

AVANT PROPOS	21
INTRODUCTION	23
L'EXEMPLE : UN MODELE TRES SIMPLE	28
1 CHAPITRE 1 : NOTATIONS ET DÉFINITIONS	31
1.1 LE MODELE CONSIDERE COMME UN ENSEMBLE D'EQUATIONS	31
1.2 LES ELEMENTS D'UN MODELE	31
1.2.1 LES VARIABLES: ENDOGENES ET EXOGÈNES	31
1.2.2 LES EQUATIONS	34
1.2.2.1 Équations de comportement	34
1.2.2.2 Identités	36
1.2.3 LES PARAMETRES	38
1.2.4 LE TERME ALEATOIRE	39
1.2.4.1 Résidus et erreurs	41
1.3 LES FORMULATIONS	42
1.3.1 LA DIMENSION TEMPORELLE	42
1.3.1.1 Conséquences de la discrétisation	43
1.3.1.2 La saisonnalité	43
1.3.1.3 Modèles statiques et dynamiques	44
1.3.1.4 Un cas particulier: les anticipations rationnelles	48
1.3.1.5 Un autre cas: les modèles en temps continu	49
1.3.2 LA LINEARITE	49

2 LA MODELISATION ECONOMETRIQUE, PRINCIPES ET TECHNIQUES

1.3.2.1	Cas pratiques de non-linéarité	50
1.3.3	AUTRES PROPRIETES	53
1.3.3.1	Continuité	53
1.3.3.2	Differenciabilité	53
1.3.3.3	Existence d'une solution	54
1.3.3.4	Unicité de la solution	57
1.3.3.5	Convexité (or concavité)	58
1.3.4	LES CONTRAINTES QUE LE MODELE DOIT RESPECTER	58
1.3.4.1	Compatibilité globale	59
1.3.4.2	L'homogénéité	65
1.3.4.3	Avec dimension	66
1.3.5	NORMALISATION ET IDENTIFICATION	66
1.3.6	CONCLUSION	70
2	CHAPITRE 2: LES USAGES DES MODÈLES	71
2.1	LES DIAGNOSTICS OPERATIONNELS	71
2.1.1	LES DIFFÉRENTS TYPES DE DIAGNOSTICS: SCENARIOS ET CHOCS	71
2.1.2	LES AVANTAGES DES MODELES	72
2.1.3	UNE REMISE EN CAUSE	73
2.2	MODELES THEORIQUES	75
2.3	PETITS MODELES QUANTIFIES	75
2.3.1	A BUT EDUCATIF	75
2.3.2	A BUT SCIENTIFIQUE	76

3	CHAPITRE 3 : LES TYPES DE MODÈLES	77
3.1	LE CHAMP	77
3.2	LA TAILLE	78
3.2.1	LES DÉTERMINANTS DE LA TAILLE	78
3.2.2	UNE CLASSIFICATION	79
3.3	L'HORIZON	80
3.3.1	POUR LA PREVISION	80
3.3.2	POUR L'ANALYSE	80
3.3.3	UNE CLASSIFICATION	81
3.4	LA PERIODICITE	81
3.5	UNE DIMENSION SUPPLEMENTAIRE ?	82
3.6	LES AUTRES MODELES	82
4	CHAPITRE 4: LE PROCESSUS DE MODÉLISATION: PRÉSENTATION ET GÉNÉRALITÉS	83
4.1	LES ETAPES DU PROCESSUS	83
4.1.1	LA PREPARATION DU MODELE	83
4.1.2	L'ESTIMATION	84
4.1.3	LA RESOLUTION ET LES TESTS SUR LE PASSE	84
4.1.4	LA RESOLUTION ET LES TESTS SUR LE FUTUR	84
4.1.5	L'UTILISATION DU MODELE POUR LES PREVISIONS ET LA POLITIQUE ECONOMIQUE	85
4.2	LE CHOIX DU LOGICIEL	85
4.2.1	COMBINER UN LOGICIEL DE MODELISATION AVEC D'AUTRES ELEMENTS	85
4.2.1.1	Pour la gestion des données	85

4 LA MODELISATION ECONOMETRIQUE, PRINCIPES ET TECHNIQUES

4.2.1.2	Pour les estimations	86
4.2.1.3	Pour la résolution	86
4.2.1.4	Pour la présentation des résultats	87
4.2.2	UTILISER UN LOGICIEL SPECIALISE DANS LA MODELISATION	87
4.2.3	UTILISER UN LOGICIEL D'USAGE GÉNÉRAL	88
4.2.4	CONCLUSION	89
4.3	COMMENT ORGANISER LE DEVELOPPEMENT DU MODELE	89
5	CHAPITRE 5: LA PRÉPARATION DU MODÈLE	93
5.1	LA PREPARATION DU MODEE: LE CADRE	93
5.2	LA PREPARATION DU MODELE: ASPECTS SPECIFIQUES DES DONNEES	95
5.2.1	TYPES DE DONNEES	96
5.2.2	L'ACCES AUX DONNEES	97
5.2.2.1	Le mode de transmission	97
5.2.2.1.1	La transmission physique	97
5.2.2.1.2	La transmission par e-mail	97
5.2.2.1.3	Transmission par Internet	97
5.2.2.1.4	Autres media	98
5.2.2.2	Changement de format	98
5.2.2.3	Aspects institutionnels	100
5.2.2.4	Comment traiter des sources disparates	100
5.2.3	LA PREPARATION DES DONNEES POUR LE TRANSFERT	100
5.2.4	LE TRAITEMENT PRELIMINAIRE DES SERIES	101

5.2.4.1	Transformations temporelles	102
5.2.4.1.1	Agrégation	102
5.2.4.1.2	Désagrégation	103
5.2.4.1.3	Lissage	103
5.2.4.1.4	Désaisonnalisation	104
5.2.4.2	Changement de nomenclature	104
5.2.4.3	Transformations formelles	105
5.2.5	MISES A JOUR	105
5.2.6	SUPPRESSIONS	106
5.2.7	LA DOCUMENTATION	106
5.2.8	LA COMPARAISON DES DONNEES SOUS EVIEWS 8	108
5.2.9	CONSEQUENCES SUR L'ORGANISATION DU TRAVAIL	109
5.2.10	LES OPTIONS PRATIQUES	110
5.2.11	RETOUR A NOTRE EXEMPLE	111
5.2.12	APPLICATION A NOTRE EXEMPLE	115
5.2.12.1	Les formules de calcul	124
5.2.12.2	le programme EVIEWS	127
5.2.12.3	Les données	129
5.2.12.4	Le modèle	131
5.2.12.4.1	Introduction des équations.	134
5.2.13	UN PREMIER TEST: LA VERIFICATION RESIDUELLE DES IDENTITES	142
5.2.13.1	Les types d'erreurs rencontrés	144
5.2.13.2	Le traitement des erreurs	146

5.2.13.3	Retour à l'exemple	146
5.2.13.4	Une astuce: comment établir les groupes d'identités et de relations de comportement	149
5.3	ÉLÉMENTS DE PROGRAMMATION SOUS EViews	152
5.3.1	LES GROUPEs 6	152
5.3.2	LES BOUCLES	154
5.3.3	LES PROGRAMMES: NOUVEAUTES EViews 8	156
6	CHAPITRE 6: L'ESTIMATION DES ÉQUATIONS	159
6.1	LE PROCESSUS D'ESTIMATION	159
6.1.1	LES METHODES D'ESTIMATION	160
6.1.1.1	Moindres carrés ordinaires	160
6.1.1.2	La prise en compte de l'autocorrélation: une première technique	161
6.1.1.3	L'homoscédasticité	163
6.1.1.4	Les retards échelonnés	163
6.1.1.5	Les moindres carrés non linéaires	164
6.1.1.6	Les estimations simultanées	166
6.1.1.7	Les estimations en données de panel	168
6.1.1.7.1	Les «pools»	168
6.1.1.7.2	Les systèmes	171
6.1.2	LES TESTS	174
6.1.2.1	Qualité globale	176
6.1.2.1.1	L'autocorrélation temporelle	176
6.1.2.1.2	La statistique du R ² ou R-carré	179

6.1.2.1.3	Le R2 en question	181
6.1.2.1.4	Autres tests de qualité globale	183
6.1.2.1.5	Le test de rupture de Chow	185
6.1.2.2	La qualité des influences individuelles	188
6.1.2.3	Le terme constant	191
6.1.3	UN SUJET SPÉCIFIQUE: LA STATIONNARITE, LA COINTEGRATION ET LES MODELES A CORRECTION D'ERREUR 193	
6.1.3.1.1	La stationnarité et le test de Dickey-Fuller	193
6.1.3.1.2	Un premier modèle à correction d'erreur	195
6.1.3.2	Les avantages de cette formulation	199
6.1.3.3	La cointégration	200
6.2	APPLICATIONS: NOTRE MODELE	202
6.2.1	LES VARIATIONS DE STOCKS	204
6.2.1.1	Les éléments de base de l'estimation sous EViews	205
6.2.1.2	Une technique alternative: utiliser la fenêtre des commandes	211
6.2.1.3	D'autres spécifications possibles	211
6.2.1.3.1	L'introduction des tendances dans les équations	214
6.2.1.4	La préparation des équations pour le modèle	218
6.2.1.4.1	Utiliser un programme.	219
6.2.1.5	Les fonctionnalités introduites par EViews 8	220
6.2.2	L'INVESTISSEMENT: LA NECESSITE D'ETABLIR UN CADRE THÉORIQUE COHERENT AVANT TOUTE ESTIMATION 221	
6.2.3	L'EMPLOI: LA STATIONNARITE, LES MODELES A CORRECTION D'ERREUR, LE TEST DE RUPTURE	229
6.2.3.1	Le cadre économique	229

6.2.3.2	Les formules: stationnarité et correction d'erreurs	231
6.2.3.3	Les premières estimations	232
6.2.3.3.1	Une variable muette vertueuse	245
6.2.4	LES EXPORTATIONS: LES PROCESSUS AUTORÉGRESSIFS, LA COINTÉGRATION, LA STABILITÉ DE LONG TERME	248
6.2.4.1	Introduire un processus autorégressif	250
6.2.4.1.1	L'application de la cointégration sous EViews	256
6.2.5	LES IMPORTATIONS: ALLER PLUS LOIN DANS LA COINTÉGRATION ET LA STABILITE DE LONG TERME	268
6.2.6	RETOUR A LA VERIFICATION RESIDUELLE	289
6.2.7	LE MODÈLE ACTUEL	291
6.2.8	ASPECTS LOGICIELS	292
Z	<u>CHAPITRE 7: LE TEST DU MODÈLE PAR DES SIMULATIONS SUR LE PASSÉ</u>	295
7.1	LA RESOLUTION	296
7.1.1	GAUSS-SEIDEL	296
7.1.2	RITZ-JORDAN	298
7.1.3	NEWTON ET SES VARIANTES	298
7.1.3.1	La méthode de Broyden	302
7.1.3.2	Le cas d'une forme identifiée	302
7.1.4	LES ITÉRATIONS ET LE TEST DE CONVERGENCE	303
7.1.4.1	Les options générales	304
7.1.4.2	Les options EViews	307
7.1.5	ETUDE DE LA CONVERGENCE	307
7.1.5.1	La matrice d'incidence	307

7.1.5.2	Application à notre modèle	308
7.1.5.3	Accélérer Gauss-Seidel: les facteurs de relaxation	324
7.1.5.3.1	Des modèles plus complexes	329
7.1.5.4	Comparaison des efficacités	332
7.1.6	LA RÉOLUTION DU MODÈLE: LES TECHNIQUES EViews ELEMENTAIRES	342
7.1.6.1	La syntaxe	342
7.1.6.1.1	Le suffixe	343
7.1.7	LES TECHNIQUES DE RESOLUTION DES PROBLEMES	344
7.1.7.1	Les outils proposés par EViews	344
7.1.7.2	Autres outils	348
7.1.7.2.1	Utiliser des facteurs de relaxation	348
7.1.7.2.2	Réordonner le modèle	349
7.1.7.2.3	modifier les valeurs des paramètres	349
7.1.7.2.4	modifier la spécification des équations	349
7.1.7.3	L'application de ces outils	350
7.2	UNE PREMIERE VALIDATION	359
7.2.1	SIMULATIONS EX POST	359
7.2.1.1	Réévaluer ce critère	364
7.2.1.1.1	Essais partiels	366
7.2.2	LES PREVISIONS EX POST	367
7.2.3	LA RESOLUTION DES MODELES: LES SCENARIOS	369
7.2.4	APPLICATION A NOTRE EXEMPLE	371
7.2.5	CHOCS ANALYTIQUES	374

7.2.5.1	Le choix des hypothèses modifiées	376
7.2.5.2	Le choix des résultats examinés	377
7.2.6	NOTRE EXEMPLE	377
7.2.6.1	Les résultats	379
7.2.6.2	Les défauts des chocs sur le passé	382
7.2.7	ASPECTS LOGICIELS	383
7.2.7.1	Deux solutions alternatives	385
7.2.7.1.1	Les tableurs	385
7.2.7.1.2	La programmation directe	387
8	CHAPITRE 8: LES TESTS DU MODÈLE SUR LE FUTUR	389
8.1.1	LES HYPOTHESES	391
8.1.1.1	Un cas particulier: les estimations	392
8.1.2	ADAPTER LES FORMUMATIONS	392
8.1.2.1	comment améliorer les chances (et la vitesse) de convergence	394
8.1.3	RESOUDRE DES PARTIES DU MODELE	395
8.1.4	VERIFIER L'EXISTENCE D'UNE SOLUTION DE LONG TERME	395
8.2	LES PROBLÈMES DE CONVERGENCE DANS LE COURT TERME	398
8.3	LES PROBLÈMES DE CONVERGENCE DANS LE MOYEN ET LONG TERME	398
8.4	LE TESTS DES RESULTATS	399
8.4.1	LES RESULTATS DES SIMULATIONS	399
8.4.2	LES RESULTATS DES CHOCS	399
8.5	LES ASPECTS EViews	400

8.5.1	LA PRODUCTION DES SIMULATIONS	400
8.5.1.1	La préparation du fichier de travail («Workfile»)	400
8.5.2	LA PRODUCTION D'UNE SIMULATION DE BASE	403
8.5.3	LA PRODUCTION DES CHOCS	404
8.5.4	LA MODIFICATION DES SPÉCIFICATIONS DU MODELE	405
8.6	LES ANTICIPATIONS RATIONNELLES	407
8.6.1	LE CADRE	407
8.6.2	LES CONSEQUENCES POUR LES SIMULATIONS DU MODELE	408
8.6.2.1	à l'intérieur de la période de prévision	408
8.6.2.2	au-delà de la période de prévision	409
8.6.3	ÉLÉMENTS TECHNIQUES	409
8.6.4	APPLICATION A NOTRE EXEMPLE	409
8.6.5	LE TEST	411
8.6.6	LES RESULTATS	411
8.7	LES SIMULATIONS STOCHASTIQUES	414
8.7.1	DES ERREURS PUREMENT STATISTIQUES	414
8.7.2	DES ERREURS SUR LES FORMULATIONS ET LES HYPOTHÈSES	415
8.7.3	LE TRAITEMENT DES ERREURS	415
8.7.4	L'INTERET DE CETTE TECHNIQUE	418
8.7.5	RETOUR A NOTRE EXEMPLE	418
8.7.6	LES RESULTATS	419
8.8	ALLER PLUS LOIN : L'ÉTUDE DES PROPRIETES DU MODELE	422
8.8.1	L'ANALYSE DES VALEURS PROPRES	422

