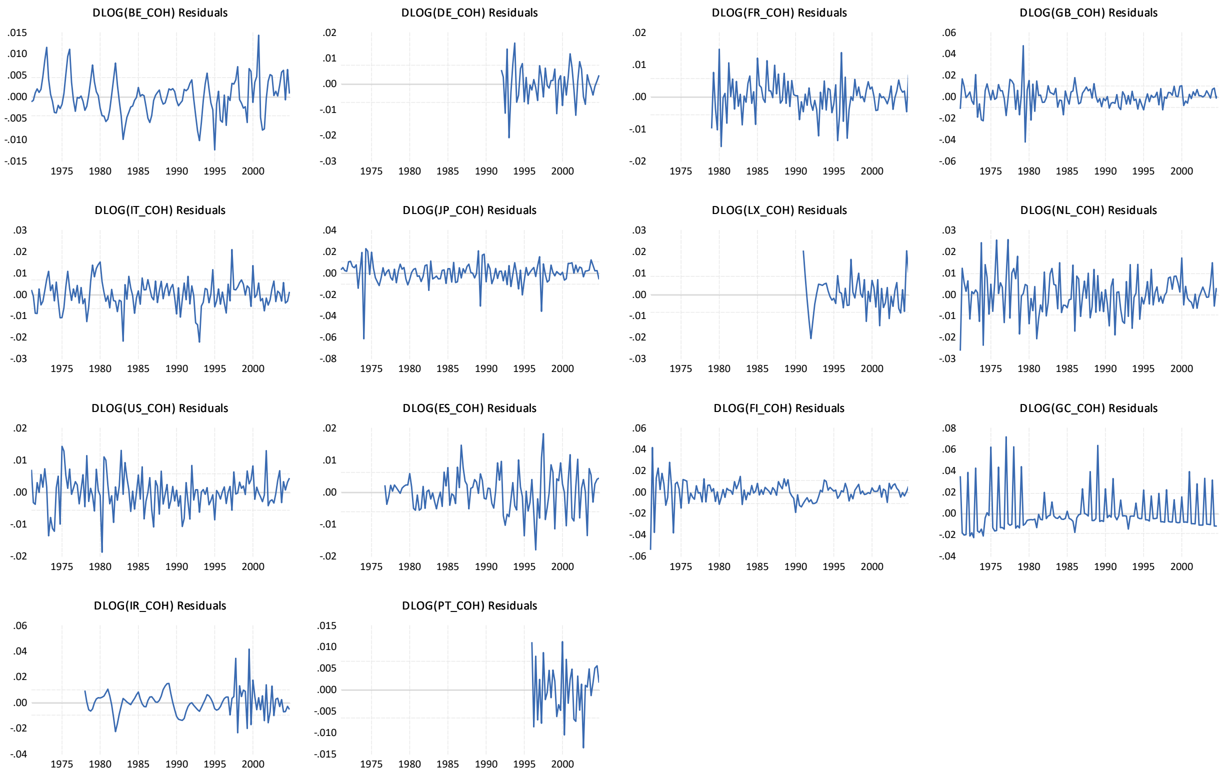
R-squared	0.382124	Mean dependent var	0.003249
Adjusted R-squared	0.314963	S.D. dependent var	0.008780
S.E. of regression	0.007267	Sum squared resid	0.002429
Durbin-Watson stat	2.389686		

Equation: DLOG(FR_COH)=C_FR_COH(1)*.25*LOG(FR_HRDI/FR_HRDI(-4))+C_FR_COH(2)*D(FR_UNR)+C_FR_COH(3)*LOG(FR_PC/FR_PC(-4))+C_FR_COH(4)*LOG(FR_COH(-1)/FR_HRDI(-1))+C_FR_COH(5)+C_FR_COH(6)*(FR_T-2008)*(FR_T<=2008)+FR_COH_EC

Observations: 104

R-squared	0.138231	Mean dependent var	0.004755
Adjusted R-squared	0.094264	S.D. dependent var	0.005932
S.E. of regression	0.005646	Sum squared resid	0.003123
Durbin-Watson stat	2.559371		

and the residuals (we can see that for Greece the data is not seasonally adjusted):



We can observe that most coefficients are significant with the right sign.

Now we can decide to force the dynamic coefficient to take the same value across countries; we just have to modify the equation, replacing the country code by “to” for this coefficient:

```
coef(10) c_to_coh
_coh.append dlog({%p}_coh)=c_to_coh(1)*.25*log({%p}_hrdi/{%p}_hrdi(-4))+c_{%p}_coh(2)*d({%p}_unr)
+c_{%p}_coh(3)*log({%p}_pc/{%p}_pc(-4))+c_{%p}_coh(4)*log({%p}_coh(-1)/{%p}_hrdi(-1))+c_{%p}_coh(5)
+c_{%p}_coh(6)*({%p}_t-2008)*({%p}_t<=2008)+{%p}_coh_ec
```

We get:

	Dymamic	Unempl	Inflation	Error Cor.	Constant	Trend	See/DW
Belgium	0.274	-0.523	-0.044	-0.018	-0.001	-0.000	0.004
	7.472	*-3.794*	*-2.238*	*-1.574*	*-0.879*	*-4.074*	*0.816*
Germany	0.274	-0.170	-0.360	-0.338	-0.025	0.000	0.008
	7.472	*-0.263*	*-2.585*	*-2.106*	*-2.070*	*0.159*	*2.386*
France	0.274	-0.390	-0.003	-0.036	-0.000	-0.000	0.006
	7.472	*-1.359*	*-0.089*	*-0.983*	*-0.001*	*-0.015*	*2.549*
Great Brtain	0.274	-0.636	-0.094	-0.047	0.002	-0.000	0.010
	7.472	*-1.680*	*-3.217*	*-1.453*	*0.890*	*-2.480*	*2.292*
Italy	0.274	-0.322	-0.018	-0.017	-0.001	-0.000	0.007
	7.472	*-1.614*	*-1.223*	*-2.511*	*-0.313*	*-1.625*	*1.438*
Japan	0.274	-1.173	-0.108	-0.076	-0.004	-0.000	0.011
	7.472	*-1.546*	*-3.201*	*-1.966*	*-1.529*	*-0.509*	*2.593*
Luxemburg	0.274	-0.895	-0.264	-0.130	-0.037	-0.000	0.009
	7.472	*-0.881*	*-2.527*	*-1.886*	*-1.405*	*-1.094*	*1.550*
Netherlands	0.274	-0.723	0.059	-0.095	0.001	0.000	0.010
	7.472	*-2.162*	*1.357*	*-3.110*	*0.472*	*0.899*	*2.401*
USA	0.274	-0.472	-0.115	-0.035	0.005	-0.000	0.006
	7.472	*-2.968*	*-3.858*	*-1.158*	*2.341*	*-1.159*	*1.946*
Spain	0.274	-0.428	-0.012	0.007	0.007	0.000	0.006
	7.472	*-3.085*	*-0.502*	*0.493*	*3.004*	*0.276*	*1.725*
Finland	0.274	-0.567	-0.031	-0.023	0.004	-0.000	0.011
	7.472	*-2.998*	*-0.944*	*-0.701*	*1.686*	*-1.150*	*2.178*
Greece	0.274	0.238	0.058	-0.007	0.011	0.001	0.019
	7.472	*0.588*	*1.439*	*-1.172*	*1.178*	*0.652*	*2.399*
Ireland	0.274	0.548	-0.127	-0.038	0.012	0.000	0.010
	7.472	*2.784*	*-3.674*	*-3.358*	*4.739*	*0.989*	*1.973*
Portugal	0.274	0.218	-0.465	0.008	0.011	-0.001	0.007
	7.472	0.774	-2.412	0.157	1.452	-2.744	2.808

with an average value for this particular coefficient, and similar results for the others.

Finally, we can apply the constraint to all four economic elements:

	Dymamic	Unempl	Inflation	Error Cor.	Constant	Trend	See/DW
Belgium	0.268	-0.379	-0.031	-0.011	-0.000	-0.000	0.005
	7.758	*-5.964*	*-3.539*	*-3.403*	*-0.462*	*-4.997*	*0.767*
Germany	0.268	-0.379	-0.031	-0.011	-0.002	-0.000	0.009
	7.758	*-5.964*	*-3.539*	*-3.403*	*-0.523*	*-1.335*	*2.581*
France	0.268	-0.379	-0.031	-0.011	0.002	-0.000	0.006
	7.758	*-5.964*	*-3.539*	*-3.403*	*1.243*	*-1.671*	*2.556*
Great Brtain	0.268	-0.379	-0.031	-0.011	0.005	-0.000	0.010
	7.758	*-5.964*	*-3.539*	*-3.403*	*2.623*	*-0.778*	*2.287*
Italy	0.268	-0.379	-0.031	-0.011	0.000	-0.000	0.007
	7.758	*-5.964*	*-3.539*	*-3.403*	*0.260*	*-2.986*	*1.458*
Japan	0.268	-0.379	-0.031	-0.011	-0.001	-0.000	0.011
	7.758	*-5.964*	*-3.539*	*-3.403*	*-0.245*	*-2.826*	*2.619*
Luxemburg	0.268	-0.379	-0.031	-0.011	0.007	0.000	0.009
	7.758	*-5.964*	*-3.539*	*-3.403*	*1.987*	*1.267*	*1.635*
Netherlands	0.268	-0.379	-0.031	-0.011	0.003	-0.000	0.010
	7.758	*-5.964*	*-3.539*	*-3.403*	*1.637*	*-1.044*	*2.433*
USA	0.268	-0.379	-0.031	-0.011	0.006	-0.000	0.006
	7.758	*-5.964*	*-3.539*	*-3.403*	*4.956*	*-0.535*	*1.831*
Spain	0.268	-0.379	-0.031	-0.011	0.005	-0.000	0.006
	7.758	*-5.964*	*-3.539*	*-3.403*	*3.441*	*-0.577*	*1.671*
Finland	0.268	-0.379	-0.031	-0.011	0.004	-0.000	0.011
	7.758	*-5.964*	*-3.539*	*-3.403*	*1.758*	*-1.686*	*2.121*
Greece	0.268	-0.379	-0.031	-0.011	0.021	0.001	0.020
	7.758	*-5.964*	*-3.539*	*-3.403*	*3.398*	*1.530*	*2.218*
Ireland	0.268	-0.379	-0.031	-0.011	0.008	0.000	0.011
	7.758	*-5.964*	*-3.539*	*-3.403*	*3.155*	*0.172*	*1.990*
Portugal	0.268	-0.379	-0.031	-0.011	-0.001	-0.001	0.007
	7.758	-5.964	-3.539	-3.403	-0.364	-2.128	2.642

Now all economic coefficients are considered significant, profiting from the number of observations attributed to them.

The associated programs, including table production, are available from the author.

10.4.3.8 The sequence of production

To produce a multi-country model, one should proceed through the following sequence:

10.4.3.8.1 Deciding on the model framework

We shall suppose that we build our model from scratch, which in this case should lead to a set of country models using the same economic framework but estimated separately, with limited variations in the estimated formulas.

In this case one should consider:

- The data available.
- The type of behaviors the model should describe.
- The expected degree of detail.

This should allow deciding:

- The variables used by the model (they have to be useful and available).
- The full form of the identities.
- The type of behaviors which should be estimated, including the explanatory elements.

10.4.3.8.2 Producing a set of country models

Then one should build a set of country models, using the technique described earlier. Of course, a global data set has to be produced, containing series for each single model, with the same name except for an identifier. For instance, French exports can be called FR_X.

As usual, concerning behavioral equations, one should start with declaration of intent before moving to estimations.

Several sequences can be considered. One can build models one by one, or apply simultaneously each phase (data production, framework, estimation, testing) to the whole set. Probably the best solution is to start with one country, go through the whole process, and apply it to the set of remaining countries, this time phase by phase simultaneously.

This allows measuring the difficulties, to solve the general problem on a single case, and defining at a limited cost the full procedure, which will then be applied to the global case. In particular, if problems were to be met at a later stage, this would probably call for reconsidering the initial ones. Having to modify the full set of programs can then become quite costly.

The choice is not so simple, however, as:

- Running estimations for the same equations on a set of countries can be organized quite easily using EViews loops.

But the estimation processes must succeed together, as the equations are not allowed to differ too much. An interesting direction found for a given country, but not associated to its specific features, can be applied only if it works also for other countries.

This means more iterations will be needed than for a single country model. The only issue is to limit their number.

Two additional elements:

- After the estimation process has been applied to the whole set, it is possible that results proved acceptable for a sizable share of countries (perhaps above 50%). Then if the remaining countries do not present too specific characteristics (like having a different political system, or a different structure of production), their formulas can be calibrated using the average of the successful results. This is what we have done for the first version of the MacSim system.
- If it proves difficult to individualize a given equation, one can resort to panel estimation, estimating the equation for all countries at the same time, and forcing some coefficients (at least one) to take the same value for all countries. This technique has been presented earlier, including a practical example. This is what we have done for the second version of the MacSim system.

10.4.3.8.3 Assembling the models

Once satisfying single country versions have been produced, assembling them is essentially technical.

10.4.3.8.3.1 COMPUTING THE TRADE FLOWS

If formalizing the trade between countries uses explicit individual flows (our preferred option) one needs a matrix of transfers. It can represent either the share of exporters in each country's imports, or the share of clients in each country's exports. Both matrixes are needed, but one can be built from the other in an exact way. The matrix elements should be considered as series. If it is only available for some periods, it should be extrapolated, but in any case, the modeler should be allowed to make it change with time in the future. This is particularly essential for transition economies: the share of Russia in Polish exports has decreased following the breakup of the Soviet Union, but it is not stabilized and could increase again in the future. Also, establishing a trade agreement between countries will increase the share of their mutual trade.

From these matrixes the individual trade flows can be computed. If the matrix is known only at current (or constant) prices, a global assumption has to be made, the simplest being that the same export (or import) price applies to one country whatever its destination (or origin).

Of course, if we use official statistics to add up the result in one dimension, the other dimension will not give the actual value. For instance, if we start from imports and share them between exporters, the sum of individual exports of one country will show a difference with the national accounts. The issue is the size of this difference: it should not be higher than the residual of the equation if it had been estimated, otherwise a correction could be made (but this would bring back an inconsistency at the world level).

Of course, the model simulations and forecast results will be consistent, as the model system is consistent by construction. And the consequences of shocks too, as the difference between two consistent systems. We are using the error observed on the past as an indication of the global precision of the model on the future, not of its internal consistency.

10.4.3.8.3.2 FORMALIZING THE TRADE FLOWS AT CONSTANT PRICES

We can do it using the framework defined earlier. We have only to consider the following technical issues:

- The export equations of single models will disappear.
- The trade flows have two dimensions and call for an additional loop.
- If the trade flows are not available as series, the coefficients will have to be calibrated. They must be common for a given market, and we cannot expect the choice between import origins to use the same price elasticity as the choice between imports and local products. We can suppose that the role of relative prices us more important for deciding between producing locally and importing and choosing between different providing countries.

Again, one advantage of our proposal is that trade flows will add up to imports without any correction, as the weighted differences to the weighted mean cancel out naturally.

10.4.3.8.3.3 FORMALIZING THE TRADE DEFLATORS

It imports are decided by importers, the transaction price is decided by the exporter.

This time we can estimate the global export price of each country (depending on local costs and the price applied by competitors). Then we can apply the results to each export market, using as competitors' price the one specific of this market.

Apart from this, we have to suppose that for each exporter the same formula applies to all its clients, including coefficient values.

10.4.3.8.3.4 INTRODUCING IDENTITIES

Finally, we need to compute identities, like global country imports and exports at current prices.

10.4.3.8.3.5 PRODUCING COMMON ELEMENTS

The other issue is the presence of an agreement between countries, defining a variable common to each of them, based on one or several common elements.

For instance, under purchasing power parity, the exchange rate of the Euro will apply to all Euro zone countries, based on Euro zone inflation.

Equations will have to be built:

- Summarizing the explanatory indicators in the group.
- Computing the common element, generally giving a choice between several options (real interest rate or Taylor rule, fixed real or nominal exchange rate).

The formula must allow changing the list of countries concerned. This is very easy:

- For computing the indicators, the element of each country in the model will be multiplied by an indicator (1=in, 0=out) to indicate if it participates in the weighted sum.
- For the use of the common result, an option must be present in each country's equation, using the same indicator (if 1, use the common value, if 0, use a country specific formula).

Even in the absence of agreement, it can be useful to compute common indicators (for instance the GDP of the full European Union, or of ASEAN).

10.4.3.8.3.6 THE REST OF THE WORLD

The Rest of the World "model" will need the usual trade flows equations, and in addition:

- An equation linking GDP and local demand to exports.
- A global imports equation linked to RoW demand and price competitiveness (its prices can be affected by the other countries inflation). Introducing an influence of the rate of use should depend on the size of the RoW.

10.4.3.8.3.7 SOLVING THE MODEL

The procedure for solving the model is the same. The number of equations is just larger, and the testing procedure more complex.

Another problem is the convergence: with less exogenous variables the solution is less anchored on assumptions, and the probability of divergence might increase. But one can also consider that an individual problem can become diluted in the global equilibrating process. Nothing is clear on that issue.

The most important issue is perhaps the price system: with less anchoring, even if each country's inflation converges in the long run to a common value, this can happen at highly heterogeneous levels. If a country's price indexes are higher than the others, the inflation created normally by a demand shock can be offset by the increase in the share of its lower priced imports. The overall effect can be deflationary, introducing properties which should be considered abnormal, especially as they are due to the distance to the base year.

This can be solved by managing the system of residuals in such a way that the long-term prices remain rather close to each other. The targeting technique we will present later can be used here[164].

10.4.3.8.3.8 USING THE MODEL

Of course, if this model is more expensive to produce, it also allows a wider range of studies, and makes them more reliable compared to the single country case. As we have said earlier, the demand multiplier for the first MacSim version is 20% higher for France using the whole system than if we run the French sub-model by itself.

In addition to a better study of changes in the local assumptions of a single country, for this country and abroad, this type of model allows in particular to study consequences of:

- Global events such as a slowing down of the world growth, or a change in the price of oil, taking into account its consequences for all countries. We no longer need to make assumptions consistent for the rest of the world. This adds feedback to the method we have presented earlier.
- Global trade agreements.
- Global policies (for instance a decrease in VAT rates in the European Union).

These elements can be assessed as a whole, and by comparing their consequences for each country.

All this can be done taking into account the role of common rules, such as the establishment of the Euro Zone, which can be changed at will.

10.4.4 THE EVIEWS PROGRAM

Again, the sequence of EViews programs is too complicated and too problem-specific to be presented here. However, a model is working, and is used in operational studies. Its elements are available and can be provided through a special agreement.

10.5 A REGIONAL MODEL

Building a regional model presents the same features as in the multi-country case, with some differences:

- There is no currency issue, nor any difference in the definition of monetary rules. But the level of interest rates can vary.
- Variables such as various deflators are not independent from one region to the other.
- Variables such as unemployment have to be computed at the country level, or at least regional variables will influence strongly other regions, especially if we consider migrations.
- “Importing” from other regions is much easier. This means in our sense that instead of a sequence: definition of imports then separation into providers, all forms of demand from one region should be separated immediately into providers, the region itself being treated like the others. But of course this region must be given a higher potential share, at the same level of competitiveness. Some products will have to come from the region, in particular services, even if intermediate consumption can be ”imported”. For instance, restaurants in Paris can propose seafood from Brittany.

[164] We use it quite often in practice.

- But this information has less chance of being available, in the absence of regional trade measurements.
- Migrations are much easier, and should be taken into account, through both their determinants and their effects. Transfer of household revenue is also easier and more natural.
- One should also consider the difficulty in transporting goods from one region to the other.

All these elements have been taken into account in one of our projects.

10.6 A MULTI COUNTRY, MULTI PRODUCT MODEL

Basically, the issues combine the features (and the difficulties) of both dimensions: countries and products.

- In addition to combining the features described earlier, the only problem lies in our opinion in the much higher complexity of the problem. Even if the EViews programs can be compacted through the use of loops, the actual number of variables is much higher, which means that the control of model properties will be more complex. In practice it will have to be conducted first at a global level, then as a detailed one.
- Also, the assumptions for forecasts and response to shocks will be more difficult to produce, especially if one wants to profit from the additional detail (which should be expected).
- Also, the requirements in terms of data are of course more important, maybe even unrealistic.

Nevertheless, we have produced a working version for 3 regions and 3 products.

10.6.1 THE EVIEWS PROGRAM

Although such a program has actually been created, its size is too important to present if here.

The technology is not too complex, as it just combines the elements from its two dimensions.

CHAPTER 10: QUASI ACCOUNTING MODELS

Until now, our proposals on model building supposed

- The availability of a large set of numerical information both in scope and length, with a short periodicity if possible.
- A good insight in the way the local economy is working, both on the whole and for each specific behavior.

In that situation, one can:

- Devise a global framework for the model.
- Establish a consistent set of formulas which describe accurately the local behaviors, possible considering several options.
- Test each of these formulas using currently accepted methods, based on enough information to make the results conclusive one way or the other.

Once this is done, the model can be tested, and options reconsidered. Testing is applied again, until convergence is obtained to a satisfying version, in both economic and statistical terms.

Unfortunately, these conditions are not always present, in particular for developing countries for which:

- The statistical system is not efficient, and has not been set in place for a long time.
- The national accounts system is not precisely established.
- There is no consensus on the way the economy is working.
- Anyway, the behaviors evolve very quickly.
- The unmeasured informal sector plays an important role.

Even when enough information is available, one might want to get results in a detail for which information is too limited and statistical testing cannot be considered.

One can resort to using a Quasi-Accounting model.

The basic idea is replacing behavioral equations by "Quasi-Accounting" ones, in which the behavior is still present but simplified, to a given degree.

11.1 THE ORIGINALITY OF THE APPROACH

In terms of complexity, in particular the number of equations, a QAM is not necessarily less developed. The lower sophistication can be used to develop two aspects: the number of products can be increased, and the number of operations too.

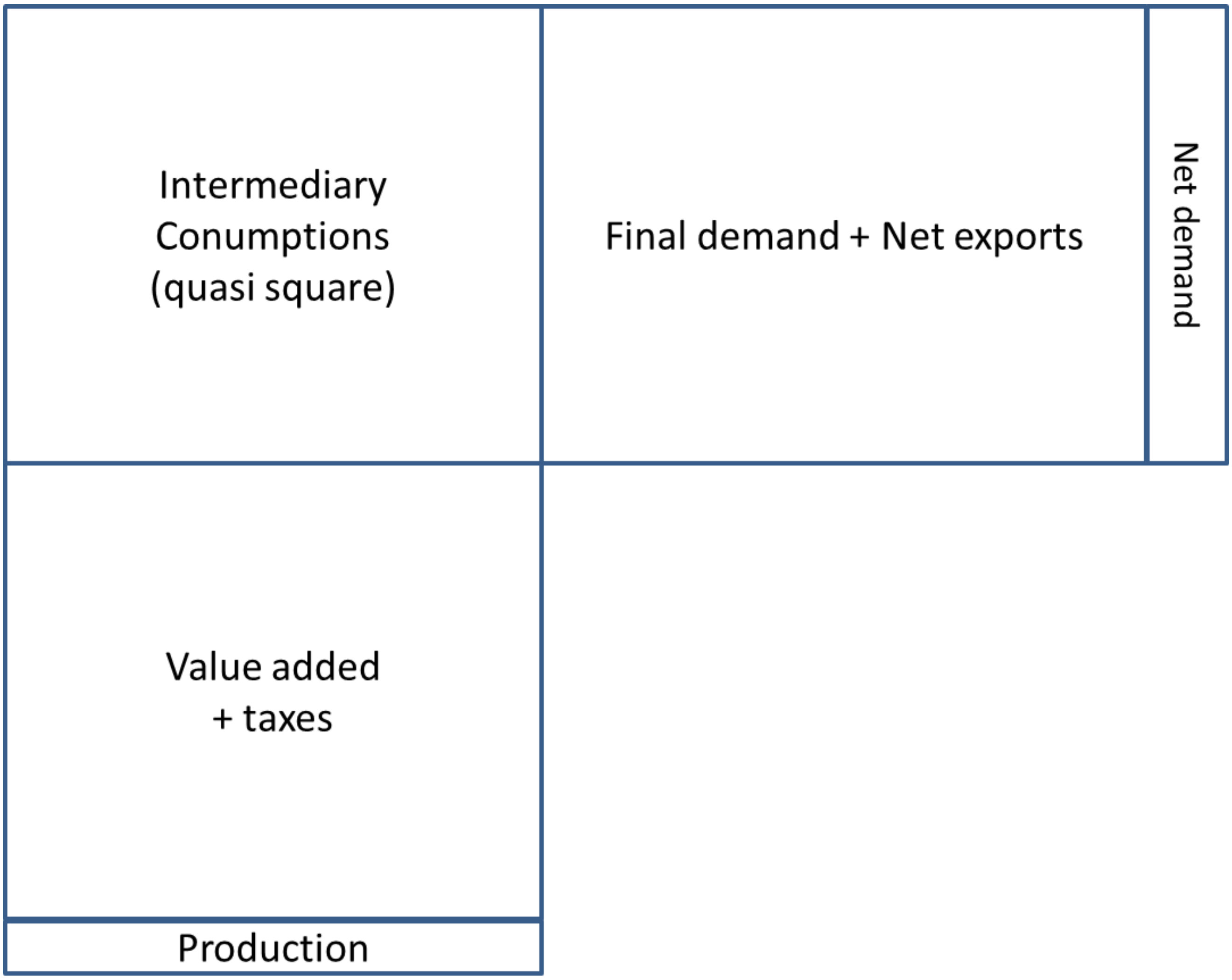
11.1.1 THE PRODUCT DETAIL

As long as econometrics are not involved, the number of products treated is only limited by data availability.

For instance, employment by product can be defined as an ratio to its value added, and imports a share of its total demand.

11.1.2 THE OPERATIONS DETAIL

To complete the specifications, one clearly needs to define an input-output table, around which the other operations will be defined, through the following framework.



This framework ensures that supply and demand are balanced for each product. Of course this calls for a complete Input-Output Table for each period (a sparse availability will introduce residuals). TThis description can be completed by a full economic table, with agents in columns and operations in lines, balanced also in transfers by identifying for each of them the emitter and the receiver.

Finally, as with Stock-Flow Consistent models, we can consider developing the financial account, both in flows and in stocks.

Two issues can be raised:

- The above table is defined at current prices. It looks difficult to consider only current elements. The economic diagnosis relies for a large part on real elements, if only the specific supply – demand balance and the productive process (investment – capital – GDP).

One simple solution would be to identify a system of exogenous prices, and use it to compute the real values. But this poses two problems: the supply and demand balances would no longer hold true, and the consequences of shocks would not take into account the difference in price evolutions (for instance the import prices are not so sensitive to local inflation).

Some behaviors apply to elements at constant prices, like the quantity of one product needed to provide a unit of another.

- These tables do not consider supply, and there is no control over the feasibility of production, and the sensitivity of trade values to available productive capacities, let alone the impact on prices.

However, once capital and employment are defined, a kind of production function can be introduced. For instance, the calibrated investment equation can stabilize the ratio of GDP to capital.

11.2 THE LEVEL OF SIMPLIFICATION

The simplification can be applied at several levels.

The simplest solution is to consider only ratios. For instance imports can represent an exogenous share of total demand, and consumption a share of household revenue One has just to specify correctly the base. Imports should be a ratio to total demand (including intermediate consumption) and VAT a ratio to final demand (including imports and excluding exports). For employment, the model should divide output by an exogenous labor productivity. In a model freed from econometric constraints, it is essential to define causalities using the relevant elements.

Proceeding a little further, one could define investment, not as a share of GDP, but as a ratio to capital, depending on the growth of GDP. Coefficients could be calibrated, with capital defined as the sum of investment and the non-discarded share of the previous value.

Even further, capacity could be defined by applying a ratio to capital (a complementary factors option) or even a function of capital and labor (à la Cobb-Douglas). Then the output gap could enter the investment formula, ensuring its disappearance in the long run.

This output gap could also influence both exports and imports.

We can also define the change in inventories as proportional to the change in GDP of the same product. This option corresponds to yet unsold products. For the intermediary goods not yet proceeded, a specific variable should be constructed, using as weights the vector of technical coefficients associated with the product.

As to endogenizing the prices, this is a much more complex problem…

11.3 THE GOOD POINTS

The Quasi Accounting model is not statistically validated, and its level of simplification will be generally very high. But is has its own advantages.

- No econometric constraint has to be applied. This frees the economist from a long, often tedious and risky process (remember David Hendry). It leaves him a choice, which can come from:
  - Theoretical considerations
  - Its own experience in building structural econometric versions.
  - The literature.

For a multi-product model, coefficients can be set at the same value, or be modulated according to the product. For instance, the production of minerals can be independent from capacity. Estimating coefficients consistent with the nature of the product is quite an impossible task, even if the data is available.

Formulations can even be different. Production can be exogenous, or follow a complementary factors or Cobb-Douglas (even CES?) formula.

- An error-correction specification can be applied, without going through the testing process of stationarity and cointegration.
- With no econometrics, one can use shorter series.
- The data needed is clearly identified, plus capital and labor. If one uses the framework described above, the data should be directly available from statisticians working on national accounts and their tables, and the model results presented in a the very format the users are accustomed to, including advisors to decision makers.
- The chosen specifications are easier to understand, and to explain. Summarizing the choices can be done using a single equation in most cases. In EViews, the equation can be specified once in a loop over products.
- Properties are also easier to control and modify if needed.
- It is probable (but not provable) that the simplification will limit the probability of convergence problems. The model should be more linear in practice, and the Newton algorithm should prove more efficient (Gauss-Seidel too, actually).
- Once you have built a successful model, you can export the programs to another case. The main difference should relate to the number of products (a single parameter) or maybe the detail of operations.

11.4 TWO EXAMPLES

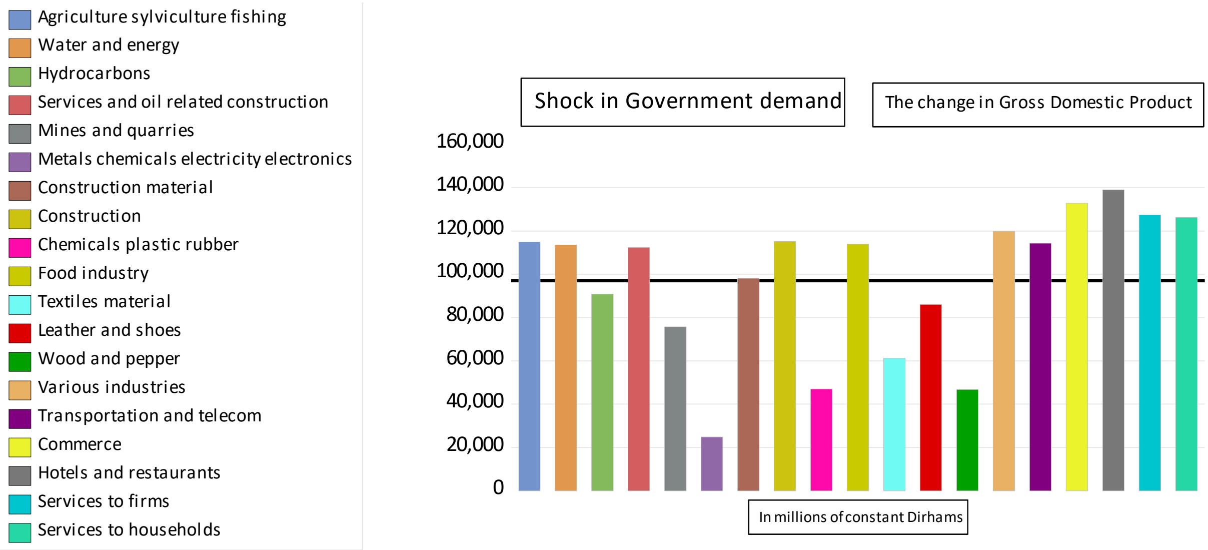
A little less than 10 years ago, we built our first two Quasi Accounting models.

The first one was part of a statistical project improving the Algerian economic information system (project AMECO). The statistical office provided a series of annual tables, according to the above framework, with a detail of 19 products. This led to the production of a Quasi-Accounting model, along the lines presented above.

In the following tables, we present the consequences of a 1% shock on demand, separately for each product (we use the exogenous Government demand, so the shock should be considered technical rather than realistic in terms of policy).

Of course, these results should call for a specific study. We will content ourselves with showing the interest of using a Quasi-Accounting model, compared to a structural version. The elements we will present are rather straightforward.

The first table shows the consequences for GDP itself, after 4 years, measured in millions of constant dirhams. The black line represents the ex-ante shock, which allows us to observe the Keynesian multiplier.

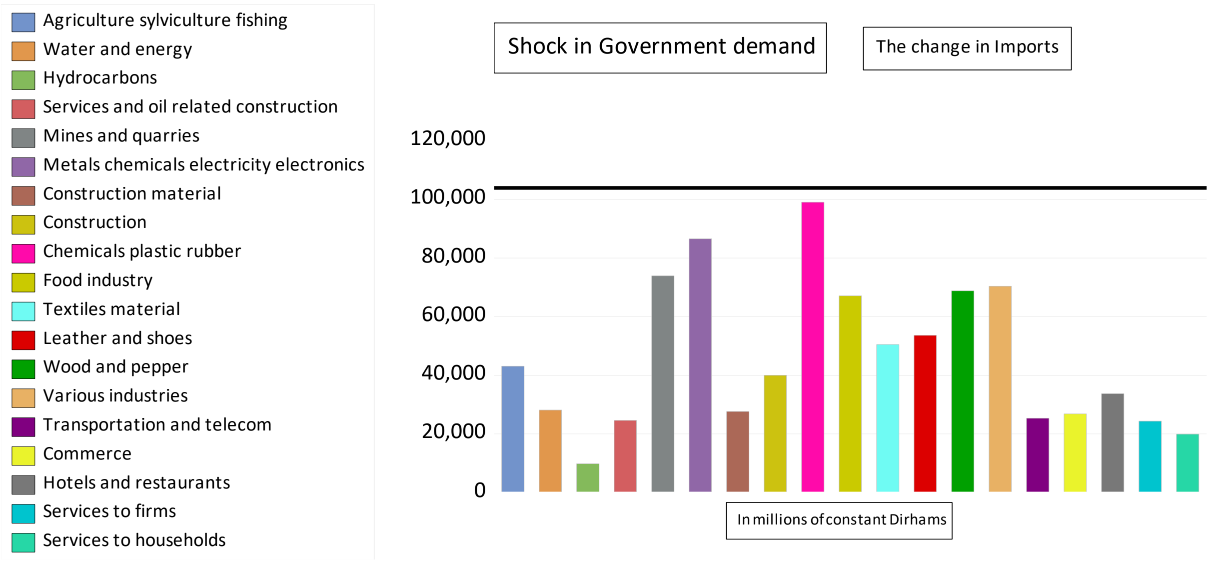


Not surprisingly, the highest values correspond to elements produced locally, and which call for additional local production, either as intermediate goods or investment.

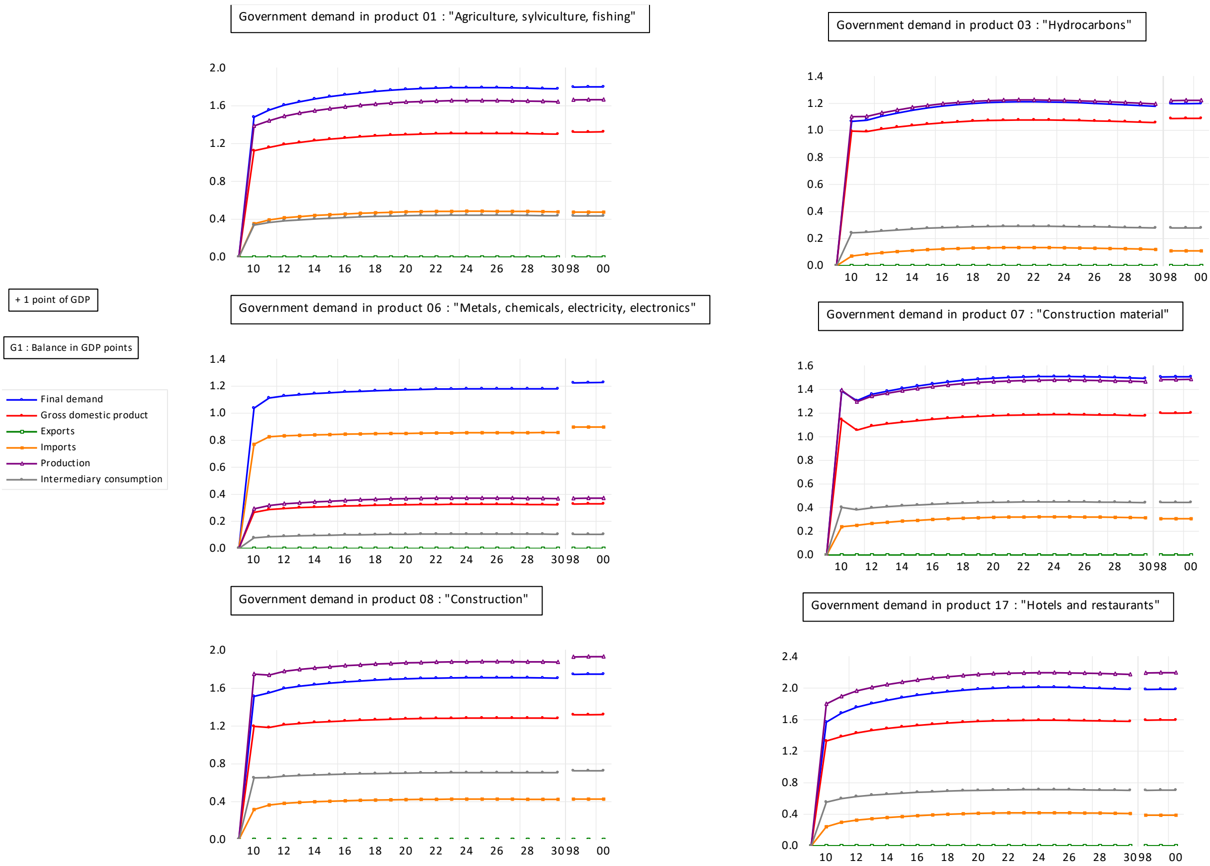
This corresponds to primary goods procuced locally (on the left) or tertiary services (on the right). Services present the highest values, especially “Hotels, cafes and restaurants” a product which, in addition to being produced locally, consumes a lot of intermediary goods and also calls for construction (which itself is produced locally, even though some of its intermediary consumption is imported…). We could probably travel along the chain further, but the model computation does the whole job.