8.8.1.1	Premier exemple: un modèle très simplifié	424
8.8.1.2	Deuxième exemple: les effets de la règle de Taylor	427
8.8.1.3	Les sentiers de croissance équilibrée	430
8.8.2	LE CAS DES MODELES A CORRECTION D'ERREUR: UN EXEMPLE SIMPLE	434
9	CHAPITRE 9: L'UTILISATION DES MODÈLES	439
9.1	LES DIAGNOSTICS OPERATIONNELS	439
9.1.1	LES SCENARIOS ET LEURS DIFFERENTS TYPES	439
9.1.2	LA GESTION DES PREVISIONS: LE CIBLAGE DES SIMULATIONS	442
9.1.2.1	Sur le passé	442
9.1.2.2	En prévision	442
9.1.3	LES DIFFERENTS TYPES DE CHOCS	443
9.1.3.1	Chocs analytiques et chocs complexes	443
9.1.3.2	Chocs externes et chocs de politique économique	444
9.1.4	LA PREVISION: ASPECTS TECHNIQUES	447
9.1.5	LE CHANGEMENT DE SPECIFICATIONS	448
9.1.5.1	Les raisons	448
9.1.5.1.1	Identifier les pays en transition	449
9.1.5.1.2	En ce qui concerne les données	450
9.1.5.2	La technique	451
9.1.6	LE CONTROLE OPTIMAL	451
9.1.7	L'ENSEIGNEMENT PAR LES MODELES	451
9.2	LA PRESENTATION DES RESULTATS	452

9.2.1	ELEMENTS GENERAUX	452
9.2.2	LES TABLEAUX	453
9.2.3	LES GRAPHIQUES	454
10	CHAPITRE 10: L'APPLICATION DES PRINCIPES AUX CAS OPERATIONNELS	479
10.1	LE CADRE COMPTABLE	480
10.1.1	LES AGENTS: UNE PREMIERE CLASSIFICATION	481
10.1.2	LES OPERATIONS	481
10.1.3	LES COMPTES ECONOMIQUES INTEGRES	482
10.1.4	SECTEURS, BRANCHES ET PRODUITS	483
10.1.5	LES SUBDIVISIONS DES AGENTS	484
10.1.6	UN MODELE A PLUSIEURS PAYS	484
10.2	UN MODEE À UN SEUL PAYS ET UN SEUL BIEN	485
10.2.1	LES ASPECTS ECONOMIQUES	485
10.2.1.1	Le processus productif	486
10.2.1.1.1	Une définition du coût relatif.	488
10.2.1.1.2	Les éventuelles modifications de la technologie	488
10.2.1.1.3	Un problème spécifique: la détermination statistique des capacités de production	489
10.2.1.2	Les variations de stocks	490
10.2.1.3	Le chômage	490
10.2.1.4	Le système de prix	492
10.2.2	LE COMPTE DES ENTREPRISES	499

10.2.2.1	Le compte des ménages	500
10.2.2.2	Le commerce extérieur	505
10.2.2.3	Le budget de l'État	510
10.2.2.4	Les éléments financiers et monétaires	511
10.2.3	LES PROGRAMMES EVIEWS	513
10.2.3.1	La production du modèle	514
10.2.3.2	La production des données: l'exemple de la base OCDE	531
10.2.3.3	La création du modèle et la vérification de la cohérence équations - données	546
10.2.3.4	L'estimation des équations: exemples	550
10.2.3.4.1	Une fonction de production à facteurs complémentaires	551
10.2.3.4.1.1	Estimation de l'investissement	551
10.2.3.4.1.2	L'estimation de l'emploi	554
10.2.3.4.2	Une fonction de type Cobb-Douglas	561
10.2.3.4.2.1	le cadre théorique	562
10.2.3.4.2.1.1	la maximisation des marges	562
10.2.3.4.2.1.2	Cibles et valeurs observées	564
10.2.3.4.2.1.3	La capacité de production	564
10.2.3.4.2.1.4	Le facteur temps	565
10.2.3.4.2.1.5	L'inertie des facteurs	565
10.2.3.4.2.1.6	les coûts relatifs	566
10.2.3.4.2.2	L'estimation	567
10.2.3.4.3	Les variations de stocks	573
10.2.3.4.4	Le chômage	575

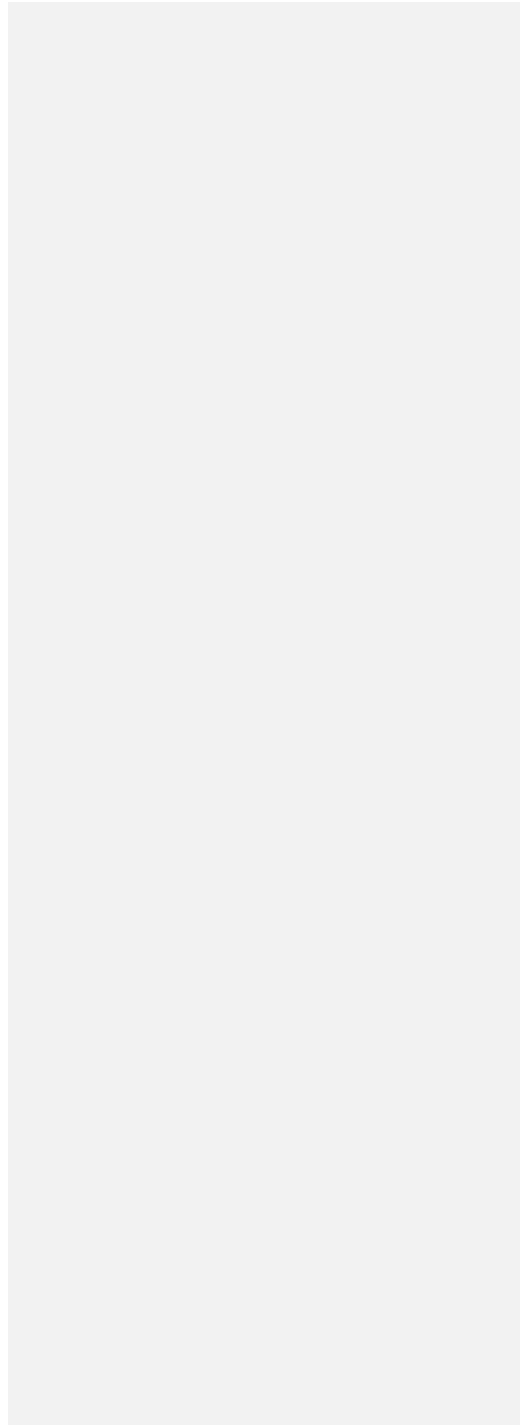
10.2.3.4.5	Le déflateur de la valeur ajoutée	578
10.2.3.4.6	Le taux de salaire	583
10.2.3.4.7	Les déflateurs des échanges extérieurs	594
10.2.3.4.8	La consommation des ménages	600
10.2.3.4.9	Les exportations	603
10.2.3.4.10	Les importations	608
10.2.3.5	Une nouvelle vérification de la cohérence équations - données	613
10.2.3.6	La résolution du modèle sur le futur	615
10.2.3.6.1	Éléments généraux	619
10.2.3.6.2	Favoriser la convergence à long terme: de nouveaux éléments	619
10.2.3.6.3	Faire converger le modèle à court terme	621
10.2.3.6.4	Faire converger le modèle à moyen et long terme	622
10.2.3.6.5	Éléments spécifiques aux prévisions	622
10.2.3.7	La production d'une prévision	622
10.2.3.8	La production des chocs d'hypothèse	623
10.2.4	L'APPLICATION DES PROGRAMMES	628
10.2.4.1	La production du cadre du modèle	628
10.2.4.2	La production des données	634
10.2.4.3	La production des groupes et la vérification de la cohérence modèle-données	635
10.2.4.4	L'estimation des équations	636
10.2.4.5	La résolution du modèle sur le futur	636
10.2.4.6	La production d'une prévision	636
10.2.4.7	Les chocs d'hypothèses: le cas des facteurs complémentaires	642

10.2.4.7.1	Une augmentation de la demande de l'État	643
10.2.4.7.2	une dévaluation (du franc français !)	648
10.2.4.7.3	une diminution des droits de douane étrangers	651
10.2.4.7.4	une diminution des droits de douane locaux	655
10.2.4.8	Le cas Cobb-Douglas	661
10.2.4.8.1	une augmentation de la demande publique	663
10.2.4.8.2	une dévaluation (du franc français)	667
10.2.4.8.3	une diminution des droits de douane étrangers	667
10.2.4.8.4	une diminution des droits de douane locaux	668
10.3	UN MODELE MONO PAYS, MULTI PRODUITS	671
10.3.1	LES DIFFERENCES STRUCTURELLES	672
10.3.2	LES DIFFERENCES DE SENSIBILITE	673
10.3.3	LES DIFFERENCES DE COMPORTEMENT LOGIQUE	673
10.3.3.1	Le processus de production	673
10.3.3.2	Le commerce extérieur	675
10.3.3.3	Les conséquences pour la décomposition par produits	675
10.3.4	INTRODUIRE LES CONSOMMATIONS INTERMEDIAIRES	676
10.3.5	LES ASPECTS SECTORIELS SPECIFIQUES	678
10.3.6	LES DONNEES	680
10.3.7	CONSEQUENCES POUR LES ESTIMATIONS	681
10.3.8	LA FONCTION DE PRODUCTION	682
10.3.8.1	Le produit primaire	682
10.3.8.2	Le produit industriel	684

10.3.9	LE CHOMAGE	684
10.3.10	LES VARIATIONS DE STOCKS	684
10.3.11	LA CONSOMMATION DES MENAGES	684
10.3.12	LE COMMERCE EXTERIEUR	685
10.3.13	LES SALAIRES	685
10.3.14	LES PRIX	686
10.3.14.1	Le déflateur de la valeur ajoutée	686
10.3.14.2	Le prix à la production	686
10.3.14.3	Le prix des échanges	686
10.3.14.4	Le prix de la demande	686
10.3.14.5	Le budget de l'État	687
10.3.14.6	Le programme EViews	687
10.3.14.7	Une liste d'éléments	687
10.4	UN MODELE MULTI PAYS, MONO PRODUIT	696
10.4.1	PREMIER SUJET: LES MODELES	696
10.4.1.1	Les modèles pays	696
10.4.2	DEUXIEME SUJET: LES ECHANGES ENTRE PAYS	698
10.4.3	UNE METHODE COHERENTE: MACSIM	699
10.4.3.1	Le système de prix	699
10.4.3.2	Les taux d'utilisation étrangers	700
10.4.3.3	Les importations globales	701
10.4.3.4	Les flux bilatéraux	701
10.4.3.5	Troisième sujet: les comportements communs	704

10.4.3.6	Le reste du monde	704
10.4.3.7	La séquence de production du modèle	704
10.4.3.7.1	Décider de l'architecture	705
10.4.3.7.2	Produire un ensemble de modèles mono pays	705
10.4.3.7.3	Assembler les modèles	706
10.4.3.7.3.1	Calculer les flux bilatéraux	706
10.4.3.7.3.2	Formaliser le commerce À prix constants	707
10.4.3.7.3.3	Formaliser les déflateurs des échanges	707
10.4.3.7.3.4	Introduire les identités	707
10.4.3.7.3.5	produire les éléments communs	708
10.4.3.7.3.6	Le reste du monde	708
10.4.3.7.3.7	Résoudre le modèle	708
10.4.3.7.3.8	Utiliser le modèle	709
10.4.4	LE PROGRAMME EViews	710
10.5	UN MODELE REGIONAL	710
10.6	UN MODÈLE MULTI PAYS, MULTI PRODUITS	711
10.6.1	LE PROGRAMME EViews	711
11	CHAPITRE 11 : NOUVEAUX ÉLÉMENTS D'EViews 8 POUR LA MODÉLISATION	713
11.1	LA GESTION DES SERIES	713
11.2	LES PROGRAMMES	715
11.3	LA GESTION DES MODELES	715
11.4	LA GESTION DE LA RESOLUTION DU MODELE	717

11.5 L'AFFICHAGE	718
<u>ANNEXE : UNE METHODE SIMPLE POUR INVERSER LE STATUT ENDOGENE – EXOGENE DES VARIABLES DU MODELE</u>	719
LE PROBLEME	719
LE CAS LINEAIRE	720
LE CAS NON-LINEAIRE	721
QUELQUES INTERROGATIONS	722
UN TEST SUR UN TRES PETIT MODELE.	724
LES RÉSULTATS	735
UN TEST SUR DE PLUS GRANDS MODELES	737
<u>BIBLIOGRAPHIE</u>	743
ECONOMETRIE ET STATISTIQUE	743
MACROECONOMIE	745
MODELES ET MODELISATION	746



AVANT PROPOS

Le but de ce livre est un peu spécial.

Tout d'abord, bien sûr, par son sujet: nous devons admettre que la modélisation économétrique structurelle n'est plus aussi populaire, après avoir perdu du terrain face aux modèles d'équilibre général dynamiques et en particulier à leurs versions stochastiques.

Nous allons soutenir que si cela peut être vrai dans le domaine académique (il suffit de considérer le programme des congrès et colloques) il y a encore beaucoup de place pour les modèles structurels. En effet de nombreuses institutions les utilisent encore et en construisent des nouveaux, aussi bien dans les pays développés qu'en développement. Nous allons chercher à montrer que cette position est tout à fait justifiée, et que, pour une grande partie des applications de modélisation, en particulier l'analyse et l'interprétation des interactions macroéconomiques, l'appel à des modèles structurels demeure une bonne stratégie, sans doute la meilleure disponible.

Mais nous ne nous arrêterons pas à prouver l'utilité de ces modèles. Pour les personnes que nous avons convaincues, ou qui l'étaient déjà, nous allons fournir un ensemble d'outils facilitant toutes les tâches du processus de modélisation. À partir d'éléments de base, nous conduirons par étapes l'utilisateur à un niveau auquel il devrait être en mesure de construire, gérer et utiliser un modèle opérationnel dans son cadre professionnel.

Cela signifie que ce livre, comme son titre l'indique, se concentrera essentiellement sur les caractéristiques appliquées et même techniques, ce qui ne signifie pas pour autant que son approche sera simpliste.

Après une description du champ nécessaire, nous allons consacrer la plus grande partie du livre à montrer au lecteur comment construire son propre modèle, de la stratégie générale aux détails techniques. Pour cela nous nous appuyerons sur un exemple spécifique, présenté au début, et que nous allons suivre à travers toutes les étapes de l'élaboration du modèle. Lorsque la situation deviendra plus complexe (avec l'ajout des dimensions produit et internationale), nous trouverons encore ce modèle à la base de notre explication.

Compte tenu du fait que le logiciel EViews devient la norme pour les travaux de modélisation (presque autant qu'Excel pour les tableurs¹), nos exemples seront basés sur ce logiciel. Cela nous permettra d'être

¹ Et contrairement à ce dernier, on ne trouve pas de version gratuite.

plus utile aux utilisateurs d'EViews, en nous concentrant sur sa pratique (y compris par la présentation de quelques astuces).

Enfin, de façon tout aussi importante, sinon plus, nous fournirons un ensemble de fichiers qui permettront au lecteur de pratiquer la modélisation (seul ou dans le cadre d'un cours). Et pour les utilisateurs plus avancés, nous donnerons accès à des programmes permettant de produire des petits modèles opérationnels, qu'ils pourront adapter à leurs propres idées, en évitant les tâches fastidieuses, qui auront déjà été déjà préparées: la production des données, la définition du cadre comptable et l'organisation des simulations sur le futur. Tous ces éléments seront fournis gratuitement et téléchargeables sur un site spécifique.

Enfin, une dernière remarque très importante. Le lecteur ne trouvera pas ici un véritable manuel d'économétrie, même limité aux pratiques usuelles. De même les éléments de macroéconomie ne prétendront-ils pas à la qualité d'un ouvrage spécifique. L'auteur ne se considère d'ailleurs pas comme un économètre ou un économiste de haut niveau. Plutôt que de se confronter à des auteurs reconnus, il se contentera de donner leurs références. D'ailleurs, dans la version anglaise distribuée par EViews, la partie «économétrie» a quasiment été éliminée, au profit de l'utilisation du manuel officiel.

Par contre il ose prétendre que sa très² longue expérience en modélisation structurelle, qu'il a pu formaliser au cours des ans en une stratégie cohérente, sera fort utile à tout chercheur abordant ce domaine, et que même les utilisateurs chevronnés y découvriront sans doute des éléments nouveaux. Cette stratégie, à la fois globale et détaillée, devrait lui permettre de construire plus rapidement un meilleur outil, respectant un certain nombre de critères. Il pourra alors se concentrer sur les éléments originaux de son projet.

² Trop ?

INTRODUCTION

Depuis une date située au début du vingtième siècle, les économistes ont cherché à produire des outils mathématiques qui, appliquées à un problème pratique, formalisent une théorie économique permettant d'y associer une image numérique fiable. L'application la plus naturelle est bien sûr de prévoir l'avenir, et cet objectif était présent dès le début. Mais on peut aussi envisager de connaître les conséquences d'un événement imprévu, ou de mesurer l'efficacité d'un changement dans la politique actuelle, ou même d'améliorer la compréhension d'un ensemble de mécanismes trop complexes pour être saisi par l'esprit humain.

Au cours des dernières décennies, trois types d'outils de ce type ont vu le jour, qui se partagent le marché actuel de la modélisation.

- Les modèles VAR (Vector AutoRegressive). Ils cherchent à donner l'image la plus fiable de l'avenir proche, en utilisant une structure complexe estimée sur la base d'éléments retardés, reposant essentiellement sur la qualité statistique, bien que la théorie économique ne soit pas absente, principalement par le biais des contraintes sur les spécifications. La principale utilisation de cet outil est de produire des prévisions à court terme.
- Les modèles d'équilibre général calculable (Computable General Equilibrium ou CGE). Ils utilisent une structure détaillée, avec des formulations a priori et les coefficients calibrés, dans le but de résoudre un problème général local, grâce à l'application d'un ou de plusieurs comportements d'optimisation. Dans leurs premières versions, les questions abordées ont généralement concerné l'optimisation de l'allocation des ressources, ou la conséquence d'accords commerciaux. Les mécanismes décrits contiennent généralement peu d'effets dynamiques.

Mais cela n'est plus vrai pour les modèles dynamiques d'équilibre général stochastiques (Dynamic Stochastic General Equilibrium ou DSGE), qui dominent le champ actuel. Ils intègrent des comportements dynamiques et tiennent compte de l'incertitude dans les évolutions. Par rapport aux modèles traditionnels (voir plus loin), ils formalisent explicitement les équilibres d'optimisation, en se basant sur le comportement agrégé des agents individuels. Cela signifie qu'ils permettent aux agents d'adapter leur comportement aux changements des règles qui régissent les comportements des autres, y compris l'Etat, échappant ainsi en principe à la critique de Lucas. Comme le modèle ne repose pas sur des équations estimées, le calibrage sera requis pour la plupart des paramètres. Leur validation est donc assez subjective.

- Les modèles dits «structurels». Ils partent d'un cadre économique donné, définissant les comportements des agents individuels selon une théorie économique globalement cohérente.

Ils utilisent les données disponibles pour associer à ces comportements des formules fiables, qui seront liées par des identités garantissant la cohérence de l'ensemble. Ces modèles peuvent être considérés comme situés à mi-chemin entre les deux catégories ci-dessus: ils s'appuient sur des tests statistiques, en même temps que sur la théorie. Pour accepter une formule, elle doit respecter les deux types de critères.

L'utilisation de ce dernier type de modèles, qui occupait tout le champ de la modélisation au début de son histoire, est maintenant limitée à l'analyse des politiques et des prévisions à moyen terme. Pour ces dernières, ils présentent des avantages considérables: les formulations théoriques complètes procurent une image claire et compréhensible, y compris par la mesure des influences individuelles. Ils permettent également d'introduire des contraintes de stabilité conduisant à des équilibres de long terme identifiés, et de séparer cet équilibre des fluctuations dynamiques qui y conduisent.

Par rapport aux modèles CGE et DSGE, les comportements d'optimisation sont présents (comme nous le verrons plus tard) et introduits dans les équations estimées. Mais ils y sont gelés, dans un état associé à une période, et au comportement des autres agents. Si ces conditions ne changent pas, la validation statistique est un avantage important. Mais la sensibilité aux chocs est imparfaite, dans une mesure difficile à évaluer.

Une très bonne (et objective) description du problème se trouvera dans:

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

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

Il nous semble que le principal critère de choix entre les modèles structurels traditionnels et les DSGE se situe dans le compromis entre la validation statistique et l'adaptabilité des comportements.

Ces dernières années, la popularité de la modélisation économétrique structurelle semble s'être stabilisée. Un indice personnel (ce n'est pas une preuve réelle) est la demande constante d'aide à la modélisation structurelle de la part des différentes institutions nationales.

Un autre problème est qu'étant le premier outil de production (apparu dans les années trente du siècle dernier), il a été appliqué immédiatement à la tâche ambitieuse de produire des prévisions fiables. La complexité de l'économie, et la présence d'un grand nombre de chocs aléatoires internes ou externe ont rendu cet objectif complètement irréaliste (et cela est encore plus vrai aujourd'hui). Pendant les années dorées de la modélisation structurelle, quand l'économie progressait à un rythme régulier (et élevé), la prévision était aussi facile que mener un cheval docile sur un chemin rectiligne: tout le monde pouvait le faire. Mais quand le cheval est devenu sauvage, les faiblesses du cavalier sont devenues visibles, et il n'est pas resté en selle trop longtemps. A défaut de réussir dans une tâche trop difficile pour n'importe quel outil (y compris les modèles VAR et EGC, qui n'ont pas à prévoir à moyen terme), ceci a jeté le discrédit

sur tous les modèles structureaux et leurs usages, y compris l'analyse des politiques et même la compréhension et l'interprétation des mécanismes économiques complexes, applications pour lesquelles ni VAR ni CGE ne peuvent les affronter à notre avis.

Cependant, même avec des ambitions limitées, produire un bon modèle économétrique structurel n'est pas une tâche simple. Même un économiste professionnel, ayant une excellente connaissance de la théorie économique (mais pas nécessairement une image complète et cohérente) et des techniques économétriques (mais pas nécessairement de leur application pratique) aura des difficultés à produire un modèle économétrique fiable et opérationnel.

Le but de ce livre est de raccourcir le processus d'apprentissage, de plusieurs façons.

Après une présentation globale des modèles économiques:

- les notations, les définitions, les caractéristiques mathématiques (dynamique, linéarité, continuité, identifiabilité ...).
- les applications: la théorie économique, la prévision, la pédagogie.
- la classification des modèles existants.

Nous allons décrire comment organiser la séquence des tâches de modélisation, de construction de modèle, de production des données et de spécification d'un cadre économique cohérent, jusqu'à la réalisation d'études opérationnelles.

Pour chaque tâche, nous donnerons tous les éléments de méthodologie nécessaires.

Nous présenterons les principales options économiques disponibles, avec quelques explications théoriques.

Toutes ces explications seront basées sur un exemple concret, la production d'un modèle très simple de l'économie française. La taille ne nous interdira pas de traiter la plupart des problèmes rencontrés dans le déroulement du processus.

Les méthodes, les techniques et les solutions proposées seront basées sur le logiciel EViews, à notre avis la meilleure option actuelle pour ce genre de tâche, si l'on considère son prix (et peut-être même dans l'absolu). Cela nous permettra de présenter quelques fonctionnalités utiles et astuces, et de fournir une séquence de programmes complets. L'utilisateur pourra les modifier à volonté, mais pas nécessairement trop lourdement, comme tous les modèles de cette catégorie présentent un nombre important d'éléments communs. L'élément principal est bien sûr le processus d'estimation, chaque cas conduisant généralement à une version nouvelle de chaque équation de comportement.

Un ensemble de programmes documentés sera fourni, en suivant les principes ci-dessus

- Pour le petit exemple,
- Pour un produit plus détaillé, un modèle associé à un seul pays, assez proche d'une version opérationnelle.

Ces programmes permettront:

- d'importer les données d'origine
- de construire le cadre du modèle
- de transformer les données en les rendant cohérentes avec les éléments du modèle.
- d'estimer un ensemble d'équations, en commençant par des comportements types, éventuellement modifiés par l'utilisateur.
- de vérifier la cohérence technique et théorique du modèle résultant.
- de produire des prévisions et des études de politique économique.

Dans chaque cas, nous allons présenter des programmes qui fonctionnent réellement. Une solution économétrique sera trouvée, fiable tant en termes statistiques qu'économiques. Et les propriétés des modèles seront plutôt satisfaisantes, avec une solution à long terme et une dynamique raisonnable qui y conduit.

Enfin, nous allons aborder les problèmes plus complexes: les modèles multi-produits et multi-pays (et les deux options combinées). Les questions spécifiques seront décrites, et un cadre associé à un modèle à plusieurs produits sera proposé, en suivant les mêmes lignes que l'exemple précédent.

Le but de ce livre est donc à la fois limité et ambitieux. Sans entrer dans des développements théoriques complexes, il devrait donner aux lecteurs tous les éléments nécessaires pour construire leur propre modèle. Étant libérés des tâches les plus techniques (et fastidieuses), ils pourront se concentrer sur les plus intelligentes (et stimulantes).

Les lecteurs doivent être conscients qu'ils ne trouveront ici ni une description complète des méthodes économétriques et statistiques disponibles, ni un cours de théorie économique. Nous allons nous limiter dans ces domaines à des éléments de base, et nous concentrer sur leurs liens avec le processus de modélisation lui-même. Pour de plus amples informations, on peut se reporter aux éléments de la liste de références fournie à la fin du volume.

L'EXEMPLE : UN MODELE TRES SIMPLE

Pour présenter les éléments et le cadre d'un modèle économétrique structurel, nous allons utiliser un exemple précis, que nous aborderons de façon permanente au cours de notre présentation. En dépit de sa taille limitée, nous pensons qu'il reste tout à fait représentatif de la classe des modèles que nous étudions dans ce manuel.

Au début de tout processus de construction du modèle, il est nécessaire de préciser d'une manière assez large sa logique globale, et les comportements qu'il doit décrire. Aucune équation n'a à être établie à ce moment. Nous allons nous placer dans cette situation.

Dans notre exemple, un économiste a décidé de construire un modèle très simple de l'économie française. Comme nos tests seront basés sur des données réelles, un pays devait être choisi, mais les principes présentés s'appliquent à tout pays industrialisé de taille moyenne (et dans la plupart des cas à tout pays quelconque).

Notre modèle inclut les éléments suivants.

- Sur la base de leurs prévisions de production et de connaissances sur la productivité des facteurs, les entreprises investissent et embauchent des travailleurs afin d'y adapter leur capacité de production. Toutefois, ils exercent une certaine prudence dans ce processus, car ils ne veulent pas se retrouver avec trop d'éléments inutilisés.
- Les niveaux atteints dans la pratique permettent de définir le potentiel de production.
- Les entreprises constituent également des stocks.
- Les ménages obtiennent des salaires, sur la base de l'emploi total (comprenant les fonctionnaires), mais aussi une part du produit intérieur brut. Ils consomment une partie de ce revenu.
- La demande finale est définie comme la somme de ses composantes: la consommation, l'investissement productif, l'investissement logement, les variations de stocks, et la demande publique.
- Les importations représentent une part de la demande locale («la demande intérieure»). Mais moins il reste de capacités disponibles, plus une augmentation de la demande devra faire appel aux importations.

- Les exportations suivent la demande mondiale, mais les producteurs sont également limités par les capacités disponibles, et leur priorité sera de satisfaire la demande locale.
- L'offre est égale à la demande.
- Le capital productif augmente avec l'investissement, mais il est soumis à l'amortissement.

Le cadre ci-dessus semble assez simple, et certainement simpliste. De toute évidence, il manque de nombreux éléments, tels que les prix, les concepts financiers et les impôts. Cette question sera traitée lors de développements ultérieurs.

N'allons pas plus loin pour le moment. On peut observer que, si nous n'avons pas encore construit une seule équation, quelques-unes sont déjà implicites dans le texte ci-dessus.

CHAPITRE 1 : NOTATIONS ET DEFINITIONS

Avant de commencer la présentation du processus de construction d'un modèle, il nous faut définir les concepts que nous utiliserons. Ils seront basés sur des exemples isolés tirés de notre (futur) modèle.

1.1 LE MODELE CONSIDERE COMME UN ENSEMBLE D'EQUATIONS

D'une manière générale, un modèle sera défini comme un ensemble de formules complètement définies qui décrivent les liens entre un ensemble de concepts.

Formellement, un modèle peut être spécifié comme un vecteur de fonction de variables.

$$f(\dots) = 0$$

Nous examinerons tour à tour :

- La nature des éléments figurant dans la fonction.
- La nature de la fonction elle-même.

1.2 LES ELEMENTS D'UN MODELE

1.2.1 LES VARIABLES: ENDOGENES ET EXOGENES

De toute évidence, un modèle sera utilisé pour obtenir une mesure de concepts économiques, en fonction d'autres éléments.

Deux types de variables apparaissent dans un modèle :

- les variables endogènes, ou de résultats, dont la valeur sera obtenue en résolvant le système d'équations ;
- les variables exogènes, ou d'hypothèses, dont la valeur est déterminée par des considérations extérieures, et qui conditionnent la solution.

Si le modèle est résolu sur les périodes passées, cette valeur est en principe connue. Mais dans les tâches de prévision, elle devra être choisie par le constructeur du modèle (ou son utilisateur).

Pour que le système puisse être résolu, le nombre de variables endogènes doit correspondre au nombre d'équations.

Notre formulation est la suivante:

$$f(x, y) = 0$$

avec

- x vecteur des variables exogènes
- y vecteur des variables endogènes (avec la même dimension que f).