The second table presents the consequences for imports. The values present as expected a negative correlation to the first table (a statistical test would certainly confirm it). But the correlation is not always strong: hydrocarbons call for little imports but also little intermediate consumption, investment and labor. And the food industry is largely imported in Algeria but calls for high local inputs (first of all agricultural products).

Again, the black line stands for the ex-ante increase in demand.



Finally, we present the decomposition of the supply-demand equilibrium for specific products, used in the above arguments. They confirm the conclusions.



Now, we shall consider a Vietnamese model built for the Ministry of Agriculture and Rural Development, in 2012.

This model starts again from full tables (including an I/O matrix) identifying this time 138 products and produced by the General Statistical Office.

Due to the purpose of the model (and limitations on memory size of computers[165]), we have compacted the tables into 52 products, with a focus on agriculture and related categories (like packaged food, leather garments, and veterinary services). Considering an I/O matrix alone (at constant and current prices, with the associated deflators) would call for 3 x 139 x 139 = 57963 equations.

At present, the model lacks some of the possible improvements

- Prices are still exogenous.
- The output gap has an impact on investment but not external trade.

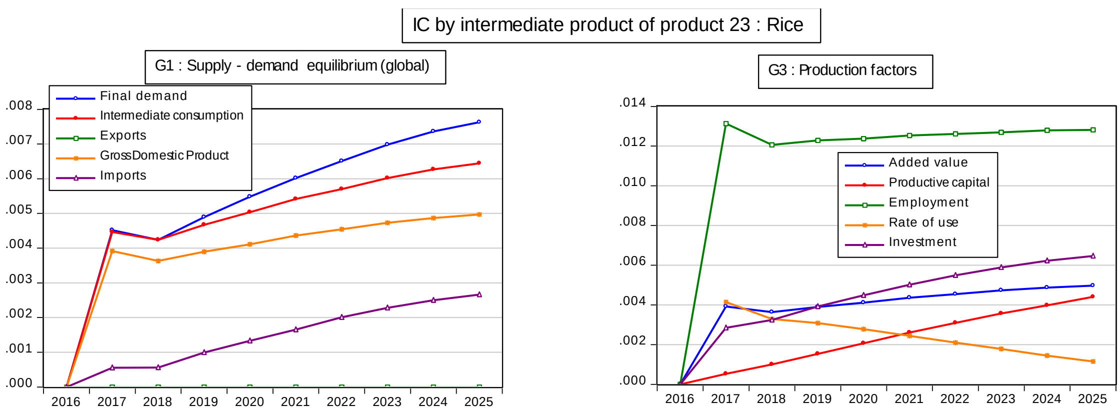
This means that:

- Supply oriented policies have no impact.
- The multiplier is higher than usual.

Of course, this could be improved.

Let us shock the demand for product 23: processed rice (patty is product 1). At the global level, we get the following results (in percentage).

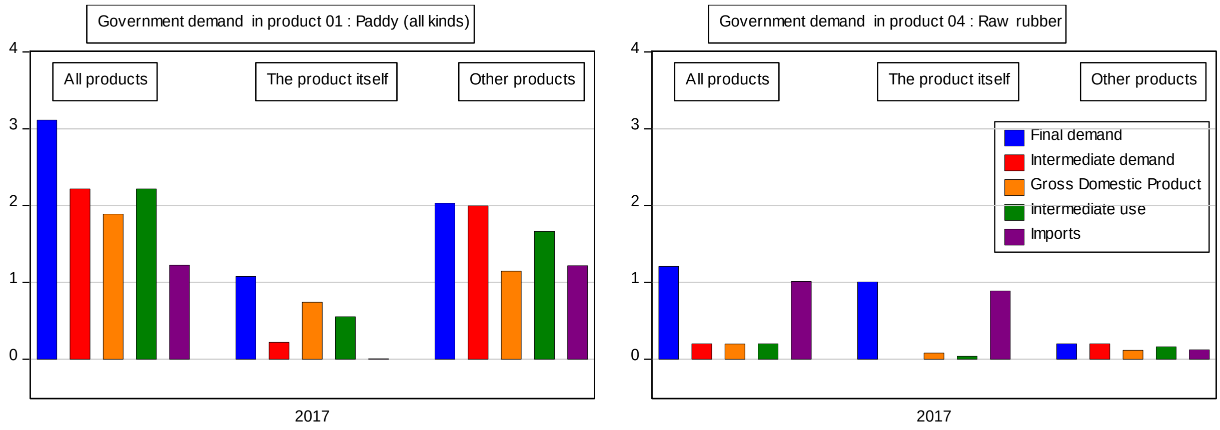
We can observe the increase in global demand and the decrease in the output gap.



Now let us compare the consequences of a demand shock for two primary products: paddy (with low intermediate consumption and almost no imports) and raw rubber (even if Vietnam has planted hoveas, it isalmost entirely imported, with very low intermediate consumption). We separate the ex-post impact on the product itself and the others. We use the same scale for both graphs.

[165] This would probably no longer be true.

We can observe that only the first shock extends to local production and demand (intermediate but also final), with a very high multiplicative effect. Of course, other products would give fewer extreme results.



To complete the diagnosis, let us observe the consequences of an increase in exogenous demand by 1% of GDP, for different products. To introduce this decision, we used Government demand, even though in real life Government does not necessarily have a use for the particular product.

We shall present the relative change in supply-demand elements, and (more explicitly in our opinion) the changes in GDP points of the same elements, and of a decomposition of local demand (total and final).

We have concentrated on products showing a panel of reactions as wide as possible.

Differences can come from:

- Most importantly, the share of imports in the original demand.
- The quantity of intermediary goods needed, and their own import share.
- The investment needed to produce it and the intermediary goods.
- The number of jobs created, and the consumption associated to the additional household revenue.

All the subsequent waves of influence are considered, as with the multiplier effect.

Although our purpose is not to produce a study on the subject, we can observe that:

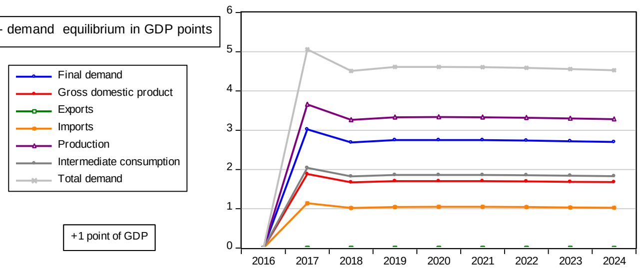
The diagnosis calls for relative variations, but the decomposition in absolute terms can provide more information. It can concern the share on imports in demand, the importance of intermediary consumption and the decomposition of demand.

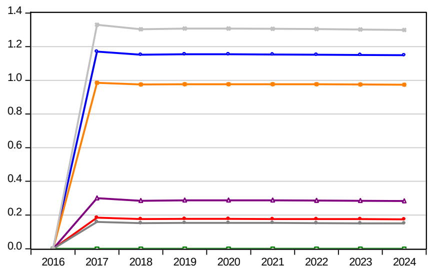
For instance:

Government demand in product 01 : Paddy (all kinds)

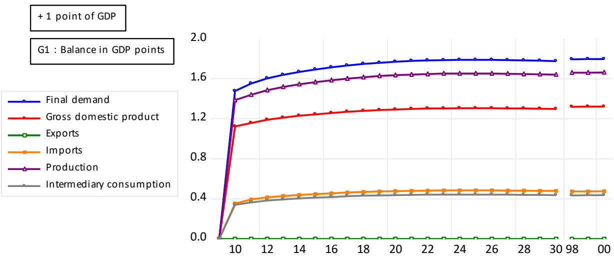
Government demand in product 04 : Raw rubber

G1 : Supply - demand equilibrium in GDP points

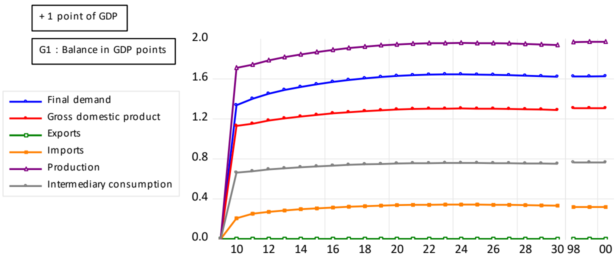




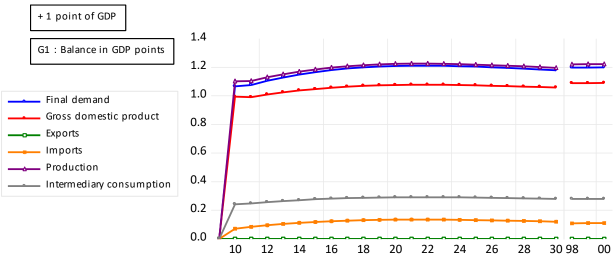
Government demand in product 01 : "Agriculture, sylviculture, fishing"



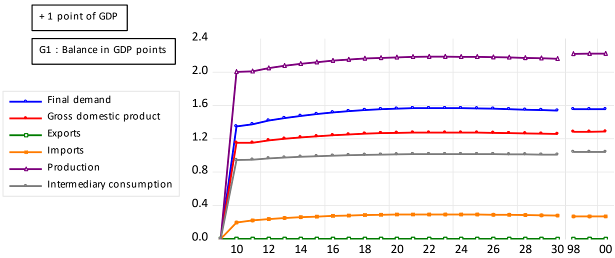
Government demand in product 02 : "Water and energy"



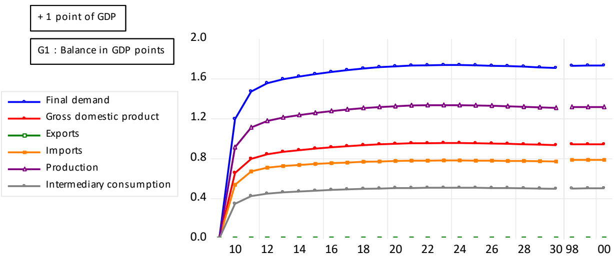
Government demand in product 03 : "Hydrocarbons"



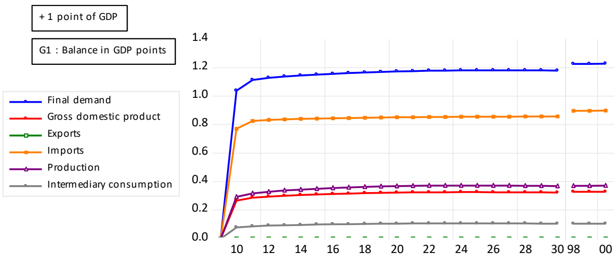
Government demand in product 04 : "Services and oil related construction"

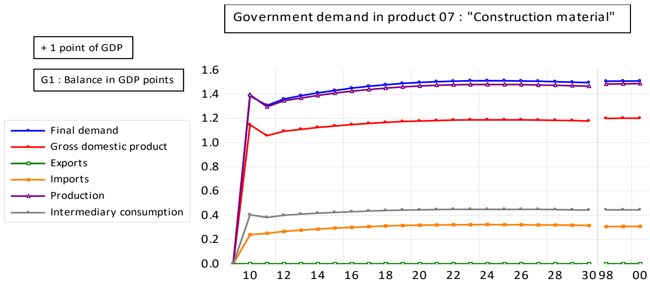


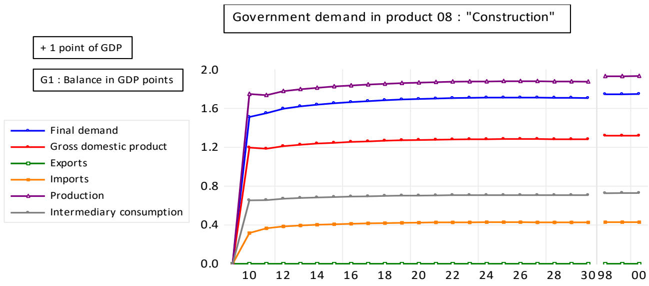
Government demand in product 05 : "Mines and quarries"

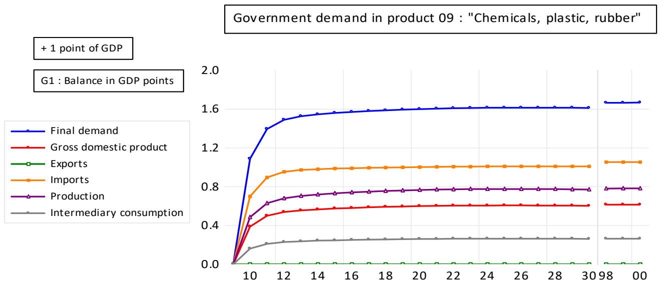


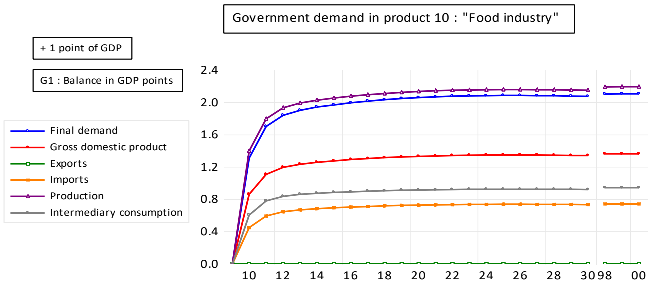
Government demand in product 06 : "Metals, chemicals, electricity, electronics"

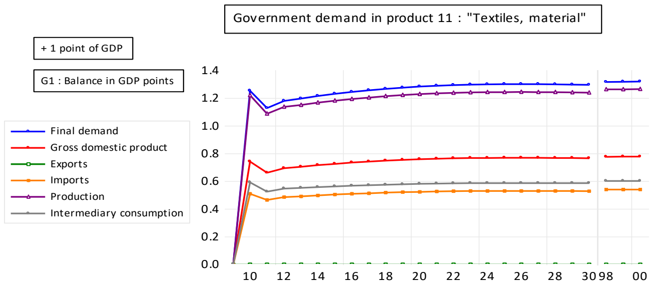


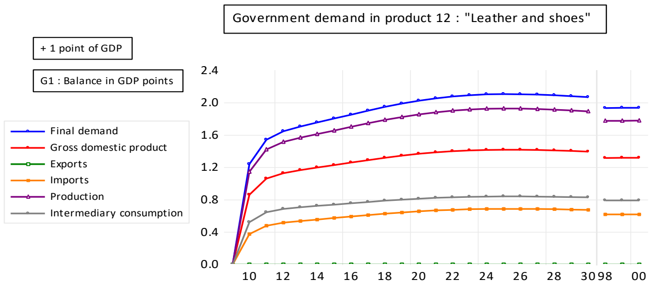

Government demand in product 07 : "Construction material"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption

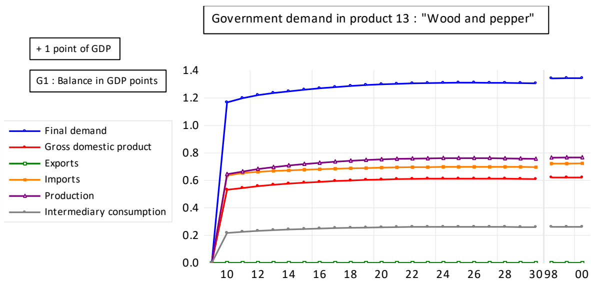

Government demand in product 08 : "Construction"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption

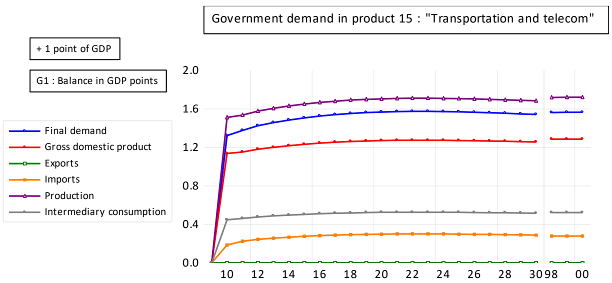

Government demand in product 09 : "Chemicals, plastic, rubber"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption

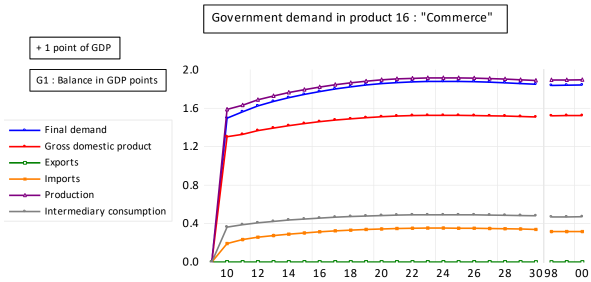

Government demand in product 10 : "Food industry"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption

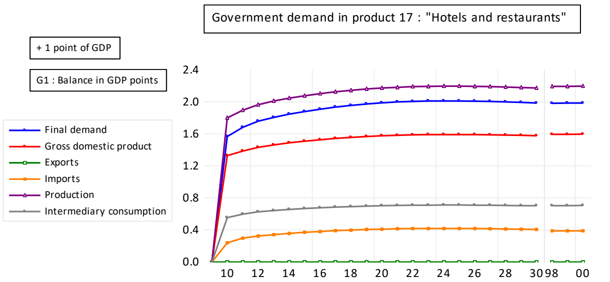

Government demand in product 11 : "Textiles, material"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption

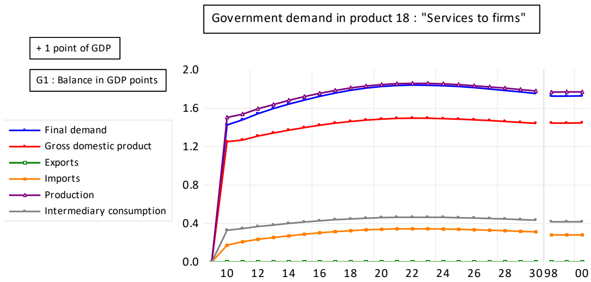

Government demand in product 12 : "Leather and shoes"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption

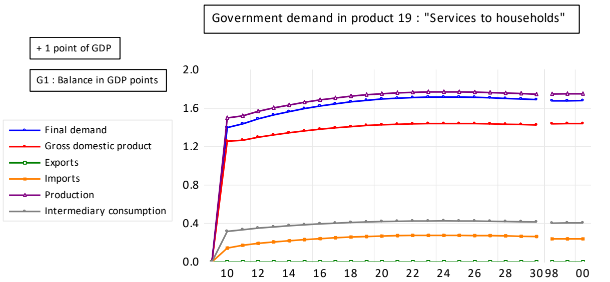

Government demand in product 13 : "Wood and pepper"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption


Government demand in product 15 : "Transportation and telecom"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption


Government demand in product 16 : "Commerce"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption


Government demand in product 17 : "Hotels and restaurants"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption


Government demand in product 18 : "Services to firms"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption


Government demand in product 19 : "Services to households"
+ 1 point of GDP
G1 : Balance in GDP points
Final demand
Gross domestic product
Exports
Imports
Production
Intermediary consumption

11.4.1 WORKING WITH BOTH TYPES

We have already stated that in some cases, using a SEM (Structural Econometric Model) is impossible, and a Quasi Accounting model is de facto the only option available.

It is also possible for some economists or institutions to decide that both options are possible, but a QAM gives the most interesting results (detail versus sophistication) or the cheapest (this quite clear, provided the data us available). But if a structural model is already operational, would a QAM help?

We can consider two options.

- The first is to use the two models separately in the same study.
  - The structural model will give more sophisticated information, in particular on the short and medium-term dynamics.
  - The QAM will give more detailed information, and assert better the structural implications, as we have just seen.

The main problem lies in the differences in the global diagnostic. One can force the two solutions to give the same global results, but this is not feasible for the response to shocks.

- The second is to use the QAM as help to improve the results of the SEM. Many options can be considered.
  - One can use the QAM as a prologue, aggregating the detailed assumptions as global elements. This can produce new ex-ante assumptions. But one can also run the QAM and aggregate the results to provide the same assumptions, now ex-post according to the QAM. This can concern:
    - * The external trade: World demand by product / importer, the foreign deflators
    - * Production: capital productivity, labor productivity, the technical coefficients
    - * The tax rates: VAT, other indirect taxes, social security, subsidies
  - But one can also consider running both models simultaneously.

Let us take employment as an example. One can:

* Estimate detailed equations for employment, formalizing labor productivity and using an error correction framework giving the employment levels.
* Compute global GDP using the SEM.
* Separate GDP into products, possible ex post from the QAM.
* Apply the estimated employment system.
* Aggregate employment and feed it into the SEM to run it again.

This process profits from the fact that employment is not an accounting variable. And the set can be run in a single EViews, program.

  - Finally, one can use the QAM to disaggregate the results of the SEM, without any feedback.

CHAPTER 11: STOCK-FLOW CONSISTENT MODELS

While the author has a long experience of building, using and maintaining structural models, and even Quasi-Accounting ones, his interest in Stock-Flow Consistent models initiated just two years ago, in the course of a project for Vietnam. Actually, this is the reason why our example will consider this country.

So this chapter will be even less affirmative than usual, and give even more the priority to technical issues, compared to economic theory. One should not be surprised if we do not describe in detail the foundations of this field, or provide a survey of existing models. This is done much better by specialists, in the references (concentrating on actual surveys) provided at the end of the book.

What we intend here is to provide the reader with the information on how to build his own version of an SFC model, leaving to him the choice regarding the formulations, the causalities, and even the identified elements, as the decision to consider some variables as negligible is not always straightforward. The associated EViews programs are available on demand and can be extended easily to a new case.

In the previous chapters, we have described the causalities and the balances in the "real" model, and addressed the optimization behavior of agents, for instance in the choice of factors by firms, and the bargaining in the determination of the wage rate. What we intend to do here is to extend this approach to the financial field. By upgrading its description, our goal is to obtain a fully consistent and homogenous model, without favoring one of the aspects.

This means that our approach will not change very much. We will establish the accounting framework for an SFC model (or the SFC part of a traditional model), and provide the reader with the practical tools required to build his own version. We will leave to him the development of the theoretical aspects, either on his own or using the principles developed in the referenced papers.

In the following paragraphs, with a growing level of affirmation, we will provide a general description of the field, describe the technical framework including a full table presentation, propose a set of accounting causalities between concepts using exogenous keys, and consider possible endogenizations of these keys.

Then we will proceed a little farther, by showing how the real and financial sides can be merged, and their links taken into account. We will reproduce some of our usual shocks, and observe how even a simplistic SFC approach modifies and completes the diagnosis.

This chapter should be considered as a simple (maybe sometimes naïve) initial proposal. The comments we will hopefully receive will help to update the present version. An actual cooperation with more advanced specialists would even be a better solution and would be quite welcome.

12.1 THE GENERAL APPROCH

The Stock-Flow Consistent field of models presents four interesting characteristics compared to traditional macroeconomic versions.

- They describe full economic tables, in which columns (agents) and lines (operations) are completely balanced.

This means that for each transfer (such as wages or interests on loans) the model must describe the origin (firms or households) and the destination (households or banks).

For goods and non-financial services, they must be part of vertical balances, and supply must be equal to demand. However, lines cannot be balanced, as the transfer between buyer and supplier is not identified. For instance, we do not know if an imported good will be bought by a firm or a household (the problem can be partially solved by introducing an artificial intermediary agent).

Even for a single country model, this calls for a detailed financial account, describing cash ("high powered money"), loans, deposits (both in local and foreign currency), bills and bonds, and equities.

This calls also for describing the impact of these elements on the real side itself.

Obviously, this feature guarantees model consistency, and the absence of "black holes". Actually, it should be required from any model, and has been present in most truly operational models, included the MESANGE model managed by the French Ministry of Finance.

- They consider stocks. Each transfer must reduce a stock for an agent and increase a stock for another. Financial stocks are not impacted by time, but not goods and non-financial services. Consumption goods disappear as they are bought, and capital is subject to depreciation.

These stocks generate revenue, and influence behaviors, in particular ratios stabilizing mechanisms. The stability of the model is improved and can be controlled.

- These stocks are consistent across agents. For instance, the amount of bonds issued by Government must be equal to amount detained. But this is obviously guaranteed by the two previous conditions.

- Each transfer actually appears four times (a feature called "quadruple book-keeping"). A transaction by an agent impacts its account, and (twice) the account of its partner in the transaction. For instance, if a household runs a negative cash account it can ask for an equivalent bank loan. This transaction will appear in the "cash" and "loans" lines in the Households column, and twice in the same lines in the Banks column.

In addition, the financial account introduces several behaviors of agents, maximizing their utility. In particular:

- Households decide how to spend their savings: housing investment, deposits, equities, bonds.
- Non-Financial Firms decide to borrow or issue equities.
- Government decides to borrow or issue bonds.
- The Rest of the World trades its currency, makes foreign direct investment (building and participations, financial or non-financial).

This feature improves the model, including the non-financial blocks, as the revenue from financial assets follows different rules.

But a survey of the present field of SFC models shows that they keep rather simple concerning the "real" part, and in consequence use little or no econometrics (which is not generally present in financial blocks, anyway). We propose to show how to improve this shortcoming. We shall give an example for this, admittedly limited as to SFC features.

Even if our elements are not always pertinent, we hope to lay out a framework which can be used to produce a fully consistent version, combining the real and financial aspects.

12.2 THE ACCOUNTING FRAMEWORK

Although the number of variables and equations can vary, the first condition is clear and mandatory: the model must produce a table in which the lines (the operations) and columns (the agents) are both systematically balanced.

This condition should actually apply to all models, and has actually been met by some of them, for instance the MESANGE model of the French economy.

12.2.1 THE AGENTS

Let us start with the agents. As usual, the classification must adapt to the behaviors. A difference in behavior (including the management of economic elements) must lead to the creation of a different category.

Taking into account basic operations calls at least for the following five:

- Households who receive wages and consume.
- Non-Financial firms which produce goods and non-financial services.
- Financial firms which manage loans and deposits.
- Government which provide non-market services and collect taxes.
- The rest of the world which trades with the country.

Now, this crude framework is not precise enough.

The Central Bank has a specific institutional role, and must be separated from other financial agents.

A large share of non-financial production comes from households (especially in Vietnam). But the associated behavior is difficult to separate from that of actual firms. Identifying this production will lead to accounting problems, starting with the transactions between them. For instance, the model will have to define the intermediary consumption of the “household” good by the “firms” branch (like the rice bought by shops from farmers). The practical solution (followed by most models) is to separate the productive role of households from its other operations and merge it with firms. This introduces minor problems which we will present (and solve) later.

However, there is a domain which should remain attached to households: the housing services, which include the “fictitious” rents paid by real estate owners to themselves.

Finally, this classification presents an obvious dissymmetry: the rest of the world is described as a whole, grouping all of its agents. Some SFC models will make the categories symmetric by identifying the same elements for the “rest of the world”. Unfortunately, this is difficult to apply to an operational model, for an obvious lack of information.

12.2.2 THE OPERATIONS

Concerning the operations, the degree of detail is not as clear. In practice, it must be made at three levels:

- The nature of the operation itself: for example social security benefits can be separated into unemployment, sickness…

- The currency: one can consider local and foreign, for which a single currency will be used (the US Dollar?).
- The horizon: short-term and long-term loans, or bills and bonds.

Our proposal will be associated to a medium case.

12.3 A PROPOSAL

We will now propose a framework taking into account the above considerations.

Compared to the single product model we have presented earlier, we have to:

- Adapt the real sector to a separation into three products.
- Create from scratch a financial block.
- And finally modify the "real" blocks to account for the new financial elements.

12.3.1 THE FRAMEWORK FOR THE NEW PRODUCTION BLOCK

The supply – demand equilibrium at constant prices is obtained for each product by:

GDP + intermediate consumption by the branch + imports

= Final local demand + intermediate consumption of the product + exports.

or using mnemonics.

$$(1)\ \ GDP_i + IC_{.i} + M_i = \ FD_i + IC_{i.} + X_i$$

The index "i" stands for each of the three products (and branches) : non-financial firms, Firms and Households. The dot (".") stands for the sum across the particular dimension (first product used, second branch using the product).

12.3.1.1 At constant prices

- Final demand FD is the sum of its elements:

$$(2)\ \ FD_i = COH_{.i} + COG_i + IP_i + \cdots ..$$

- Exports and imports come from the rest of the model.
- GDP balances equation (1)
- IC is given by applying a "technical coefficient" to value added:

$$(3)\ IC_{j,i} = \ tc_{j,i}\ .\ Q_i$$

- Production is given simply by addition:

$$(4)\ \ PR_i = GDP_i + IC_{.i}$$

But we must consider that GDP is the sum of Value Added, VAT and tariffs (both at constant prices):

VAT at constant prices is obtained by applying a "constant" rate to final demand.

(5) $VAT0_i = r\_vat0_i * FD_i$

For tariffs, we apply a rate to imports:

(6) $TAR0_i = r\_tar0_i * M_i$

Now value added can be obtained from GDP:

(7) $Q_i = GDP_i - VAT0_i - TAR0_i$

The "constant" rates can be approximated by the rate in the base year. However, it should also include the change in the weightings across time. For instance, the VAT rate grows naturally as the weight of basic products in household consumption decreases. The rate is actually called "using the regulations of the base year".

As Government (non-market) GDP measures essentially the compensations of civil servants, its link to intermediate consumption of Government is not so clear as other types. It is generally considered exogenous at constant prices.[166]

Finally, the above makes it clear that the data for VAT must be separated from other indirect taxes if one wants to identify GDP.

12.3.1.2 At current prices

The situation is different here, as we know the deflators for GDP (a decision of the local producing agent), and the deflators for imports and exports (a decision of the local or foreign exporting agent). As we know all the quantities, we can apply deflators to these three elements, and we get the value of demand for each product.

(8) $XV_i = PX_i * X_i$

(9) $MV_i = PM_i * M_i$

(10) $GDPV_i = PGDP_i * GDP_i$

(11) $ICV_{j,i} = tc_{j,i} \,.\, QV_i$

---

[166] (definition taken literally from the OECD website) General government final consumption can be broken down into two distinct groups. The first reflects expenditures for collective consumption (defense, justice, etc.) which benefit society as a whole, or large parts of society, and are often known as public goods and services. The second relates to expenditures for individual consumption (health care, housing, education, etc.), that reflect expenditures incurred by government on behalf of an individual household. This category of expenditure is equal to social transfers in kind from government to households and so includes expenditure by government on market goods and services provided to households. As goods and services produced by government usually do not have a market price, the relevant products are valued at the sum of costs needed to produce these goods and services. These costs mainly consist of compensation of employees, intermediate consumption and depreciation. Final consumption of government can then be estimated as the difference between on the one hand government output, and on the other hand payments made for goods and services produced by government and the relevant output that is used for fixed capital formation.

This means that the equilibrium comes from FDV and its deflator:

(12) $FDV_i + ICV_{i.} + XV_i = GDPV_i + ICV_{.i} + MV_i$

(13) $PFD_i = FDV_i \ / \ FD_i$

(14) $VAT_i = r\_vat_i * FDV_i$

(15) $TAR_i = r\_tar_i * MV_i$

(16) $QV_i = GDPV_i - VAT_i + TAR_i$

12.3.2 THE FINANCIAL BLOCK

12.3.2.1 The financial operations

We shall follow a traditional separation of the way each agent distributes its (positive of negative) financial capacity. However, it includes some simplifications, and the user is welcome to introduce or suppress minor elements (which can have some impact on causalities).

- High Powered Money or cash issued by the Central Bank, and used by the agents (households, firms) for their transactions.
- Deposits made by households
  - In local currency in local financial firms
  - In foreign currency in foreign banks (or local subsidiaries of foreign banks).
- Loans
  - in local currency obtained by households, firms or Government from local financial firms.
  - in foreign currency obtained by firms or Government from the rest of the world.
- Changes in bank reserves placed at the Central Bank
- Bonds and bills issued by Government, and bought by households, financial firms, the Central Bank and the rest of the world.
- Equities issued by firms and bought by households, financial firms and the rest of the world.
- Foreign direct investment by the rest of the world in local firms of both types.

This leads to the following framework.

Financial account												
Change in reserves in foreign currency						DFOR_X				DFOR_C		
Change in High Powered Money				DHPM_C			DHPM_H	DHPM_P	DHPM_F			
Change in deposits, local currency			DDPL_F				DDPL_H					
Change in deposits, foreign currency						DDPX_X	DDPX_H					
Change in loans, local currency	DLOL_H	DLOL_P			DLOL_G				DLOL_F			
Change in loans, foreign currency		DLOX_P	DLOX_F		DLOX_G							DLOX_X
Change in bank reserves				DRES_C					DRES_F			
Change in bills and bonds					DBBO_G		DBBO_H		DBBO_F	DBBO_C		DBBO_X
Change in equities		DEQU_P					DEQU_H		DEQU_F			DEQU_X
Foreign direct investment		FDIV_P	FDIV_F									FDIV_X

We hope the simplifications are not harmful.

12.3.3 THE ADDITIONAL CHANGES LINKED TO THE FINANCIAL ACCOUNT.

Taking into account these changes calls for two additions.

- The associated interests and dividends:

Primary allocation of income												
Interests on deposits, local currency			IDPL_F				IDPL_H					
Interests on deposits, foreign currency						IDPX_X	IDPX_H					
Interests on loans, local currency	ILOL_H	ILOL_P			ILOL_G				ILOL_F			
Interests on loans, foreign currency		ILOX_P	ILOX_F		ILOX_G							ILOX_X
Interests on bills and bonds					IBBO_G		IBBO_H		IBBO_F			IBBO_X
Interests	NI_H	NI_P	NI_F		NI_G	NI_X						
Dividends		DIV_P					DIV_H		DIV_F			DIV_X

- The stocks completing the “S” in SFC.

Balance sheet												
Debt	DEB_H	DEB_P	DEB_F		DEB_G	DEB_X						
Productive capital	KV_H	KV_P			KV_G							
Stock of reserves in foreign currency				FOR_C								FOR_X
Stock of High Powered Money				HPM_C			HPM_H	HPM_P	HPM_F			
Stock of deposits, local currency			DPL_F				DPL_H					
Stock of deposits, foreign currency						DPX_X	DPX_H					
Stock of loans, local currency	LOL_H	LOL_P			LOL_G				LOL_F			
Stock of loans, foreign currency						LOX_X		LOX_P	LOX_F		LOX_G	
Stock of bank reserves				RES_C					RES_F			
Stock of bills and bonds					BBO_G		BBO_H		BBO_F	BBO_X		BBO_C
Stock of equities		EQU_P					EQU_H		EQU_F			EQU_X
Foreign direct capital		KFDIV_P	KFDIV_F									KFDIV_X

12.4 EVOLVING TOWARDS A FULL MACROECONOMIC MODEL

- In the “real” part of our model, we will need the elements presented above.
- An optimization of productive factors taking into account the relative costs of labor and capital.
- An impact of the output gap on investment and the production deflator, introducing again an optimization process (long- and short-term respectively).
- An implicit choice of the optimal trade deflators by the exporting agent.
- Traditional Keynesian and price-wage loops fully formalized using the most logical framework available.
- A full table for the six agents, balanced in lines and columns, and producing the financial capacities as the last line.

As an example, we will build a macro-econometric model based using Vietnamese data, but in the very general framework we have presented earlier. To comply with the SFC conditions, this model will consider three productive sectors (and products): Financial, Non-Financial and Households (considering only the housing sector).

We will start with a version in which the splitting by agents (such as the partition of households’ savings) uses exogenous distribution keys.

We will consider incentives for deciding on these keys. The main one could be the rate of return of each category, which would have to be made explicit and actualized. But we leave for now to specialists the choice and formalizations.

Even without this feature, one of the key properties of this model will be the combination of stabilizing processes coming from the real side (correcting the output gap, or maximizing profits by considering the tradeoff between output and unitary margins) and the financial side (for instance considering the debt ratios, or the tradeoff between the rates of return of instruments and their demand level).

The properties of this model will be observed through a series of shocks on:

- First on the real exogenous assumptions (such as Government demand).
- Some financial assumptions (such as the Central bank’s base rate).

And in future versions, we will shock:

- More financial assumptions.
- The basic keys (such as the share of equities bought in households' savings).

12.4.1 A FULL PROPOSAL

In summary, our model will propose the following framework.

Noms	Hhlds	Firms	Fin F	Ctl Bk	Govmt	RoW	Hhlds	Firms	Fin F	Ctl Bk	Govmt	RoW
Production account												
Exports of goods and services	DLOL_H					XV					FCAP_G	
Imports of goods and services												MV
Intermediate consumption	ICBV_H	ICBV_P	ICBV_F		COGV		ICPV_H	ICPV_P	ICPV_F			
Gross value added / GDP							GDPV_H	GDPV_P	GDPV_F		GDPV_G	
Ext. trade in goods and services						XV-MV						XV-MV
Operating account												
Gross wages and salaries	WAGE_H	WAGE_P	WAGE_F		WAGE_G		WAGE					
Employers social contributions	SC_H	SC_P	SC_F		SC_G						SCR_G	
Workers social contributions	SC_W										SC_W	
Tariffs		TAR									TAR	
Taxes on production and imports	OIT_H	OIT_P	OIT_F								OIT	
Subsidies					SUBS		SUBS_H	SUBS_P	SUBS_F			
Gross operating surplus							EBE_H	EBE_P	EBE_F		EBE_G	EBE_X
Gross mixed income		REVQ					REVQ					
Primary allocation of income												
Interests on deposits, local currency			IDPL_F				IDPL_H					
Interests on deposits, foreign currency						IDPX_X	IDPX_H					
Interests on loans, local currency	ILOL_H	ILOL_P			ILOL_G				ILOL_F			
Interests on loans, foreign currency		ILOX_P	ILOX_F		ILOX_G							ILOX_X
Interests on bills and bonds					IBBO_G		IBBO_H		IBBO_F			IBBO_X
Interests	NI_H	NI_P	NI_F		NI_G	NI_X						
Dividends		DIV_P					DIV_H		DIV_F			DIV_X
Secondary allocation of income												
Value Added Tax	VAT_H	VAT_P	VAT_F								VAT	
Taxes on income and wealth	TP_H+...	TP_P	TP_F								TP+ICT	
Social benefits					SOCB		SOCB					
Current international cooperation						COOP					COOP	
Gross disposable income							RDB_H	RDB_P	RDB_F		RDB_G	RDB_X
Allocation of income												
Final consumption expenditure	COHV				COGV							
Gross savings							GSV_H	GSV_P	GSV_F		GSV_G	
Capital account												
Gross fixed capital formation	IV_H	IV_P	IV_F		IV_G							
Change in inventories		CIV										
Financing capacity (+)							FCAP_H	FCAP_P	FCAP_F			FCAP_X
Financial account												
Change in reserves in foreign currency						DFOR_X				DFOR_C		
Change in High Powered Money				DHPM_C			DHPM_H	DHPM_P	DHPM_F			
Change in deposits, local currency			DDPL_F				DDPL_H					
Change in deposits, foreign currency						DDPX_X	DDPX_H					
Change in loans, local currency		DLOL_P			DLOL_G				DLOL_F			
Change in loans, foreign currency		DLOX_P	DLOX_F		DLOX_G							DLOX_X
Change in bank reserves				DRES_C					DRES_F			
Change in bills and bonds					DBBO_G		DBBO_H		DBBO_F	DBBO_C		DBBO_X
Change in equities		DEQU_P					DEQU_H		DEQU_F			DEQU_X
Foreign direct investment		FDIV_P	FDIV_F									FDIV_X
Balance sheet												
Debt	DEB_H	DEB_P	DEB_F		DEB_G	DEB_X						
Productive capital	KV_H	KV_P			KV_G							
Stock of reserves in foreign currency						FOR_X				FOR_C		
Stock of High Powered Money				HPM_C			HPM_H	HPM_P	HPM_F			
Stock of deposits, local currency			DPL_F				DPL_H					
Stock of deposits, foreign currency						DPX_X	DPX_H					
Stock of loans, local currency	LOL_H	LOL_P			LOL_G				LOL_F			
Stock of loans, foreign currency		LOX_P	LOX_F		LOX_G							LOX_X
Stock of bank reserves				RES_C					RES_F			
Stock of bills and bonds					BBO_G		BBO_H		BBO_F	BBO_C		BBO_X
Stock of equities		EQU_P					EQU_H		EQU_F			EQU_X
Foreign direct capital		KFDIV_P	KFDIV_F									KFDIV_X

Of course, the model will also describe:

- The deflators.

- The transactions on goods and services at constant prices.
- Capital and labor.

12.4.2 THE CAUSALITIES

The above elements do not solve anything. They just map the domain on which we will work.

Two tasks remain.

- Defining the causalities between the financial elements (remember we are dealing with elements at current prices, which must not remain exogenous). Exogenous (unidimensional) keys can be used here.
- Endogenizing these keys.

From now we feel like an explorer travelling in uncharted territory, using the efficient instruments we can possess. Help from local scouts would be great.

We would nevertheless make proposals, if only to show that addressing the first point does not lead to completely unreasonable results.

The reader can use the following elements as he desires. The EViews programs are available for using and also editing. We will welcome comments and maybe cooperation.

12.4.3 THE CAUSALITIES (KEYS)

Our next step is defining the inner causalities and the closures. This will be done using what we think to be the most logical sequence, but this can be questioned of course.

12.4.3.1 High Powered Money.

We will suppose that households and firms are able to obtain the cash they need from the Central Bank, depending on their economic situation. This gives DHPM_H, DHPM_P and DHPM_F and the sum DHPM, closing the line.

12.4.3.2 Households

We will suppose that households ask for (and obtain) loans, depending on their financial capacity, consumption and housing investment, and they separate the remaining balance into buying bonds and equities, and making deposits, both in local and foreign currency.

This closes the household's column.

12.4.3.3 Deposits in foreign currency

We will suppose that these deposits are obtained from the rest of the world, with local banks as a simple intermediary.

This closes the foreign deposits line.

12.4.3.4 Deposits in local currency

We will suppose that these deposits are made by households and held by local banks.

This closes the local currency deposits line.

12.4.3.5 Foreign reserves

We will suppose that in addition to providing HPM, the Central Bank buys foreign currency (or sells the local one) to finance the foreign deficit, but also as a policy instrument.

The line is offset by the Rest of the World.

This closes the foreign reserves line.

12.4.3.6 Foreign Direct Investment

We will suppose that it is made by the Row in both financial and non-financial firms.

This closes the FDI line.

12.4.3.7 Government

We suppose that Government finances its deficit by obtaining loans and issuing bonds/bills.

This closes the Government column.

12.4.3.8 Foreign loans

We will suppose that they are made by the RoW to firms (both Financial and Non-Financial) and Government.

This closes the line.

12.4.3.9 Non-financial firms

We will suppose that they close their account by borrowing and issuing equities.

12.4.3.10 Bank reserves

The banks transfer reserves to the Central Bank in proportion to their loans.

This closes the reserves line.

12.4.3.11 The Central Bank

The Central Bank closes its account by buying bonds.

12.4.3.12 The rest

The remaining elements correspond to the Bills/Bonds and equities bought by financial firms and the Rest of the World. The sum by agent is now fixed, and determining one element decides on the remaining three.

12.4.4 EXPLICITING THE BEHAVIORS

We shall now try to improve the above by proposing explanations to the elements.

As usual, one can consider two categories: behavioral or identity.

The identities normally represent closures, except maybe for the reserves held by the Central Bank, transferred from Financial firms.

The behavioral equations can be separated again, into:

- Decisions depending only on non-financial variables. They are exogenous to the block.
- Decisions implying other financial elements. A causal sequence must be defined.
- Splitting financial decisions with a given sum. The share(s) are defined first, then the elements.

12.4.4.1 "Exogenous" decisions

One can consider 7 elements:

- Loans to households, depending on their consumption, their housing investment and (negatively) their savings.

We will suppose that this decision is made before the other uses of the resources available.

- Loans by the RoW to firms (both types) and Government.
  The decision should be made at constant (foreign) prices, depending possibly on the perceived present health and prospects of the local economy.

  Non-Financial Firms: they can depend on trade and financial capacity

  Financial Firms: same

  Government:: they can depend on financial capacity, but a share can come from international cooperation, both international institutions or individual countries.

- Foreign Direct Investment.

This element is already present in the initial model. It depends on relative profitability and demand (local and foreign) addressed to the country. It can be shared between the types of firms according to both criteria.

It influences total factor productivity and export potential.

- The change in reserves of the Central Bank
  To the automatic impact of financial capacity can be added a behavioral element aiming at controlling the exchange rate.
- The dividends paid to equities holders by the firms
  As we consider only equities issued by non-Financial firms, dividends will come only from them. They will represent a share on previous gross disposable income, and will be distributed to financial firms, households and rest of the world according to their share of total equities.

12.4.4.2 First sharing

One can now apply three sharing processes.

12.4.4.2.1 Households

Households will share their savings (plus money from loans) between:

- Investment in housing

- Deposits in local currency
- Deposits in foreign currency
- Buying bills and bonds
- Buying equities

Of course, all these changes can actually be negative.

With such a number of opportunities, the sharing is rather complex.

The explanation could come from:

- The relative interest rates (and the expected returns from equities). The interest on housing investment plays an opposite role.
- The expected evolution of equities.
- The expected evolution of the exchange rate.

The household account is now closed.

12.4.4.2.2 Foreign direct investment

Sharing between financial and non-financial firms should depend on relative profitability, and demand potential for the non-financial sector.

12.4.4.3 First closures

As stated before, we can now determine total High-Powered Money (provided by the Central Bank), deposits in local currency (in banks) and deposits in foreign currency (in banks but the money comes from the RoW).

We know enough now to determine the amount in the financial firms account which is subject to reserves, and the change in reserves in the Central Bank. However, the ratio applied to stocks, so we must also consider the change in the value of the previous stock

12.4.4.4 Second sharing

12.4.4.4.1 Government

The government deficit (minus the foreign loans) can be financed either by borrowing of issuing bills and bonds.

The choice depends mostly on the interest rates.

Bills can be separated from bonds with different rates (depending on the horizon).

The Government account is closed.

12.4.4.4.2 Non-Financial Firms

They can finance the remaining sum by issuing equities or borrowing.

The choice is difficult, but one can consider:

- That the consequences of borrowing are clear.
- That the more you borrow the higher the interest rate.

- That issuing equities you will lose assets ownership, future profits and dividends (all unclear)
- That if your firms regain health, it is easier to pay back debt than buy back equities (at a higher price).

12.4.4.4.3 New closures

Now we know all elements in the Central Bank account except bills and bonds, which balance it accordingly.

Also, the loans by Financial Firms balance the line.

12.4.4.4.4 Final sharing

As stated earlier, the remaining bills/bonds and equities must be shared between local Financial Firms and the RoW (mostly foreign financial firms).

This can be done in several ways, but once an agent has decided on the sharing, the behavior of the other is fixed.

This balances the two remaining lines and columns

12.4.4.5 The list of equations in the financial block

The identities are completely defined, the logic of the behaviors state the potential explanations as a sum, enclosed in parentheses, and following a "f", as we have done for the real model.

<u>Foreign Reserves</u>

Equation [1] Foreign reserves: Central bank

dforr_c = fcap_x + dforra_c

Equation [2] Foreign reserves : Rest of the World

dforr_x = - dforr_c

<u>High Power Money</u>

Equation [3] High Power Money : Households

dhpm_h = f * (hdi_h)

Equation [4] High Power Money : Non Financial Firms

dhpm_p = f * (gdpmv_p)

Equation [5] High Power Money : Financial Firms

dhpm_f = f * (gdpmv_f)

Equation [6] High Power Money : Central Bank

dhpm_c = dhpm_h + dhpm_p + dhpm_f

<u>Foreign Direct Investment</u>

Equation [7] Foreign Direct Investment : Non Financial Firms

fdiv_p = f * (ur_p + rprob_p)

Equation [8] Foreign Direct Investment : Financial Firms

fdiv_f = f * (rprob_f)

Equation [9] Foreign Direct Investment : Rest of the World

fdiv_x = fdiv_p + fdiv_f

Deposits in local currency

Equation [10] Deposits in local currency : Households

ddpl_h = r_ddpl_h * (fcap_h + dlol_h)

Equation [11] Deposits in local currency : Financial Firms

ddpl_f = - ddpl_h

Deposits in foreign currency

Equation [12] Deposits in foreign currency :]Households

ddpx_h = r_ddpx_h * (fcap_h + dlol_h)

Equation [13] Deposits in foreign currency : Rest of the World

ddpx_x = - ddpx_h

Loans in local currency

Equation [14] Loans in local currency:]Households

dlol_h = f * iv_h

Equation [15] Loans in local currency : Non Financial Firms

dlol_p = r_dlolp * (dhpm_p + fdiv_p + dlox_p)

Equation [16] Loans in local currency : Government

dlol_g = (1 - r_dbbo_g) * (fcap_g + dlol_g)

Equation [17] Loans in local currency : Financial Firms

dlol_f = dlol_h + dlol_p + dlol_g

Loans in Foreign currency

Equation [18] Loans in foreign currency : Non Financial Firms

dlox_p = f * (@pch(gdpv_p) + @pch(ur_p) + fcap_x / gdpv_p)

Equation [19] Loans in foreign currency : Financial Firms

dlox_f = f * (@pch(gdpv_f) + fcap_f / gdpv_f)

Equation [20] Loans in foreign currency : Government

dlox_g = f * (fcap_g / gdpv)

Equation [21] Loans in foreign currency : Rest of the World

dlox_x = dlox_p + dlox_f + dlox_g

Bills and Bonds

Equation [22] Bills and Bonds : Households

dbbo_h = r_dbbo_h * (fcap_h + dlol_h)

Equation [23] Bills and Bonds : Government

dbbo_g = r_dbbog * (fcap_g + dlol_g)

Equation [26] Bills and Bonds : Rest of the World

dbbo_x = dbbo_g - (dbbo_h + dbbo_f - dbbo_c)

Equation [27] Bills and Bonds : Financial Firms

dbbo_f = r_dbbo_f * (fcap_f + fdiv_f + ddpl_f + dres_f - (dhpm_f - dlol_f - dlox_f))

Equation [32] Bills and Bonds : Central Bank

dbbo_c = dforr_c + dbbo_c - dhpm_c

Equities

Equation [24] Equities :]Households

dequ_h = (1 - r_ddpl_h - r_ddpx_h - r_dbbo_h) * (fcap_h + dlol_h)

Equation [25] Equities : Non Financial Firms

dequ_p = (1 - r_dlolp) * (dhpm_p + fdiv_p + dlox_p)

Equation [28] Equities : Rest of the World

dequ_x = dequ_p - dequ_h - dequ_f

Equation [29] Equities : Financial Firms

dequ_f = (1 - r_dbbo_f) * (fcap_f + fdiv_f + ddpl_f - (dhpm_f - dlol_f - dlox_f))

Reserves

Equation [30] Reserves : Financial Firms

dres_f = r_dres_f * dlol_f

Equation [31] Reserves : Central Bank

dres_c = dres_f

12.5 THE SIMULATIONS

We shall present the properties of the present version, by applying shocks to assumptions. If our comments will be somewhat detailed and technically sophisticated, one must not conclude that we are proposing this version as final.

In practice we shall make eight shocks, the first two traditional shocks concentrating on demand and supply features of the model.

The shocks will be represented and commented for:

1. An increase in government investment, by 1 point of market GDP.
2. A decrease in the VAT rate, by one point of the rate.
3. A devaluation of the Vietnamese Dong by 1%

We have also produced, but we will not show the results, for:

4. An increase in Foreign Direct Investment by one point of the capital level.
5. A decrease in the rate of tariffs applied to foreign products, by one point of the rate.
6. A decrease on the rate of tariffs applied by the rest of the world to Vietnamese products, by one point of the rate.
7. An increase in quotas applied to foreign products.
8. An increase in quotas applied by the rest of the world to Vietnamese products.

None of these shocks apply to elements in the financial part of the model, although some influences transit through it, like the interests from various financial objects purchased by the agents.

All these shocks start in 2021, and are sustained over the whole period.

For each shock we shall produce 6 graphs:

1 - The supply demand-equilibrium: GDP, Final demand, exports and imports at constant prices, plus the GDP deflator.

2 - The production process: Investment, capacity, firms employment, rate of use of capacities.

3 – Some important ratios: rate of use, unemployment rate, real wage cost, capital-labor ratio.

4 - External trade at current prices (1): Exports, imports, export – import ratio

5 - External trade summary (2): Export - import ratio at constant and current prices, terms of trade.

6- Prices : of demand, of GDP, of imports, of exports, wage rate

THE SHOCKS

1 A demand shock: An increase in government consumption

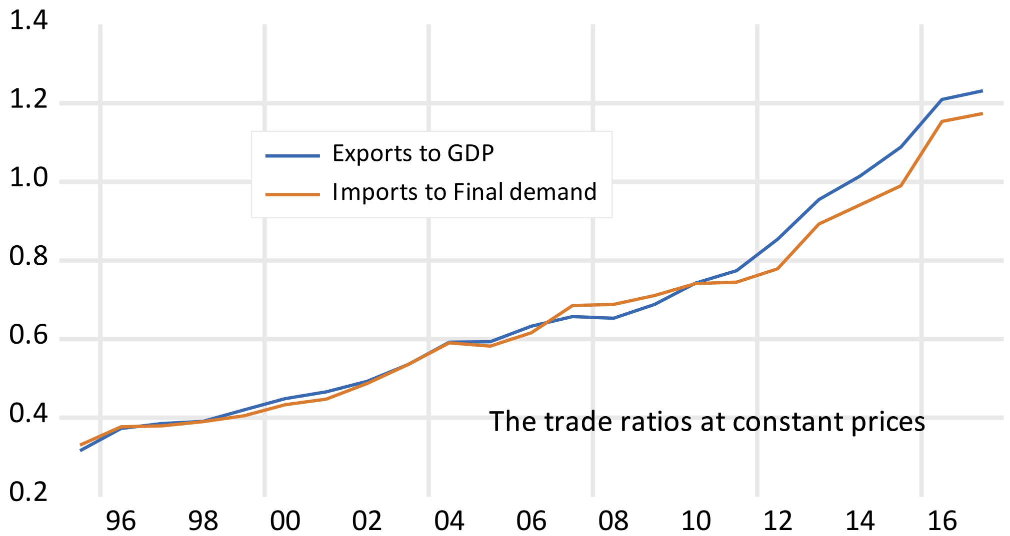
Starting in the year 2021 we increase Government demand by 1% of the GDP forecasted for the same period. This will make interpretation easier than for a constant shock; in particular concerning the "Keynesian multiplier".

The definition of this element, in practice the most frequently used to characterize a model, is quite simple. In the presence of a given shock on demand, coming for instance from Government consumption, it measures the ratio between the change in GDP produced by the shock and its initial size. If the ratio is higher than one, it means that the economic effects of the shock on the economy are higher than the negative ones. Of course, this figure has also to be interpreted.

In our case, the "Keynesian multiplier" will be obtained naturally by considering the relative percent change of GDP itself. If GDP increases by more than 1%, it means that the response is higher than the shock.

We observe in Graph 1 that the multiplier presents a limited cycle (with a period of around 40 years) converging to around 0.35 in the long run. This is consistent with other models associated with a very open economy.

.



Of course, using Production and Total demand as a denominator would bring the value down (to around 0.55 on 2017).

Final demand presents the same evolution, around a higher level, with an average value of about 1%, which means that the positive effect of endogenous demand is offset by the loss on external trade.

In the short run, the change in imports follows the same dynamics, at a lower level. Exports presents a very limited decrease.

These elements can be easily explained, considering the model as a whole.

As local demand and GDP are of the same order of magnitude (their difference is net exports, relatively small compared to GDP), the initial (or ex-ante) shock increased demand by roughly 1%. This means that the impact of model mechanisms reduces the ex-post value of GDP growth. Let us detail this aspect.

Graph 2 shows that for the non-financial product, both factors (capital and employment) grow significantly, requiring a strong effort on investment in the short run. As explained earlier, firms expect production to represent a "normal" share of capacities, defining a target on their rate of use. If production increases, a gap will appear, which firms will try to close by increasing capacities through investment and labor. However, this process is slowed by the cautiousness of firms, and the technical difficulty in implementing their decisions. This inertia applies especially to capital, for which the consequences of wrong decisions are more costly. It takes a long time (around 10 years) for capacities to reach. But then the inertia on decisions (particularly capital) leads to overshooting, starting a cyclical process with a decreasing amplitude.

However, capital and employment evolutions diverge from capacity. In the beginning, employment grows more, as adaptation to the target is faster (the change in capital comes only from investment, itself quite inert). In the medium run, the order is reversed by the increase in the relative cost of labor, coming from the reduced unemployment and the partial indexation on the value-added deflator, the element with the highest increase among prices (see Graph 6). In the end the increase in employment is quite small, and capital increases by 0.3% (representing in the long run an increase in investment of the same amount).

Graph 3 shows that the changes in the labor-capital ratio and their relative cost converge, a result consistent with our Cobb-Douglas assumption.

Consumption grows even slower (around 0.25%). There are several reasons for this inertia.

- Employment adapts gradually to the production increase.
- The purchasing power of wages is affected initially by non-unitary indexation on prices, the gap building up with additional inflation. This loss will disappear as the price increase stabilizes.
- The short-term sensitivity of consumption to revenue is lower than one.
- Inflation reduces the purchasing power of current savings.
- A large share of household revenue is not indexed on activity.

The real wage rate profits from the decrease in unemployment, but this has an adverse moderating effect on employment itself, as we have seen.

The initial decrease in unemployment is the main explanation for the evolution of prices (Graph 6). In the short and medium terms, the higher inflation comes from tensions on the rate of use (with higher sales perspectives, firms feel that they can apply higher prices), and on the labor market through the unemployment rate.

One moderating element is the increase in labor productivity, due to the inertia facing an increase in production. But we have seen that the adaptation of labor to production was rather fast. [167] This element appears essentially in the first periods, creating very temporary deflation.

[167] As to capital, the effect is more complex: the investment effort, destined to generate capacities for the future, is higher in relative terms than the present increase in GDP. But in the cost we only consider the share of the present period, through a correcting coefficient.

The loss in competitiveness will reduce quantities sold. Let us now consider the trade balance (Graphs 4 and 5). We know that the trade prices have a different (calibrated) sensitivity to local inflation (0.2 for imports, 0.7 for exports). And the short- and long-term competitiveness elasticities are lower than one: 0.53 for exports, 0.46 for imports. But the change in local prices is not negligible, especially in the long-term.

The evolution of exports is the simplest to explain they are essentially sensitive to the loss in price competitiveness, with the highest elasticity. They also follow the rate of use of capacities, but the effect is limited by the (calibrated) coefficient and the evolution of the rate itself, which returns to the base value at the end of the sample. In addition, exports are favored by the higher share of FDI capital

Imports are more affected: they follow local demand, with the addition of losses in market shares, coming first from the limits on available capacity (decreasing with time) then from the gradual loss in price competitiveness, which compares the local production price to rather stable import prices. Both effects roughly offset each other during the period, and the profile of imports follows demand. The level is intermediate between final and intermediate demands (the latter roughly indexed on GDP).

Concerning the balances, the terms of trade are little affected, and the loss is similar on both nominal and real balances.

Now, as to the Government budget, the additional expenditures lead to some revenue, coming mostly from the taxes on production. But as borrowing finances this measure, the interests paid will also increase.

We can see that the revenue effect wins in the short run, but the accumulation of debt reverses the order quickly, to a level which could be considered unsustainable (limiting the reliability of the shocks' message).

In the above explanation, we have evoked FDI. It actually plays an important (and original) role.

First, it accounts for the low decrease in exports, as it creates capacities oriented in that direction.

It also explains the initial deflationary effect, as the higher productivity of both factors reduces costs.

In the medium-term, the effect decreases, but remains significantly favorable.

12.5.1.1.1 The financial elements

Now, let us observe the consequences for the financial account, the main originality of our model. One must be aware that although the financial account is completely consistent and balanced, and follows an explicit causal sequence, the agents take decisions by applying exogenous shares. Applying truly endogenous behaviors will be introduced in a later version.

For this, we will use the global economic table, concentrating on the financial account.

One must also realize that elements appear at current prices. For shocks prioritizing deflators (such as a decrease in the VAT rate, or a devaluation) this diverges from the usual comments on real values.

In the following we have neglected the smallest elements, but one can check that the table is balanced both in lines and columns.

We shall first consider the results for the first period of the shock, 2021. We shall then move to the medium-term (2025) and the long-term (2050).

The agent directly concerned is Government.

Its investment is the most changed element in the table (1.023). As it must be financed without compensation, interests grow already (0.123). Some revenue increases: social contributions (0.005), tariffs (0.046), direct taxes (0.007). The resulting deficit (0.973) has to be compensated by issuing bonds (0.438), local borrowing (0.438) and foreign (0.097), the latter allowed by the trade deficit.

For firms, the situation improves with GDP (0.315), reduced by wages (0.063), taxes (0.053): The gross operating surplus increases by 0.201, 0.058 of which goes to individual workers. Taking into account VAT (0.116), and the increased investment and inventories (0.076), financial capacity decreases by 0.044. But the funds from foreign loans (0.036) and FDI (0.050) allow equities and local loans to decrease (0.027).

The rest of the world uses the trade surplus (0.863) to buy bills and bonds (0.830), equities (0.234), and to loan (0.133). The difference comes from the increase in reserves (0.198), and FDI (0.048).

Households use the additional revenue (0.130), interests (0.003), to consume (0.088), to invest in lodging (0.008), to buy money (0.012), equities (0.005), bills (0.002), and to increase deposits in both currencies (0.020).

Financial agents (including the Central Bank) have a mostly passive behavior.

The above elements are consistent with our framework and causalities. They can be criticized by specialists of the field for:

- The causalities themselves.
- The values used for shares.
- Their exogeneity.

However, all these choices can be modified, provided the consistency of the accounting system remains.

Let us move now to the two later tables. As we have seen earlier, for a demand shock, most of the consequences are obtained at the first period.

The main dynamic effects come from:

- The increase in investment, with an initial inertia and a later decrease as capacities adapt to the target.
- The increase in household consumption as employment adapts gradually and the wage rate gains purchasing power at every period.
- The accumulation of the Government debt, as the investment effort is made every year, with the addition of growing interests.

All these elements are present in the tables, including an increase in the debt by 19 points of GDP which looks quite unsustainable.

The situation should change when we endogenize the shares, as they should consider the new relative profitability.

2 A supply shock: decreasing the VAT rate by one point

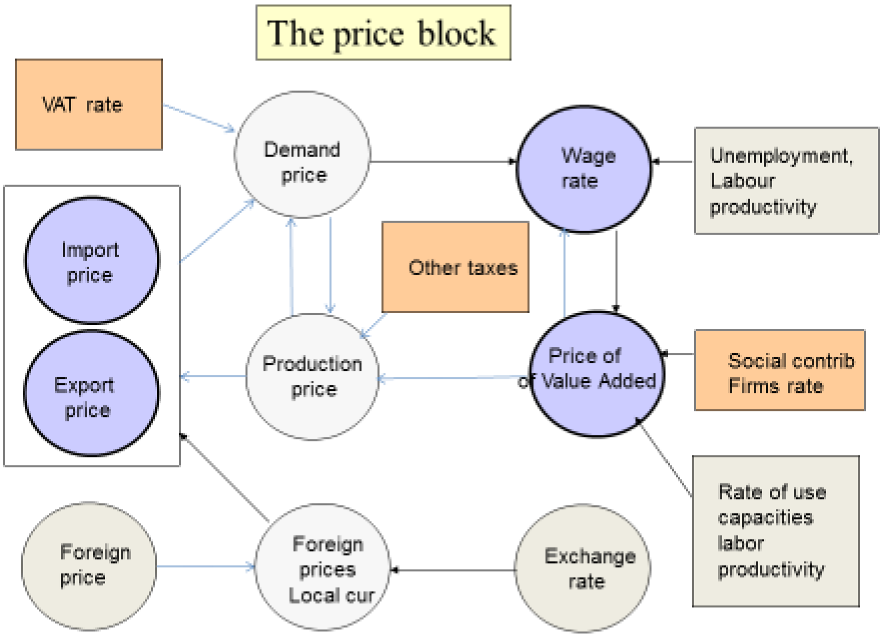
We suppose that the Government decreases the VAT rate by 1 point.

The results are quite similar to the previous shock on the French model.

Graph 6 shows that all prices decrease significantly. Of course, the highest impact is on the demand price (fully depending on VAT). The wage rate (partially depending on it, and gaining in purchasing power through the decrease in unemployment).

But all other prices decrease too. The value-added deflator is indexed on the cost of factors, not only wages but also capital. The production price follows, and the prices of trade, depending on the latter.

We will use again the following graph:



One can observe the loop between the wage rate and the value-added deflator, moderated by trade deflators (the import price is less sensitive to the GDP deflator). Compared to a complementary factors framework, the margins target includes also the capital cost, moderating the role of wages but introducing an additional loop.

The gain in competitiveness favors exports (Graph 1), once capacities have started to adapt, creating growth through demand in investment and consumption goods. The subsequent increase in imports (coming also from intermediate goods and energy requested by exports) is accentuated at first by the initial bottlenecks. Then competitiveness increases, capacity adapts (and even over adapts) reducing also the demand for equipment goods. But the variation of imports remains significantly positive, higher in fact than exports. But the loss on trade in real terms is much lower than for the previous shock, for a comparable gain on GDP.

In the medium run, the deflationary effect keeps increasing moderately (Graph 6), as well as the gains in competitiveness.

Concerning the factors (Graph 2) capital and investment are again favored by the increase in the real wage rate, as the previous case, but the main impact comes from the immediate decrease in the demand price. As employment adapts faster than capital (which closes slowly with its target), capacity takes some time in increasing, and the rate of use remains high.

The impact on the real trade balance (Graph 5) is negative, as the increase in GDP is followed by demand, both final and intermediate. With the global loss on the terms of trade, the impact on the current balance is even more negative.

As to the Government budget, its evolution is similar to the previous case, except for the interests, with a growing importance of interests paid, which decrease at first with the rate itself (remember it is indexed on the evolution of CPI, which decreases for the first two years). However, the cost is lower, and the gain on GDP is higher: this instrument seems more efficient on that aspect.

As to FDI, its effect is more favorable here, as profitability increases: the numerator (profits) is not sensitive to VAT (it is even positively affected as wages are), and the denominator (the value of capital at cost of renewal) decreases. The profitability ratio increases.

Its contribution to the improvement of exports both through factor productivity and competitiveness explains the persistence of the GDP increase.

12.5.1.1.2 The financial elements

Now, considering the financial account, we face perhaps the most important difficulty in the SFC approach: the main channel of this shock is through prices, and appreciating the efficiency of the decision takes essentially into account the evolution of quantities.

The table is presented at current prices, and each of these prices changes a lot following the decision, with different intensities depending on the role of unchanged foreign prices. This is particularly visible in the ending paragraph giving the decomposition of GDP: we see that if GDP decreases a little (0.016) this comes from a balancing between an increase in value added (0.797), tariffs (0.107) and a decrease in VAT itself (-0.919).

Of course, a presentation at constant prices could be clearer, but this will not be easy as the change in the value of deflators is not homogenous at all.

3 A financial shock: devaluating the Vietnamese Dong by 1%.

Let us first consider Graph 6.

We can see that all prices increase gradually, converging monotonously to a 1% value (we have checked that this will be achieved a few years later). Only wages grow faster, due to the improved activity and the subsequent gains in purchasing power

The hierarchy is quite logical: the foreign price increases by 1%. The import price follows as the most influenced, then comes the export price, the consumption price (an average of local production and imports), and the value-added price (the most "local").

Graph 6 shows that the gains in competitiveness are important, if we compare the export price with the 1% line (representing the increase in the foreign price in Dongs), and the import price with the production price. This brings as usual exports, production, demand and imports.

These gains bring growth (Graph 1) coming mostly from local demand, and a small gain in net exports.

Concerning the trade balance, the change at constant prices is small with a changing sign, compared to the loss on the terms of trade (Graph 5).

However, all these effects will disappear with time at all prices converge to the target.

12.5.2 CONCLUSION ON MODEL PROPERTIES

We can draw the following conclusions:

The model converges over as long a period as we want, with stabilizing growth rates. The consequences of changes in assumptions stabilize too.

The properties look reasonable in the short and long run, including the role of instruments (sign and size), and especially:

- The stabilization of the rate of use of capacities (convergence of value added and capacities).
- The variation of prices, relative to each other.
- The link between factors and their costs.
- The contribution of factors to capacity growth.
- The presence of cycles, converging after a reasonable period.
- The differentiated role of policy instruments.

The introduction of a detailed financial account (both in variations and stocks) brings interesting features and provides a better stability, even with exogenous shares.

The choices made in the values of shares and the causalities will need to be checked, and the shares remain to be endogenized, following an optimizing framework.

Name	Role	Definition	Agent
DBBO_C	Eq32	Bills and Bonds	Central Bank
DBBO_F	Eq27	Bills and Bonds	Financial Firms
DBBO_G	Eq23	Bills and Bonds	Government
DBBO_H	Eq22	Bills and Bonds	Households
DBBO_X	Eq26	Bills and Bonds	Rest of the World
DDPL_F	Eq11	Deposits in local currency	Financial Firms
DDPL_H	Eq10	Deposits in local currency	Households
DDPX_H	Eq12	Deposits in foreign currency	Households
DDPX_X	Eq13	Deposits in foreign currency	Rest of the World
DEQU_F	Eq29	Equities	Financial Firms
DEQU_H	Eq24	Equities	Households
DEQU_P	Eq25	Equities	Non-Financial Firms
DEQU_X	Eq28	Equities	Rest of the World
DFORR_C	Eq1	Reserves in foreign currency	Central Bank
DFORR_X	Eq2	Reserves in foreign currency	Rest of the World
DHPM_C	Eq6	High Power Money	Central Bank
DHPM_F	Eq5	High Power Money	Financial Firms
DHPM_H	Eq3	High Power Money	Households
DHPM_P	Eq4	High Power Money	Non-Financial Firms
DLOL_F	Eq17	Loans in local currency	Financial Firms
DLOL_G	Eq16	Loans in local currency	Government
DLOL_H	Eq14	Loans in local currency	Households
DLOL_P	Eq15	Loans in local currency	Non-Financial Firms
DLOX_F	Eq19	Loans in foreign currency	Financial Firms
DLOX_G	Eq20	Loans in foreign currency	Government
DLOX_P	Eq18	Loans in foreign currency	Non-Financial Firms
DLOX_X	Eq21	Loans in foreign currency	Households
DRES_C	Eq31	Reserves at the Central Bank	Rest of the World
DRES_F	Eq30	Reserves at the Central Bank	Financial Firms
FDIV_F	Eq8	Foreign Direct Investment	Financial Firms
FDIV_P	Eq7	Foreign Direct Investment	Non-Financial Firms
FDIV_X	Eq9	Foreign Direct Investment	Rest of the World
DFORRA_C	Exog	Policy correction to Foreign Reserves	Central Bank
FCAP_F	Exog	Financial Capacity	Financial Firms
FCAP_G	Exog	Financial Capacity	Government
FCAP_H	Exog	Financial Capacity	Households
FCAP_X	Exog	Financial Capacity	Rest of the World
GDPMV_F	Exog	Market GDP	Financial Firms
GDPMV_P	Exog	Market GDP	Non-Financial Firms
GDPV	Exog	GDP	Total
GDPV_F	Exog	GDP	Financial Firms
GDPV_P	Exog	GDP	Financial Firms
HDI_H	Exog	Household Disposable Income	Households
IV_H	Exog	Investment	Households
R_DBBO_F	Exog	Share of Bills and bonds in the residual	Financial Firms
R_DBBO_G	Exog	Share of Bills/Bonds versus Local loans	Government
R_DBBO_H	Exog	Share of Bills and bonds in the residual	Households
R_DDPL_H	Exog	Share of Deposits in Dongs in the residual	Households
R_DDPX_H	Exog	Share of Deposits in Foreign Currency in the r	Households
R_DLOL_P	Exog	Share of Loans in Dongs in the residual	Non-Financial Firms
R_DRES_F	Exog	Ratio of Reserves in Loans	Financial Firms
RPROB_F	Exog	Profits rate	Financial Firms
RPROB_P	Exog	Profits rate	Non-Financial Firms
UR_P	Exog	Rate of use of capacities	Non-Financial Firms

SOLVING A MODEL FOR TARGET VALUES

In the previous version of the book, we have presented a method allowing to switch the exogenous / exogenous status of model variables without editing the actual code, using a quasi-Newton algorithm.

After a detailed description of the methodology, we provided some practical examples and a subprogram allowing to apply it in a simple way.

Although this remains an option, starting with EViews 12, one can perform a similar task using simple and official menus and commands.

We shall start with a detailed presentation of our method, then give the syntax of both new EViews options, which are not yet documented from a theoretical point. We shall also provide examples.

Unfortunately, the User’s Guide provides few elements on the methodology. So I have chosen the following approach. I will:

- Develop my methodology, assuming it is the same as the one used by EViews 12.
- Illustrate it on practical examples.
- Show how to use EViews commands to solve the same problem.
  This is also the easiest approach for me of course. It will also allow users of other packages to replicate the method.

13.1 THE METHOD USING A PROGRAM

13.1.1 INTRODUCTION

We start by presenting a general method allowing making a model reach given values for a set of endogenous variables, by freeing the values of an equivalent set of exogenous elements.

This method can treat single or multi period simulations, and rational expectations can be considered at a small price in complexity (but a higher one in computation time, especially in the last case)

At first, we shall consider solving a backward-looking model for a single period.

13.1.2 THE PROBLEM

Let us consider the model:

$$y = f(y, x)$$

with:

- y vector of endogenous with dimension m
- x vector of endogenous with dimension n

We have suppressed the time index, as past (and future) values do not change in the solving process.

Let us suppose we want to obtain:

$$y^s = y^{s*}$$

For a subset of y of dimension p

through the endogenization of a subset of x with the same dimension, called x^s.

13.1.3 THE LINEAR CASE

If f was a linear function:

$$y = A \cdot y + B \cdot x + b$$

we would have

$$y = (I - A)^{-1}(B \cdot x + b)$$

Let us call C the sub-matrix $(I - A)^{-1} \cdot B$ (dimension (p x n) limited to the lines of y^s et $C^s$the sub-matrix (p x p) of C associated to the elements ofx^s.

We have:

$$y^{s*} = C^s \cdot (x^{s*} - x^{s0}) + C \cdot x^0 + c$$

The values allowing reaching $y^{s*}$are obtained immediately through:

$$x^{s*} = x^{s0} + (C^s)^{-1}(y^{s*} - C \cdot x^0 - c))$$

$$x^{s*} = x^{s0} + (C^s)^{-1}(y^{s*} - y^{s^0})$$

The C matrixes can be obtained easily by a matrix transformation, once A, B and b are known.

13.1.4 THE NON-LINEAR CASE

Let us now consider the non-linear case. We still have:

$$y = f(y, x)$$

and we still want to get:

$$y^s = y^{s*}$$

We shall start by supposing that the model is quasi-linear and apply the above method. The C matrix becomes:

$$\partial g(y)/\partial g(x)$$

Where g represents the function $y = g(x)$ where the y vector has been eliminated from the right term.

It is unfortunately clear that identifying the g function is difficult if not impossible in practice. On the other hand, the values of the Jacobian are easy to obtain, or its approximation through finite differences.

One just has to solve the model for starting values of x, apply in sequence small shocks to each of the exogenous, and compute the associated solution. We get:

$$\partial g(y_i)/\partial g(x_j) \approx (g_i(x + \Delta x_j) - g_i(x))/\delta x_j$$

where $\Delta x_j$is a vector of dimension n with null values except for x_j, changed by δx_j.

To get the equivalent of matrix C, we just have to select in the Jacobian the columns and lines associated with $x^s$and y^s.

The solution uses the same formula as the linear case:

$$x^{s*} = x^{s0} + (C^s)^{-1}(y^{s*} - y^{s0})$$

One shall notice that solving the model and computing the changes are only required for the exogenous concerned, and the inversion applies to a much smaller (p x p) matrix.

The only difference with the linear case is that the solution is not exact, due to the linear approximation. But one just has to extend the process, starting from the new values obtained for x (actually limited to x^s). The process remains exactly the same and can be repeated as many times as necessary. And as the linear approximation becomes more and more accurate (in term of properties) when we get closer to the solution (especially if the model is locally convex), the share of error eliminated each time becomes higher and higher, and convergence should be reached after a few iterations.