Par exemple, dans notre modèle:

- Les importations seront endogènes, car elles dépendent de la demande locale. Les exportations également, fonction de la demande mondiale.
- La demande mondiale sera exogène, car nous construisons un modèle pour un seul pays, et nous allons négliger l'impact des variables locales sur l'économie mondiale. Bien sûr, cet impact existe, car la France est (encore) un pays important, et sa croissance a une certaine influence sur l'économie mondiale. Mais l'amélioration correspondante, relativement limitée, ne peut être obtenue qu'au prix très élevé de la construction d'un modèle mondial. Cette simplification serait moins acceptable pour un modèle des États-Unis, de la Chine ou de l'Union Européenne dans son ensemble (nous aborderons cette question en détail plus tard).

Techniquement, on peut contester le fait que les exportations sont endogènes pour notre modèle. Comme nous avons décidé de ne la faire dépendre que de la demande mondiale exogène, elles sont *de facto* prédéterminées, à l'exception d'un résidu statistique inconnu. Elles sont logiquement exogènes et techniquement endogènes. Nous allons utiliser le second critère (le problème va disparaître avec l'introduction d'autres éléments explicatifs, ce qui rendra les exportations réellement endogènes).

Quant à la demande de l'Etat, les modèles de ce type les garderont aussi exogènes, mais pour des raisons différentes:

- L'objectif de tout modèle est de montrer à son utilisateur (qui peut être le gouvernement ou un conseiller du gouvernement, un économiste indépendant qui joue le rôle du gouvernement ou un étudiant à qui l'on a posé une question d'économie appliquée) les conséquences de ses décisions. Ces décisions doivent donc être laissées libres, et non imposées à celui-ci.

- Le comportement de l'État est presque impossible à formaliser, car il a seulement quelques cibles (principalement la croissance, l'inflation, le chômage, les soldes budgétaire et commercial) et un nombre beaucoup plus élevé d'instruments. Si leurs valeurs de base sont plus ou moins fixes, il peut s'en écarter arbitrairement, assez rapidement. Pour atteindre le même objectif, les gouvernements français successifs ont utilisé différentes approches globales, et faisant appel à des ensembles d'instruments complexes¹.
- L'État seul a assez de pouvoir individuel pour influencer de manière significative l'économie nationale.

Chacun des deux types d'éléments exogènes est caractéristique d'une catégorie plus large:

- Les variables considérées comme extérieures au champ modélisé, sur lesquelles les agents économiques pris en compte par le modèle ont peu ou pas d'influence. En dehors de la situation dans d'autres pays, cela peut représenter la population², ou les conditions météorologiques³, ou la surface disponible pour l'agriculture. Le cadre théorique du modèle peut aussi supposer exogènes certains éléments structurels, tels que le taux d'intérêt réel, l'évolution de la productivité des facteurs, ou le taux de dépréciation du capital.
- Les variables contrôlées par un agent identifié, mais dont le modèle ne décrira pas le processus de décisions. Même si c'était formellement possible, le constructeur du modèle veut maîtriser leur valeur, et mesurer leurs conséquences sur l'équilibre économique. Celles-ci seront appelées variables de décision ou «instruments».

La modification des hypothèses concernant ces deux types de variables, par conséquent, se rapporte à des questions d'esprit très différent:

- Qu'est-ce qui se passera si ... peut-être ? (Le prix du pétrole augmente brusquement.)
- Qu'est-ce qui arrivera si je (l'État) décide ... ? (de diminuer le taux de TVA sur les produits de consommation courante⁴)

¹ Par exemple, pour réduire le chômage, un gouvernement peut accroître la demande ou réduire les impôts des entreprises, et l'instrument fiscal peut être les cotisations de sécurité sociale, ou les subventions.

² Dans les modèles de long terme la croissance est susceptible d'affecter la mortalité et les naissances donc la population.

³ Qui peut dépendre de la croissance (effet de serre).

⁴ Pour autant que la Commission européenne le permette.

1.2.2 LES EQUATIONS

Les liens entre les variables sont traités par des formules mathématiques appelées «équations». On en distingue deux types: équations de comportement et identités. Nous allons d'abord décrire leur rôle. Nous traiterons plus tard de leurs caractéristiques.

1.2.2.1 Équations de comportement

Le premier rôle de notre modèle est de décrire des «comportements»: le constructeur du modèle, appliquant la plupart du temps une théorie économique courante, définira une forme fonctionnelle décrivant le comportement d'un agent donné, et utilisera l'économétrie pour choisir une formulation précise, comportant des paramètres estimés.

Pour décrire la consommation, on peut supposer que sa part dans le revenu des ménages est déterminée par:

- leur niveau de revenu (un revenu plus élevé rendra moins attrayante ou nécessaire la consommation, par rapport à l'épargne)⁵ ;
- les variations récentes de leur revenu (les consommateurs mettent du temps à s'adapter à leur nouvelle situation) ;
- l'évolution du chômage: s'il se développe, la perspective de perdre leur emploi amènera les ménages à accroître leurs réserves ;
- l'inflation: elle définit la contribution nécessaire pour maintenir le pouvoir d'achat de l'épargne financière accumulée jusqu'ici.

Une fois identifiés, tous ces éléments seront réunis dans une formule, ou plutôt un ensemble de formules possibles (les ménages peuvent considérer l'inflation actuelle, ou la moyenne au cours de la dernière année, l'augmentation du chômage peut être représentée par son niveau ou son pourcentage de variation). Ces formules seront confrontées aux données disponibles, pour identifier une spécification statistiquement acceptable dans son ensemble, chaque élément devant participer de façon significative à l'explication, et les valeurs des coefficients obtenus être cohérents avec la théorie économique. Une fois les paramètres estimés, chaque élément de la formulation résultante contribuera à la logique de comportement de l'agent associé.

Mais le processus n'est pas toujours aussi simple. Deux autres cas peuvent être envisagés.

⁵ Rappelons que l'investissement en logement est considéré comme de l'épargne.

- Le comportement peut être formalisé, mais pas directement sous forme de formules prêtes à l'estimation. Un cadre doit d'abord être établi, puis traité par des transformations analytiques, y compris éventuellement des dérivations et maximisations, conduisant finalement à l'équation (ou l'ensemble d'équations) à estimer. Ce sera le cas pour notre fonction de production Cobb-Douglas pour laquelle nous calculerons la combinaison du travail et du capital qui maximise les bénéfices pour un niveau de production donné. Ou de la définition d'un taux de salaire résultant de négociations entre les syndicats de travailleurs et les gestionnaires des entreprises, en fonction de leur pouvoir de négociation respectif.
- Souvent, le constructeur du modèle ne sera pas en mesure de formuler précisément l'équation, mais considérera un ensemble d'éléments explicatifs potentiels, en attendant les diagnostics économétriques pour faire un choix définitif entre les formulations (en principe linéaires par rapport aux coefficients). Par exemple, le taux de change peut dépendre de la comparaison de l'inflation locale et étrangère, et du solde commercial.

Dans tous les cas, même si l'intensité exacte des influences est inconnue du constructeur du modèle⁶, la théorie économique définit généralement un intervalle de validité, et surtout un signe. Quelle que soit la qualité de l'explication statistique, elle sera rejetée si son signe ne correspond pas à la théorie. Dans l'exemple ci-dessus, l'augmentation de la demande de travail doit générer des gains de pouvoir d'achat du taux de salaire.

La formulation de ces équations théoriques fait souvent appel à des opérateurs spécifiques, permettant des calculs alternatifs: variables booléennes, opérateurs «maximum» et «minimum». Par exemple, dans les modèles de déséquilibre, l'équation théorique peut inclure une contrainte. On peut considérer aussi le cas d'une fonction de production à facteurs complémentaires, où le niveau de chaque facteur détermine une contrainte individuelle:

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

avec CAP la capacité de production, L l'emploi, K le capital, et pl et pk les productivités associées.

⁶ Sinon, il ne l'aurait pas estimée.

1.2.2.2 Identités

Un modèle composé uniquement d'équations de comportement n'est généralement pas utilisable en tant que tel. Des équations supplémentaires seront nécessaires, mais cette fois avec des formes incontestables.

Plusieurs cas peuvent être identifiés, qui peuvent s'appliquer simultanément:

- Certains concepts sont liés par une formule comptable, et nous devons veiller à leur cohérence numérique. Par exemple, une fois que le modèle a défini le revenu des ménages, on ne peut pas estimer l'épargne et la consommation séparément, car la somme des deux est connue⁷. Un seul élément est estimé: cela peut être l'épargne, la consommation, le taux d'épargne ou le rapport de la consommation au revenu, et les autres éléments vont suivre, en utilisant des identités.
- Certains concepts sont liés par une séquence causale d'éléments, et des éléments de la chaîne ne sont pas définis par des comportements. Par exemple, si nous estimons l'emploi des entreprises et la consommation des ménages, il faut formaliser les revenus des ménages (comme une somme incluant les salaires) pour que la création d'emplois améliore la consommation. Et dans notre exemple, la définition de la demande finale (comme la somme de ses composants) garantit que les importations suivront la consommation.

Bien sûr, on peut envisager d'éliminer ces identités en remplaçant chaque élément par la formule correspondante. Ce n'est pas toujours techniquement possible, mais en tout cas cela :

- conduit à des formulations trop complexes, difficiles à interpréter et plus longues à calculer ;
- élimine des informations potentiellement intéressantes.

En outre, on sera amené à introduire:

- des variables intermédiaires simplifiant les formulations (et en accélérant les calculs). Même si le taux de croissance du taux de salaire réel, qui utilise une expression un peu complexe, n'a pas été considéré comme un élément intéressant sur le plan économique, il sera utile de le définir, s'il apparaît comme un élément explicatif dans de nombreuses équations.

⁷ Ce serait également absurde en termes de comportement des ménages.

- Les éléments purement descriptifs: le ratio du solde budgétaire au PIB est un élément crucial dans l'évaluation de la santé financière de l'État (et l'un des «critères de Maastricht» pour l'entrée dans l'Union Monétaire Européenne).

Enfin, la théorie économique n'est pas toujours absente de ce type d'équation: l'équilibre offre-demande doit être respecté:

$$Q \text{ (offre des producteurs locaux)} + M \text{ (offre étrangère utilisée par le pays)} \\ = FD \text{ (demande des agents locaux)} + X \text{ (demande étrangère au pays).}$$

Et le choix de la variable qui équilibre aura une forte incidence sur les propriétés du modèle théorique.

Si les exportations et les importations proviennent de comportements, et si la demande est la somme de ses composantes, nous devons calculer Q comme:

$$Q \text{ (offre locale)} = (FD-M) \text{ (demande locale fournie par les producteurs locaux)} + X \text{ (demande étrangère fournie par les producteurs locaux)}$$

Cela signifie que la production va s'adapter à la demande (qui elle-même peut dépendre de la disponibilité des produits).

Mais on peut aussi supposer que:

- Les producteurs ont choisi de limiter leur production à un niveau inférieur à la demande, parce qu'une production plus élevée générerait des bénéfices marginaux négatifs. Dans ce cas, Q sera fixé, et nous pourrions avoir:

$$Q = \text{fixé}, X = f(\text{demande étrangère}), FD = f(\text{économie}), M = FD - Q + X$$

- Ou le pays ne peut importer qu'en devises, qu'il obtient par le biais des exportations.

$$X = f(WD), M = f(X), Q = \text{fixé}, FD = Q + (M - X)$$

1.2.3 LES PARAMETRES

Les paramètres peuvent être définis comme des scalaires dont la valeur peut varier d'une simulation à l'autre. La seule différence formelle avec des variables exogènes est qu'ils n'ont pas de dimension temporelle⁸.

Deux types de paramètres peuvent être pris en considération, en fonction de la manière dont leur valeur est établie:

- Ceux qui sont estimés par référence au passé: à partir d'une formule théorique, bien définie, mais comprenant des paramètres inconnus, le constructeur du modèle cherchera les valeurs qui rendent la formulation la plus proche de la réalité observée, selon une certaine distance. Cela signifie que l'on applique l' «économétrie».
- Ceux qui sont décidés par le constructeur du modèle: la théorie économique peut fournir des hypothèses *a priori* sur un comportement particulier. Par exemple, si une banque centrale décide de son taux d'intérêt réel selon une règle de Taylor, celle-ci dicte pour la sensibilité au niveau de l'inflation une intensité de 0,5. Un cas particulier est représenté par une variable de commande, qui donne (sans modification de la formulation) le choix entre plusieurs types de comportements séparés.

La distinction n'est pas aussi claire que 'on pourrait croire: en particulier, si l'estimation ne fournit pas un résultat économiquement cohérent, le constructeur du modèle peut être amené à intervenir sur les valeurs de certains paramètres.

Avec a comme vecteur de paramètres (à estimé), le système devient:

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

Et dans notre exemple, on peut estimer l'influence de la demande mondiale sur les exportations, par exemple en supposant que les variations relatives sont proportionnelles (ou de façon équivalente que l'élasticité des exportations à la demande mondiale est constante).

⁸ Dans EViews, modifier une valeur de paramètre s'applique au modèle courant, et en tenir compte réclame une nouvelle compilation qui le mettra à jour. Ceci est à la fois fastidieux et source d'erreurs. On pourrait envisager de remplacer les paramètres par des séries avec une valeur constante, ce qui facilitera la gestion des scénarios.

$$\Delta X / X = a \cdot \Delta WD / WD$$

où a devrait être proche de l'unité, si la part du pays sur le marché mondial est stable⁹.

Mais si le coefficient estimé n'est pas significatif, nous pouvons revenir à :

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

Ce choix aurait également pu être fait dès le départ pour des raisons théoriques.

Clairement, pour estimer un paramètre, il est nécessaire de définir entièrement la formule associée.

1.2.4 LE TERME ALEATOIRE

Dans la pratique, le comportement des agents ne répond pas exactement à une équation spécifiable, et la formulation obtenue par estimation de reproduira pas vraiment la réalité. Elle ne fera que s'approcher de ce comportement, en utilisant des éléments conformes à une théorie économique, chacun d'eux apportant une contribution suffisamment importante à l'évolution de la variable expliquée. Le nombre de paramètres estimés alors sera généralement beaucoup plus faible que la taille de l'échantillon, ou le nombre de valeurs observées. Dans la pratique, l'ajout d'éléments à l'explication peut :

- Dans les bons cas, améliorer la qualité de l'explication donnée par les éléments déjà présents, qui peuvent maintenant se concentrer sur leur rôle naturel, au lieu de chercher à participer à l'explication d'autres mécanismes où leur efficacité est limitée¹⁰.

⁹ Dans notre modèle WD désigne le commerce mondial (y compris son expansion), et non la demande globale des pays.

¹⁰ Comme un travailleur qui doit consacrer son temps à deux tâches, et est vraiment qualifié pour une seule. Par exemple, si un excellent musicien, mais parolier moyen est associé avec un bon parolier, la qualité des chansons (paroles et musique à la fois) va s'améliorer.