We identify here a Newton type method, which is characterized by a supralinear convergence property. An interesting feature is that we know the actual precision of the solution, contrary to other algorithms which consider only the speed of convergence.

13.1.5 A FEW QUESTIONINGS

Now, several questions appear:

The method we suggest looks quite simple for a package which does not provide an automatic process (such as Troll), and represents an efficient alternative to more complex (and less efficient) tools. One can question why it is not proposed right away by packages and applied more often by modelers (unless we are not up to date as to the present fashions).

Several answers can be considered:

- It does not work in practice, or only in very simple (non-operational) cases. We shall prove it works, at least for simple but operational models. There is no reason (but it remains to be proved) that it does not work also for the majority of reasonable cases.
- It does not work always. But then we can at least try, and maybe improve it in “difficult” cases. We can expect two main cases for failure:

  - The model is highly nonlinear. This is one of the problems with the Newton method, which can lead the solution to absurd values if the Jacobian computed is very far from its “exact” value. It is essential to start from values rather close to the solution, which should be the case in operational cases, if the forecasting quality of the model is high enough to provide spontaneously reasonable, if not acceptable, solutions.
  - The Jacobian is highly collinear, and cannot be inverted, or suggests once inverted very high changes in the exogenous (at least two of them). This case is less problematic, as another method (« by hand » included) will face the same problems, and here at least the problems will be evidenced immediately, and maybe interpreted through the observation of the matrix. The time gained can be used to look in other directions.

- It does not apply to the problem. Indeed, the method is limited to reaching of a given set of n endogenous values by freeing n exogenous elements. A professional modeler might want to spread the charge of the process on more exogenous elements. But if the weight of the contributions is known, the method can still be applied: several exogenous will be affected, but the algorithm looks only for one value. For instance, if the rates of social contributions by employers and workers are used simultaneously with the same variation, this variation becomes the unique endogenous element.

More problematic is the case of limits on changes, set by the user[168]: a given exogenous is used until it reaches a limit, in which case the computation switches to a new element. This case can also be treated automatically, but the program become more complex, and the solution is not guaranteed within the acceptable range (but neither "by hand" anyway).

- Finally, if the method supposes backward looking processes, it can also be applied to forward looking models. Two techniques can be considered, associated to the current methods for the simple case:

  - Fair Taylor: we iterate on the period-by-period solution, initializing future values by thus of the last iteration (at the beginning, the historical, theoretical, or previously computed values).
  - Lafarge: we solve the global models where equations are repeated along the time dimension.

But now we have to consider two nested loops, and the process can become unacceptably slow. Specific improvements could be found.

13.1.6 A FEW IDEAS ON CHOSING THE INSTRUMENTS

In choosing both targets and instruments, a few principles can be established, based on the general idea that the Jacobian matrix should be as regular as possible.

- The targets should be as independent as possible. For instance, using two correlated deflators, or demand + imports, will be dangerous.
- Even more important, the instruments should have different impacts. In the same way, using two instruments increasing household revenue (like social security payments and the income tax rate) can lead to problems.
- The difference between the base solution and the targets must not be contradictory. For instance, it will be difficult to improve the budget balance and increase GDP.
- The instruments should have an important impact. Reaching a realistic target must not need unrealistic variations.
- And obviously the target itself should be realistic….

13.1.7 A TEST ON A VERY SMALL MODEL.

We shall present here the conditions and results of a test conducted on a very small model, called pic_b, with 18 equations (5 estimated) used by Jean Louis Brillet for his course in modeling techniques.

We have supposed a change in GDP (called Q) and imports at constant prices (M) of respectively +5% and -5% compared for the historical values for the same period. The change in values goes much beyond usual cases, and their direction is rather opposed: an increase in GDP generates imports, even if all things being equal an increase in imports reduces GDP.

---

[168] Mostly for sociological reasons : it not reasonable to expect an increase in the income tax by 20% to accepted by households.

We shall use as targeting variables the real wage rate WR and world demand WD. These elements are rather independent, first from each other (this should increase their efficiency and accuracy, as a team with different skills), and also from the targeted elements.

We will give:

- The EViews program, clearly short in statements (but not in comments). It can apply to any model, the only information requested (in addition to the access to the mode itself) being the period concerned, the lists of targeted variables (with their values) those of targeting elements.
- The results from iterations. In this list, the "_cur" suffix is associated with current values (of endogenous and exogenous) and the "_star" suffix to the endogenous targets.

The « delta » values represent the difference between target and present solution, for each targeted endogenous. We observe that we need in practice 4 or rather 3 iterations (the first one consisting in simulating the model with the initial values). The relative higher gain is generally obtained at the second.

As an indication, running the program takes 0,003 seconds on a portable computer of average power.

We shall the replicate this test on a larger model.

13.1.7.1 The EViews program

```
'============================================
'       This program computes for a set of exogenous variables
'        the values
'       which allow a set of endogenous to reach given solutions

'       The two sets must have the same dimension
'       The pic_b model is used as an example
'        but the program is quite general
'       The present program applies to a single period
'       but it can be generalized quite easily
'============================================
'      We set the directory
cd "c:\program files\eviews5\pic\"
'      Results will be dumped as text to a file called jacob.txt
'      overriding any previous elements
output(t,o) jacob
'
'     We create a specific workfile called pic_star
'     from the original pic_b
'     We make sure no file of this name is open at present
'
'
close pic_b.wf1
close pic_star
wfopen pic_b
wfsave pic_star
'      We create a new model
'      Basically, the purpose is to present the model used as an example
delete(noerr) _pic_b
model _pic_b
_pic_b.append cap=pk*k(-1)
_pic_b.append ur=q/cap
_pic_b.append q+m=fd+x
_pic_b.merge eq_i
_pic_b.append log(prle_t)=c_prle(1)+c_prle(2)*(t-2002)+c_prle(3)*(t-t1)*(t<t1)+c_prle(4)*(t-t2)*(t<t2)
_pic_b.append led=q/prle_t
_pic_b.merge eq_le
_pic_b.append lt=le+lg
_pic_b.append rhi=wr*lt+xhr*q
_pic_b.append co=rhi*(1-sr)
```

```
_pic_b.merge eq_ic
_pic_b.append fd=co+i+gd+ic+rfdx*q
_pic_b.append res_m=log(m/(fd+0.5*x)) -1.322108*log(ur)+0.419518*log(compm)-
0.012582*(@trend(60:1)*(t<=2002)+@elem(@trend(60:1),"2002S2")*(t>2002))
_pic_b.merge eq_m
_pic_b.append res_x=log(x/wd)+0.686672*log(ur)-4.87E-
05*(@trend(60:1)*(t<=2002)+@elem(@trend(60:1),"2002S2")*(t>2002))
_pic_b.merge eq_x

_pic_b.append k=k(-1)*(1-dr)+i
'      Now we start the actual program
'      We shall use the period 2001S1
smpl 2001S1 2001S1
'      We define the lists of elements which change status
'       exogenous to endogenous
'       endogenous to exogenous
'       The numbers must be the same
'       If not, the program stops with a message
group g_vexog1 wd wr
group g_vendo1 q m
!t1=g_vexog1.@count
!t2=g_vendo1.@count
if !t1<>!t2 then
statusline !t1  exogenous  !t2  endogenous: stop
stop
endif
'      nstar is the common number
scalar nstar=g_vexog1.@count
'      We delete Scenario 1 (it exists)
_pic_b.scenario(d) "Scenario 1"
'      We build Scenario 0 as the base
'      It will contain the current solution
'      The associated suffix will be "_0"
'      We declare all the overriden exogenous
'      (with the currently best solution)
'      (note: we start with a blank list and fill it one by one)
_pic_b.scenario(n) "Scenario 0"
_pic_b.append assign @all _0
_pic_b.override
for !i=1 to g_vexog1.@count
%1=g_vexog1.@seriesname(!i)
_pic_b.override(m) {%1}
next
```

```
'     Now we create the alternate scenarios
'     for computation of the Jacobian
'     They are numbered 1 to nstar (here 2)
'     The suffix is "_n"
'     The exogenous are the same as Scenario 0
'     Note: all exogenous have to be overriden
'     but only one will change from the _0 value
'     for a given scenario
for !i=1 to nstar
_pic_b.scenario(n,a={!i},i="Scenario 0") "Scenario {!i}"
_pic_b.append assign @all _{!i}
next
'     Now we give the values for the targets
'     We used arbitrary numbers
smpl 2000S1 2001S1
genr q_star=q*(1+0.05*(t=2001))
genr m_star=m*(1-0.05*(t=2001))

'     Now we define the Jacobian
matrix(2,2) jacob
'     Now we initialize the exogenous instruments
'     used to reach the target
for !i=1 to g_vexog1.@count
%1=g_vexog1.@seriesname(!i)
genr {%1}_cur={%1}
genr {%1}_0={%1}
for !j=1 to g_vexog1.@count
genr {%1}_{!j}={%1}
next
next
'      And the endogenous target variables
for !i=1 to g_vendo1.@count
%1=g_vendo1.@seriesname(!i)
genr {%1}_0={%1}
genr {%1}_cur={%1}_0
next
'     We set the maximum number of iterations
'     and the convergence criterion
scalar nitmax=20
scalar sconv=1e-6
'     We initialize as control variables
'      the number of iterations
```

```
'      And the variable controlling the convergence
!niter=1
!iconv=0
'     We double the elements so they can be displayed
scalar niter=0
scalar iconv=0
'    We display several elements for control

print g_vexog1 g_vendo1
print *_cur *_star
'---------------------------------------------------------
'      Now we start the loop
'---------------------------------------------------------
smpl 2001S1 2001S1
'     It will run while convergence is not achieved
'     and the maximum numbre of loops is not reached
while !niter<=nitmax and !iconv=0
'       We test for convergence
'       We suppose it is achieved (!iconv=1)
'       Then we look at the relative difference
'        between the present solution and the target
'        for each targetted endogenous
'        If the relative difference is higher than the criterion
'        for any target
'        convergence is not yet reached (!iconv=0)
'        else we stop
!iconv=1
scalar iconv=1
for !j=1 to g_vendo1.@count
%1=g_vendo1.@seriesname(!j)
genr delta_{!j}=({%1}_star-{%1}_cur)/{%1}_cur
scalar delta=@elem(delta_{!j},"2001S1")

if @abs(delta)>sconv then
!iconv=0
scalar iconv=0
endif
next
if !iconv=1 then
stop
endif
'     We increment the iteration counter
```

```
!niter=!niter+1
scalar niter=niter+1
'      We display the differences
print delta_*
'       We compute the base solution
'        defined as Scenario 0
_pic_b.scenario "Scenario 0"
solve _pic_b
'       The solution values are set as current
for !i=1 to g_vendo1.@count
%1=g_vendo1.@seriesname(!i)
genr {%1}_cur={%1}_0
next
'      We display the current solution (exogenous + endogenous) and the targets
print niter
print *_cur  *_star
'      Now we compute the Jacobian
'      We start a loop on the shocked instruments
for !i=1 to g_vexog1.@count
'       For each exogenous, we compute the whole set of overriding instruments
'       shocking only the current one by 0.1 %
'       while the others keep their base value
'       This is necessary to implement the current changes
for !j=1 to g_vexog1.@count
%2=g_vexog1.@seriesname(!j)
genr {%2}_{!i}={%2}_cur*(1+.001*({!i}={!j}))
next
'       We solve the model under Scenario !i
_pic_b.scenario "Scenario {!i}"
_pic_b.solve
'       Now we compute the relative change in the target endogenous
'       This will give a column of the Jacobian matrix
for !j=1 to g_vendo1.@count
%2=g_vendo1.@seriesname(!j)
_z=(({%2}_{!i}-{%2}_cur)/{%2}_cur)/0.001
jacob(!j,!i)=@elem(_z,"2001S1")
next
next
'      This ends the computation of the Jacobian

'       We compute its inverse
matrix jacob_inv=@inverse(jacob)
```

```
'	We apply to each exogenous a change equal to
'	the inverse of the jacobian
'	times the remaining error.
'	This gives a new solution for the exogenous
'	 We start the process again until convergence (hopefully)
for !i=1 to g_vexog1.@count
%1=g_vexog1.@seriesname(!i)
genr {%1}_old={%1}_cur
for !j=1 to g_vendo1.@count
%2=g_vendo1.@seriesname(!j)
genr {%1}_cur={%1}_cur+{%1}_old*jacob_inv(!i,!j)*({%2}_star-{%2}_cur)/{%2}_cur
next
genr {%1}_{!i}={%1}_cur
genr {%1}_0={%1}_cur
next
wend
```

13.1.7.2 The results

```
Iteration 0: Initial values
====================================================
    obs         2000S1      2000S2      2001S1
====================================================
    WD         411615.4    432806.5    439347.6
    WR         0.028897    0.029101    0.029511
     Q         1344432.    1365342.    1385977.
     M         356344.0    383838.0    389639.0

====================================================
====================================================
    obs         2000S1      2000S2      2001S1
====================================================
   M_CUR       356344.0    383838.0    389639.0
   Q_CUR       1344432.    1365342.    1385977.
  WD_CUR       411615.4    432806.5    439347.6
  WR_CUR       0.028897    0.029101    0.029511
  M_STAR       356344.0    383838.0    370157.0
  Q_STAR       1344432.    1365342.    1455275.

====================================================
==========================
```

```
    obs          2001S1
==========================
   DELTA_1      0.050000
   DELTA_2     -0.050000

==========================
Iteration 1: solution with the initial exogenous
==========================
    obs          2001S1
==========================
    M_CUR       409107.1
    Q_CUR       1383326.
   WD_CUR        439347.6
   WR_CUR        0.029511
   M_STAR       370157.0
   Q_STAR       1455275.

   DELTA_1      0.052012   The model solution gives larger errors.
   DELTA_2     -0.095208   We have not started targetting yet.

==========================
Iteration 2
==========================
    obs          2001S1
==========================
    M_CUR       365029.1
    Q_CUR       1441045.
   WD_CUR        887589.2
   WR_CUR        0.012164
   M_STAR       370157.0
   Q_STAR       1455275.

   DELTA_1      0.009875
   DELTA_2      0.014048

==========================
Iteration 3
==========================
    obs          2001S1
==========================
    M_CUR       370184.2
    Q_CUR       1455269.
```

```
  WD_CUR      893720.4
  WR_CUR      0.012746
  M_STAR      370157.0
  Q_STAR      1455275.

  DELTA_1     4.35E-06
  DELTA_2    -7.34E-05

========================
Iteration 4: Convergence!
========================
   obs        2001S1
========================
   M_CUR      370157.0
   Q_CUR      1455275.
  WD_CUR      893915.9
  WR_CUR      0.012738
  M_STAR      370157.0
  Q_STAR      1455275.

  DELTA_1     2.25E-07
  DELTA_2     5.42E-08
```

13.1.8 A TEST ON A LARGER MODEL

We reproduced the test on cases closer to operational.

Note: applying the program to a new model is immediate, provided it is available in an EViews workfile along with the required data. One has only to update the two lists of targets and instruments, set the values for the targets, and call for the workfile. In all it should take less than ten minutes.

First case: the quarterly model of the French economy, which we have just used to illustrate modelling under EViews. It contains 88 equations, 11 of which are behavioral. Its specifications are quite traditional, and all behavioral estimations are validated by econometrics. It features a Keynesian and a price-wage loops.

We have also selected a more operational case. We ask the model to reach.

- An unemployment rate of 6%
- An inflation rate of 2%.
- A null trade balance (exports = imports at current prices).
- A budget deficit of 1% of GDP.

For this we shall use the following instruments:

- The rates of social contributions by both firms and workers.
- Government investment at constant prices.

- The number of civil servants.

The algorithm converges again, this time in 5 iterations. Computation time is 0.231 seconds.

Second case: a multi country yearly model for the Andean community (5 countries including Venezuela and a simplified rest of the world). It includes 770 equations, 11 x 5 of which are behavioral. Its specifications are again quite traditional, and all behavioral estimations are calibrated this time. It formalizes completely the trade flows.

This time we set as targets the Government deficit of each country (at 3% of GDP) and as instruments in sequence:

- Government investment
- The VAT rate
- The rate of social security contributions by workers
- The rate of social security contributions by firms.
- The exchange rate
- An increase in the long-term interest rate.

THz system converged in all cases, taking between 3.7 and 5.2 seconds (5 iterations most of the time).

and finally:

- targets: the Government deficit of each country (at 3% of GDP) and balanced trade (net exports=0)
- instruments: Government investment and subsidies to firms.

converges in 7 .2 seconds.

13.2 USING THE EVIEWS INTEGRATED COMMAND

Of course, this new feature is better integrated in the software (but limited to EViews). As usual, you can work through menus and programs.

The syntax for the program version is:

```
Name-of-the model.control(options) control target trajectory
```

You can either:

- give the full name of the control, target and trajectory series. You do not need to define options.
- Specify suffixes for the trajectory and (computed) control series.

options = (create, tsuffix=*suffix*, csuffix=*suffix*)

In case of 2 or more sets, you can use:

```
Name-of-the model.control(options) control1 control2 target1 target2 trajectory1 trajectory2

Name-of-the model.control(options) control1 target1 trajectory1 control2 target2 trajectory2

Name-of-the model.control(options) control1 control2 target1 trajectory1 target2 trajectory2 …..
```

13.2.1 AN EXAMPLE BASED ON OUR SMALL MODEL

We will present the method using the small described in chapter 13.

We will use Q (GDP) as target, and EC_M (the residual of imports) as instrument.

First, we will solve the model, and set the trajectory target (Q) as 5% higher than the result.

```
pic_b.append assign @all _b

pic_b.solve
genr q_t=q_b*1.05
```

Then we apply the method, specifying "_c" as the suffix for the instrument series found.

```
pic_b.control(create,csuffix=_c,tsuffix=_t)  q  ec_m
```

Then we solve the model overriding the instrument with the value found.

```
pic_b.append assign @all _c
 pic_b.override ec_m
 pic_b.solve
```

The result is indeed equal to the trajectory.

The case is favorable as the change in the instrument is small (and the staring point was close to the target).

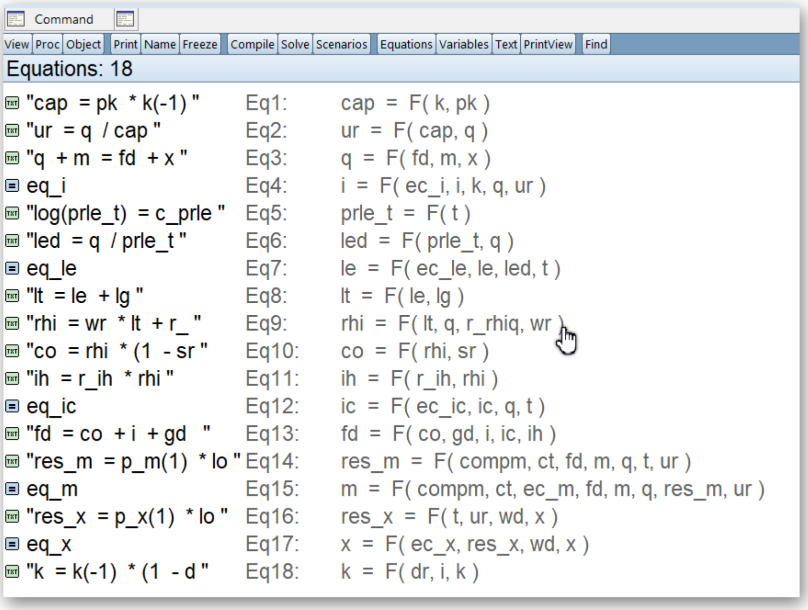
	2005Q1	2005Q2	2005Q3	2005Q4	2006Q1	2006Q2
Q	3.6200983E+11	3.6418841E+11	3.6638011E+11	3.6858500E+11	3.7080316E+11	3.7303467E+11
Q_B	3.6016430E+11	3.6149600E+11	3.6311300E+11	3.6494000E+11	3.6691620E+11	3.6905170E+11
Q_T	3.7817252E+11	3.7957080E+11	3.8126865E+11	3.8318700E+11	3.8526201E+11	3.8750429E+11
Q_C	3.7817250E+11	3.7957080E+11	3.8126870E+11	3.8318700E+11	3.8526200E+11	3.8750430E+11
EC_M	-2.3526954E-03	-2.3526954E-03	-2.3526954E-03	-2.3526954E-03	-2.3526954E-03	-2.3526954E-03
EC_M_C	-1.6698030E-01	-5.4794633E-03	-1.5024555E-02	-1.5654600E-02	-1.5059421E-02	-1.7597416E-02

In our first experiments, the algorithm can stall if the gap is high.

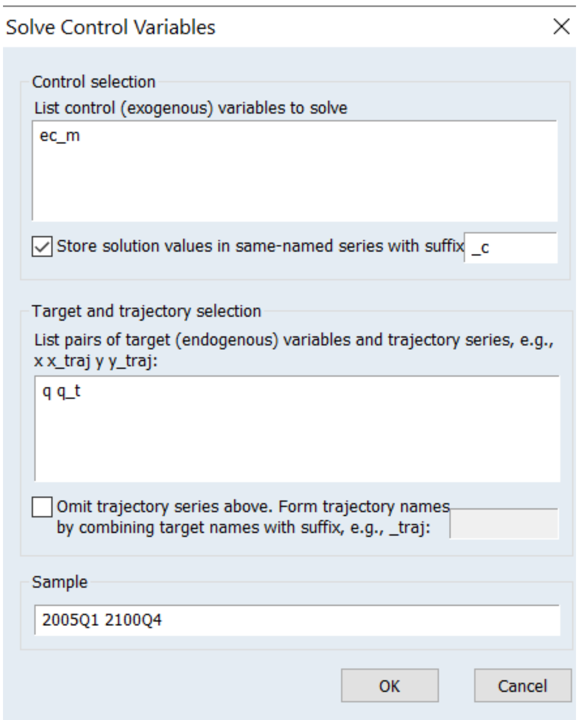
13.2.2 USING A MENU

Using a menu, the logic is exactly the same.

First, you access the model page.



In the “Proc” item, you select “Solve Controls for Target”. You obtain a menu in which you specify the elements:

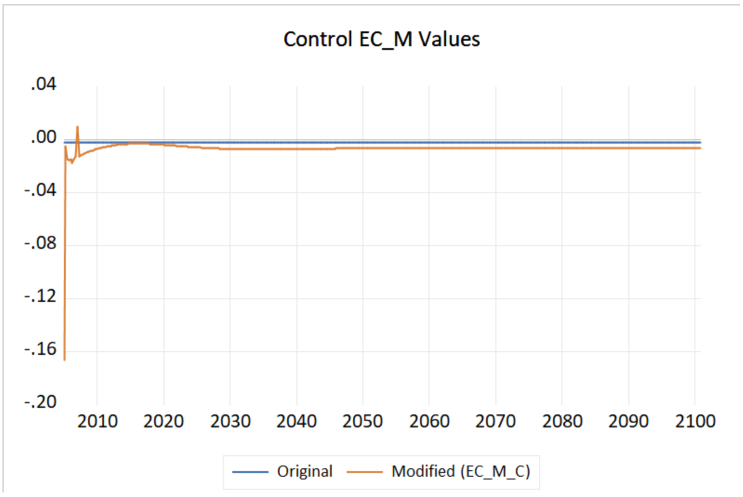


Control Value Percentage Change Summary

Control	Minimum	Q1	Q2	Q3	Maximum
	Percentage change (%)				
EC_M	-514.192	194.434	196.482	196.789	6997.404

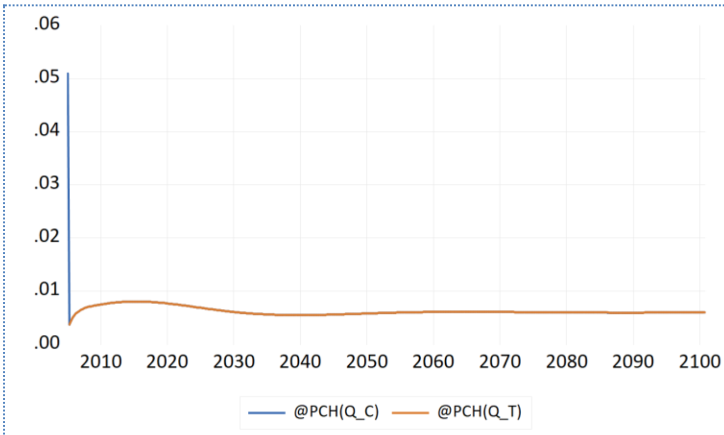
Control Value Level Change Summary

Control	Minimum	Q1	Q2	Q3	Maximum
	Level change				
EC_M	-0.165	-0.00463	-0.00462	-0.00457	0.0121



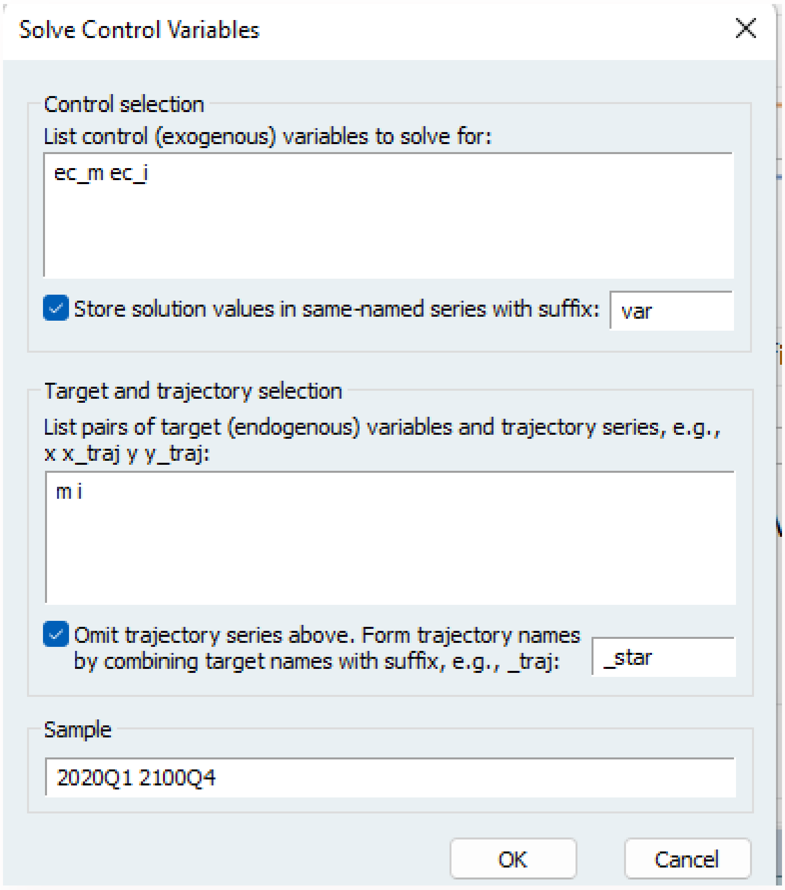
We solve the model again, using the value found for the instrument.

Obviously, it allows to meet the target.



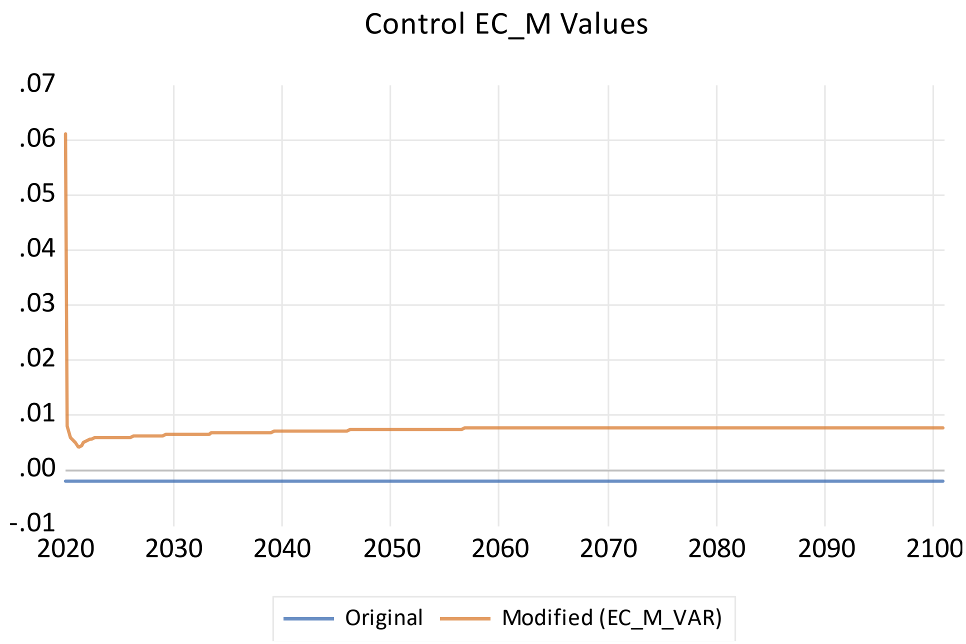
13.2.3 WITH TWO TARGETS

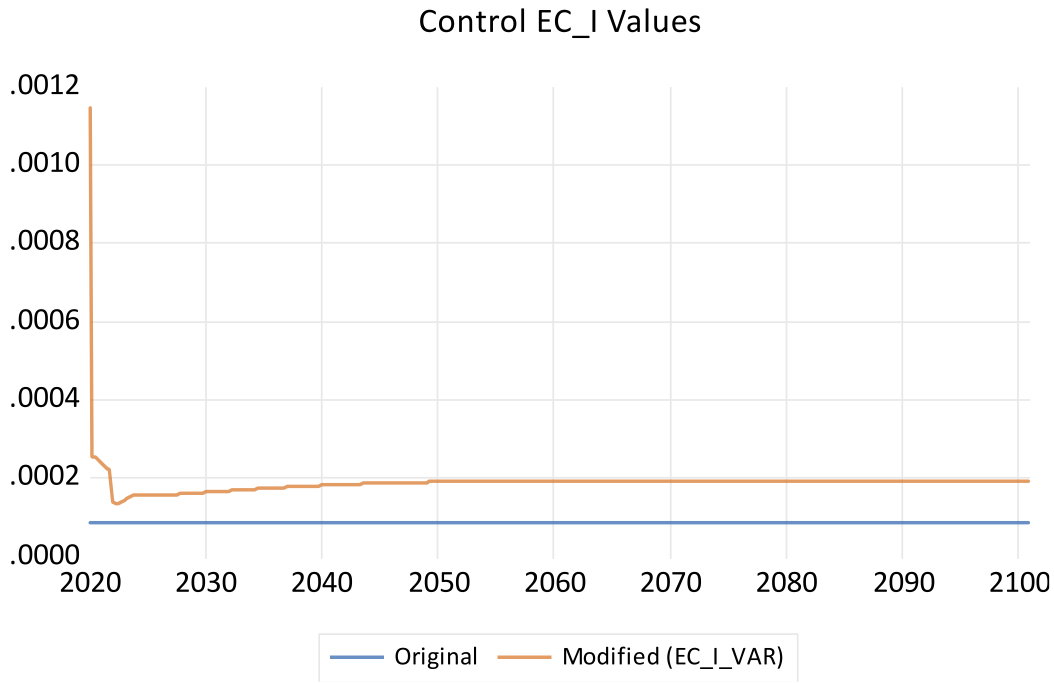
Let us now suppose we want to reach given values to investment M and investment I, through the determination of residuals EC_I and EC_M, over the 2020 – 2100 period, we will access the model item, then “Procs” and “Solve Controls for Targets"



Control	Percentage change (%)				
	Minimum	Q1	Q2	Q3	Maximum
EC_M	-2698.312	-417.858	-415.299	-390.636	-263.469
EC_I	53.399	110.343	120.146	121.203	1211.488

Control	Level change				
	Minimum	Q1	Q2	Q3	Maximum
EC_M	0.0062	0.00919	0.00977	0.00983	0.0635
EC_I	4.67e-05	9.64e-05	0.000105	0.000106	0.00106





13.3 RESPECIFICATION OF ENDOGENOUS AND EXOGENOUS VARIABLES

As we have not considered using this option in practice, we will only rely on the EViews integrated command.

This feature provides you with the edited version of the model.

The syntax is:

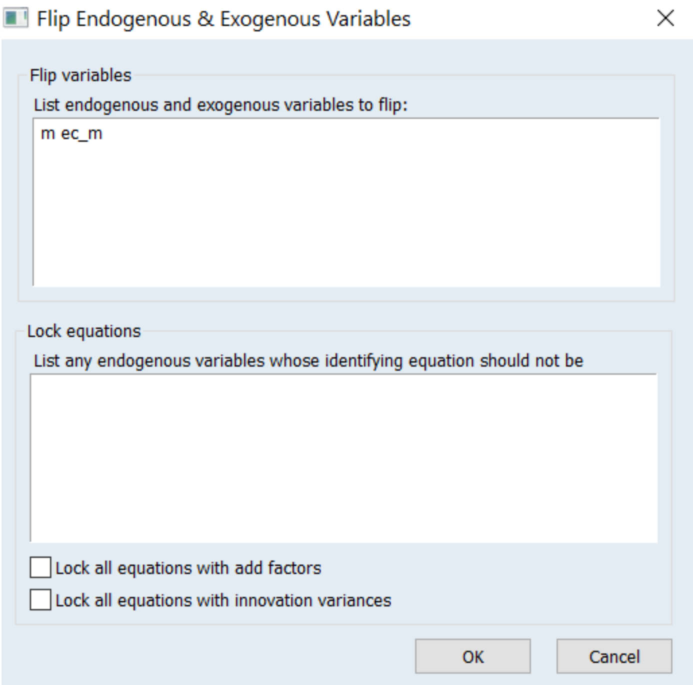
```
model_name.fliptype(dest-name=name) variables [@lock fixed-variables] [@lockafs] [@lockinnovs]
```

- The “dest-name” parameter allows to create a new model with the given name.
- The “@” modifiers forbids the method to modify
  - A list of variables
  - The equations with add-factors
  - The equations with innovations. This is useful for eliminating estimated equations, which generally imply a given causality.

In menu mode, one accesses the model menu, and uses the item ”Flip Endog/Exog Variables” under “Proc”.

13.3.1 THE FIRST EXAMPLE

It switches imports and its residual. This exogenizes M and computes the residual which between its value and the result from the model formula.



Model: PIC_B
Date: 06/11/20   Time: 20:04
Flipped the endogenous/exogenous type of 2 variables.

Respecification Summary

Variables made endogenous: EC_M
Variables made exogenous: M
Dropped innovation variance for variables: M
Broke link to embedded objects: EQ_M
Number of equations modified: 1

Equation Details

Equation DLOG(M) = @COEF(1) * DLOG(FD + CT * Q) + @COEF(2) * DLOG(UR) + @COEF(3) * DLOG(COMPM) + @COEF(4) * RES_M(-1) + @COEF(5
Rewritten as EC_M = 0.27576095249772596+0.102809064088758*RES_M(-1)+0.105751945795573*DLOG(COMPM)-1.67641497179222*DLOG(FD + CT*
Dropped innovation variance.

One can also reach a given GDP by endogenizing Government demand. The supply-demand equilibrium gives now Final demand FD, and GD as the residual substracting Household consumption, Investment, and the Change in inventories. All these elements come from the model solution, making it economically consistent.

```
Model: PIC_C
Date: 06/11/20   Time: 20:26
Flipped the endogenous/exogenous type of 2 variables.

Respecification Summary
-----------------------
Variables made endogenous:  GD
Variables made exogenous:  Q
Number of equations modified: 2

Equation Details
----------------
Equation Q  + M  = FD  + X
Rewritten as FD = M+Q-X

Equation FD  = CO  + I  + GD  + IC  + IH
Rewritten as GD = -CO+FD-I-IC-IH
```

13.3.2 A MORE COMPLEX EXAMPLE

Considering two switches, one can endogenize capital productivity PK and the wage rate WR to reach values of Imports M and Household consumption CO.

The two new chains are independent (one linked to supply, the other to demand). Imports give the rate of use UR, which gives capacity CAP and capital productivity PK. CO gives Household revenue RHI which gives the wage rate WR. This is rather logical.

```
Model: PIC_K
Date: 06/11/20   Time: 22:51
Flipped the endogenous/exogenous type of 4 variables.

Respecification Summary
-----------------------
Variables made endogenous:  PK  WR
Variables made exogenous:  CO  M
Dropped innovation variance for variables:  M
Broke link to embedded objects:  EQ_M
Number of equations modified: 5

Equation Details
----------------
Equation UR  = Q  / CAP
Rewritten as CAP = Q/UR

Equation DLOG(M) = @COEF(1) * DLOG(FD + CT * Q) + @COEF(2) * DLOG(UR) + @COEF(3) * DLOG(COMPM) + @COEF(4) * RES_M(-1) + @COEF(5) ...
Rewritten as UR = EXP(4.68709871606174*(0.27576095249772596 - EC_M + 0.102809064088758*RES_M(-1) + 0.105751945795573*DLOG(COMPM) ...
Dropped innovation variance.

Equation CO  = RHI  * (1  - SR)
Rewritten as RHI = CO/(1 - SR)

Equation CAP  = PK  * K(-1)
Rewritten as PK = CAP/K(-1)

Equation RHI  = WR  * LT  + R_RHIQ  * Q
Rewritten as WR = (-Q*R_RHIQ + RHI)/LT
```

13.3.3 A LESS LOGICAL ONE

Now, if we endogenize capital productivity PK and World demand WD to reach values of GDP Q and exports X, the process becomes less logical, as it transits through imports:

```
Model: PIC_F
Date: 06/11/20   Time: 22:21
Flipped the endogenous/exogenous type of 4 variables.

Respecification Summary
-----------------------
Variables made endogenous:  PK  WD
Variables made exogenous:  Q  X
Dropped innovation variance for variables:  M  X
Broke link to embedded objects:  EQ_M  EQ_X
Number of equations modified: 5

Equation Details
----------------
Equation UR  = Q  / CAP
Rewritten as CAP = Q/UR

Equation DLOG(M) = @COEF(1) * DLOG(FD + CT * Q) + @COEF(2) * DLOG(UR) + @COEF(3) * DLOG(COMPM) + @COEF(4) * RES_M(-1) + @COEF(5) ...
Rewritten as UR = EXP(4.68709871606174*(0.27576095249772596 - EC_M + 0.102809064088758*RES_M(-1) + 0.105751945795573*DLOG(COMPM) ...
Dropped innovation variance.

Equation Q  + M  = FD  + X
Rewritten as M = FD-Q+X

Equation CAP  = PK  * K(-1)
Rewritten as PK = CAP/K(-1)

Equation DLOG(X) = @COEF(1) * DLOG(WD) + @COEF(2) * RES_X(-1) + @COEF(3) + EC_X
Rewritten as WD = EXP(LOG(WD(-1)) + 2.024220093943912*(-0.0196006933825688 - EC_X + 0.091424035217868912*RES_X(-1) + LOG(X) - LO ...
Dropped innovation variance.
```

But of course one cannot object if the goal is only technical: to reach a solution.