- Mais l'élément nouveau peut rivaliser avec les autres dans l'explication d'un mécanisme pour laquelle ils ont tous une certaine compétence, ce qui limite l'amélioration et laisse le partage de l'explication plutôt indéterminé (et donc limite la significativité des coefficients)^{11, 12}.
- Dans la pratique, ces problèmes de corrélation apparaissent toujours, parfois très tôt, et généralement avant le cinquième ou le sixième élément. Au-delà de ce chiffre l'identification des coefficients individuels diminuera généralement, et la qualité globale s'améliorera de moins en moins.

Cela signifie qu'une équation économétrique typique contiendra un maximum de quatre paramètres, même si les variables sont connues sur 50 à 100 trimestres.

Il sera donc nécessaire, pour formuler un modèle exact, d'accepter la présence de termes non nuls supplémentaires (les résidus). Si l'on croit au modèle, ce résidu doit être interprété comme une perturbation aléatoire sans signification économique. Mais si l'équation est mal spécifiée, il pourra également provenir d'autres sources: l'omission d'une variable, son remplacement par une autre moins pertinente, le choix d'une forme d'équation inadéquate¹³.

La faute ne reposera pas toujours sur le constructeur du modèle, qui pourrait ne pas avoir été en mesure d'appliquer ses idées de départ. Les variables dont il a besoin peuvent ne pas être mesurées avec précision, ou seulement avec une définition légèrement différente, ou elles peuvent ne pas être disponibles du tout, comme par exemple les objectifs ou les anticipations d'un agent donné.

En pratique, on va souvent supposer que ce résidu suit une distribution aléatoire, avec une moyenne nulle, un écart-type constant, et des résidus indépendants entre périodes.

Notre formulation devient donc, dans le cas général, en notant u le vecteur des résidus:

¹¹ Ceci peut être un problème pour le modèle si les deux éléments concurrents ont une sensibilité différente à une variable particulière. Par exemple, si l'un est sensible à un taux d'imposition, l'autre non: alors le rôle du taux d'imposition risque d'être mal déterminé.

¹² Si deux travailleurs ayant le même profil accomplissent une tâche ensemble, il est difficile d'évaluer leur contribution individuelle.

¹³ Bien sûr, comme nous l'avons déjà dit, on n'est jamais en mesure d'estimer la « vraie » équation. Cette remarque doit s'appliquer à une erreur conceptuelle importante, qui conduit à des comportements très différents d'une approximation acceptable de la réalité.

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

Dans l'exemple, si nous voulons représenter les variations de la consommation des ménages comme un pourcentage constant des variations de la production totale, on écrira :

$$CO = a \cdot Q + b + u$$

ou plutôt, si l'on veut avoir une influence relative constante :

$$CO/Q = a + u$$

1.2.4.1 Résidus et erreurs

C'est probablement le moment de présenter un point important concernant la nature de l'économétrie.

Quand il considère une équation de comportement, l'économiste peut avoir deux positions extrêmes.

- Il croit que le comportement peut être exactement défini selon une formule, qui est affectée par un terme d'erreur avec une distribution donnée (peut-être un bruit blanc, ou une loi normale). Avec un nombre infini d'observations nous obtiendrons une mesure exacte des paramètres, et de la distribution de l'erreur.
- Il pense que le concept qu'il veut décrire est bien lié à d'autres éléments économiques, mais la relation n'est qu'une application, dont toute formule représente au mieux une approximation. À cette application, un terme aléatoire peut également être ajouté, si l'on croit que la répétition des mêmes éléments explicatifs apportera un résultat différent. Des observations supplémentaires peuvent seulement fournir une meilleure topologie.

Le débat est rendu plus complexe par plusieurs faits :

- Les données sur lesquelles il veut baser son estimation ne sont pas mesurées correctement. On ne peut pas s'attendre à ce que les statisticiens produisent une information parfaitement exacte, pour de nombreuses raisons : erreurs de mesure, d'échantillonnage, de concepts...
- Même si les séries sont mesurées correctement, les concepts qu'elles vont utiliser ne sont pas nécessairement les bons. Par exemple, un comportement donné doit être appliqué aux seules

entreprises qui font des profits, ou une séparation pertinente (par exemple la taille des entreprises) n'est pas disponible au niveau macroéconomique.

- Les décalages discrets qu'il appliquera à ces concepts ne sont pas les bons non plus. Par exemple, on pourra savoir que l'agent tient compte de l'indice des prix du mois dernier, mais seules des données trimestrielles sont disponibles.
- La période d'estimation n'est pas homogène, et cela ne peut pas être expliqué par les données. Par exemple, l'état d'esprit des consommateurs (et leur comportement de consommation) peut évoluer sans aucun lien avec des éléments mesurables économiquement¹⁴.

À partir des éléments ci-dessus, la conclusion logique devrait être:

- La première position est illusoire, et à un point qui est impossible à mesurer (bien sûr).
- Mais nous devons la prendre si nous voulons appliquer des méthodes économétriques.

Cela signifie que, dans le texte qui suit, nous allons adopter la première position, mais nous garderons toujours à l'esprit la situation réelle, et donnerons à la différence entre le concept et son estimation le nom moins ambitieux de «résidu».

1.3 LES FORMULATIONS

Nous allons maintenant examiner la forme des équations. Examinons d'abord la dimension temporelle.

1.3.1 LA DIMENSION TEMPORELLE

Les variables des modèles économiques ont généralement une dimension temporelle, ce qui signifie qu'elles sont connues par des valeurs discrètes, presque toujours avec une périodicité constante: généralement annuelle, trimestrielle ou mensuelle. Cela signifie que nous allons étudier nos différents modèles en temps discret.

Il y a toutefois des exceptions. Les plus fréquentes s'appliquent aux modèles microéconomiques décrivant par exemple le comportement d'un panel d'entreprises ou de ménages, et la dimension correspondra à des éléments d'un ensemble. Parfois, ils vont être ordonnés selon le niveau d'une variable, comme le niveau de revenu pour un ensemble de ménages. Le temps peut être présenté comme une dimension supplémentaire, mais peut-être avec un intervalle variable, soit prédéterminé (phases de la lune) ou imprévisible (les périodes de froid intense).

¹⁴ Par exemple, si les citoyens réagissent à une privation de libertés, ou craignent pour leur sécurité.

1.3.1.1 Conséquences de la discrétisation

La discrétisation temporelle des variables sera introduite de plusieurs façons, ce qui conduit à définir trois types:

- les variables vraiment instantanées, mesurées à un point donné dans le temps: dans un modèle annuel, le niveau du capital productif le 31 décembre à minuit (défini comme une variable de stock) ;
- les moyennes: le niveau moyen de l'emploi observé au cours d'une période ;
- les flux: la quantité de marchandises produites au cours d'une période.

Le même concept économique peut apparaître sous plusieurs formes: l'inflation et le niveau des prix, la dette et le déficit pour la période, la moyenne et les niveaux d'emploi en fin de période, et les créations nettes d'emplois. Pour un ménage, on peut considérer le revenu, sa variation annuelle, et le revenu global accumulé depuis sa création.

1.3.1.2 La saisonnalité

Lorsque les modèles ont une périodicité inférieure à l'année, certaines séries peuvent présenter une distorsion spécifique en fonction de la sous-période considérée. Cela peut provenir du climat: en hiver, la consommation d'électricité va augmenter avec le chauffage et l'éclairage, mais les travaux du bâtiment seront la plupart du temps arrêtés. Cela peut être dû à des comportements humains: la concentration des vacances dans les mois d'été peut réduire la production, et l'approche de Noël augmentera la consommation (dans les pays chrétiens). Nous allons ici fournir un schéma de base du problème, laissant une description plus précise aux ouvrages spécialisés, comme Ladiray et Quenneville (2001).

Partir des données brutes peut entraîner des problèmes: par exemple, le niveau de production en été sera inférieur à ce que nous pouvions attendre de la main-d'œuvre et des capitaux installés. Cela perturbera les estimations et donnera aux modèles des solutions plus difficiles à interpréter.

Deux solutions peuvent être envisagées:

- introduire dans les équations estimées des variables «muettes» associées à chaque sous-période ;
- extraire de la série leur composante saisonnière pour produire un nouvel ensemble de valeurs («désaisonnalisées»).

Bien sûr, il ne faut pas mélanger les deux types de techniques dans la même équation (ou modèle).

La deuxième méthode sera privilégiée, car elle résout également le problème d'interprétation.

Plusieurs techniques sont disponibles, la plus connue étant Census-X12-ARIMA, développé par le US Census Bureau et Statistique Canada¹⁵. Mais Tramo-Seats¹⁶ est également un choix courant. Les deux méthodes sont disponibles sous EViews.

Il faut être conscient que ce processus réduit souvent la qualité statistique des estimations. Par exemple, si la demande est particulièrement forte dans le dernier trimestre de chaque année, et les importations suivent, désaisonnaliser les deux séries rendra le lien moins évident, conduisant à des résultats moins précis. De façon encore plus claire, la relation entre la demande de chauffage et la température va perdre de son explication.

Ces exemples mettent en évidence la question principale: dans un modèle à une seule équation, la transformation est essentielle si l'explication intègre une composante saisonnière, en plus des caractéristiques vraiment économiques. Par exemple, la production agricole sera plus faible en hiver, même si le même niveau de travail, la terre, les engrais, les machines est disponible ... En vérité, dans le même temps, l'utilisation d'engrais diminue, et probablement celle du travail aussi, mais à un moindre degré. Cela signifie que l'ensemble des variables, aussi bien expliquées et explicatives, doit être corrigé des variations saisonnières (mais surtout la première).

Au contraire, si la seule justification de la saisonnalité vient des éléments explicatifs, la désaisonnalisation n'est pas nécessaire, et réduit même la qualité des estimations (avec la variabilité des éléments). On pourra utiliser des séries brutes pour estimer une équation d'importations, en utilisant la demande, le taux d'utilisation des capacités et la compétitivité des prix comme éléments explicatifs.

Mais ce qui est vrai pour une équation ne s'applique pas à l'ensemble du modèle. On ne peut pas mélanger les deux types de séries, ce qui signifie que la désaisonnalisation prévaudra dans la pratique.

1.3.1.3 Modèles statiques et dynamiques

Pour déterminer l'équilibre pour une période donnée, certains modèles utiliseront uniquement les variables de cette période: nous les appellerons «statiques». Ils correspondent à la formulation:

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

¹⁵ <http://www.census.gov/srd/www/x12a/>

¹⁶ <http://www.bde.es/webbde/es/secciones/servicio/software/econom.html>

Le cas le plus fréquent est celui de modèles entrées-sorties (ou «input-output»), qui appliquent une matrice de coefficients techniques à une décomposition donnée de la demande en catégories de biens (qui ne dépend elle-même que d'éléments instantanés), pour calculer la production détaillée associée.

$$Q = A \cdot FD$$

(A représentant une matrice carrée n par n)

$$FD = f(Q)$$

Au contraire, les modèles dynamiques utilisent des variables d'autres périodes.

Les raisons sont assez nombreuses. Elles peuvent être:

- théoriques: certains agents seront censés fonder leur comportement sur l'observation du passé. Les entreprises vont baisser leurs prix si la demande du trimestre précédent a été trop faible. Ou elles vont baser leurs anticipations de croissance de la demande sur les évolutions antérieures de la même variable. Ces deux exemples illustrent les principaux domaines d'application: utiliser le passé pour établir une image de l'avenir (anticipations), ou pour mesurer l'écart précédent entre valeurs réelles et cible, que l'agent va chercher à réduire au cours de la période.
- institutionnelles: l'impôt sur le revenu payé par les ménages peut être basé sur leur revenu de la période précédente (c'est le cas en France, pour le moment).
- techniques: si un modèle prend en compte à la fois une variable et son taux de croissance, le calcul de l'un à partir de l'autre fait intervenir le niveau précédent.

On constate que chacune de ces justifications suppose que les influences viennent seulement des périodes précédentes: on parlera d'influences retardées.

La formulation devient donc:

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

Revenons à notre modèle. Nous pouvons déjà observer une influence retardée incontestable: le niveau actuel du capital contiendra la partie non déclassée de son niveau précédent. Pour les autres comportements, la situation n'est pas encore claire. Cependant, sans aller trop loin dans la théorie économique, on peut penser à plusieurs directions possibles:

Pour la consommation des ménages, nous avons déjà considéré que l'adaptation à un niveau de revenu nouveau prend un certain temps. Cela signifie qu'elle dépendra des niveaux antérieurs. Si nous détaillons les produits, le niveau précédent peut avoir une influence positive (certaines consommations reposent sur l'habitude) ou négative (en général, les gens n'achètent pas une nouvelle voiture tous les trois mois):

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

Les entreprises investissent pour adapter leurs capacités au niveau de la production nécessaire dans le futur. On peut supposer qu'elles basent leurs anticipations sur les valeurs du passé.

$$I_t = f(Q_t, Q_{t-1}, Q_{t-2}, \dots)$$

Il est intéressant de noter que les formulations précédentes peuvent être simplifiées, ce qui élimine tout retard supérieur à un par l'ajout de variables intermédiaires:

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

(Où y_i et y_j représentent des variables, indexées par le temps t)

est équivalent à

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

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

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

.....

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

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

dans lequel un décalage de k périodes sur une seule variable a été remplacé par k décalages d'une période sur plusieurs variables (dont certaines sont nouvelles).

La même méthode permet clairement d'éliminer les variables retardées exogènes.

Sur l'équation d'investissement de l'exemple, cela donnerait:

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

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

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

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

Mais si cette méthode simplifie la formulation théorique, elle présente l'inconvénient évident d'augmenter artificiellement la taille du modèle et de réduire sa lisibilité, sans produire d'information supplémentaire. Son intérêt est réservé à des études spécifiques. Par exemple, l'évaluation de la dynamique du modèle peut faire appel à la linéarisation du modèle en fonction de variables instantanées et retardées. La transformation ci-dessus va limiter les matrices du système à deux éléments (avec des décalages 0 et 1), ce qui rendra les calculs formels plus faciles, et indépendants du nombre de retards.

Elle nous permet également d'utiliser une formulation simplifiée dans les présentations qui vont suivre:

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

1.3.1.4 Un cas particulier: les anticipations rationnelles

Il est apparu naturel, dans les exemples précédents, de ne considérer que des décalages négatifs. Ce sera le cas si l'on suppose que l'anticipation des agents ne repose que sur l'observation du passé (et du présent)¹⁷.

Pour justifier des formulations à retards positifs, il est nécessaire de supposer:

- que les agents ont la possibilité, en tenant compte de leurs décisions actuelles, de déterminer les valeurs futures de certaines variables (et que le comportement associé peut être formalisé) ;
- que les agents anticipent parfaitement l'avenir (anticipations parfaites) ;
- que l'attente par les agents d'évolutions spécifiques a pour conséquence la réalisation de ces évolutions (anticipations auto-réalisatrices) ;
- que les agents basent leurs anticipations sur les comportements des autres agents, dont ils peuvent déterminer les réalisations (anticipations rationnelles). Fondamentalement, cela signifie qu'ils sont capables d'appliquer (sans forcément le connaître) le modèle régissant l'économie, et le processus de décision définissant ses hypothèses. Par exemple, ils peuvent prévoir le programme d'investissements du gouvernement (en fonction des conditions économiques), ils peuvent déterminer comment les entreprises et les ménages vont réagir, et ils connaissent les liens entre ces éléments (ils sont capables de tenir compte de l'équilibre offre-demande).

Cependant, ils ne connaissent pas nécessairement la partie inexplicée des comportements (qui peut être associée au terme aléatoire). S'ils ne le font pas, mais s'ils connaissent leur distribution, nous parlerons d'anticipations rationnelles stochastiques. EViews ne fournit pas cette fonction à l'heure actuelle (seulement l'une ou l'autre).

Ils n'ont pas besoin de connaître les formules réelles, seulement être capables de les calculer.

Vous n'avez pas besoin de croire aux anticipations rationnelles pour les appliquer. Produire des simulations alternatives utilisant différentes hypothèses sur les anticipations permettra d'améliorer considérablement la connaissance d'un modèle particulier et des mécanismes économiques en général. Nous présenterons ceci plus loin sur un exemple.

¹⁷ Cette utilisation est rendue nécessaire par l'absence de mesure directe des anticipations. Exceptionnellement, elles peuvent être obtenues par des enquêtes, conduisant à une estimation spécifique.

1.3.1.5 Un autre cas: les modèles en temps continu

Ces modèles évolueront rarement au-delà d'un stade théorique, si ce n'est que par manque d'informations statistiques.

Mais certains modèles opérationnels, décrivant par exemple les opérations boursières, peuvent réduire leur périodicité à un intervalle quotidien ou bien plus court.

1.3.2 LA LINEARITE

Nous considérons ici la linéarité par rapport aux variables. La linéarité par rapport aux coefficients apparaîtra dans le chapitre sur l'estimation.

La linéarité éventuelle d'un modèle représente une propriété très importante pour sa résolution et son utilisation. Mais nous devons d'abord définir la notion de linéarité, qui peut être plus ou moins stricte. Nous considérerons quatre cas.

La linéarité la plus forte sera représentée par:

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

mais on peut autoriser les éléments des matrices à évoluer avec le temps:

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

Une définition encore moins restrictive supposera la linéarité par rapport aux variables endogènes seules:

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

ou même par rapport aux endogènes de la période:

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

C'est aussi un domaine très spécifique: certains modèles théoriques seront formulés comme un système d'équations où les variables apparaissent en fonction d'un temps continu, et les variations (ou taux de croissance) deviennent des dérivées exactes. On aboutit alors à un système d'équations différentielles, que l'on peut être amené à intégrer.

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

En utilisant le multiplicateur à titre d'exemple, on peut déjà montrer que ces propriétés affectent le calcul de la dérivée des solutions d'un modèle. Nous détaillerons plus tard les conséquences sur les propriétés de convergence.

- La première propriété indique que cette dérivée ne dépend pas de l'équilibre initial, ou de la période considérée. En multipliant le choc par un facteur donné, le résultat évoluera proportionnellement. Il suffit de le calculer une fois pour le connaître définitivement.
- Dans le deuxième cas, le multiplicateur ne dépendra que de la période. Un changement des hypothèses de départ n'affectera pas les conséquences de leur variation.
- Dans le troisième cas, le multiplicateur dépend seulement des valeurs exogènes (et des coefficients). Il doit être recalculé chaque fois que ces éléments changent (ou ont changé dans le passé).
- Le dernier cas est semblable au troisième. Mais la convergence sera affectée (voir plus loin).

1.3.2.1 Cas pratiques de non-linéarité

Il est assez clair que la présence d'une seule équation non linéaire rend le modèle non-linéaire, selon l'une des définitions précédentes. Les raisons de la non-linéarité sont multiples ; on trouvera en particulier :

- des expressions mesurées en taux de croissance (donc éventuellement linéaires par rapport aux endogènes de la période). Par exemple, le taux de croissance des salaires peut dépendre de l'inflation ;
- des expressions formulées comme des élasticités (généralement intégrées en logarithmes). On supposera par exemple que les importations et la demande intérieure présentent des variations relatives proportionnelles ;
- les ratios intervenant dans les équations de comportement ;
- les équations utilisant des éléments à prix courants, calculés comme le produit d'une quantité par un déflateur (qui traduit l'évolution du prix par rapport à une année de référence). Par exemple, le solde commercial est obtenu comme la différence entre les produits des exportations et des importations à prix constants par leurs déflateurs respectifs ;

Cette liste n'est pas exhaustive, mais elle représente les cas les plus courants. La fonction quadratique, très présente dans les calculs mathématiques et d'optimisation, l'est rarement dans les modèles économétriques. (ou alors par sa dérivée linéaire).

Parfois, cette distinction est purement formelle, et un changement de variable adéquat permettra le retour à une formulation linéaire. Cependant, si l'on prend en compte l'ensemble du modèle, remplacer par son logarithme une variable calculée en élasticités ne fera que transférer le problème si le niveau apparaît également dans le modèle.

Ainsi, dans notre exemple général, si l'on utilise pour l'équation des exportations de la formulation :

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

On peut très bien introduire les variables $LX = \text{Log}(X)$ et $LWD = \text{Log}(WD)$, ce qui rendra l'équation linéaire :

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

Mais il sera nécessaire d'introduire les exportations dans l'équilibre de l'offre-demande :

$$Q + M = FD + X$$

en ajoutant l'équation non linéaire

$$X = \text{Exp}(LX)$$

Par conséquent, la plupart des modèles économiques présentant un minimum de réalisme ne seront pas linéaires. Mais les calculs numériques montrent en général que, même pour les modèles contenant de nombreuses formules non linéaires, l'approximation par une forme linéarisée autour d'une solution du modèle (signalée ici par un astérisque) :

$$\left(\frac{\partial f}{\partial y_t}\right)(y_t - y_t^*) + \left(\frac{\partial f}{\partial y_{t-1}}\right)(y_{t-1} - y_{t-1}^*) + \left(\frac{\partial f}{\partial x_t}\right)(x_t - x_t^*) = 0$$

est acceptable en général.

D'autre part, la stabilité des dérivées au cours du temps est beaucoup plus discutable.

Supposons que l'équation d'importations ait la forme suivante:

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

La linéarisation autour d'une solution particulière (notée *), nous donne

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

ou

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

ce qui représente une bonne approximation linéaire de la relation entre M et FD , à condition que M et FD ne s'éloignent pas trop de leur valeur de base¹⁸. Celle-ci peut représenter une trajectoire de référence, dont les valeurs simulées diffèrent en raison d'une modification des hypothèses.

Mais, si l'on restreint davantage l'expression à une influence constante (linéarité à coefficients constants),

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

l'approximation ne peut être acceptée que si le rapport M / FD n'évolue pas trop avec le temps. Cela n'est pas vrai en général: l'expansion du commerce international a conduit et conduit encore à une croissance soutenue de la part des importations dans la demande intérieure, pour la plupart des pays. Le rapport M^* / FD^* va augmenter régulièrement, et la dernière formulation est tout à fait inadaptée pour les prévisions.

1.3.3 AUTRES PROPRIETES

1.3.3.1 Continuité

Nous considérons ici la continuité de l'ensemble des variables endogènes par rapport aux hypothèses (variables exogènes, paramètres). Elle n'est presque jamais vérifiée formellement, mais ne doit être considérées que l'ensemble des solutions (et hypothèses) acceptables.

Par exemple, la plupart des modèles utilisent des ratios, ce qui est acceptable si le dénominateur ne peut jamais devenir nul (comme la productivité du travail mesurée par le ratio de la production à l'emploi). Ou l'utilisation des logarithmes pour relier les importations à la demande exige (logiquement) que ces éléments sont strictement positifs. Et un modèle entièrement linéaire peut produire un PIB négatif, mais cela ne l'invalide pas si de telles valeurs sont associées à des hypothèses ou des paramètres absurdes.

Ainsi, même si tous les modèles présentent des discontinuités potentielles, elles ne devraient jamais se produire dans la pratique. Nous pouvons imaginer que trois cas possibles:

- Le cadre du modèle est correct, mais quelque chose ne va pas dans ses éléments: les hypothèses numériques, les coefficients estimés.
- L'algorithme utilisé pour la résolution du modèle conduit à des valeurs absurdes (voir plus loin).
- Les équations de comportement sont mal spécifiées. Comme nous le verrons plus loin, il peut être dangereux d'établir ces éléments sans une évaluation préalable des mécanismes associés (il est par exemple dangereux d'utiliser systématiquement les logarithmes comme une solution naturelle).

Il est nécessaire, cependant, de distinguer ces cas de ceux où la discontinuité s'applique à la dérivée d'une équation différentiable par morceaux, comme nous allons le voir dans le paragraphe suivant.

1.3.3.2 Différenciabilité

Elle est moins nécessaire, mais son absence peut conduire à des problèmes dans la phase de résolution, ainsi que dans l'interprétation des résultats.

La séparation des critères précédents n'est pas toujours simple, car la non-dérivabilité d'une variable peut correspondre à la discontinuité d'une autre: une productivité marginale discontinue peut rendre la production associée non dérivable aux points de discontinuité.

Pour revenir à notre exemple, on pourrait définir la consommation des ménages de la manière suivante:

- Ils reçoivent une part constante a de la production Q .
- Au-delà d'un seuil de revenu $-R$ – ils consomment une part c_0 .
- Sur le supplément ils consomment une part c_1 .

L'équation de la consommation deviendra:

$$CO_i = c_0 \cdot a \cdot Q_i + (c_1 - c_0) \cdot \max(0, (a \cdot Q_i - R_i))$$

Au point $Q = R / a$, CO n'est pas dérivable (la dérivée à gauche est $c_0 \cdot a$, à droite $c_1 \cdot a$). Et la sensibilité de la consommation au revenu n'est pas continue.

Cette dérivée n'est pas purement formelle: elle définit la propension marginale à consommer (consommation associée à une augmentation du revenu unitaire), qui peut elle-même apparaître dans le modèle, au moins comme un élément descriptif.

Dans le compte des ménages, l'évolution de l'impôt sur le revenu en fonction de celui-ci (avec des taux associés à des tranches) représentera un autre exemple, dans la détermination du revenu disponible des ménages.

1.3.3.3 Existence d'une solution

Il est évidemment nécessaire que le modèle ait une solution, du moins quand lui sont proposées des hypothèses acceptables. Mais l'absence potentielle d'une solution est présente dans de nombreux systèmes formels, y compris pour les modèles linéaires. Cette absence de solution est généralement équivalente à l'existence d'une solution absurde, que l'on peut illustrer sur le cas suivant.

Considérons un modèle à $n + 1$ variables endogènes: X (dimension n) et x (une seule variable). Nous le décrivons comme f , un vecteur de formules (dimension $n + 1$), où x apparaît comme un argument d'un logarithme,

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

Si aucune des valeurs positives de x n'assure la solution du modèle complet, il n'a pas de solution.

En d'autres termes, en prenant l'argument du logarithme en tant que paramètre

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

et en faisant varier α dans \mathbb{R}^+ , la résolution du modèle associé à x et X ne fournira jamais de valeur de x égale à ce paramètre.

Le modèle n'a évidemment pas de solution.

Donnons un autre exemple, plus fruste mais aussi plus clair. Le modèle:

$$\begin{aligned} y &= a \cdot x^2 \\ x &= \text{Log}(-y) \end{aligned}$$

n'a évidemment pas de solution.

Mais si le constructeur du modèle a utilisé une formulation en logarithmes, il n'a normalement pas considéré comme acceptable que l'argument prenne des valeurs négatives. En remplaçant le logarithme par une autre expression donnant des valeurs voisines, nous aurions probablement obtenu une solution. Mais si la variable était restée négative, cette solution aurait été inacceptable.

Pour illustrer ce cas, nous allons réduire le modèle habituel à une version à trois équations.

La production s'adapte à la demande corrigée par les importations et les exportations, cette dernière étant exogène:

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

Pour la demande, on suppose que ses variations relatives sont proportionnelles à celles de la production:

$$[2] \quad FD = a \cdot Q + FD_0$$

et les importations représentent une part de la demande

$$[3] \quad \text{Log}(M) = b \cdot \text{Log}(FD) + c$$

Supposons que l'on a obtenu par estimation sur le passé: $a = 1,10$, valeur justifiée par des importations croissant plus vite que la demande, évidemment associée à des importations croissant plus vite que les exportations.

Nous allons maintenant produire une prévision.

Le modèle peut être réduit en:

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

(de (1) et (3))

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

de (2)

et

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

De toute évidence, si Q augmente (comme $a - 1 = 0,05$), l'élément négatif deviendra finalement plus élevé que le positif, ce qui signifie que X / Q va devenir négatif, ce qui est impossible car il entre dans un logarithme de l'équation (2). Le modèle n'a pas de solution.

Bien entendu, ces observations mathématiques ont une contrepartie économique. À long terme, la demande finale ne peut pas croître continuellement plus vite que la production, si les importations sont une part de la demande et les exportations sont fixées. Les hypothèses, par conséquent, ne sont pas compatibles avec la formule estimée.

On remarquera que l'absence de solution est due ici à l'adoption d'une condition implicite vérifiée numériquement sur le passé, mais pas systématiquement. Ce sera dans la pratique le cas le plus fréquent.

1.3.3.4 Unicité de la solution

L'unicité de la solution, pour des valeurs raisonnables des paramètres et des hypothèses, est également très importante. En effet, nous ne voyons pas comment on pourrait utiliser un modèle qui laisse le choix entre plusieurs résultats, sauf peut-être si cette liberté a un sens économique précis.

Dans la pratique, la simple observation des équations montre que la plupart des modèles sont fortement non linéaires, mais que l'approximation linéaire est assez précise à l'intérieur du domaine des solutions économiquement acceptables. Cela limite la possibilité d'équilibres multiples.

Si le système était entièrement linéaire et les matrices associées régulières, il aurait en effet une solution unique. Cependant, comme nous nous éloignons de ce domaine, la quasi-linéarité disparaît, et nous ne pouvons pas éliminer la possibilité d'autres solutions, sans doute assez éloignées de la solution raisonnable et même probablement tout à fait absurdes. Heureusement, il est facile de commencer les calculs à l'intérieur d'un domaine «raisonnable» (par exemple en prenant les valeurs historiques, ou en prévision les valeurs de la solution obtenue pour la période précédente). Un algorithme efficace convergera alors vers l'équilibre acceptable, et nous ne soupçonnerons même pas l'existence des autres.

L'exception la plus importante sera celle des modèles d'optimisation, qui recherchent les valeurs des variables exogènes qui donnent le meilleur résultat pour un objectif donné (par exemple l'ensemble des réductions d'impôts qui produisent la plus forte baisse du chômage, pour un certain coût) : si plusieurs combinaisons de valeurs donnent un résultat égal en qualité¹⁹, cette indétermination ne compromettra

¹⁹ Par exemple, si le modèle est trop simple pour différencier le rôle des deux taxes.

pas le sens de la solution. L'existence de plusieurs (ou une infinité de) solutions représentera un diagnostic économique, qui devra être interprété comme tel²⁰.

Un autre cas où le problème apparaît est celui où la formule représente l'inversion d'une autre formule donnant une valeur constante, au moins sur un certain intervalle. Par exemple, si au-delà d'un certain seuil de revenu les ménages en épargnent la totalité.

$$CO = \min(f(Q), CO^*)$$

Le niveau de revenu associée à CO^* représentera l'ensemble complet des valeurs supérieures à ce seuil.