It remains to be seen how this method affects the change of convergence for both Newton and Gauss-Seidel. This can be the goal of a further study.

13.4 THE CONSEQUENCES FOR THE SOLVING PROCESS

This is here that the choice of method has the highest impact.

- The "multivariate" method" uses a quasi-Newton algorithm, probably similar to the one detailed in the above paragraphs. The process is not accessible to the user. It should be fast (it is on the following example).

If it does not converge (a proven possibility), the remedies are limited. The main option consists in a change in the instruments, or a reduction of the gap between the target and the base solution, by using a less ambitious target or making a better guess.

One should remember that the problems should come from non-linearities, and the difference between the Jacobians (or rather their inverse) at the present values and the solution.

As the user has no control over the process, it is quite difficult to interpret the problems and ask for additional information, as in a normal application.

- The "respecification" method create a new model version. The solving process is the responsibility of the user, who can choose between Gauss-Seidel, Broyden or Newton.

For information on the differences between methods and their problems, please refer to Chapter 7.1.

Basically, solving problems rely on the linearity of the model, which has a different impact according to the algorithm. The Newton method solves a succession of linearized versions, and a quasi-linear model will converge quickly. Whereas the Gauss-Seidel process can lead away from the solution of a linear model, in a "cobweb" manner.

This means that for Gauss-Seidel, the problems can come from the eigenvalues of the Jacobian, which can be quite different from the original ones. It the new model framework is not checked for logic, is quite possible that the highest modulus is higher than one. If that happens, one should examine the whole set of new causalities, and try a different respecification. As we shall see, blocking some equations from the process is available as an option.

For Newton (and maybe Broyden) one should check that the process does not create equations with non-linear properties. Here logical interpretation can help. This will happen for instance if employment depends on unemployment, a volatile element.

A LIST OF USEFUL SERIES FOR PRODUCING A SINGLE COUNTRY, SINGLE PRODUCT MODEL

We shall give now a list of series normally useful for producing a simple but operational model. Of course, this is only a proposal, designed as a starting point to which the user is welcome to apply as many changes as he wishes.

Just as we have done for model specifications, we think it is easier to start from a proposal, even rather far from what the modeler has in mind, than from nothing. Convergence to an acceptable version can start immediately, by a series of acceptations, rejections and changes based on an actual solution.

14.1 THE PERIODICITY

It should be **quarterly**, annual if quarterly data is not available. This is not so important for population, or for stock variables: capital, housing, credits and debts.

14.2 GENERAL ELEMENTS

If several series are available for the same concept (as unemployment) the first priority is to get one series, the second to get all available series with their definition. Of course, if several sets of series are available for the same domain (like exports and imports in both local currency and dollars) the priority is to make the whole set as consistent **as possible**. This means for instance that being part of the major set of data used can be considered more important than perfect adaptation to the concept required.

We are not considering **product decomposition.** But oil could be separated in terms of production, exports, and relation to the State budget.

The agent decomposition is:

Households
Firms (private and public)
Financial institutions (non-state)
State (including local administrations and Social Security)
Rest of the World.

14.3 THE SERIES

The most important series are underlined.

Some of the variables can be computed from others. They can be provided or not.

We need

14.3.1 THE SUPPLY - DEMAND VARIABLES AT CONSTANT PRICES (WHOLE ECONOMY)

Total local demand (including intermediary consumption)
Intermediary consumption by use and product
Total local demand
Household consumption at constant prices
Housing investment by households at constant prices
Housing investment by the State at constant prices
Productive investment by firms at constant prices
Government investment at constant prices
Government consumption at constant prices
Changes in inventories at constant prices.

Exports at constant prices (constant terms)
Imports at constant prices (constant terms)
GDP at constant prices

14.3.2 THE SAME ELEMENTS AT CURRENT PRICES

14.3.3 THE SAME ELEMENTS DECOMPOSED INTO PRODUCTS

All elements have an additional dimension (3) except for

intermediate consumption (3 x 3)
productive investment (3 x 3)

14.4 PRODUCTIVE CAPACITY

Productive capital at constant prices
If not, capital depreciation at constant prices
Productive capital (material) at constant prices
Productive capital (building) at constant prices
Rate of use of capacities (ratio)

14.5 EMPLOYMENT

Employment of firms (global)
Employment of households
Employment of financial institutions
Employment of the State (including local administrations)
Self employed
Family workers
Unemployment
Unfilled job vacancies
Weekly work duration
Population
Population in age of working (or 15-65 years)

14.6 PRICE DEFLATORS

The deflators associated to the supply - demand equilibrium (but we can compute them ourselves if we have elements at both current and constant prices).

The consumer price index
The production price index
The yearly wages (global)
The yearly wages (firms)
The yearly wages (State)

The yearly wages (employed by households)
The yearly wages (by product)

14.7 HOUSEHOLDS ACCOUNT

All in current terms

Revenue before income tax
Revenue after income tax
Revenue from wages
Individual workers revenue
Pensions
Social security benefits (global)
Social security benefits (illness).
Social security benefits (family).
Revenue from housing
Interests received.
Income tax
Social security contributions
Interests paid.
Housing capital

14.8 FIRMS ACCOUNT (ALL TYPES OF FIRMS)

All elements in current terms

Value added
Intermediary consumption
Tax on production
Wages
Social security contributions paid by firms
Revenue of individual workers
Dividends paid
Dividends received
Interests paid
Interests received
Tax on profits
Productive investment
Change in inventories
Balance

14.9 REST OF THE WORLD

Trade balance
Financing capacity
Balance of services

Exchange rate to the dollar

Exchange rate to the partners of the country.

Balance of payments

Capital movements

Wage transfers of the country's expatriate workers.

Demand in the Rest of the World (weighted by the share of countries in local exports).

14.10 GOVERNMENT ACCOUNT (ELEMENTS NOT DESCRIBED EARLIER)

All elements in current terms

VAT

Other indirect taxes

Tariffs

14.11 FINANCIAL DOMAIN

The following applies to a traditional model. For Stock-Flow Consistent versions, the requirements are much more important. They are developed in the specific chapter.

Short-term interest rate

Long-term interest rate

Interest rate on consumption

Interest rate on housing investment

All elements in current terms

Stock of debt for firms (global)

Stock of debt for firms (private)

Stock of debt for firms (public)

Stock of debt for the State

Stock of debt for households

Capital movements

Foreign direct investment (in current terms)

Foreign direct investment (in constant terms)

A LIST OF USEFUL SERIES FOR A MODEL WITH SEVERAL SECTORS

15.1 THE PERIODICITY

Again, it should be **quarterly**. But this is not so important for population, or for stock variables: capital, housing, credits and debts.

15.2 GENERAL ELEMENTS

If several series are available for the same concept (as unemployment) the first priority is to get one series, the second to get all available series with their definition. Of course if several sets of series are available for the same domain (like exports and imports in both local currency and dollars) the priority is to get one consistent **set**.

The agent decomposition is:

Households

Firms (private and public)

Financial institutions (non-state)

State (including local administrations and Social Security)

Rest of the World

In addition to non-market services (separated in the first version) the **product decomposition** is:

Agriculture, fishing and forestry

Manufacturing

Market services

This decomposition applies to series noted *

In addition, **separations** should be considered for:

Agricultural production into firms and family

This applies to production, labor, investment, wages, and revenue

Manufacturing trade, into energy and non-energy

This applies to local and world demand, imports and exports at constant and current prices.

15.3 THE SERIES

The most important series are noted by *.

Some of the variables can be computed from others. They can be provided or not.

We need

15.3.1 THE SUPPLY - DEMAND VARIABLES AT CONSTANT PRICES (WHOLE ECONOMY)

* Total local demand (including intermediary consumption)

* Intermediary consumption by use and product

* Total final local demand

* Household consumption at constant prices

* Housing investment by households at constant prices

* Housing investment by the State at constant prices
* Productive investment by firms at constant prices
* Government investment at constant prices
* Government consumption at constant prices
* Changes in inventories at constant prices.
* Exports at constant prices (constant terms)
* Imports at constant prices (constant terms)
* GDP at constant prices

15.3.2 THE SAME ELEMENTS AT CURRENT PRICES

All elements have dimension 3 except for

Intermediate consumption (3 x 3)
Productive investment (3 x 3)
Change in inventories (3 x 3)

15.4 PRODUCTIVE CAPACITY

* Productive capital at constant prices

If not, capital depreciation at constant prices

* Productive capital (material) at constant prices
* Productive capital (building) at constant prices.
* Rate of use of capacities (ratio).

15.5 EMPLOYMENT

* Employment of firms (global).
Employment of households
Employment of financial institutions
Employment of the State (including local administrations)
* Wage earners in firms (global).
Employment of households
Employment of financial institutions
Employment of the State (including local administrations)
* Self employed
* Family workers
Unemployment
Unfilled job vacancies
Weekly work duration
Population
Population in age of working (or 15-65 years)

15.6 PRICE DEFLATORS

The deflators associated to the supply - demand equilibrium (but we can compute them ourselves if we have elements at both current and constant prices).

The consumer price index

The production price index

* The yearly wages (global)

* The yearly wages (firms)

The yearly wages (State)

The yearly wages (employed by households)

The yearly wages (by product)

15.7 HOUSEHOLDS ACCOUNT

All in current terms

Revenue before income tax

Revenue after income tax

* Revenue from wages

* Individual workers revenue

Pensions

Social security benefits (global)

Social security benefits (illness).

Social security benefits (family).

Revenue from housing

Interests received.

Income tax

Social security contributions

Interests paid.

Housing capital

15.8 FIRMS ACCOUNT (ALL TYPES OF FIRMS)

All elements in current terms

* Value added

* Intermediary consumption

* Tax on production

* Wages

* Social security contributions paid by firms

* Revenue of individual workers

* Dividends paid

* Dividends received

* Interests paid

* Interests received
* Tax on profits
* Productive investment
* Change in inventories
* Balance.

15.9 REST OF THE WORLD

Trade balance

Financing capacity

Balance of services

*Exchange rate to the dollar

Exchange rate to the partners of the country

Balance of payments

Capital movements

*FDI

*Wage transfers of the country's expatriate workers.

*Demand in the Rest of the World (weighted by the share of countries in local exports).

15.10 GOVERNMENT ACCOUNT (ELEMENTS NOT DESCRIBED EARLIER)

All elements in current terms.

* VAT
* Other indirect taxes
* Tariffs

15.11 FINANCIAL DOMAIN

Short-term interest rate

Long-term interest rate

Interest rate on consumption

Interest rate on housing investment

All elements in current terms

Stock of debt for firms (global)

* Stock of debt for firms (private)
* Stock of debt for firms (public)

Stock of debt for the State

Stock of debt for households

* Capital movements
* Foreign direct investment (in current terms)
* Foreign direct investment (in constant terms)

RECENT ADDITIONS TO SESSION MANAGEMENT

The following features are not specially used for modelling

16.1 SNAPSHOTS

Snapshots allow you to keep the contents of a workfile as a hidden element, and compare them to the current items.

Clicking on a given item will give the actual changes (the numerical difference for a series).

Snapshots can be created automatically (with a given frequency) or on request.

Accessing a snapshot, you can have it replace the current workfile.

This feature is described in detail in the manual.

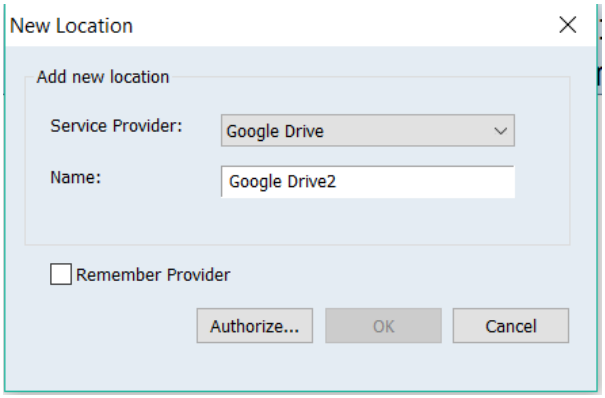
16.2 CLOUD MANAGEMENT

16.2.1 GOOGLE DRIVE

You can easily access a Cloud directory.

For instance for a Google Drive you can use

Open>Foreign Data as Workfile>New Location>



Then authorize and get access to the whole directory.

USING EVIEWS MENUS TO BUILD AND MANAGE AN ECONOMETRIC MODEL

In this chapter, we shall review the functions allowing to produce and use a model under EViews. To make the process clearer and shorter, we will use a very small model, with only two equations, one behavioral and one identity.

We will concentrate here on using menus, not the most efficient way as we have explained, and should be obvious already. Programming makes the task easier and (more importantly) allows more options. In some cases, it is the only way to work in a clear, fast and secure way.

But going through the following example, or perhaps his own similar case, one can rain in the different functions. Then comparing the following example with the same tasks addressed by the program, and keeping it as a reminder, should make the process easier.

Let us remember also that EViews provides, in the "Capture" window, a sequence of the commands performed during the session, including the program translation of the menu commands. This means that by saving this text one can reproduce all the tasks as a program, obviously after some editing.

17.1 GENERAL MENU FEATURES

17.1.1 COMMAND CAPTURE

This new function represents an important improvement in the management of sessions.

As we will see, modelling under EViews relies mostly on programming (except for the research for the best behavioral formula). However, this process calls very often for menu functions, in particular when one wants to get information in case of problems or questionings. For instance

- If the numerical profile of a computed series looks strange, one can rely on a graphical display.
- One can consider adding an explanation to an estimated equation, or changing its sample period.

This means that the user is quite frequently switching between running programs and calling for menu elements.

Before version 9, only the typed commands were retained by EViews (in the Log window). One had to note (by hand!) the accesses to menu functions.

Now the log contains also the transcription of the calls to menus, translated as commands. This means that the sequence of statements contains the entire set of function calls.

This set can be saved as a program, which describes all the tasks between two points of a particular session. In most cases, this program will give the same results as the ones obtained in the session. Of course, this is not always true, for instance as an element used in the beginning might have been modified later.

An additional good point is that one gets a transcription of menu commands as program elements. The syntax can be used later in programs, independently. This is very useful as the syntax of some commands (for instance in graph management or model solving) can be quite complex.

17.1.2 WINDOW MANAGEMENT AND DOCKING

EViews 9 provides a large number of features for managing windows. This allows to display one the same screen several elements, like a spreadsheet, a graph and a program. One particular feature is that the command window can be displayed anywhere, and can be separated from the Capture window, allowing to control both the sequence of commands and the sequence of tasks.

The window containing the messages from a program (if any) is also available directly as a clickable item.

17.1.3 SNAPSHOTS

Snapshots allow you to keep the contents of a workfile as a hidden element, and compare them to the current items.

Clicking on a given item will give the actual changes (the numerical difference for a series).

Snapshots can be created automatically (with a given frequency) or on request.

Accessing a snapshot, you can have it replace the current workfile.

This feature is described in detail in the manual.

17.1.4 LOGGING

This feature is quite interesting for controlling execution of programs.

17.2 A VERY SMALL MODEL

We will base the document on a very small model, admittedly artificial, but nevertheless based on actual data.

We will consider a self-sufficient economy, with no external trade. We will suppose that final demand is divided into public (called G) and private (called CO[169]), without separating the latter. This means in particular that we will not identify investment, capital and therefore supply.

These elements are measured at constant prices.

G will be exogenous, and CO endogenous, defined by a behavioral equation. Total final demand (equal to GDP) will be called Q.

This leads to the model:

[1] CO = f(Q)

[2] Q = CO + g[170]

---

[169] Logic would name it C, but this name is reserved by EViews for generic coefficients.

[170] When this is possible, we will restrict capital letter to endogenous.

17.3 THE DATA

Obviously, these variables cannot correspond to actual series, as any country has an external trade.

We will use World Bank data for France, for GDP and private consumption, which we will identify as demand. We will suppose that Government demand balances equation 2.

As we will only present basic features, we have not changed the presentation from the previous version of the book. Obviously, we are not producing an operational model, or even a model using actual concepts.

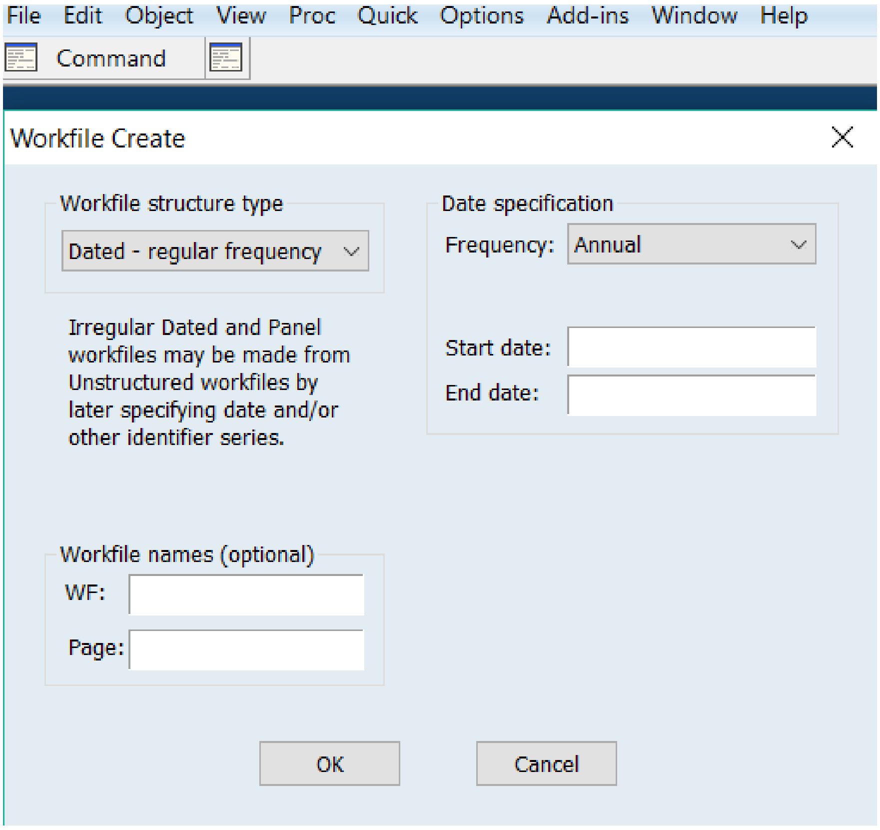
17.4 CREATING THE WORKFILE

Before taking any step, we need an EViews workfile (just as you need a desk to start working). This file will be the location where you keep the elements in your project, starting with the data itself.

You will access to the workfile creation window by calling for

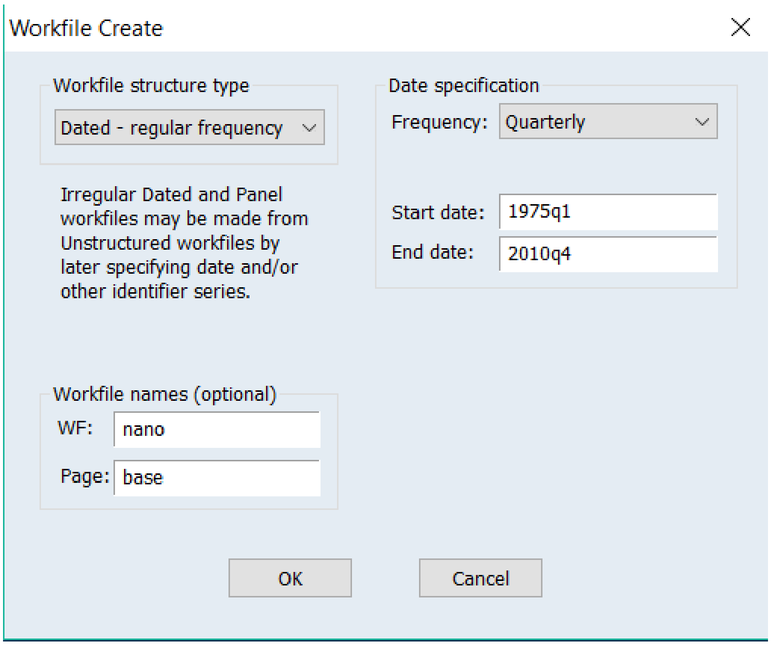
File > New> Workfile.

This opens the following menu:

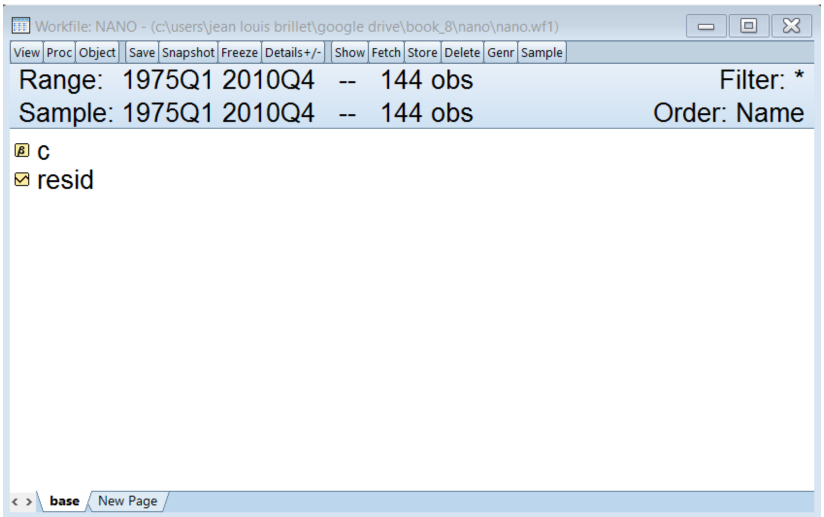


Our data will have a regular quarterly frequency, with observations from 1975 to 2010.

We will call our workfile “nano” with a page called “base”



This will create a workfile, which appears as:



We can see that:

- The current page is called “base”
- The (quarterly) data can start from 1975 to 2010.
- Any following statement will apply to the full sample.
- The file contains already two elements:
  - **c** which will contain the generic coefficient vector.
  - **resid** which will contain the residual from the last estimation process.

As we have made no estimation yet, **c** is filled with zeroes and **resid** by NAs.

These names cannot be given to another item.

17.4.1 PROGRAMMING

The equivalent statement in a program can be:

```
wfcreate(wf=nano_q,page=base) q 1975q1 2010q4
```

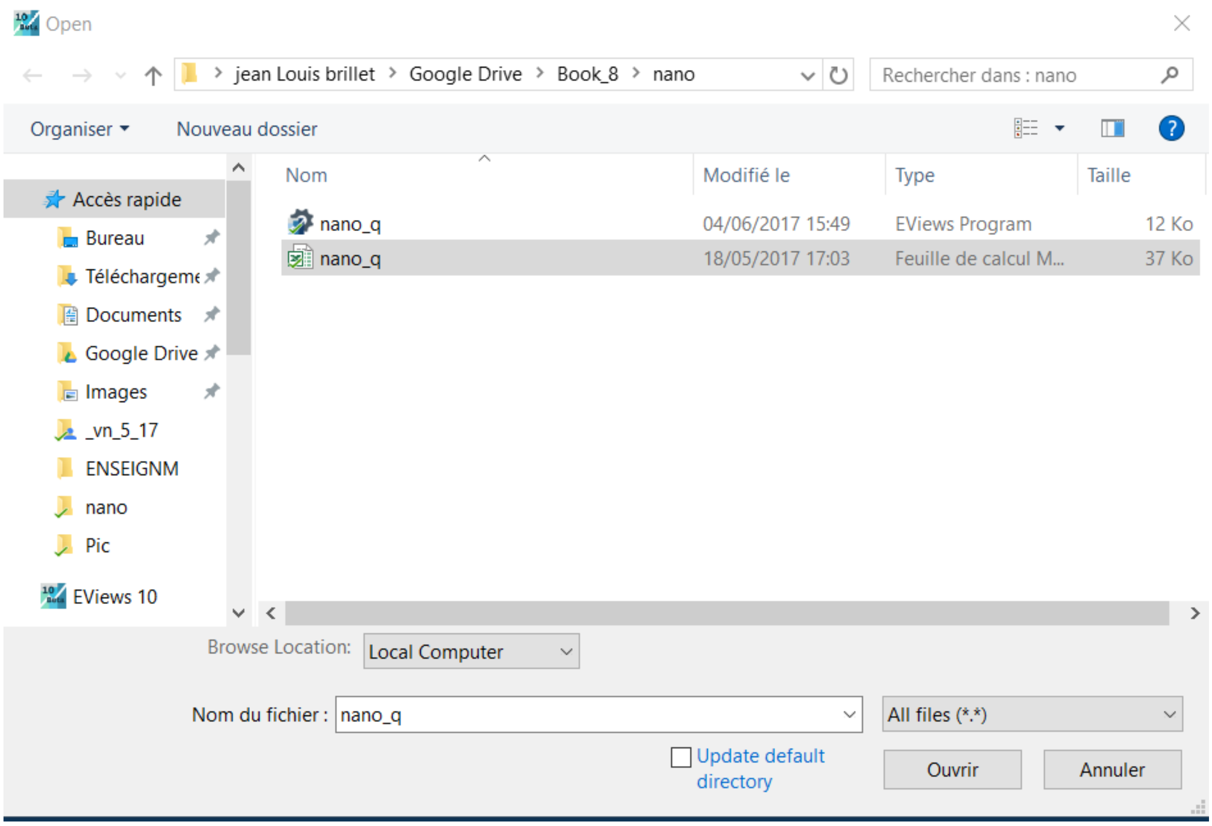
The only advantage of programming is that the stored statement can be repeated edited.

17.5 READING THE DATA

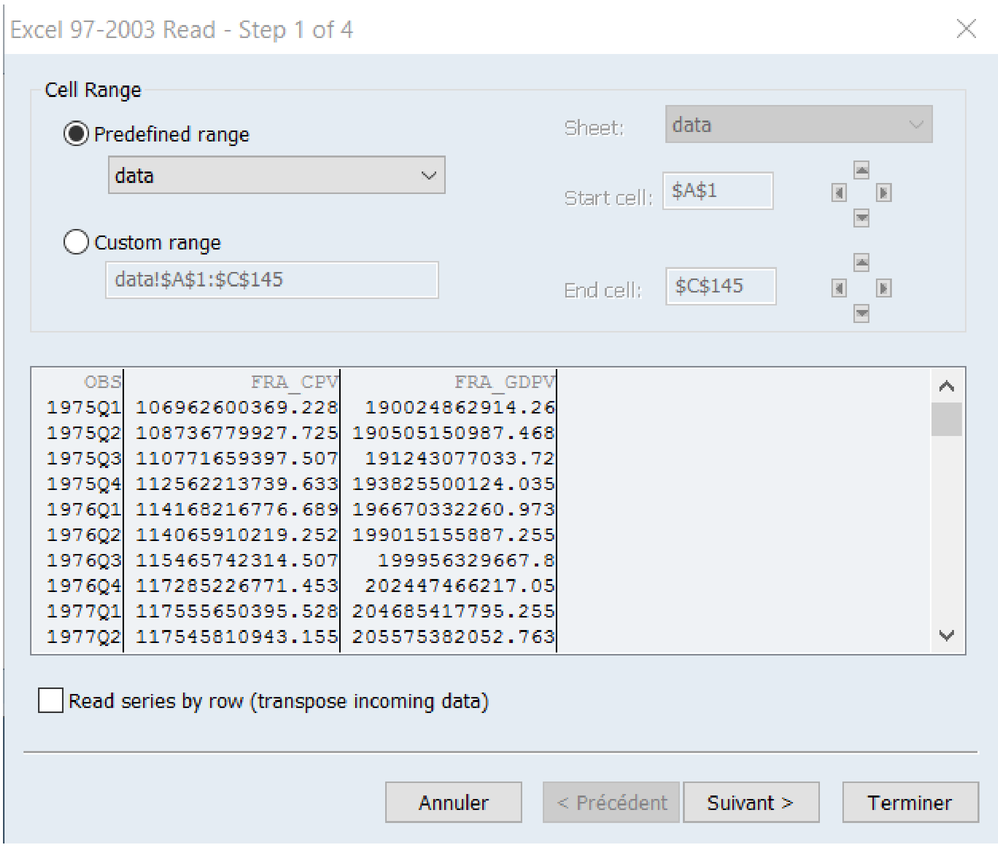
There are several ways to read the data into EViews, even in some cases from the same data set.

We will import the data from the file “nano_q.xls” by calling for:

Import>



Calling “Open” (or “Ouvrir”) gets you the following:

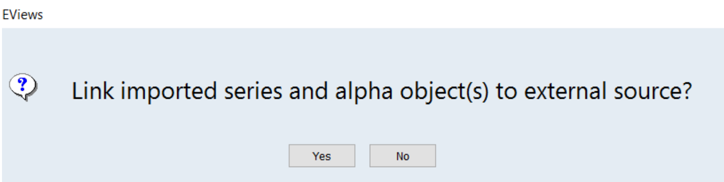


You can check that the data complies with the sample decided at workfile creation for the sheet “base”.

You get two series in columns: FRA_CPV and FRA_GDPV. These series are for private consumption and GDP respectively (V stands for constant prices). If series were in rows, the “Read series by row” should have been ticked.

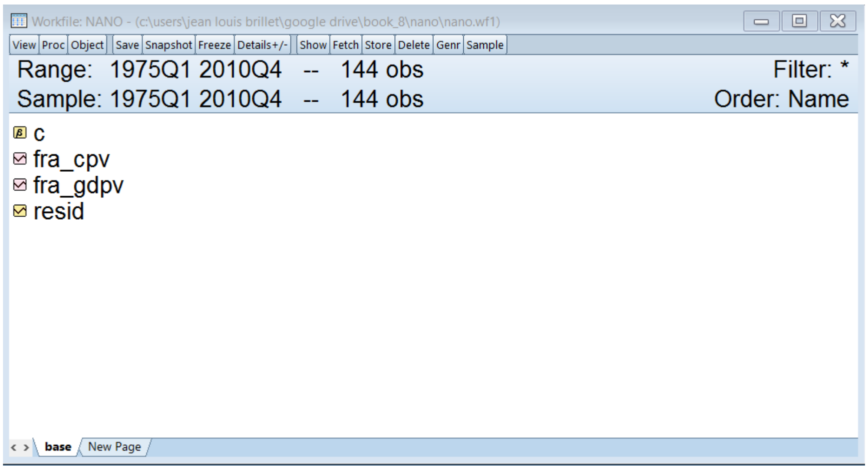
You can finish (“Terminer”).

But you get the message:

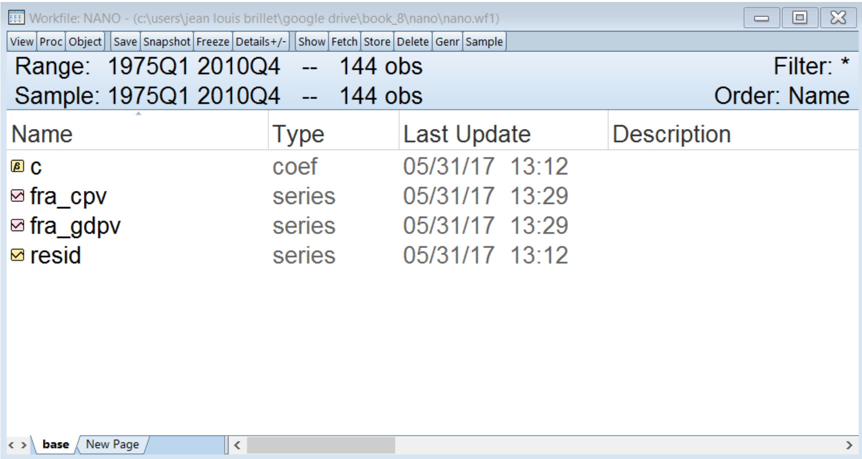


If you click “Yes” (we will do it) the series in the “data” page will be linked to their counterpart in the Excel file. This means that if the Excel file is modified, the change (if any) will apply to the workfile too.

The two series will appear in the page.

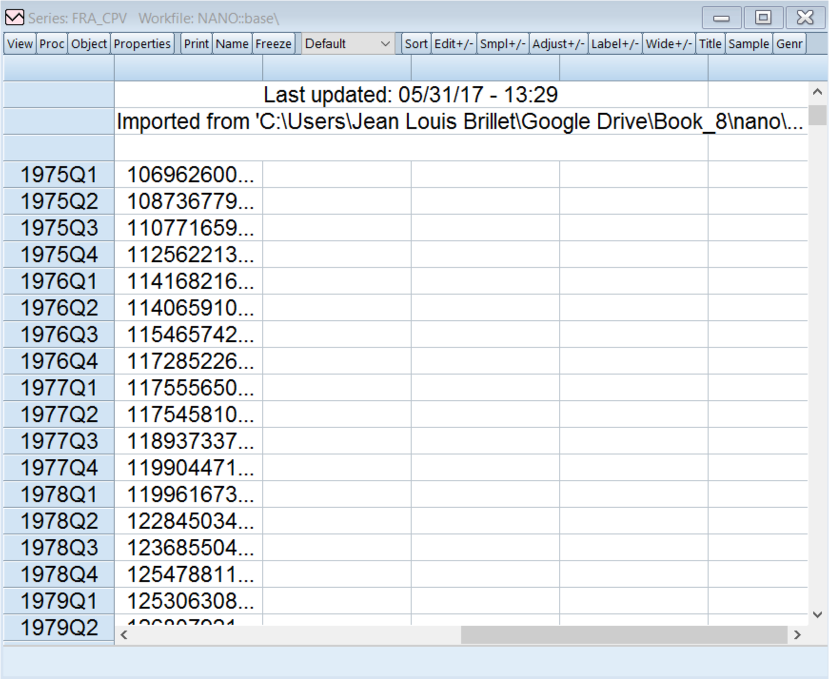


If we click on “Details +/- we get additional information:

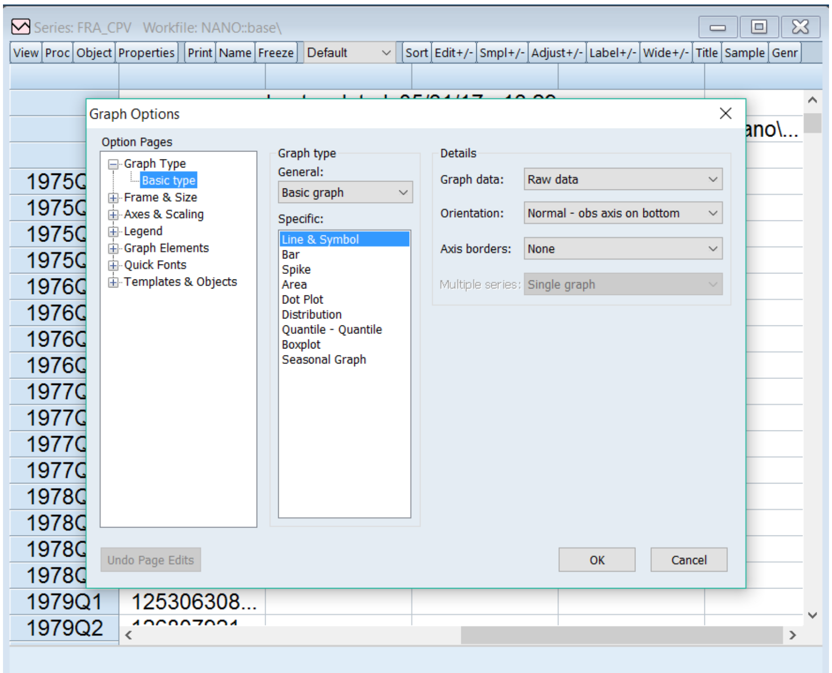


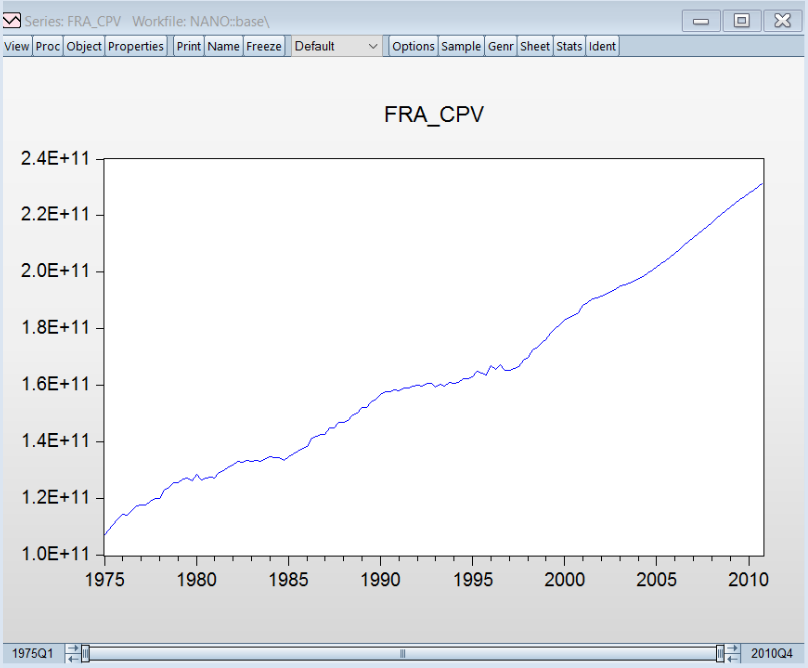
But of course the page will contain less series (not important here).