Dans le cas général, le principal danger des solutions multiples concerne les études de sensibilité: si l'on veut mesurer et interpréter les effets économiques d'une modification des hypothèses, l'existence d'une simulation de référence unique est une nécessité absolue.

Enfin, l'obtention de plusieurs solutions très proches les unes aux autres peut provenir de problèmes purement numériques, dus à l'imprécision de l'algorithme: tout élément de l'ensemble peut alors être accepté, si la différence est suffisamment faible.

1.3.3.5 Convexité (or concavité)

La convexité du système est définie par celle de l'évolution de chaque variable endogène par rapport à chaque variable exogène et paramètre pris individuellement (ou d'une combinaison linéaire d'entre elles). Elle peut être réclamée par certains algorithmes, en particulier en optimisation. En pratique, elle est très difficile à établir, et même rarement vérifiée. De toute façon, cette caractéristique est liée à la définition des variables, et un seul changement de variable peut la faire disparaître ou apparaître.

1.3.4 LES CONTRAINTES QUE LE MODELE DOIT RESPECTER

En plus de sa validité théorique, le modèle devra répondre à un ensemble de contraintes plus techniques.

²⁰ Pourvu que l'algorithme utilisé pour la résolution du modèle soit capable de gérer cette indétermination.

1.3.4.1 Compatibilité globale

Les contraintes de compatibilité porteront en pratique:

a - **Sur les endogènes entre elles**: on ne peut pas laisser le modèle calculer des variables indépendamment si elles sont liées par une relation logique, comptable ou théorique. Par exemple, si le prix à la consommation entre dans la détermination du taux de salaire, il devra également être directement influencé par le prix (estimé) de la production locale. Ou le niveau d'emploi doit affecter les revenus des ménages et leur consommation à travers une séquence de liens.

Les soldes comptables doivent être vérifiés: une fois le revenu des ménages calculé comme la somme de ses éléments, une augmentation de la consommation doit produire une diminution correspondante de l'épargne.

Le problème le plus important réside peut-être dans l'équilibre offre-demande, qui devra s'appliquer à prix constants et courants. Cela peut conduire soit à utiliser l'un de ses éléments pour équilibrer l'équation, ou à répartir le résidu global sur l'ensemble des éléments d'un côté. En formulant l'offre et la demande globales comme:

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

et

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

On pourra utiliser par exemple, soit

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

ou on corrigera l'ensemble de variables de la demande en multipliant chacun d'eux non corrigé par le rapport global O/D .

Dans la plupart des cas, l'équilibre en prix constants sera appliqué automatiquement. Il peut être écrit comme:

$$\text{Production locale} + \text{Importations} = \text{demande locale} + \text{exportations}$$

ou en identifiant les consommations intermédiaires:

$$\text{PIB local} + \text{consommations intermédiaires} + \text{Importations} \\ = \text{demande finale locale} + \text{consommations intermédiaires} + \text{exportations}$$

Avec un seul produit, la consommation intermédiaire peut être éliminée, et on utilisera généralement l'équation pour calculer le PIB, en vérifiant qu'il n'est pas plus élevé que la capacité de production²¹.

Avec plusieurs produits, il faut considérer des équations d'équilibre individuelles, dans lequel l'offre de biens de consommation intermédiaires représente les intrants nécessaires à la production du bien, et la demande de biens de consommation intermédiaires les utilisations intermédiaires du bien lui-même.

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

Si l'on suppose que les rendements d'échelle sont constants, on peut remplacer les consommations intermédiaires par l'application à la valeur ajoutée d'une matrice de coefficients techniques.

En définissant c_{ij} comme la quantité de bien i nécessaire pour produire une unité du bien j , nous obtenons:

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

Ou en termes matriciels

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

²¹ Cela peut être obtenu par une part des importations croissant avec des contraintes sur la capacité de production locale.

Ou encore

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

L'utilisation de ce cadre assure l'équilibre offre-demande pour tous les produits.

Dans la pratique, la plupart des problèmes viennent de l'équilibre en prix courants. Si les prix de demande sont calculés individuellement en utilisant des équations de comportement, il n'y a aucune chance que l'équilibre soit atteint. Le processus décrit précédemment permet dans la pratique de corriger les prix. En appelant S et D l'offre et la demande des éléments à prix constants, ps et pd les déflateurs associés, on peut calculer les valeurs globales par:

$$SV = \sum_{i=1}^n ps_i S_i$$

$$DV = \sum_{j=1}^m pd_j D_j$$

La première option calculera un prix spécifique

$$pd_m = (SV - \sum_{j=1}^{m-1} DV_j) / D_m$$

Et la seconde

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

où les « pd » sont les prix de la demande calculées indépendamment, et les « pd' » les valeurs corrigées.

Le facteur correctif:

$$r = SV / \sum_{j=1}^m pd_j D_j$$

peut aussi être écrit comme

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

ce qui avec

$$pd'_j = r pd_j$$

donne un ensemble d'équations assurant l'équilibre. Comme «r» mesure l'écart potentiel entre l'offre et la demande, il faut vérifier qu'il n'est pas trop différent de l'unité.

Les questions suivantes vont apparaître:

- Avec la première méthode, quel élément doit être utilisé pour équilibrer le système ? Le choix est entre
 - une variable d'importance faible, dans le but de réduire les conséquences pour les propriétés du modèle, ou peut-être même une variable qui n'a pas d'influence sur le reste du modèle,
 - une variable d'importance élevée, pour réduire le facteur de correction.
- La seconde méthode représente une application extrême de la première, où toutes les variables placées d'un côté sont affectées de manière proportionnelle.

En fait, aucune des solutions ne domine clairement, la pire étant quand même à notre sens la première, qui correspond à l'acceptation *de facto* d'un déséquilibre, caché, mais avec des conséquences potentiellement dommageables. En outre, la seconde pourrait être associée à un processus de convergence économique, tandis que la première ne peut avoir aucune interprétation économique que ce soit.

En fait, il faut surtout chercher à limiter l'ampleur de la correction elle-même. On peut assimiler le problème à l'élimination des déchets toxiques: au lieu de les stocker dans un emplacement spécifique

(caché ou non), ou de les disperser dans la nature, la meilleure solution est évidemment de réduire leur production autant que possible. Cela signifie en particulier que les prix individuels non corrigés doivent être conçus pour donner des valeurs de l'offre et de la demande globales naturellement proches.

b - Sur les liens exogènes -> endogènes. Ils doivent être formulés avec soin. Par exemple, si le taux des cotisations sociales est défini comme une variable exogène dans le modèle, il doit entrer dans tous les calculs des niveaux de cotisations. En particulier, il ne peut pas coexister avec une représentation exogène des cotisations, ou utilisant un coefficient estimé.

Pour éviter ce type d'erreur, une étude systématique des propriétés du modèle doit être effectuée avant toute mise en œuvre opérationnelle: dans notre exemple, cela correspond à vérifier que l'augmentation du taux de cotisations sociales a tous les effets attendus sur les recettes de l'État ainsi que sur les comptes et les comportements des autres agents .

En outre, le vrai concept exogène doit être décidé au départ. En ce qui concerne les cotisations, la variable de décision est clairement son taux, tandis que le revenu associé est influencé par les prix endogènes et l'emploi.

c - Sur les exogènes entre elles: il faut éviter de définir séparément deux variables exogènes si elles sont liées (dans n'importe quelle direction) par une relation logique. Si possible, il faut endogénéiser l'une d'elles en formalisant ce lien.

Supposons par exemple qu'un modèle construit pour la France utilise deux mesures exogènes des prix établis par ses concurrents étrangers: en devises et en euros (avec un taux de change fixe). Pour tenir compte d'une augmentation de l'inflation étrangère, ces deux variables doivent être modifiées simultanément. Ceci introduit une complexité, et peut conduire à des erreurs si l'on n'est pas assez prudent, alors qu'on peut éviter le problème simplement en endogénéisant le prix en euros comme le produit du prix en monnaie étrangère par le taux de change (éventuellement exogène).

Toutefois, l'établissement de tels liens n'est pas toujours possible. Par exemple, dans un modèle national, les prix étrangers et la production étrangère sont exogènes , mais l'un est clairement influencé par l'autre. Cependant, la nature et l'importance du lien sont très variables. Par exemple, une diminution de la production étrangère peut produire une déflation mondiale²², tandis que l'inflation peut y réduire les exportations et la production. Pour les décrire complètement il faudrait avoir recours à un modèle étranger ou mondial. Une solution intermédiaire pourrait consister à établir un ensemble de

²² C'est le cas pour le modèle mondial MacSim que nous présenterons plus tard.

multiplicateurs linéaires reliant ces éléments, mais en général, le constructeur du modèle prendra lui-même en charge le problème en produisant un ensemble d'hypothèses cohérentes (avec peut-être un peu d'aide venant de spécialistes de l'économie mondiale, ou de l'utilisation d'un modèle mondial séparé).

d - **Sur les liens endogènes-> exogènes:** ils sont évidemment proscrits, car contrairement aux liens précédents le constructeur du modèle ne peut pas les maîtriser. On les trouve dans certains modèles, cependant, à travers la présence des exogènes suivantes:

- éléments mesurés en termes constants, alors qu'ils devraient évoluer avec l'activité économique,
- déflateurs, qui doivent dépendre d'autres déflateurs,
- éléments mesurés en termes courants, pour les deux raisons.

Si le modèle associé peut éventuellement produire des estimations et même des prévisions correctes, il court le risque de présenter des propriétés de sensibilité anormales. Prenons un exemple:

Supposons que le revenu des ménages HI est composé

- des revenus salariaux, calculés comme le produit de l'emploi LT par le taux de salaire W ,
- des autres revenus exogènes .

Les salaires seront indexés parfaitement sur les prix:

$$W = WR \cdot CPI$$

On aura donc

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

Cette équation peut se comporter correctement en prévision . Mais supposons qu'un changement d'hypothèses augmente les prix. Le pouvoir d'achat du salaire total reste inchangé, mais pour le complément HIQ il sera réduit dans la même proportion que la hausse des prix:

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

On peut mettre en doute ce postulat. Certains éléments de revenus non salariaux (les prestations sociales, les loyers, les profits des entrepreneurs individuels, les revenus des travailleurs) sont plus ou moins

indexés, et peut-être même sur indexés dans le cas de paiements d'intérêts (le taux d'intérêt devrait augmenter avec l'inflation). Mais d'autres, associés à paiements différés (dividendes, impôt sur le revenu), ne changeront pas immédiatement. La sensibilité globale des prix n'est pas claire, mais une valeur nulle n'est évidemment pas correcte.

Nous allons rencontrer le même problème avec une variation du PIB:

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

où l'on ne peut pas supposer que les recettes ne changent pas avec l'activité économique. Certains éléments ne bougeront pas, ou montreront une sensibilité limitée (les retraites), mais les dividendes et les revenus des propriétaires de petites entreprises vont évoluer fortement.

En conclusion, même si une variable mesurée à prix courants n'a pas de contenu théorique, elle ne devrait pas rester exogène, surtout si elle peut être supposée croître à prix constants. Il est en général préférable de considérer comme exogène son rapport à une autre variable, censée suivre la même tendance (si l'on n'a pas d'idée particulière, on peut utiliser le PIB). L'équation du modèle va calculer la variable en appliquant le ratio exogène. Ceci peut être aussi valable pour les variables à prix constants (qui augmentent généralement avec la production), à l'exception des variables de décision identifiées comme telles (l'investissement de l'État par exemple).

Dans le cas ci-dessus, on pourrait écrire:

$$HI = WR \cdot CPI \cdot LT + CPI \cdot Q \cdot r_{hiq}$$

formule dans laquelle l'introduction de Q relie des revenus supplémentaires à la croissance globale de l'économie.

1.3.4.2 L'homogénéité

Si certaines équations dans un modèle ne répondent pas aux contraintes d'homogénéité, cela met en danger ses propriétés, notamment sa sensibilité aux chocs. Citons quelques cas:

- Des relations linéaires entre les valeurs et les quantités. L'équation:

$$CO \text{ (consommation à prix constants)} = a \text{ HRI (revenu à prix courants)} + b$$

est non seulement absurde d'un point de vue théorique, mais conduira à long terme à un niveau d'épargne

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

qui deviendra clairement négatif au-delà d'un certain niveau de prix.

- Le mélange de logarithmes et de niveaux. De la même manière, l'équation:

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

(cette fois les deux éléments seront mesurés en quantités) fait tendre le rapport CO / HRI vers 0, et donc le taux d'épargne vers 1, lorsque HRI croît indéfiniment.

Ce dernier exemple montre cependant une limite de l'argument: sur de courtes périodes, l'équation peut présenter une explication satisfaisante, car la consommation rapportée au revenu (la propension à consommer, complément à 1 du taux d'épargne) diminue de façon régulière avec le revenu. C'est la vitesse de la baisse, et son évolution à long terme, qui est ici mise en question.

1.3.4.3 Avec dimension

Le problème est identique à celui de la dimension exogène. Il invite à une étude attentive du sens théorique de la constante. En outre, comme la plupart des variables croissent avec le temps, l'influence de la constante va généralement diminuer ou même disparaître dans la pratique. Nous aborderons cette question plus tard, sur un cas pratique.

1.3.5 NORMALISATION ET IDENTIFICATION

Une fois que les équations sont estimées, le problème de la normalisation demeure. Nous avons vu que très souvent la formule estimée ne va pas expliquer une variable, mais une expression (logarithme, taux de croissance, taux, ou une expression plus complexe). Mais certains algorithmes de simulation réclameront un modèle sous une forme particulière, appelée «identifiée», dans lequel une des variables individuelles non transformées apparaît du côté gauche:

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

Cela signifie que le constructeur de modèle pourra devoir, après estimation, transformer la formulation: cette opération s'appelle la normalisation du modèle.

L'avantage est double:

- L'application de certains algorithmes de résolution est rendue plus facile. Dans certains cas (Gauss-Seidel), cette forme est effectivement réclamée.
- Ce type de formulation permet une meilleure interprétation du processus de détermination de l'équilibre, à condition que chaque équation puisse être interprétée comme une relation causale. Si l'équation décrit un comportement, l'économiste aura placé à la gauche l'élément qu'elle est censée déterminer, et les autres éléments à droite. C'est ce que nous avons fait naturellement dans notre exemple. Par exemple, l'équation qui décrit le choix par les ménages de leur niveau de consommation placera naturellement la variable «consommation» à gauche.

La grande majorité des équations prendra naturellement une forme identifiée. Parfois, une simple transformation sera nécessaire, cependant. L'opérateur non linéaire le plus fréquemment utilisé est peut-être le logarithme, associé à l'intégration d'une formule en élasticités.