By double-clicking on a series (fra_cpv for instance) we get the values:


Series: FRA_CPV Workfile: NANO::base\
View Proc Object Properties Print Name Freeze Default Sort Edit+/- Smpl+/- Adjust+/- Label+/- Wide+/- Title Sample Genr
Last updated: 05/31/17 - 13:29
Imported from 'C:\Users\Jean Louis Brillet\Google Drive\Book_8\nano\...
1975Q1 106962600...
1975Q2 108736779...
1975Q3 110771659...
1975Q4 112562213...
1976Q1 114168216...
1976Q2 114065910...
1976Q3 115465742...
1976Q4 117285226...
1977Q1 117555650...
1977Q2 117545810...
1977Q3 118937337...
1977Q4 119904471...
1978Q1 119961673...
1978Q2 122845034...
1978Q3 123685504...
1978Q4 125478811...
1979Q1 125306308...
1979Q2

Or a graph:


Series: FRA_CPV Workfile: NANO::base\
View Proc Object Properties Print Name Freeze Default Sort Edit+/- Smpl+/- Adjust+/- Label+/- Wide+/- Title Sample Genr
Graph Options
Option Pages
Graph Type
Basic type
Frame & Size
Axes & Scaling
Legend
Graph Elements
Quick Fonts
Templates & Objects
Graph type
General:
Basic graph
Specific:
Line & Symbol
Bar
Spike
Area
Dot Plot
Distribution
Quantile - Quantile
Boxplot
Seasonal Graph
Details
Graph data: Raw data
Orientation: Normal - obs axis on bottom
Axis borders: None
Multiple series: Single graph
Undo Page Edits
OK
Cancel
1979Q1 125306308...
1979Q2



17.5.1 PROGRAMMING

The equivalent statement will be :

```
smpl 1975q1 2010q4
 read(t=xls,b2) nano_q.xls
```

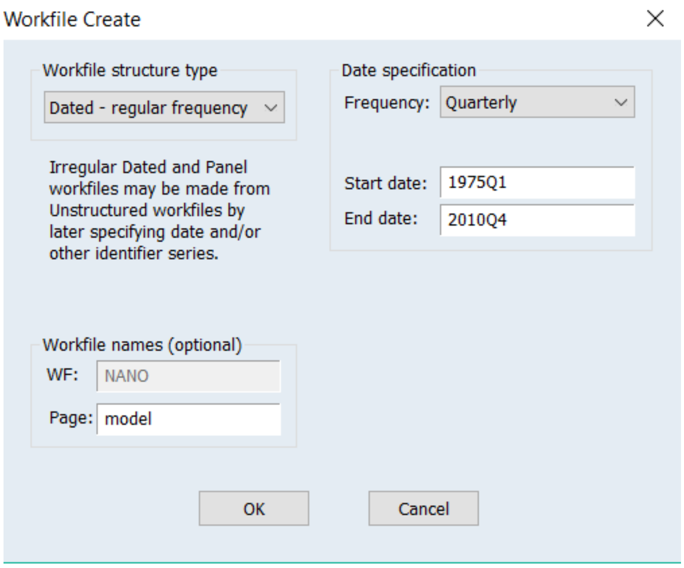
Again, the advantage is that the statement can be repeated, as an element of a program producing several tasks (perhaps creating a full model).

17.6 CREATING THE PAGE

We will not create the model (and the associated items) in the page “base”. There are two main reasons for this.

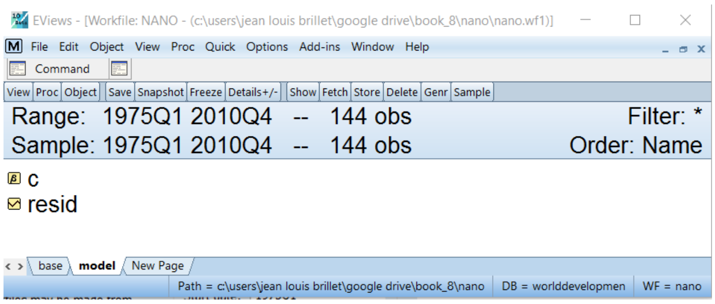
- We do not want to risk modifying the original data.
- We want to limit the number of items to the ones actually used.

The new page will be called “model”. EViews will propose the same characteristics as the original page “data”.



Of course it contains only “c” and “resid”.

We must give the page access to the original data.



17.6.1 PROGRAMMING

The programming equivalent will be :

```
pagecreate(page=base)  q 1975q1 2010q4
```

17.7 LINKING THE DATA

Of course, the simplest option would be to copy the original series in the new page. There is a cleaner and more efficient solution.

We start by specifying in the Command window:

Command

```
pageselect model
link Q
link CO
```

We get:

```
Range:  1975Q1 2010Q4  --  144 obs
Sample: 1975Q1 2010Q4  --  144 obs
c
co
q
resid
```

We can see that the two variables appear with a pink question mark.

Double clicking on CO we get

```
Series Link
Source series:
Source page:
Link by: Date (with frequency conversion)

***  Unable to Perform Link  ***
Workfile page "" not found
```

This means that the values for CO should be obtained in another source, which has not been yet defined.

The second phase will define the source.

Command

```
Q.linkto base\fra_gdpv
CO.linkto base\fra_cpv
```

```
Range:  1975Q1 2010Q4  --  144 obs
Sample: 1975Q1 2010Q4  --  144 obs
c
co
q
resid
```

And displaying the series will give (along with the values).

```
Last updated: 05/31/17 - 15:23
Imported from 'C:\Users\Jean Louis Brillet\Google Drive\Book_8\nano\nano_q.xls'
Page Link: base\fra_cpv
```

17.7.1 PROGRAMMING

The programming equivalent uses four statements

```
link CO
link Q
CO.linkto base::FRA_CPV
Q.linkto base::FRA_GDPV
```

The advantage of programming begins to appear.

17.8 DISPLAYING THE DATA

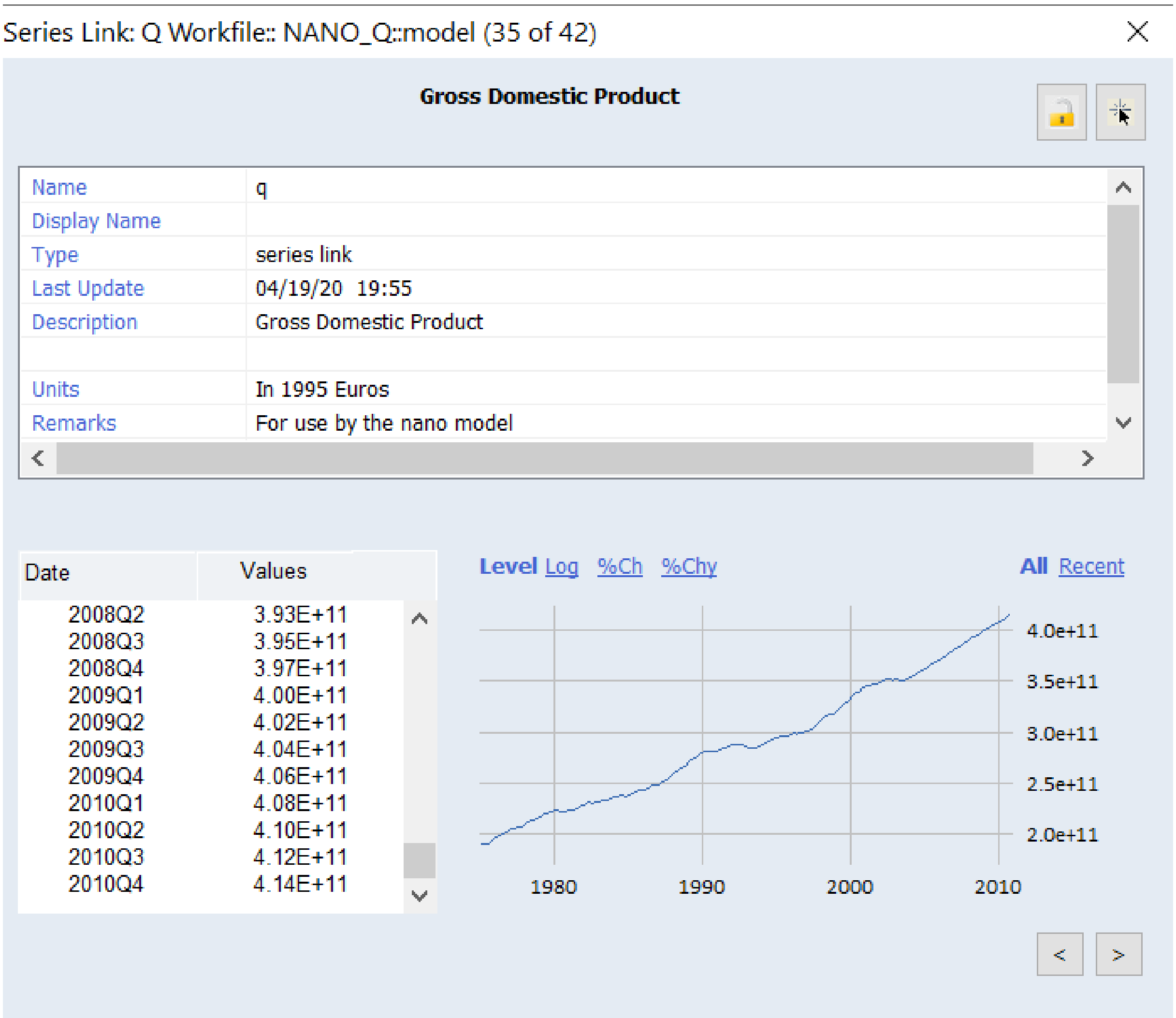
17.8.1.1 Object preview

This feature allows to display the characteristics of any element in a workfile, giving much more information than a spreadsheet or a graph presentation.

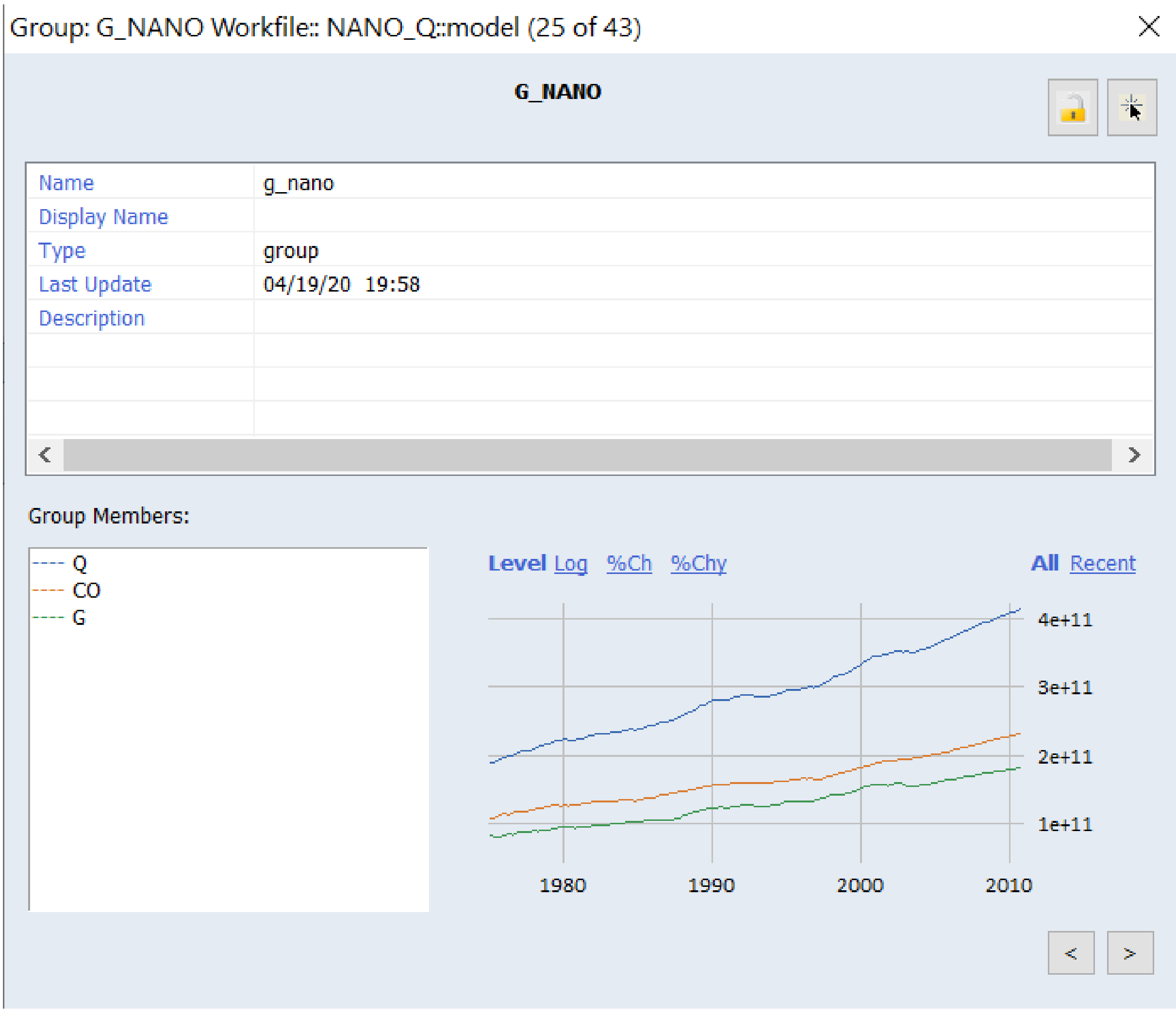
This feature allows to display the characteristics of any element in a workfile, giving much more information than a spreadsheet or a graph presentation.

It is called by selecting an element and right clicking on the second item (below “open”).

For the GDP of our model, we can get:



This feature can be applied to a group.



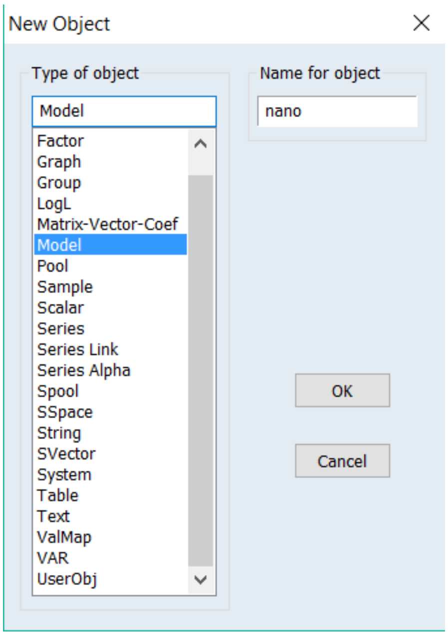
17.9 CREATING THE MODEL

We can now proceed to the model. For this we will start with a blank model.

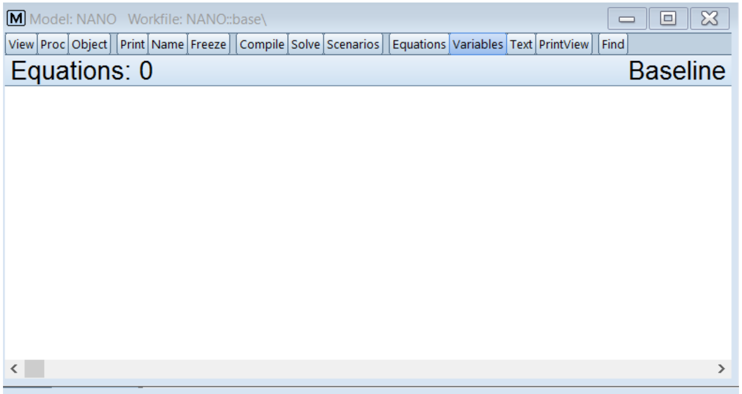
For this, we will use:

Object>New object>Model

And specify "nano" as the name.



The resulting model "nano" has no equations.



Now we will enter our equations.

But the equation for CO is not estimated yet.

Still, we want to show our intention through a tentative formula.

This could be represented by:

CO = f(Q).

This introduces two problems.

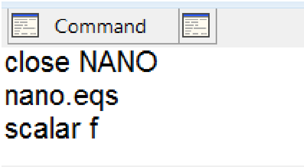
- First, it will not be accepted by EViews as we call for a non-existent function f.

A simple solution is to modify slightly the formula into:

CO = f*(Q)

- But in this case f will be counted as a (missing) exogenous.

To avoid this, we will define f as a scalar in the command window.

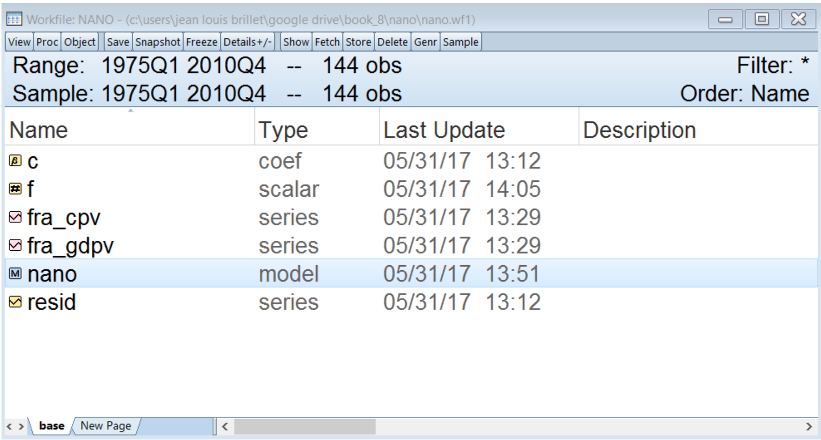


If the Command window is not present, you can call it through

Window>Display Command Capture Window

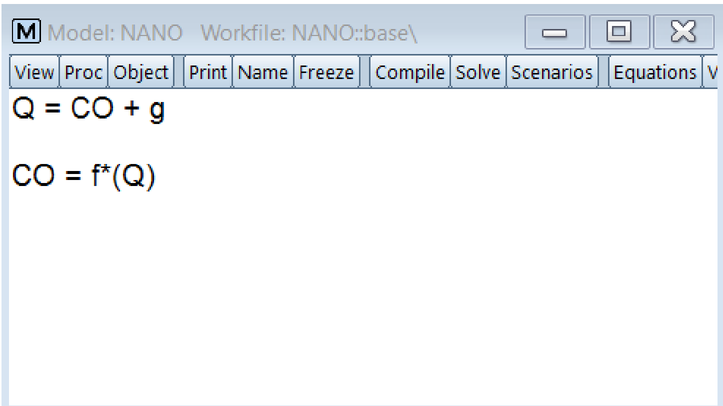
And clicking on the screen image on the left.

Now f and **nano** are present in the workfile.



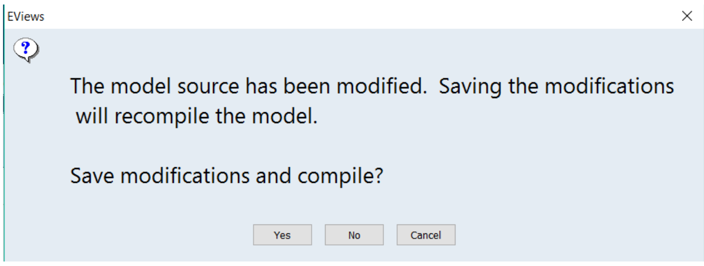
Let us go back to the model by clicking on "nano".

To access the model text, we need to click on "text" and enter the text:



Now we can view the model using "View".

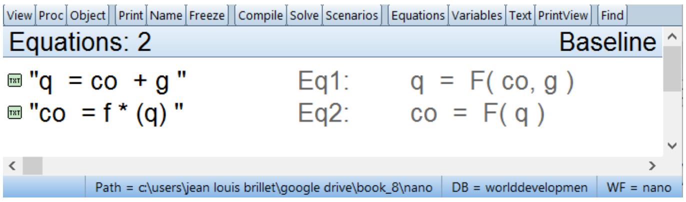
We get the following message:



Of course, we answer "Yes"

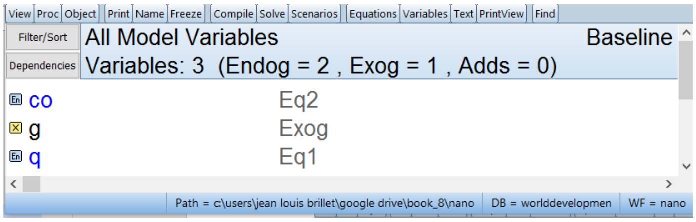
We can use the following views.

Equations.



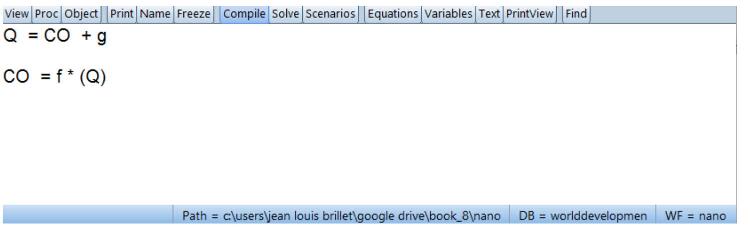
The model has indeed two equations.

Variables.



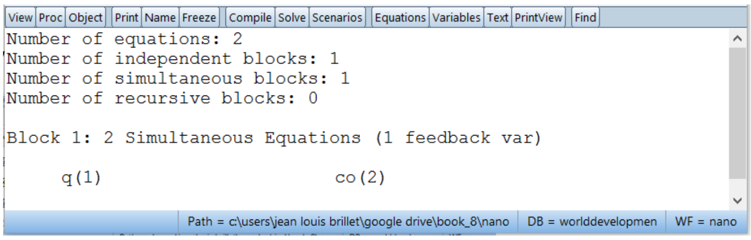
The model has three variables, two endogenous, one exogenous.

Source text



The original text.

Block structure.



The model has a single block, which is simultaneous. This means that with any ordering, the block presents at least one loop, and all variables in the block participate in this loop. The model solution cannot be obtained in one pass.

This subject is developed extensively in the book.

17.9.1 PROGRAMMING

Creating the model will use:

```
if @isobject("nano") then

delete nano

endif

model nano

nano.append Q=CO+g

scalar f=0

nano.append CO=f*(Q)
```

This option allows repetition, edition and inclusion in a wider program.

Also, can check (without testing himself) if the program exists, and delete it in this case[171].

17.10 CHECKING THE MODEL

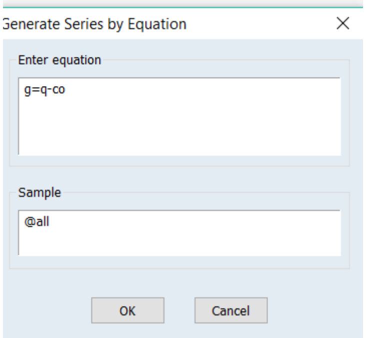
Of course, without a behavioral formula the model cannot be simulated. But we can still check that:

- The data is present
- The syntax is acceptable.
- The identity holds true.

[171] Adding the two equations to the same existing model would create a model with four equations and two endogenous variables, which would be refused by EViews.

Actually the data is not yet fully defined. We still need values for G. For this we can rewrite equation (2) as:

G = Q – CO



The three above conditions are obviously met (a syntax error would produce a message at compiling).

17.11 ESTIMATING CONSUMPTION

17.11.1 A VERY SIMPLE FORMULA

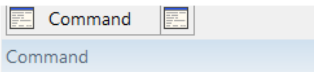
We can start by supposing that consumption shows a constant elasticity to GDP. The growth of household consumption can follow that of household revenue, itself following the growth of GDP. We will try this, using the command window.

Of course, we could also use:

Quick> Estimate Equation

Our argument here is that the Command window is easier to edit, in case of successive estimations of the same concept. Its justification was much stronger until EViews 7, as it was the only way to have access to the full (and reproductible) sequence of the session statements.

Let us try, limiting the period to the historical values (admittedly our data set is rather old).

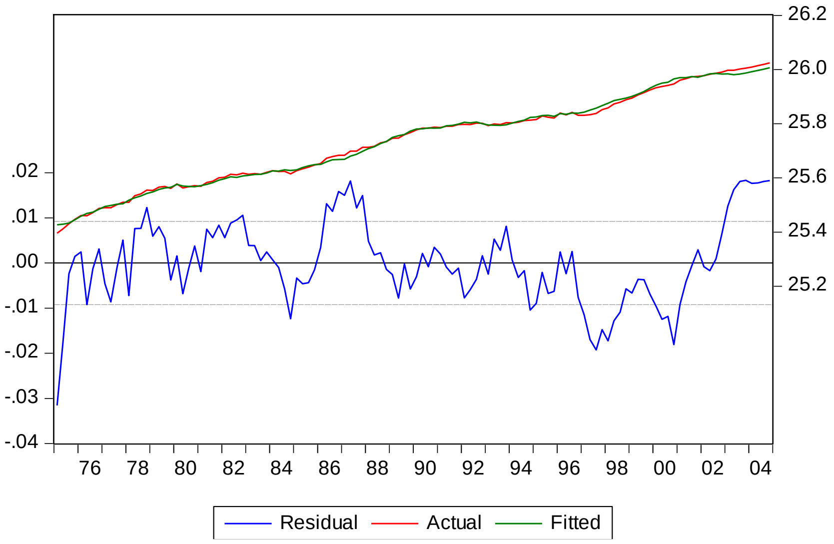



```
smpl 1975Q1 2004Q4
ls log(CO) log(Q) c
```

Dependent Variable: LOG(CO)
Method: Least Squares
Date: 05/31/17 Time: 15:28
Sample: 1975Q1 2004Q4
Included observations: 120

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(Q)	0.907456	0.004656	194.8849	0.0000
C	1.860303	0.122528	15.18271	0.0000

R-squared	0.996903	Mean dependent var	25.73854
Adjusted R-squared	0.996876	S.D. dependent var	0.165065
S.E. of regression	0.009225	Akaike info criterion	-6.517230
Sum squared resid	0.010042	Schwarz criterion	-6.470772
Log likelihood	393.0338	Hannan-Quinn criter.	-6.498363
F-statistic	37980.13	Durbin-Watson stat	0.345898
Prob(F-statistic)	0.000000		



The obvious problem comes from the autocorrelation of residuals. It can be treated using an autoregressive process.

smpl 1975q1 2004Q4

ls log(co) log(q) c ar(1)

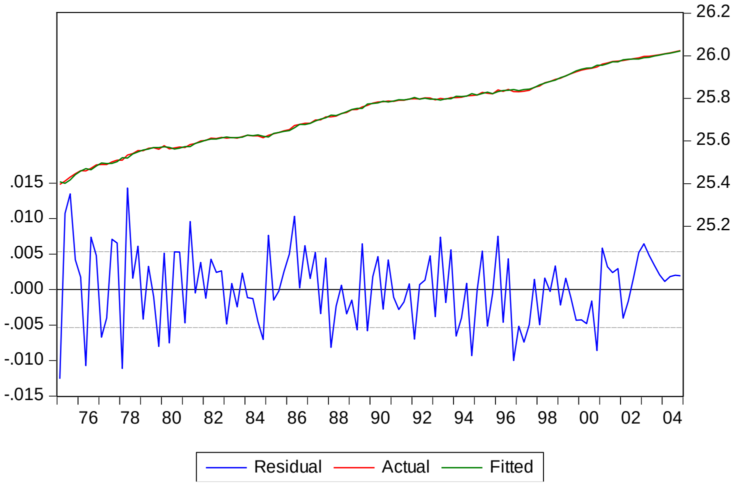
Please remember that our goal is not to present sophisticated methods but rather a sequence of simple steps.

Dependent Variable: LOG(CO)
Method: ARMA Maximum Likelihood (BFGS)
Date: 05/31/17 Time: 15:34
Sample: 1975Q1 2004Q4
Included observations: 120
Convergence achieved after 8 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(Q)	0.925049	0.013912	66.49284	0.0000
C	1.396773	0.364296	3.834172	0.0002
AR(1)	0.863752	0.046381	18.62283	0.0000
SIGMASQ	2.76E-05	3.84E-06	7.175565	0.0000

R-squared	0.998979	Mean dependent var	25.73854
Adjusted R-squared	0.998953	S.D. dependent var	0.165065
S.E. of regression	0.005342	Akaike info criterion	-7.582152
Sum squared resid	0.003311	Schwarz criterion	-7.489236
Log likelihood	458.9291	Hannan-Quinn criter.	-7.544418
F-statistic	37830.47	Durbin-Watson stat	2.226451
Prob(F-statistic)	0.000000		

Inverted AR Roots	.86



The results improve a lot, both in terms of correlation and global quality.

However, a problem remains: if we use this formula in the long run, consumption will grow steadily slower than GDP, leading to a 100% savings rate…. And introducing an additional trend does not change things.

17.11.2 IMPROVING THE FORMULA

Without resorting to sophisticated elements, we can try to improve the formula both in theory and model properties, which request consumption to grow at the same speed as revenue (and GDP) in the long-term.

At the same time, theory implies that in the first periods the short-term elasticity is lower than one, as households take time in adapting to fluctuations of their revenue.

Obviously the only option is to consider two elasticities: one unitary (the long run), one supposedly lower (the short run).

Clearly this calls for an error correction model, the simplest form being:

Dlog(CO) = a . Dlog(Q) + b Log(CO(-1)/Q(-1)) + c

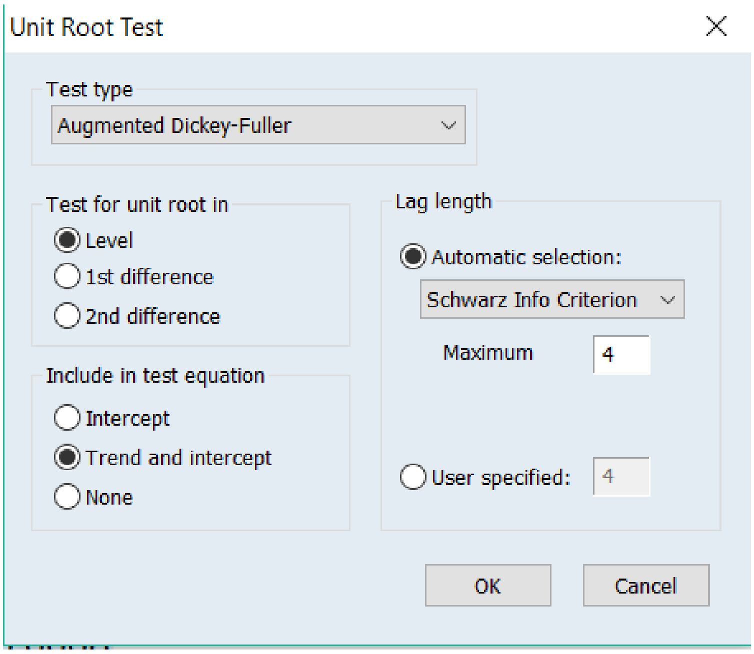
But to be allowed to follow this process, the ratio between CO and Q must be stationary.

This is checked simply by

Displaying the log of the ratio in the Command window:

```
show log(co/q)
```

and use “Unit root test” calling for a Dickey Fuller test allowing for a trend (we will limit the number of lags to 4).



		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.872243	0.0162
Test critical values:	1% level	-4.036983	
	5% level	-3.448021	
	10% level	-3.149135	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(CO/Q))
Method: Least Squares
Date: 05/31/17 Time: 17:03
Sample: 1975Q2 2004Q4
Included observations: 119

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(CO(-1)/Q(-1))	-0.203473	0.052547	-3.872243	0.0002
C	-0.110916	0.028697	-3.865001	0.0002
@TREND("1975Q2")	-0.000104	2.93E-05	-3.538965	0.0006

R-squared	0.114974	Mean dependent var	-7.78E-05
Adjusted R-squared	0.099715	S.D. dependent var	0.005485
S.E. of regression	0.005204	Akaike info criterion	-7.653892
Sum squared resid	0.003141	Schwarz criterion	-7.583830
Log likelihood	458.4066	Hannan-Quinn criter.	-7.625442
F-statistic	7.534822	Durbin-Watson stat	2.032378
Prob(F-statistic)	0.000838		

The test allows to non-stationarity a very low probability (1.6%).

We can now estimate the error correction equation.

First, to make the trend clearer (and the handling of time elements simpler) we will define a variable which takes the value of the year in the first quarter, and grows by 0.25 for each following period.

Command

```
smpl @all
genr t=1975
smpl 1975q2 2010Q4
genr t=t(-1)+0.25
```

Then we estimate the equation[172]

Equation: UNTITLED   Workfile: NANO::model\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: DLOG(CO)
Method: Least Squares
Date: 05/31/17   Time: 17:30
Sample: 1975Q2 2004Q4
Included observations: 119

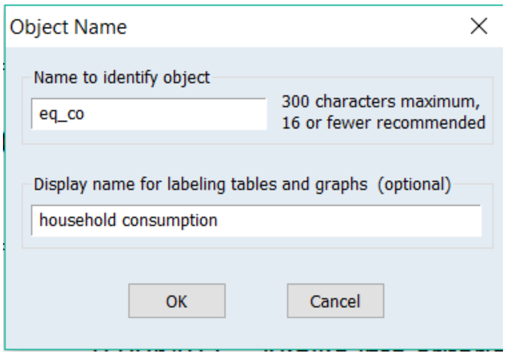
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(Q)	0.774414	0.094842	8.165351	0.0000
LOG(CO(-1)/Q(-1))	-0.201474	0.051529	-3.909895	0.0002
C	0.738497	0.203299	3.632569	0.0004
T	-0.000429	0.000115	-3.729211	0.0003

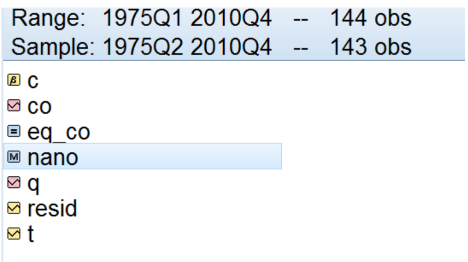
R-squared	0.422298	Mean dependent var	0.005288
Adjusted R-squared	0.407227	S.D. dependent var	0.006627
S.E. of regression	0.005103	Akaike info criterion	-7.685109
Sum squared resid	0.002994	Schwarz criterion	-7.591693
Log likelihood	461.2640	Hannan-Quinn criter.	-7.647176
F-statistic	28.02152	Durbin-Watson stat	2.176474
Prob(F-statistic)	0.000000		

The short-term elasticity is 0.77, and the gap is closed at each subsequent period by 20%.

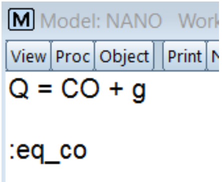
[172] Note : We could also have computed the residual from the test, and use it as a lagged element in the equation.

We can now name the equation (as eq_co) and specify a label



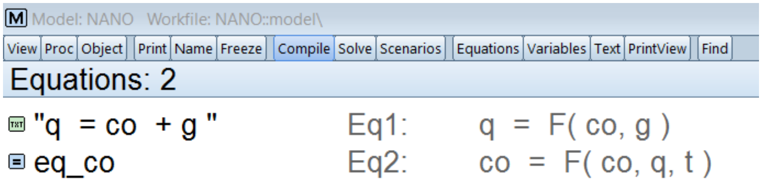


To introduce the equation, we must edit the model text, replacing the equation by the name of the equation, preceded by “:” as:

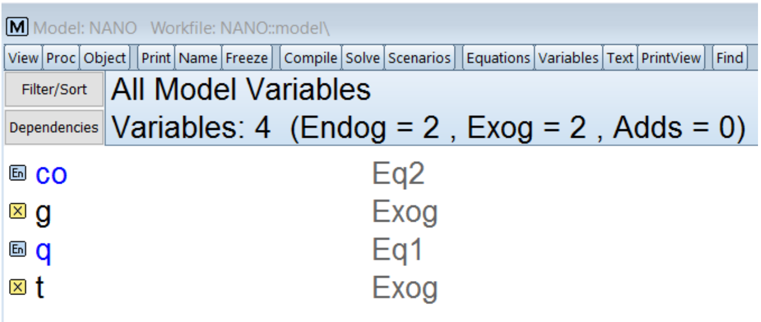


The characteristics do not really change.

- co appears (as lagged) on the right hand side of its equation.



- t is a new exogenous



17.11.2.1 Programming

The above techniques look quite fine as long as one is evaluating potential formulas. But when a choice has been made, it introduces several problems.

- The formula will be used a large number of times. Obviously, it has to be stored somewhere.
- But updating the initial version must be possible.
  - The data changes through data set extension, new methodology or new field.
  - New ideas can show up.
  - The equation is used for a new case (a different country.).
- If several equations are estimated, the coefficient vector must be characteristic of the equation, as well as its name and its residual.
  In addition:
- Knowing the name of the dependent variable, one must have access to the equation elements.
- Looking at the elements, one must be able to associate them to an equation.

This is why we propose the following framework, conditional on the "programming" option.

- Name the equation, the coefficient vector and the residual according to the dependent variable.
- Introduce these elements on the equation.

We will use for the CO equation:

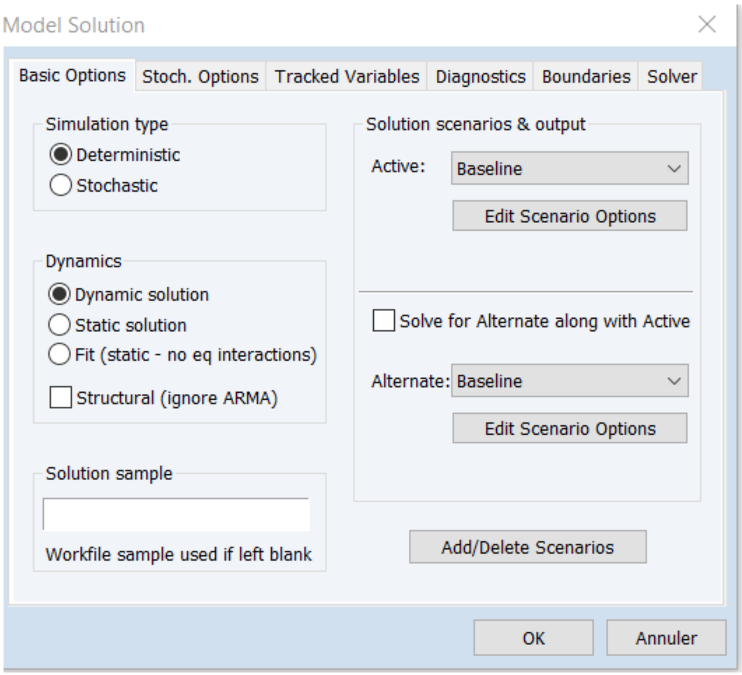
```
genr ec_co=0
coef(10) c_co
equation eq_co.ls(p) dlog(co)=c_co(1)*dlog(q)+c_co(2)*(log(co(-1)/q(-1))
    -c_co(3)*t-c    _co(4))+ec_co
genr ec_co=resid
nano.drop co
nano.merge eq_co
```

One will note that before the estimated equation is merged into the model, one has to delete it in its text form. This is done simply by dropping the dependent variable from the model.

However, to make the process clearer and limit mistakes, we suggest creating the model again from scratch, by deleting it and specifying again the whole set of equations. In programming mode, the cost is negligible.

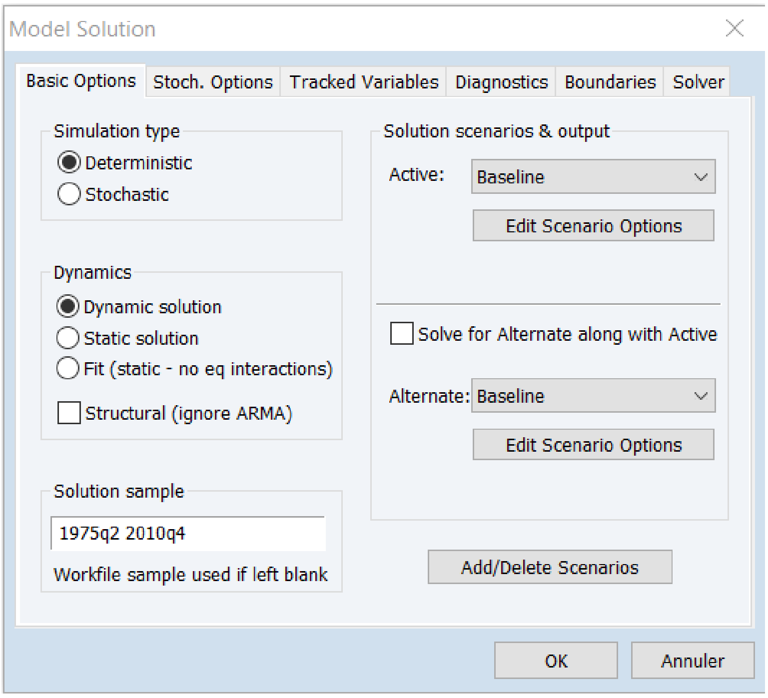
17.12 SOLVING THE MODEL

To solve the model, you need to open it (by double clicking) and click on the “solve” item:

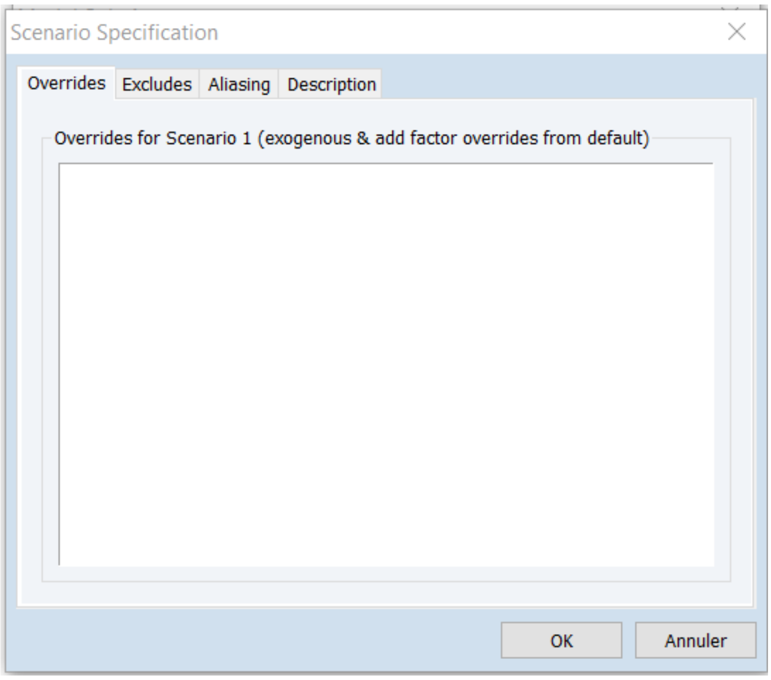


Let us solve the model over the largest period possible (restricted to 1975Q2 to 2010Q4 as there is one lag).

We shall perform a solution both deterministic (no random residuals) and dynamic (solving one period uses the simulated values for the lagged elements).

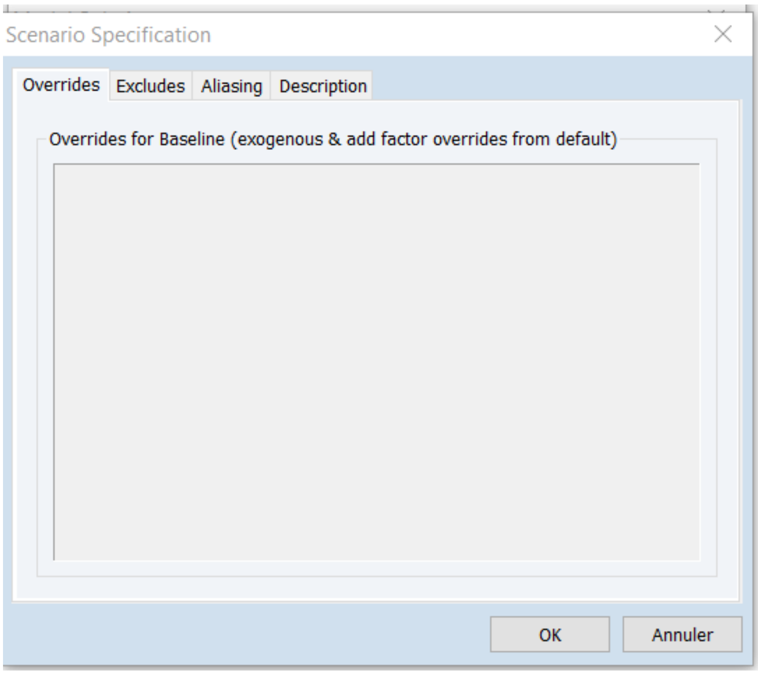


By clicking on “Edit Scenario Options” we get:

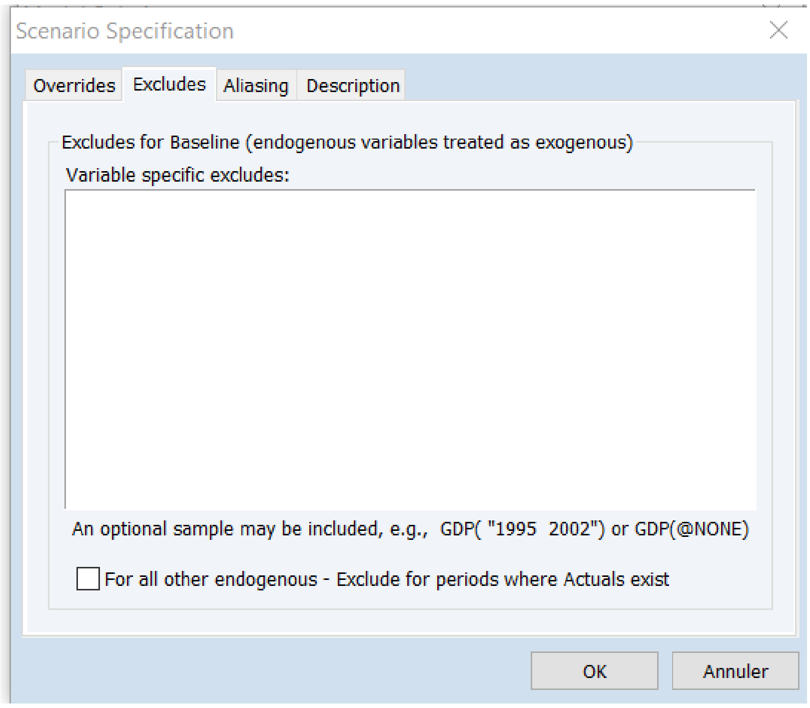


We can now access the following menus:

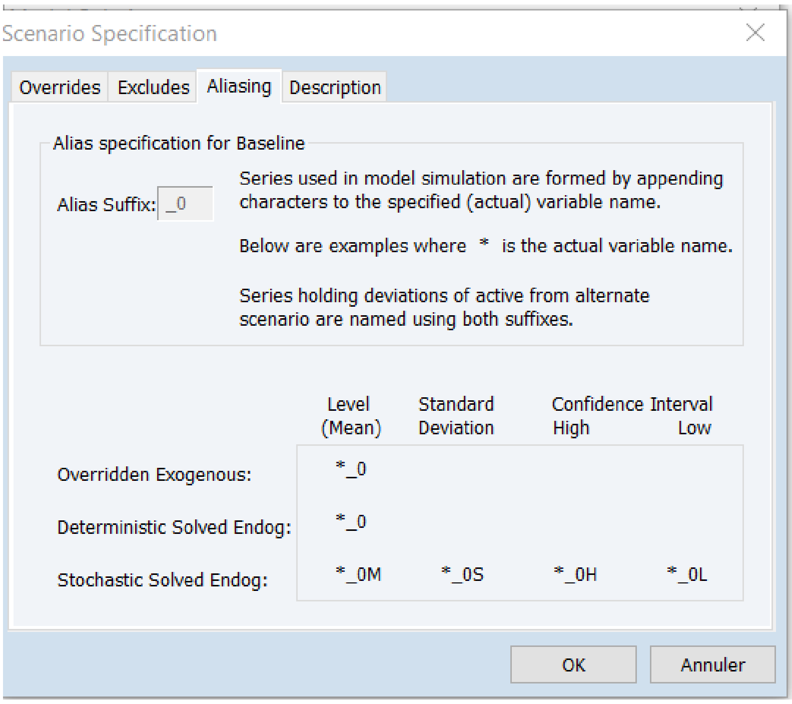
Overrides : we override no variable. This would allow to replace the values of a variable by alternative ones, stored in a series with the same name but an additional suffix (more on this later).



Excludes : we exclude no variable. This would exclude the endogenous from the solving process, and use its workfile value (which must be present).

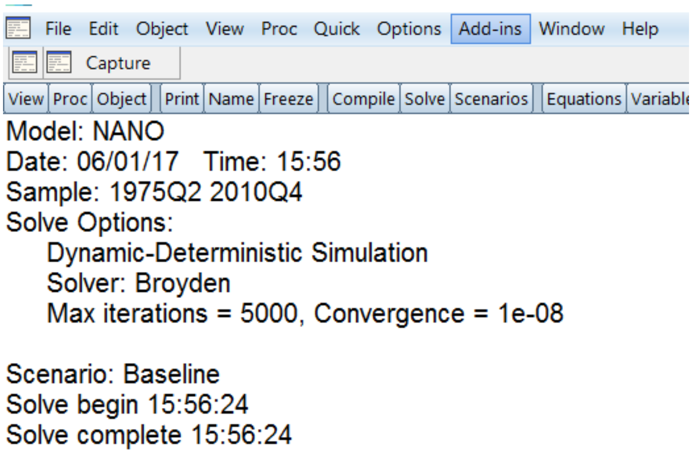


Aliasing : This sets the names of the objects created. For now, the important element is the suffix of the “Deterministic Solved Endogenous” : _0.

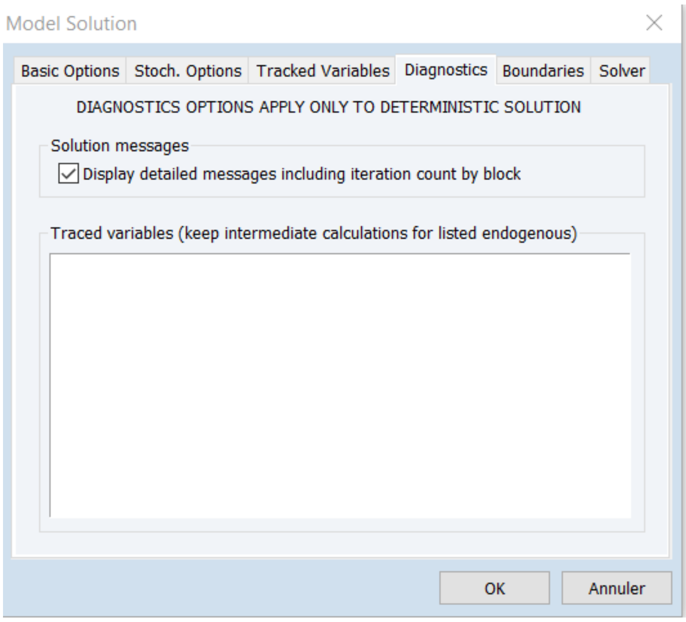


A few explanations:

- The method for solving is “Broyden”.
- It works (no error messages).
- It takes less than 5000 iterations to obtain convergence.
- At convergence, no residual is higher than 10^{-6} (in relative absiolute velue).



The number of iterations can be obtained by using « Diagnosis » and clicking on “Display……”



We get:

```
Model: NANO
Date: 06/01/17   Time: 16:16
Sample: 1975Q2 2010Q4
Solve Options:
    Dynamic-Deterministic Simulation
    Solver: Broyden
    Max iterations = 5000, Convergence = 1e-08

Scenario: Baseline
Solve begin 16:16:42
    1975Q2   Convergence after 4 iterations
    1975Q3   Convergence after 4 iterations
    1975Q4   Convergence after 4 iterations
    1976Q1   Convergence after 4 iterations
    1976Q2   Convergence after 4 iterations
    1976Q3   Convergence after 4 iterations
    1976Q4   Convergence after 4 iterations
..................................................................
    2009Q1   Convergence after 5 iterations
    2009Q2   Convergence after 5 iterations
    2009Q3   Convergence after 5 iterations
    2009Q4   Convergence after 5 iterations
    2010Q1   Convergence after 5 iterations
    2010Q2   Convergence after 5 iterations
    2010Q3   Convergence after 5 iterations
    2010Q4   Convergence after 5 iterations
Solve complete 16:16:43
```

The method needs 4 to 5 iterations to converge.

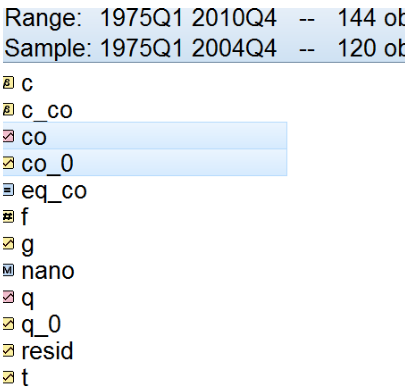
Looking at the workfile; we see that two new variables have been created: co_0 and q_0.

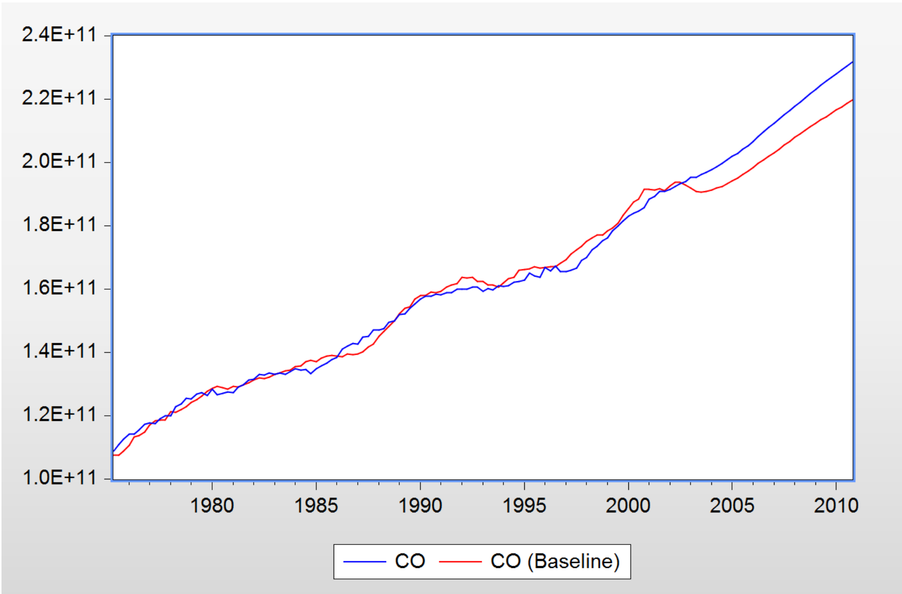
They contain the results of the simulation (remember that the suffix associated with "Baseline" is _0).

```
Range:  1975Q1 2010Q4  --  144 obs
Sample: 1975Q1 2004Q4  --  120 obs
c
c_co
co
co_0
eq_co
f
g
nano
q
q_0
resid
t
```

We can produce a graph for CO (or Q) with actual and simulated values.

We open both series as a group and call for "Graph"





The results are rather good until 2004 (the estimation period), but a gap appears afterwards.

17.12.1 PROGRAMMING

With programming, the sequence is simply ;

```
smpl 1975q2 2010q4
solve(d=d) nano
```

17.13 A SHOCK

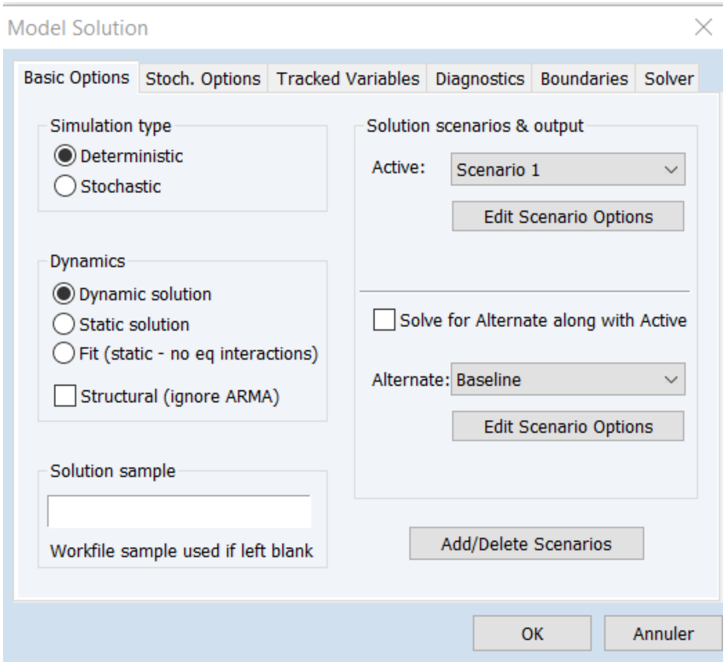
Now let us suppose we want to see what happens if an assumption changes, for instance if we increase Government demand by 1% of GDP, starting in 1981.

We need to change:

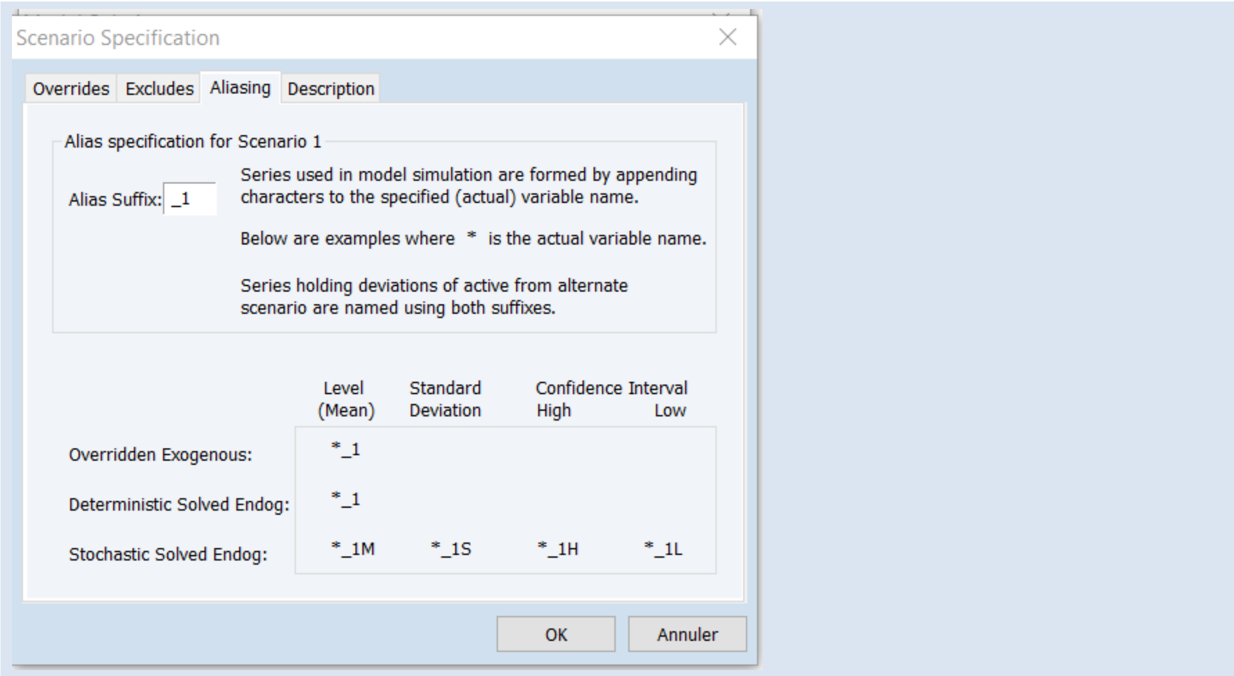
- The Scenario: Baseline does not allow to override variables. We must change it.
- The suffix associated with overridden elements. It will also be used for solution results.

We will need of course also to create the variable.

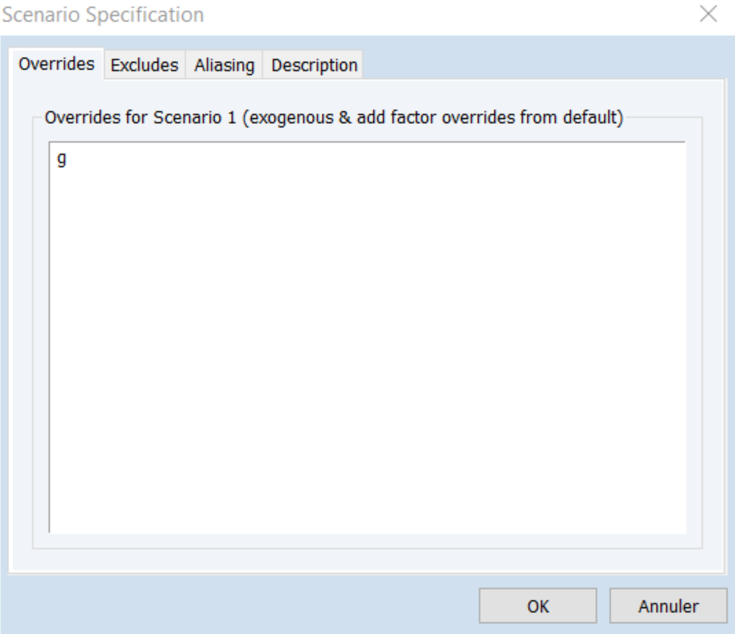
In the window “Active” we chose “Scenario 1”



In “Edit Scenario Options” we chose “Aliasing”. We can see that the suffix associated with “Scenario 1” is “_1”.

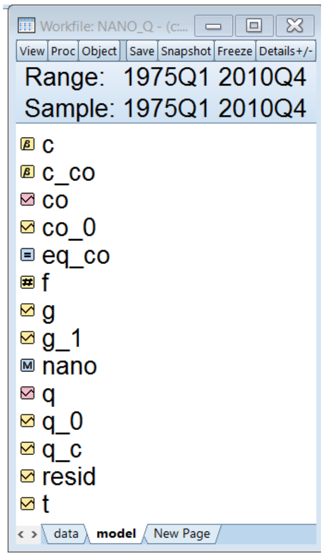


In “Edit Scenario Options” we chose “Override” and, in the window, we specify “g”:



In the “Command” window we create a variable called g_1 increased from g by 1% of GDP starting in the first quarter of 1981. This is easier controlled by the availability of “t”.

```
smpl @all
genr g_1=g+0.01*q*(t>=1981)
```

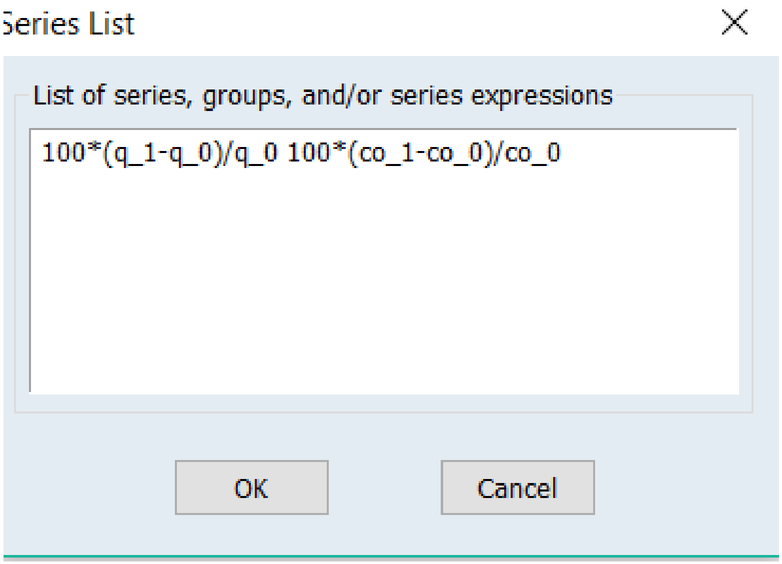


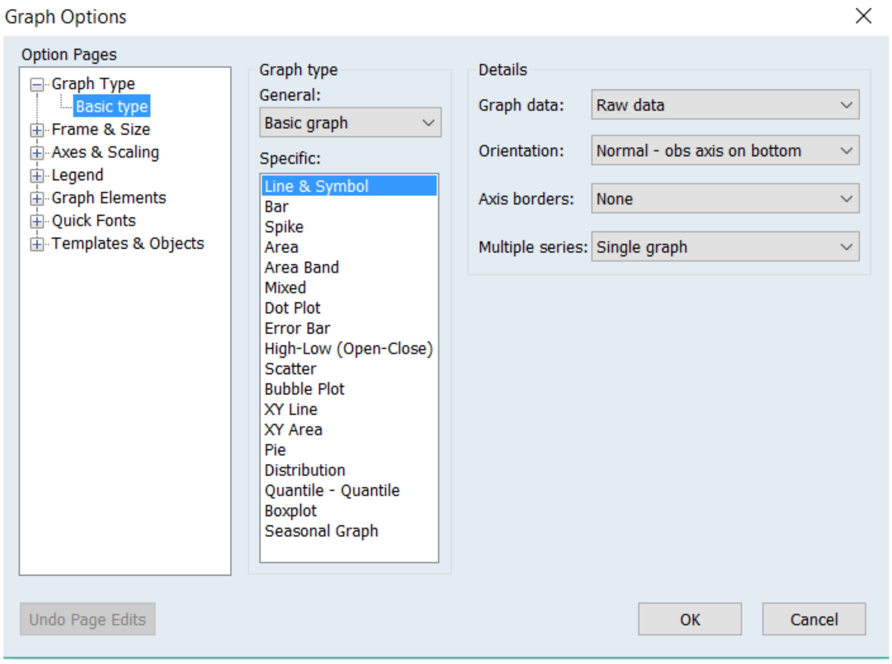
g_1 appears now in the workfile.

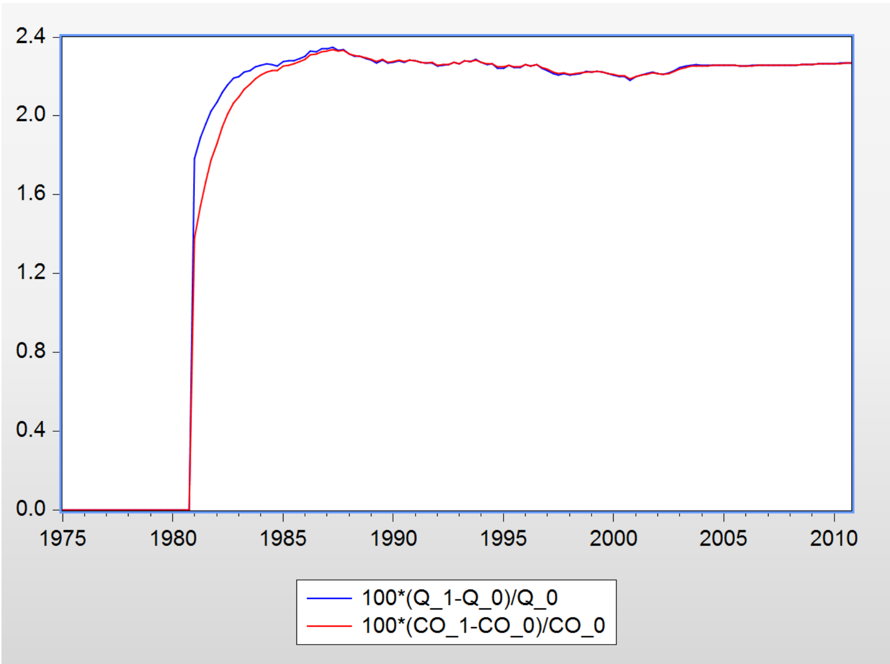
We solve the model again.

```
Model: NANO
Date: 06/01/17   Time: 18:40
Sample: 1975Q2 2010Q4
Solve Options:
    Dynamic-Deterministic Simulation
    Solver: Broyden
    Max iterations = 5000, Convergence = 1e-08
Scenario: Scenario 1
Solve begin 18:40:32
Solve complete 18:40:32
```

Now we can produce a graph for the difference.







We will leave the interpretation for later, when we solve the model over the future, its natural field.

17.13.1 PROGRAMMING

Programming a shock is easier. We shall use:

```
smpl 1975Q2 2010Q4
nano.append assign @all _b
solve(d=d) nano
genr g_v=g+q*.01*(t>=1980)
```

```
nano.scenario "scenario 1"
nano.append assign @all _v
nano.override g
solve(d=d) nano
```

To produce a graph, we will create explicit variables.

For this, we will use a loop. The statements between “for” and “next” will be executed for each value after the “for”, replacing in turn the “%1” by the value, and dropping the brackets.

```
for %1 CO Q
series dv_{%1}={%1}_v-{%1}_b
series pv_{%1}=100*dv_{%1}/({%1}_b+({%1}_b=0))
next
graph pv_co pv_q
```

17.14 FORECASTING

We shall now use the model over the period for which it has been created: the future

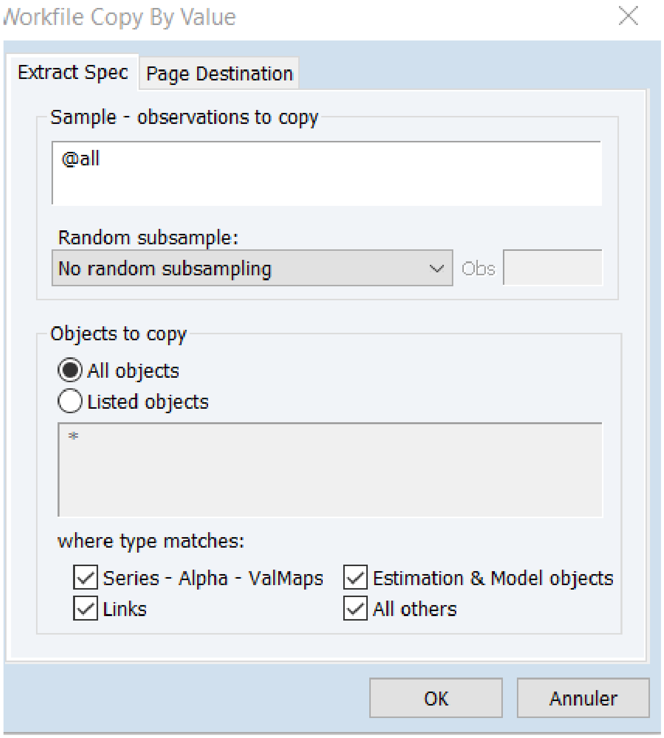
17.14.1 CREATING A PAGE

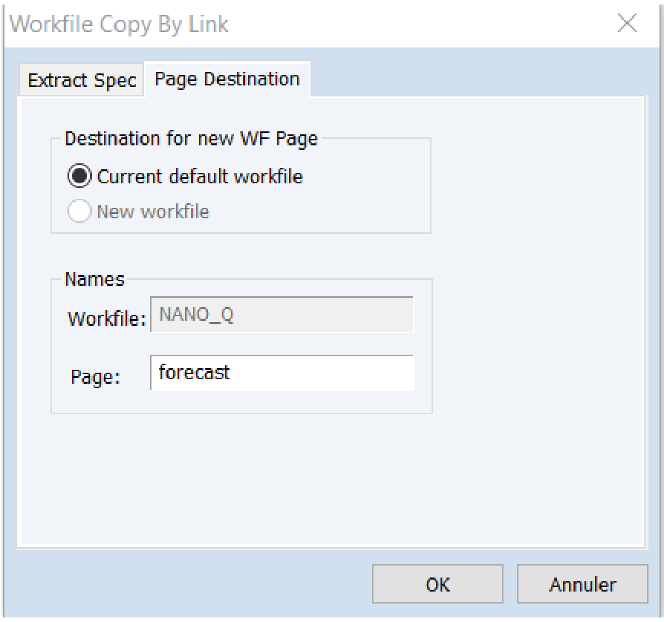
We need a new page with the future periods. We will call it “forecast” and extend the period to 2100.

We will copy the « model » page first , as « forecast ».

We shall use :

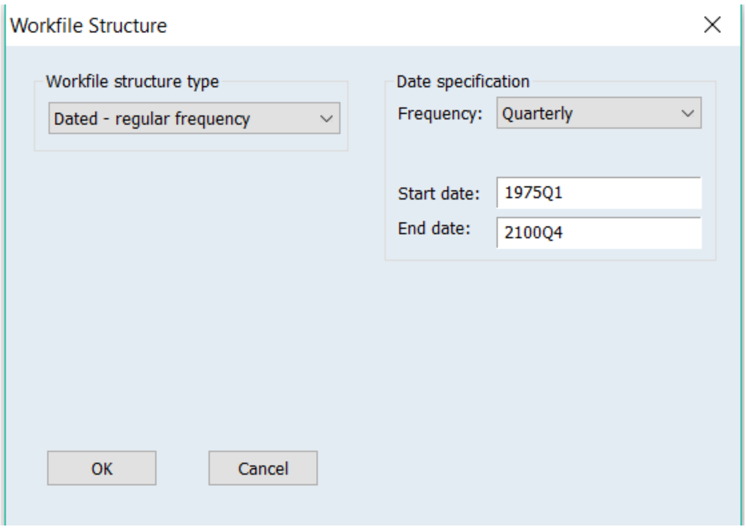
```
Proc>Copy/Extract to current page>By value to New Page or Workfile
```





Then we change the periods in the page using:

```
Proc>Structure/Resize Current Page
```



17.14.1.1 Programming

We shall use

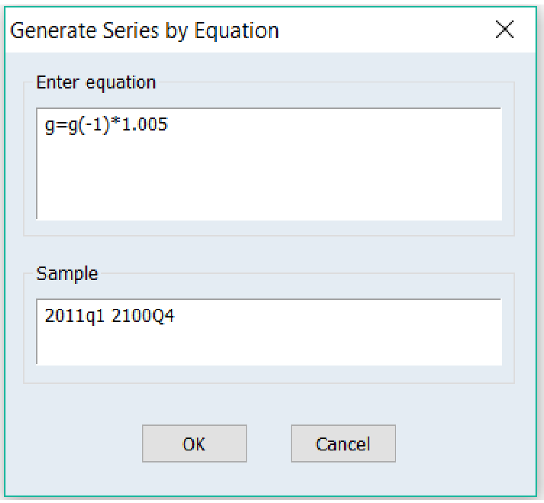
```
pagecopy(page=forecast)

pagestruct(end=2100Q4) *
```

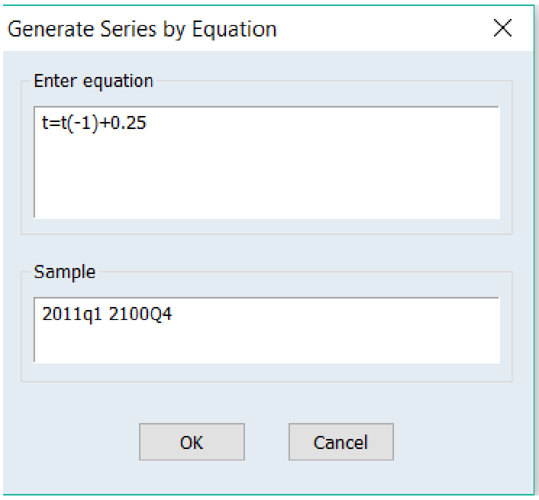
17.14.2 THE ASSUMPTIONS

The model will compute the endogenous on the future. But for this it needs values for the exogenous variables.

Our model contains two exogenous : Government Demand g and time t.

We will suppose that g grows by 0.5% at each quarter (roughly 2% per year) after 2010Q4.





17.14.2.1 Programming

In a program, we will simply use:

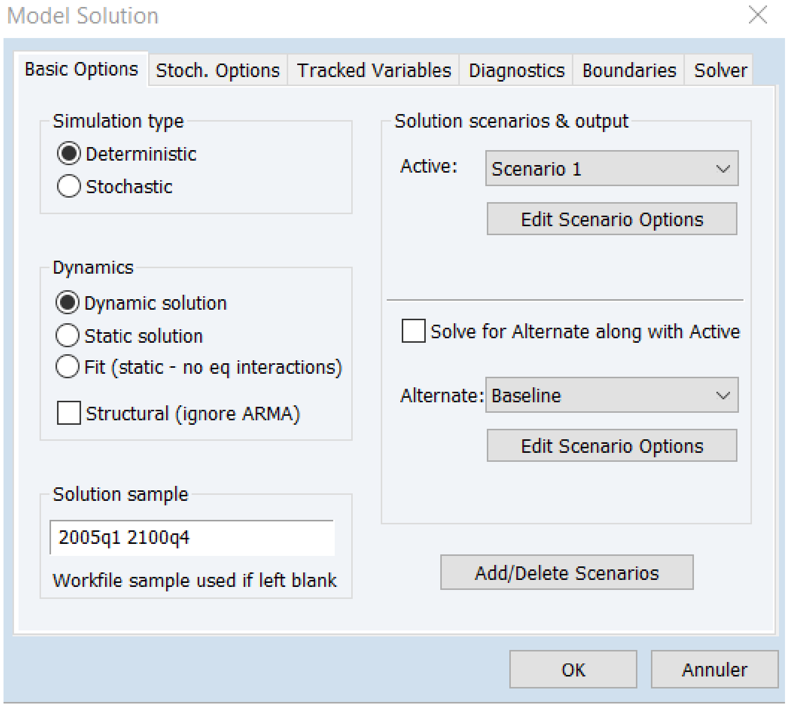
```
smpl 2010Q1 2100Q4

genr t=t(-1)+0.25

genr g=g(-1)*1.005
```

Note that we must also forecast the residual for the consumption equation. As it applies to a Log, it works as a multiplicative factor. We will keep it constant over time.

```
genr ec_co=ec_co(-1)
```

17.14.3 BASIC SIMULATION

Simulating the model follows the same statements as before.



The model solves without problem.

```
Model: NANO
Date: 06/02/17   Time: 13:07
Sample: 2005Q1 2100Q4
Solve Options:
    Dynamic-Deterministic Simulation
    Solver: Broyden
    Max iterations = 5000, Convergence = 1e-08
Scenario: Scenario 1
Solve begin 13:07:22
    2005Q1    Convergence after 4 iterations
    2005Q2    Convergence after 4 iterations
    2005Q3    Convergence after 4 iterations
    2005Q4    Convergence after 4 iterations
    2006Q1    Convergence after 4 iterations
    2006Q2    Convergence after 5 iterations
    2006Q3    Convergence after 5 iterations
    2006Q4    Convergence after 5 iterations
--------------------------------------------------------------
    2099Q1    Convergence after 4 iterations
    2099Q2    Convergence after 4 iterations
    2099Q3    Convergence after 4 iterations
```

```
     2099Q4     Convergence after 4 iterations
     2100Q1     Convergence after 4 iterations
     2100Q2     Convergence after 4 iterations
     2100Q3     Convergence after 4 iterations
     2100Q4     Convergence after 4 iterations
Solve complete 13:07:24
```

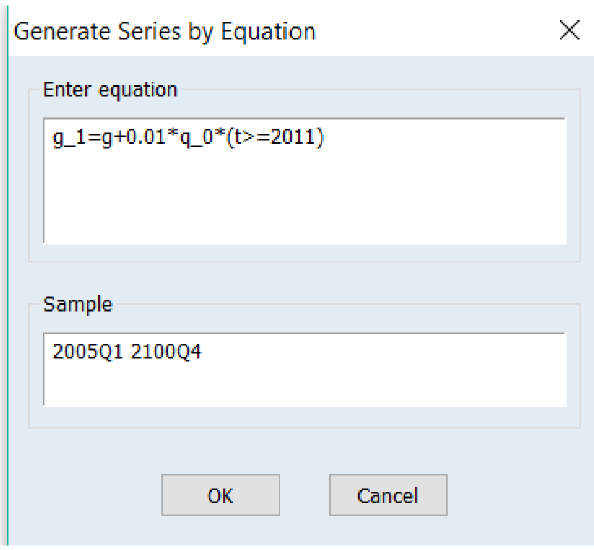
17.14.3.1 Programming

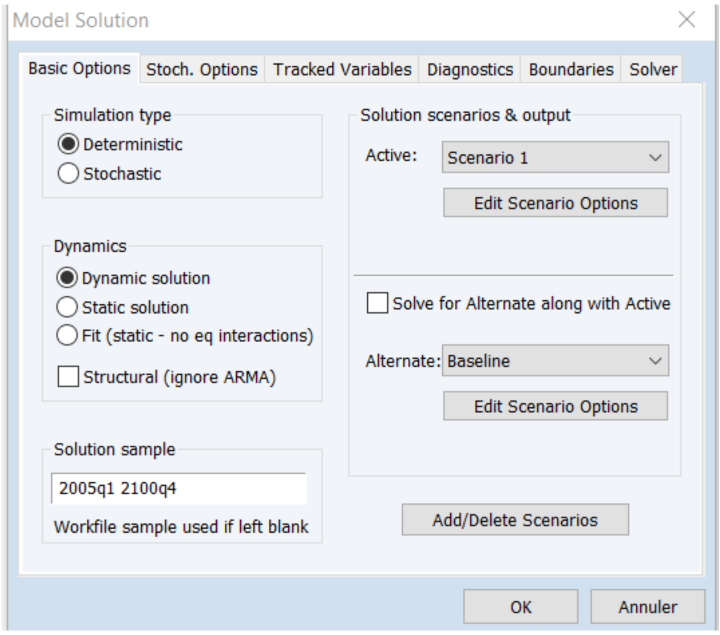
```
smpl 2005Q1 2100Q4
nano.scenario "Baseline"
nano.override
solve(d=d) nano
```

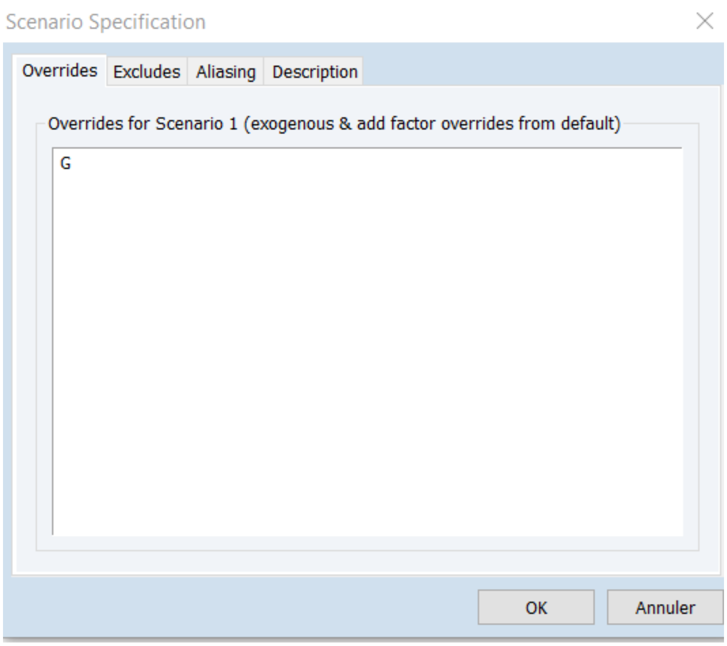
17.14.4 A SHOCK

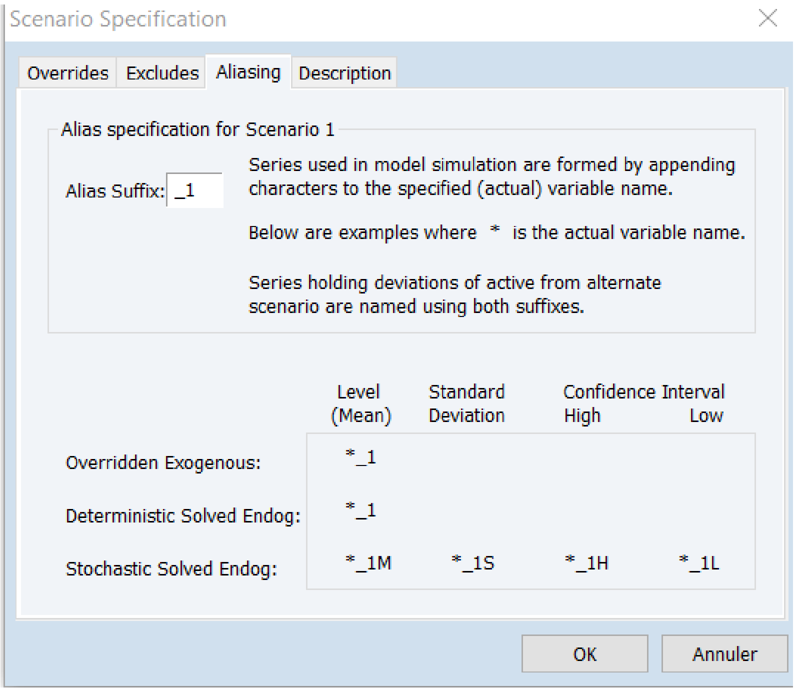
We will make the same shock as before.

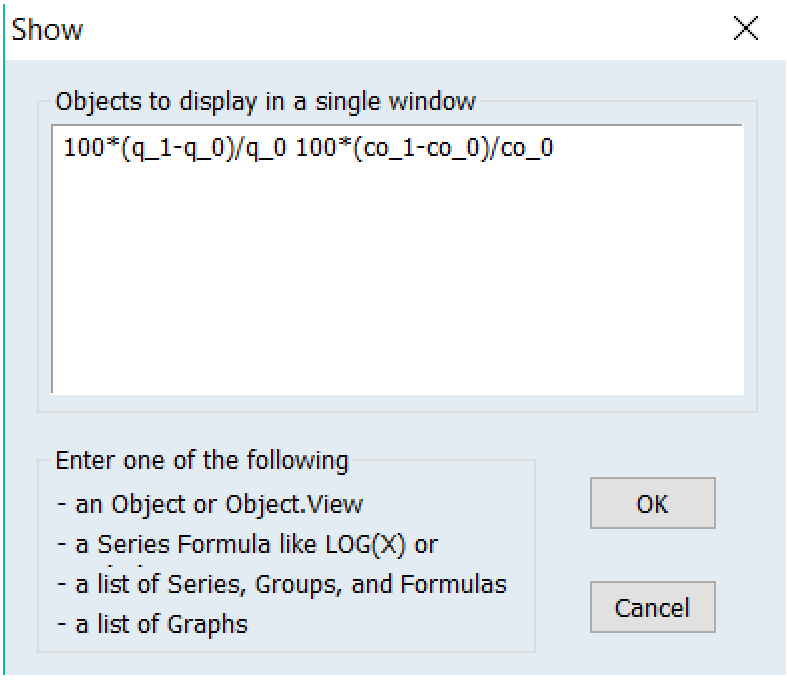
We shock again Government demand by 1% of the baseline GDP, starting in 2011 this time.

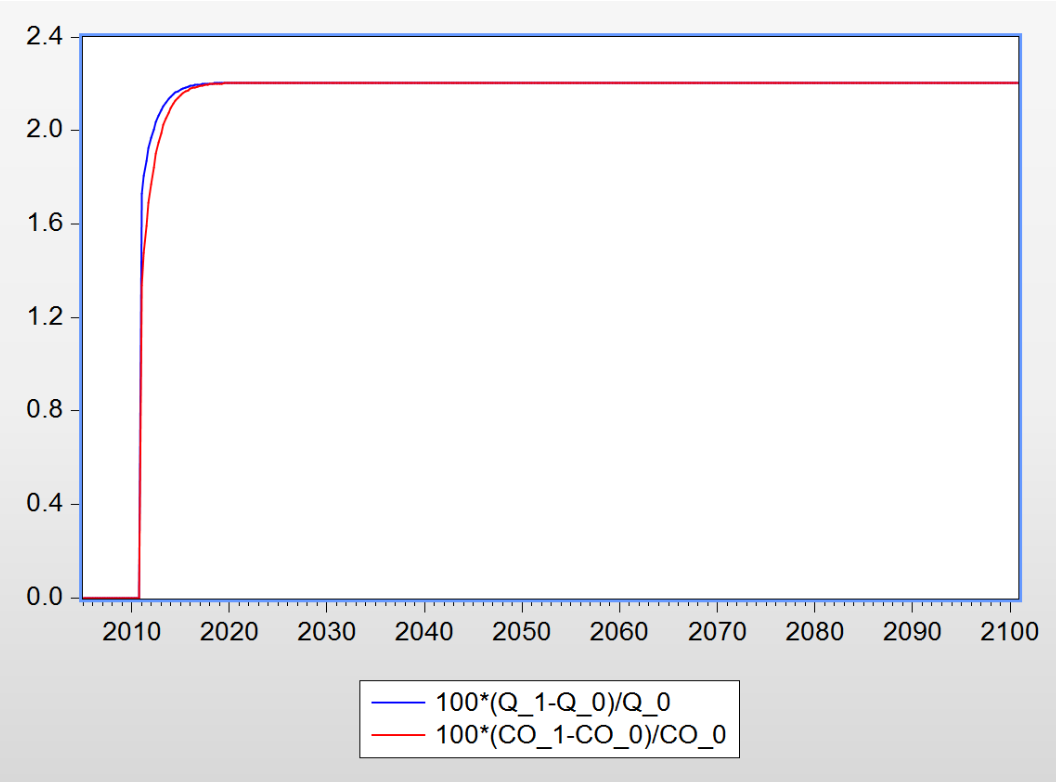



Model Solution
Basic Options
Stoch. Options
Tracked Variables
Diagnostics
Boundaries
Solver
Simulation type
Deterministic
Stochastic
Dynamics
Dynamic solution
Static solution
Fit (static - no eq interactions)
Structural (ignore ARMA)
Solution sample
2005q1 2100q4
Workfile sample used if left blank
Solution scenarios & output
Active:
Scenario 1
Edit Scenario Options
Solve for Alternate along with Active
Alternate:
Baseline
Edit Scenario Options
Add/Delete Scenarios
OK
Annuler

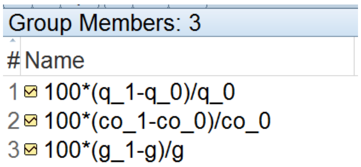

Scenario Specification
Overrides
Excludes
Aliasing
Description
Overrides for Scenario 1 (exogenous & add factor overrides from default)
G
OK
Annuler

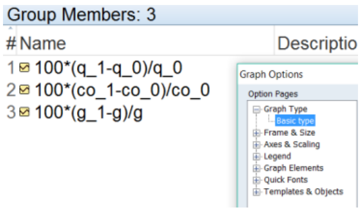

Scenario Specification
Overrides
Excludes
Aliasing
Description
Alias specification for Scenario 1
Alias Suffix: _1
Series used in model simulation are formed by appending characters to the specified (actual) variable name.
Below are examples where * is the actual variable name.
Series holding deviations of active from alternate scenario are named using both suffixes.
Level (Mean)
Standard Deviation
Confidence Interval High Low
Overridden Exogenous: *_1
Deterministic Solved Endog: *_1
Stochastic Solved Endog: *_1M *_1S *_1H *_1L
OK
Annuler


Show
Objects to display in a single window
100*(q_1-q_0)/q_0 100*(co_1-co_0)/co_0
Enter one of the following
- an Object or Object.View
- a Series Formula like LOG(X) or
- a list of Series, Groups, and Formulas
- a list of Graphs
OK
Cancel

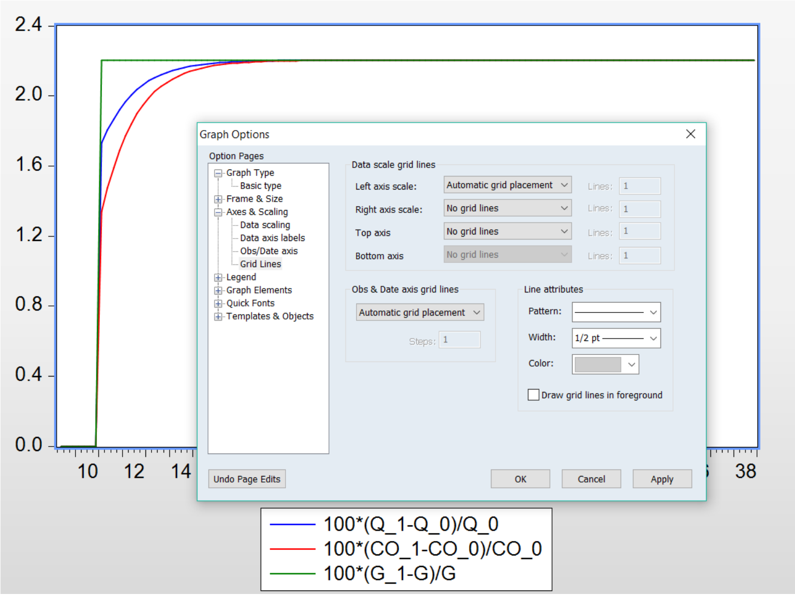


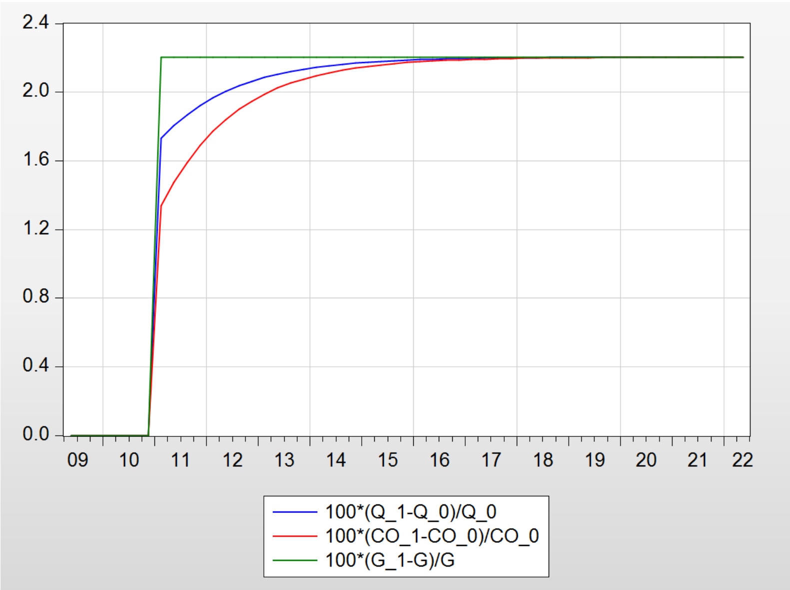
We can add the relative variation of g.

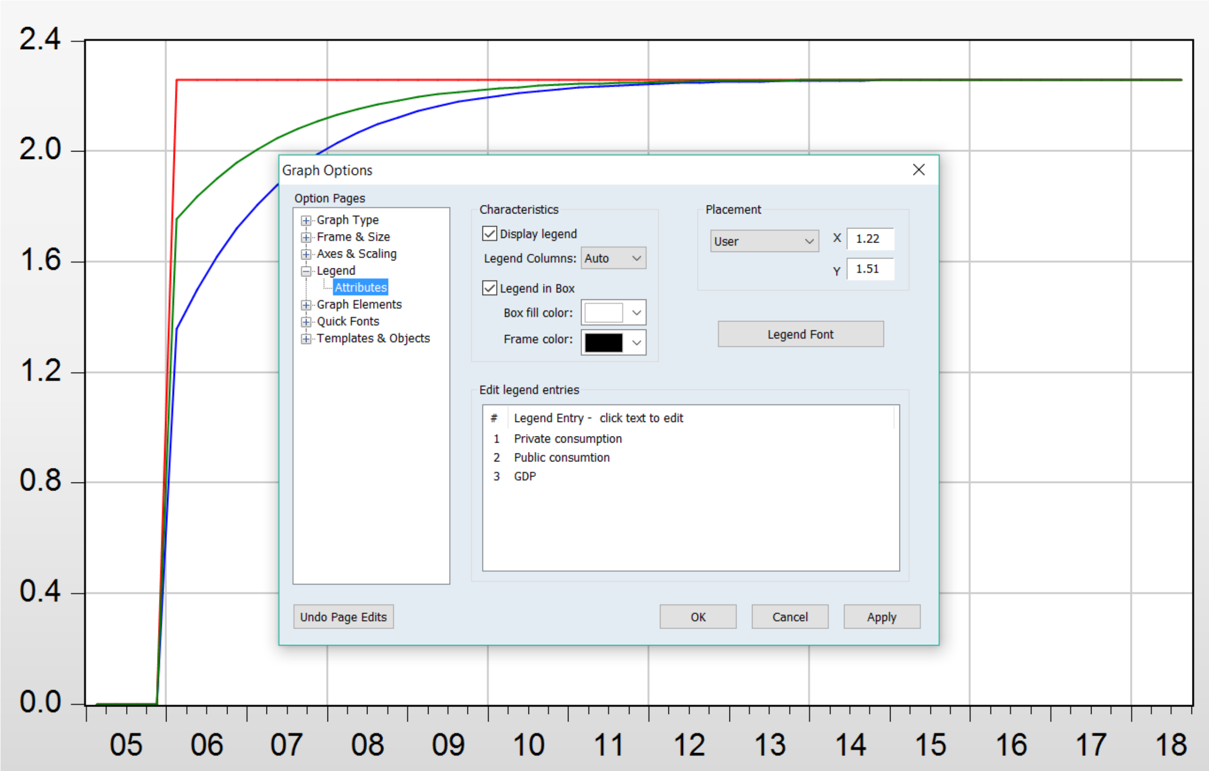


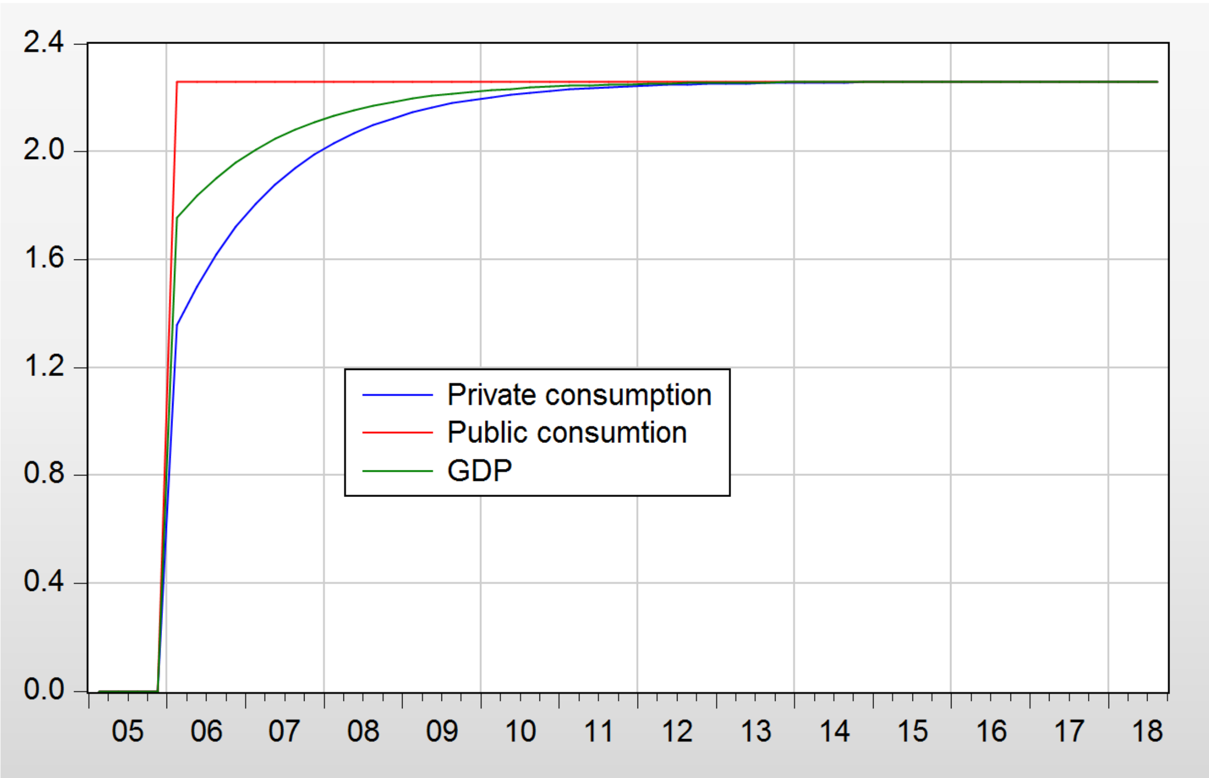


We edit the graph somewhat.


2.4
2.0
1.6
1.2
0.8
0.4
0.0
10
12
14
38
Graph Options
Option Pages
Graph Type
Basic type
Frame & Size
Axes & Scaling
Data scaling
Data axis labels
Obs/Date axis
Grid Lines
Legend
Graph Elements
Quick Fonts
Templates & Objects
Data scale grid lines
Left axis scale:
Automatic grid placement
Right axis scale:
No grid lines
Top axis
No grid lines
Bottom axis
No grid lines
Lines:
Obs & Date axis grid lines
Automatic grid placement
Steps:
Line attributes
Pattern:
Width:
1/2 pt
Color:
Draw grid lines in foreground
Undo Page Edits
OK
Cancel
Apply
100*(Q_1-Q_0)/Q_0
100*(CO_1-CO_0)/CO_0
100*(G_1-G)/G


2.4
2.0
1.6
1.2
0.8
0.4
0.0
09
10
11
12
13
14
15
16
17
18
19
20
21
22
100*(Q_1-Q_0)/Q_0
100*(CO_1-CO_0)/CO_0
100*(G_1-G)/G





17.14.4.1 Programming

Our statements will be similar to the previous case.

```
smpl 2005Q2 2100Q4
```

```
nano.scenario "scenario 1"
genr g_v=g+q_p*.01*(t>=2011)
nano.append assign @all _v
nano.override g
solve(d=d) nano
smpl 2005Q2 2100Q4
for %1 CO Q
series dv_{%1}={%1}_v-{%1}_0
series pv_{%1}=100*dv_{%1}/({%1}_0 +({%1}_0=0))
next
```

17.15 A FULL PROGRAM

The sequence of programming statements we have used can be collected into a full program, which can contain comments explaining each step in the process.

We can also use elements which are not so easy to specify by menu (especially if typing them must be repeated). For this reason, the program we propose is a little more complicated. The differences are:

17.15.1 A RESIDUAL IN THE ESTIMATED EQUATION

In the consumption equation, we shall introduce a residual call ec_co. It will be set to zero before estimation, and to the observed residual after it is performed.

Note that we have blocked the time trend after the estimation period, to allow the solution to follow a steady state path (Q and CO will grow at the same rate in the long run).

```
'   The name of the residual connects it with the variable
genr ec_co=0
'   We create a vector of coefficients
'   The name of the vector also connects it with the variable
coef(10) c_co
'   We specify the equation using coefficient and residual
'   The name of the equation again connects it with the variable
'   The equation is estimated immediately using least (not necessarily ordinary) squares
equation   eq_co.ls(p)   dlog(co)=c_co(1)*dlog(q)+c_co(2)*(log(co(-1)/q(-1))-c_co(3)*(t-2004)*(t<2004)-
c_co(4))+ec_co
'   The residual is given the value of the left side - right side
genr ec_co=resid
```

17.15.2 A RESIDUAL CHECK

We will check that the residual takes a zero value for all equations.

```
nano.scenario "scenario 1"
nano.append assign @all _s
'    We compute a zero residual with the suffix "_s"
genr ec_co_s=0
'    We declare that ec_co must use an alternate value
'     The suffixed variable must override the base one
nano.override ec_co
solve(d=d) nano

'    We compute the difference between simulation and history
for %1 CO Q
series ds_{%1}={%1}-{%1}_s
series ps_{%1}=100*ds_{%1}/({%1}+({%1}=0))
next
```

17.15.3 COMPUTING BASE FUTURE VALUES FOR THE ENDOGENOUS

We also initialize endogenous variables using the same assumption. This will give us an alternate starting point and allow to check if the solution deviates from the theoretical value.

As q and co are linked to the values in the model page (which are in turn linked to FRA_GDPV and FRA_CPV) we must "unlink" them to allow changes.

```
unlink q
unlink co
genr q=q(-1)*1.005
genr co=co(-1)*1.005
```

17.15.4 A BASELINE SOLUTION USING A SCENARIO

```
smpl 2005Q1 2100Q4
nano.scenario "scenario 1"
nano.append assign @all _p
nano.override
solve(d=d) nano
```

17.15.5 PRODUCING A SHOCK FROM SCENARIO

```
smpl 2005Q2 2100Q4
genr g_v=g+q_p*.01*(t>=2011)
nano.append assign @all _v
nano.override g
solve (d=d) nano
```

17.15.6 COMPUTING DIFFERENCES USING A LOOP

```
smpl 2005Q2 2100Q4
for %1 CO Q
series dv_{%1}={%1}_v-{%1}_p
series pv_{%1}=100*dv_{%1}/({%1}_p+({%1}_p=0))
next
```

17.15.7 THE PROGRAM

```
'
'   This program describes the functions associated to the building of a macroeconometric model under EViews
'   Although the model is extremely simple, all the main functions are treated
'  ----------- The model
'
'   (1) Private consumption CO is a function of GDP : CO=f(Q). This function will be estimated
'  (2) GDP is the sum of private and public consumption g : Q=CO+g. This equation is supposed to be exact (or at least consistent with the data)
'-------------  The program
'   First we set the directory
'   This is useful if we manage several projects at the same time
'   This guarantees that programs will access the right files
'                               files will be stored in the right place
'                               no file outside the directory will be wrongly affected
'
cd "c:\users\jean louis brillet\google drive\book_8\nano"
'   We direct the output to nano.rtf
'   This means that all "printed" output (mostly the results of tests for which the "p" option has been specified)
'   will go to the specified file
'   r calls for the RTF format (the richer, read by Word and the like). "t" would call for character text.
'   o specifies that any previous results stored under the same name will be overriden
```

```
output(r,o) nano
'
'    We create the work file nano_q  (quarterly from 1975q1 to 2010q4) with a page called "origin"
'    This page will contain the original series
'    If the workfile exists, it will be open
'    But if it exists and is open, a new version will be open in a new window
'    This is obviously dangerous, to avoid it we ask any open version to be closed
' *****  Creating the framework and the data
close nano_q.wf1
wfcreate(wf=nano_q,page=base) q 1975q1 2010q4
'    We also destroy any existing element (to start from a blank file)
delete *
'
'    We read 2 series of the French economy (OECD data) from nano_q.xls
'    The data is organized as a table with series in colums, and its upper left cell is B2
'     (second row, second colum)
'
smpl 1975q1 2010q4
read(t=xls,b2) nano_q.xls 2
'     Now we create a page for the model elements
pagecreate(page=model)  q 1975q1 2010q4
'    We start creating the data
'
'    We create a time trend : the yearly date for the first quarter, then +0.25 per quarter
smpl 1975q1 1975q1
genr T=1975
smpl 1975q2 2010q4
genr T=t(-1)+0.25
'
'    We create the model data from two orginal series :
'    CO from FRA_CPV : private consumption at 1995 prices
'    Q from FRA_GDPV : gross domestic product at 1995 prices
'    g as the difference between Q and CO
smpl 1975q1 2010q4
'    We will create CO and Q by reference to original elements
'    First we define them as "linked" variables
link CO
link Q
'    Then we create the link giving the reference to page and the series
CO.linkto base::FRA_CPV
Q.linkto base::FRA_GDPV
'    g is not available as such. We have to use the equation reversed
```

```
genr g=Q-CO
'    We create a first version of the model
'     just to check its structure
'     and the fact that it can be estilated
'
'     We ask EViews to destroy any existing version
'     to start from scratch
'     We do it only if the model exists
'      (this avoids an error message)
'
if @isobject("nano") then
delete nano
endif
'    We create a new model called "nano"
model nano
'  We add the first equation (identity)
nano.append Q=CO+g
'   For the second, we just declare our intention
'   We would like to write CO=f(Q)
'    but EViews will call for the function "f"
'   So we have to trick EViews into accepting the syntax
'   This is the closest we have found
'    f is defined as a scalar to avoid its identification as a (spurious) exogenous
scalar f=0
nano.append CO=f*(Q)
'
'   We produce a residual check
'   The model is broken into single equation ones
'   and each model is solved based on actual values
'
'    We use the largest period for which CO, Q and g are all available
'    equivalent to the availability of their product
smpl 1975q1 2004q4 if co*q*g<>na
'    We ask the results to be given the suffix "_c"
nano.append assign @all _c
'    We solve the model using the option "f"
solve(d=f) nano
'   We check that the identity holds true
'   by computing the difference between actual and computed
'   We also know that the behavioral equation can be estimated
genr dc_q=q_c-q
genr pc_q=dc_q*100/q
```

```
'     ******  Estimating the equation
'    We estimate the equation
'    First, a very simple case (linear in logs)
smpl 1975Q1 2004q4
ls(p) log(co) log(q) c
'    Equivalent to
ls(p) log(co)=c(1)*log(q)+c(2)
'    The first two positions of the "c" vector will be filled with estimated values
'     as well as the "Resid" series with the residual (actual left hand side - estimated right hand side)
'    second, with an autoregressive term
ls(p) log(co) log(q) c ar(1)
'     Now, a more sophisticated form using error correction
'     We suppose that the long-term elasticity between CO and Q is unitary
'     but we leave it free in the short run
'    First, we test the stationarity of Log(CO/Q)
uroot(trend,p ) log(co/q)
uroot(trend,p,pp) log(co/q)
'   It works (the  negative T statistic is high enough)
'    Now two techniques
'   *****  The fast one (good for the exploratory phase)
ls(p) dlog(co) dlog(q) log(co(-1)/q(-1)) t c
'     problem : estimates will be erased by the next estimation
'     and where is the equation?
'     a little better : the equation is stored
equation eq1.ls(p)  dlog(co) dlog(q) log(co(-1)/q(-1)) c t
'   ****** the slow one (good once the choice is made)
'    We create a residual, with 0 as the initial value
'     The name of the residual connects it with the variable
genr ec_co=0
'    We create a vector of coefficients
'     The name of the vector also connects it with the variable
coef(10) c_co
'    We specify the equation using coefficient and residual
'     The name of the equation again connects it with the variable
'    The equation is estimated immediately using least (not necessarily ordinary) squares
'     To delay estimation, one must drop the ".ls"
equation    eq_co.ls(p)    dlog(co)=c_co(1)*dlog(q)+c_co(2)*(log(co(-1)/q(-1))-c_co(3)*(t-2004)*(t<2004)-
c_co(4))+ec_co
'    The residual is given the value of the left side - right side
genr ec_co=resid
'     We also compute the difference between the estimated and observed variables
```

```
genr co_h=exp(c_co(1)*dlog(q)+c_co(2)*log(co(-1)/q(-1))+c_co(3)*t+c_co(4))
genr dh_co=co-co_h
genr ph_co=dh_co*100/co
'    **** creating the model
'    Now we have our model
'    We delete the old version
delete nano
'     We create a new blank one
model nano
'     We introduce the same identity as before
nano.append Q=CO+g
'      And we merge the estimated equation
'      "merging" also includes statistics
nano.merge eq_co
'     Now we have to check model consistency again
'     We shall use the suffix "_d"
nano.append assign @all _d
solve(d=f) nano
'    We will use a loop (first form)
'    It starts with "for" followed by a parameter and a list, and ends with "next"
'   The statements in between will be repeated
'    as many times as there are elements in the "for" list
'    each time the parameter is replaced by the current element
'   The brackets are just for locating the parameter
'    they are formally dropped after the replacement
'     A trick: if a variable is naturally zero (like a contemplated tax with a present zero rate) the error will be zero too
'    Computing the relative error in EViews will divide 0 by 0 generating an error, with a message
'    For us, 0/0
for %1 CO Q
genr dd_{%1}={%1}-{%1}_d
genr pd_{%1}=100*dd_{%1}/({%1}+({%1}=0))
next
'    Solving the model
'    Assessing the reliability over the past
'    Gives an idea of the reliability of forecasts
'    In the future, we will not know the residuals
'    So we set the residual at its most probable value : zero
'   ****** The "old" way
'    We store the residual value under an alternate name

genr ec_co_h=ec_co
```

```
'   We set the base value to zero
genr ec_co=0
'    We solve the model with the suffix "_s"
nano.append assign @all _s
solve(d=d) nano
'    We recover the estimation residual
genr ec_co=ec_co_h
'     This method is tedious, and runs the risk of losing the residual
'   The "scenario" way
'    We decide on a scenario, with the suffix "_s"
nano.scenario "scenario 1"
nano.append assign @all _s
'    We compute a zero residual with the suffix "_s"
genr ec_co_s=0
'    We declare that ec_co must use an alternate value
'     The suffixed variable must override the base one
nano.override ec_co
solve(d=d) nano

'    We compute the difference between simulation and history
for %1 CO Q
series ds_{%1}={%1}-{%1}_s
series ps_{%1}=100*ds_{%1}/({%1}+({%1}=0))
next
'    Now we test the economic properties of the model (!)
'    by measuring the response of the solution to a shock on assumptions
smpl 1964Q1 2004Q4
'   We still use "Scenario 1"
'   We must drop any current overriding (in our case ec_co)
'   This is done by declaring a blank list
nano.override
'    The base solution (should take the historical values)
smpl 1975Q2 2004Q4
nano.append assign @all _b
solve(d=d) nano
'     We create an alternate g, increased by one point of GDP
'      starting in 1980, to check that with a zero shock we get the base solution
genr g_v=g+q*.01*(t>=1981)
nano.scenario "scenario 1"
'   Now the suffix is "_v"
nano.append assign @all _v
'   We override g
```

```
nano.override g
solve(d=d) nano
'     We compute the difference between "_v" and "_b"
smpl 1979q1 2004q4
for %1 CO Q
series dv_{%1}={%1}_v-{%1}_b
series pv_{%1}=100*dv_{%1}/({%1}_b+({%1}_b=0))
next
'   Now we work on the future
'   We copy the page under a new name
pagecopy(page=forecast)
'     We extend the period to 2100
pagestruct(end=2100Q4) *
'     We extend the assumptions
'     First, the time trend!
smpl 2005Q1 2100Q4
genr t=t(-1)+0.25
'     The residual is kept at its last value
smpl 2005Q1 2100Q4
genr ec_co=ec_co(-1)
'     g will grow at 0.5% per quarter
genr g=g(-1)*1.005
'     We also initialize endogenous variables using the same assumption
'     This will give us an alternate starting point
'      and allow to check if the solution deviates from the theoretical value
unlink q
unlink co
genr q=q(-1)*1.005
genr co=co(-1)*1.005
'      We solve the model with the suffix "_p"
'      making sure that no overriden value is called
smpl 2005Q2 2100Q4
nano.scenario "scenario 1"
nano.append assign @all _p
nano.override
solve(d=d) nano
'       Again, we produce a shock, starting in 2006
genr g_v=g+q_p*.01*(t>=2006)
nano.append assign @all _v
nano.override g
solve(d=d) nano
'     and we compute the deviations from the base
```

```
smpl 2005q1 2100Q4
for %1 CO Q
series dv_{%1}={%1}_v-{%1}_p
series pv_{%1}=100*dv_{%1}/({%1}_p+({%1}_p=0))
next
series pv_g=q_p/g*(t>=2006)
pageselect forecast
save nano_q
```

INDEX