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

représente

$$\text{Log}(x) = \int f(\dots) \cdot dx$$

Dans ce cas, il suffit de remplacer:

$$\text{Log}(x) = f(\dots)$$

par

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

Si vous utilisez EViews^{23, 24}, le logiciel va le faire pour vous. Vous pouvez écrire l'équation en utilisant la première forme, et le logiciel va normaliser l'équation lui-même. Ceci est également vrai si l'élément de gauche contient plusieurs variables, mais permet une normalisation directe. Les cas les plus fréquents sont les suivants:

Une variation du logarithme:

$$\text{Log}(x_t / x_{t-1}) = f(\dots)$$

Un taux de croissance:

$$(x_t - x_{t-1}) / x_{t-1} = f(\dots)$$

Un ratio:

$$x_t / y_t = f(\dots)$$

Pour choisir les variables à calculer, EViews choisira la première variable de la spécification de l'équation. Cette méthode simple est appliquée même si la variable a été repérée comme étant calculée par une équation précédente. Par exemple, dans notre modèle, si nous introduisons l'estimation des de M, puis nous spécifions:

²³ Ou la plupart des logiciels du même type.

²⁴ Toutefois, cette nouvelle politique devrait respecter le cadre économique du modèle d'origine.

$$M + Q = FD + X$$

EViews donnera un message d'erreur, car M apparaît comme étant calculée deux fois.

En outre, lorsque l'on demande à EViews la prévision du résultat, on peut choisir entre le calcul du terme de gauche ou de l'élément qui détermine, par exemple M ou $\Delta \text{Log}(M)$ pour notre équation d'importations.

Cependant, EViews ne va pas établir analytiquement une équation pour la variable. Par exemple:

$$M / (Q + M) = f(\dots)$$

sera traduit par:

$$M = (Q + M) \cdot f(\dots)$$

En introduisant un processus non-récurif sur M .

En tout état de cause, la normalisation de l'équation générale

$$f(y, \dots) = 0$$

est possible par l'ajout des deux côtés de la même variable, ce qui donne:

$$y = y + f(y, \dots)$$

Cependant, la convergence d'un modèle défini de cette manière est souvent difficile à obtenir (par exemple, si « f » est lié positivement à y). Dans ce cas, on peut utiliser (la valeur de « a » peut être négative):

$$y = y + a \cdot f(y, \dots)$$

Des simplifications plus importantes sont parfois possibles et seront abordées avec le processus de résolution numérique.

L'identification n'est pas toujours économiquement très simple: dans notre exemple, lorsqu'il s'agit d'équilibrer la demande et l'offre, on peut observer que les trois dernières variables (demande finale , exportations et importations) vont être déterminées par leur propre équation (la somme de ses éléments pour la première, des équations estimées pour les autres). Cela signifie que l'équilibrage doit se faire par le biais du PIB, et nous devons écrire l'équation comme suit:

$$Q + M = FD + X$$

ou

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

ce qui rend son contenu théorique plus clair comme: la production doit (et peut) satisfaire à la fois les exportations et la partie non importée de la demande intérieure.

1.3.6 CONCLUSION

Il doit être clair maintenant que la définition formelle de l'ensemble des équations correspond, en liaison avec l'estimation des équations de comportement, à un processus itératif et simultané:

- Équations de comportement: à partir d'une formulation théorique initiale, on évoluera progressivement vers leur forme définitive par le rapprochement entre cette théorie et les résultats de l'estimation.
- Identités comptables: elles ont été définies de manière aussi précise que possible dans la phase préliminaire, afin d'établir un cadre cohérent, mais souvent elles vont devoir s'adapter à l'évolution des équations de comportement .

Supposons par exemple que les résultats de l'estimation suggèrent pour l'explication économétrique des exportations l'exclusion de leur composante agricole, la définissant comme exogène: une nouvelle équation et une nouvelle variable apparaîtront, et l'équation donnant le total des exportations deviendra une identité.

CHAPITRE 2: LES USAGES DES MODELES

Nous allons maintenant présenter un panorama des applications utilisant des modèles. Les commentaires seront centrés sur l'exemple des modèles économiques, et plus particulièrement macroéconomiques. Mais la plupart des observations sont transposables au cas général.

Pour chacune de ces applications, les détails techniques sont laissés à la partie «implémentation» (chapitre 7). Pour comprendre les aspects pratiques de l'utilisation de modèles, il faut d'abord connaître leur mode de construction, décrit plus loin dans les chapitres 4 à 8.

2.1 LES DIAGNOSTICS OPERATIONNELS

L'utilisation la plus naturelle d'un modèle semble être l'évaluation du futur économique, que ce soit son évolution la plus probable ou les conséquences de certaines décisions. Les hypothèses concernant l'avenir seront injectées dans le modèle, et sa résolution va produire le diagnostic demandé. Ainsi, on cherchera à anticiper l'évolution des principaux agrégats de l'économie française jusqu'en 2020, à partir d'hypothèses sur l'évolution de l'économie internationale.

2.1.1 LES DIFFÉRENTS TYPES DE DIAGNOSTICS: SCENARIOS ET CHOCS

Deux types de prévisions peuvent être envisagés: les scénarios et les variantes.

Dans un scénario, on s'intéresse aux résultats absolus, et on associera à un ensemble complet d'hypothèses une évolution future de l'équilibre économique. On pourra chercher à obtenir

- des prévisions sur la base des hypothèses les plus probables,
- des prévisions associées à un ensemble de décisions donné (comme le programme d'un parti),
- une évaluation de l'intervalle des évolutions possibles,
- les hypothèses permettant d'atteindre certains objectifs économiques.

Au contraire, avec une variante (ou un choc), on part d'une simulation de base (souvent appelée «prévision de référence» ou «de base»), ou d'une simulation sur la période historique, et on mesure la sensibilité de l'équilibre économique à un changement d'hypothèses. Les deux trajectoires économiques seront alors comparées (sur le passé, l'une d'entre elles peut être l'évolution historique).

Ces chocs peuvent être plus ou moins complexes, allant de la modification d'une seule hypothèse à la prise en compte d'une nouvelle politique économique.

Ces deux techniques, les scénarios et les variantes, avant la réalisation de tout exercice opérationnel, joueront un rôle important dans le processus de validation du modèle.

2.1.2 LES AVANTAGES DES MODELES

Maintenant que nous avons décrit les caractéristiques des modèles et leurs utilisations de base, nous allons présenter les avantages qu'ils apportent (et aussi de leurs défauts).

Par rapport au diagnostic fourni par un expert humain, tous les modèles permettent:

- de garantir la cohérence comptable de l'équilibre qu'ils produisent ;
- de prendre en compte un nombre pratiquement illimité d'influences interdépendantes ;
- de fournir une formalisation explicite des comportements, qui permet à un utilisateur extérieur de les interpréter ;
- de produire un calcul exact et instantané des formules associées ;
- d'adapter immédiatement la totalité du système à une variation locale de la formulation théorique ;

mais également

- de permettre la stabilité du raisonnement, pour les utilisateurs humains d'un modèle inchangé ;
- d'offrir la possibilité de comparaisons formelles avec d'autres modèles.

Cette ambition de prévision était déjà présente lors de la construction des premiers modèles. Mais ce type d'utilisation a bénéficié (depuis les années 1970) de certaines évolutions:

- les progrès de la théorie économique, permettant la formalisation des mécanismes plus sophistiqués, mieux adaptés à la réalité observée ;
- les progrès de l'économétrie, donnant accès à des méthodes statistiques produisant des formulations en association avec un problème donné, et de tester des hypothèses plus complexes ;
- l'amélioration des algorithmes numériques, à la fois pour la vitesse de calcul et la résolution de systèmes plus complexes ;
- l'amélioration simultanée des ordinateurs permettant de traiter des problèmes de taille croissante, par des méthodes de plus en plus complexes ;
- les progrès de la science de la modélisation, permettant la production de modèles mieux adaptés au problème initial, ce qui facilite la production d'hypothèses, et réduisant le coût de l'obtention de solutions acceptables ;
- la production de logiciels spécialisés dans la construction de modèles, de plus en plus efficaces, conviviaux, et connectés avec d'autres logiciels ;

- l'amélioration de la fiabilité des données, et la croissance de l'échantillon disponible, en ce qui concerne à la fois le champ et le nombre d'observations (années et périodicité)²⁵ ;
- la facilité de communication entre modélisateurs, par contact direct et forums, ce qui permet de communiquer des idées, des programmes et des méthodes, et d'obtenir la solution aux problèmes petits et grands déjà traités par d'autres.

2.1.3 UNE REMISE EN CAUSE

Cependant, l'utilisation de modèles a engendré des critiques dès leur apparition, en utilisant souvent le terme «boîte noire» qui décrit la difficulté de contrôler et de comprendre un ensemble de mécanismes souvent individuellement simples mais globalement très complexes.

Au cours des dernières décennies la critique a augmenté, au point d'appeler à un rejet global de l'utilisation des modèles traditionnels («structurels»). De façon un peu surprenante, les critiques trouvent souvent leurs arguments dans les améliorations ci-dessus. On peut trouver:

- Une critique utilitariste: les modèles se sont révélés incapables d'anticiper correctement l'avenir. Si cette observation est apparue (au milieu des années soixante-dix), ce n'est évidemment pas parce que la qualité des modèles avait diminué. Mais l'information sur la performance des modèles est devenue plus accessible (certaines études systématiques ont été produites), et les fluctuations suivant le premier choc pétrolier ont rendu les prévisions plus difficiles. En période de croissance soutenue et régulière, extrapoler la tendance est très facile pour les experts comme pour les modèles²⁶. Paradoxalement, l'émergence de cette critique a suivi, et non précédé, l'intervention de plus en plus directe des constructeurs de modèles et de leurs partenaires dans les résultats de prévision.
- Une critique économétrique: les techniques modernes exigent pour les observations une quantité et une qualité que les échantillons disponibles n'ont pas suivis. Un fossé s'est creusé

²⁵ Toutefois, la taille des échantillons ne doit pas nécessairement croître avec le temps. Dans un système de comptes nationaux, l'année de référence doit être modifiée de temps à autre, et les anciennes données ne sont pas nécessairement converties.

²⁶ On pourrait dire que ces dernières années ont présenté une dégradation régulière de l'activité. Mais en dehors du fait que cette observation est contestable, les prévisionnistes ont souvent la tentation d'anticiper la sortie de crise, ce qui conduit à un biais de prévision.

entre les méthodes d'estimation jugées par les théoriciens de l'économétrie comme les seules acceptables, et des méthodes vraiment applicables à un modèle²⁷.

- Une critique théorique: le développement de la théorie économique a conduit souvent à des formulations plus sophistiquées, que les informations disponibles ont du mal à valider. Et en tout cas de nombreux domaines présentent maintenant plusieurs théories alternatives, entre lesquelles tout choix court le risque d'être critiqué par la majorité des économistes. Egalement, dans le domaine monétaire, aller au-delà d'un cadre de base conduit à s'appuyer sur des informations indisponibles dans la pratique, ou sur des formulations trop complexes pour être estimées.
- Une critique mixte: les utilisateurs de modèles ne sont plus des clients passifs. Ils critiquent les formulations, quant à leurs spécifications estimées, ou leurs propriétés numériques. Cette évolution est paradoxalement favorisée par l'amélioration de l'interprétation logique des mécanismes économiques, elle-même engendrée essentiellement par la connaissance économique (même les articles de magazines de vulgarisation économique utilisent des relations macroéconomiques implicites) et la pratique de la modélisation (la population des clients inclut de plus en plus d'anciens constructeurs de modèles ou au moins des adeptes de cours sur la modélisation). On pourrait dire que les utilisateurs du modèle demandent à celui-ci d'aller au-delà de leur diagnostic spontané, et ils veulent que cette information additionnelle soit justifiée.

Il est clair que ces critiques sont d'autant plus pertinentes que l'objectif recherché est ambitieux. Les prévisions sont plus vulnérables que les simples projections indicatives, qui visent à couvrir le champ des évolutions possibles. Quant à l'étude des chocs de politique économique, elles ne sont pas sujettes à des erreurs sur les hypothèses de base, si l'on néglige les non-linéarités^{28,29, 30}.

²⁷ En fait, la taille de l'échantillon requis par les techniques actuelles (50 ou mieux 100 observations) limite la possibilité d'estimation d'équations contenant des déflateurs ou des variables à prix constants. Même en utilisant des données trimestrielles, séparer les valeurs en prix et en volume est tout à fait discutable à 15 ans de la période de base.

²⁸ Avec un modèle linéaire, la conséquence d'un choc ne dépend que de sa taille, et non de la simulation dont il part.

²⁹ Un exemple est l'impact d'une baisse des droits de douane locaux. *Ex ante* elle augmente les importations (un choc de demande négatif). *Ex post* elle diminue les coûts des facteurs locaux (à la fois l'investissement, et le travail, indexé sur un prix à la consommation en baisse). Cela conduit à augmenter la capacité locale et améliorer sa compétitivité, tant le marché local (limitation de l'augmentation des importations), qu'étranger. Dans la plupart des modèles, le PIB diminue puis augmente.

L'interprétation complète d'un tel choc fournit un grand nombre d'informations, même si l'on reste au niveau non quantitatif.

Cette pertinence dépend aussi du crédit accordé aux résultats. On peut utiliser les chiffres comme tels, ou se contenter d'ordres de grandeur, ou même simplement chercher à mieux comprendre les mécanismes économiques globaux en identifiant les interactions les plus importantes (impliquant éventuellement des chaînes causales complexes). À notre sens, c'est dans ce dernier domaine que l'utilisation de modèles est la plus féconde et la moins contestable.

2.2 MODELES THEORIQUES

Contrairement aux modèles précédents, les modèles théoriques peuvent être construits dans le seul but de formaliser une théorie économique. Elle peut être suffisante pour écrire leurs équations, associant à un comportement théorique un système cohérent et complet. Reproduire la réalité observée n'est pas le but principal de ces modèles, et il n'est pas obligatoire d'estimer leurs paramètres: on peut choisir une valeur arbitraire, souvent dictée par la théorie elle-même. En fait, cette estimation sera souvent techniquement impossible, lorsque certaines des variables utilisées ne sont pas disponibles sur le plan statistique (les objectifs ou les attentes des agents, par exemple).

Cependant, même en se basant sur des séries artificielles et des paramètres arbitraires, la simulation numérique de ces modèles peut être intéressante. En fait, les formules sont souvent si complexes que la résolution du modèle numérique sera nécessaire pour observer ses solutions ainsi que ses propriétés (telles que la sensibilité des solutions à des hypothèses et à des coefficients).

2.3 PETITS MODELES QUANTIFIES

2.3.1 A BUT EDUCATIF

Bien souvent du même type que les précédents, ces modèles tentent de présenter des mécanismes économiques aussi complets que possible, sur la base des données réelles, sous une forme interprétable et concise. Si nécessaire, on favorisera le message contenu dans la présentation par rapport au respect des critères statistiques.

C'est le cas du produit MacSim, ce qui permet aux étudiants d'interpréter les mécanismes de la macroéconomie internationale.

2.3.2 A BUT SCIENTIFIQUE

Ces modèles représentent un cas intermédiaire. On cherche une représentation réaliste de l'économie, adaptée à la réalité observée, mais suffisamment simple pour accepter l'application des méthodes d'analyse complexes (et l'interprétation de leurs résultats). En plus de la recherche scientifique, cette étude peut être faite pour mesurer et analyser les propriétés d'un modèle opérationnel sur une représentation simplifiée (dans les années 1980 MiniDMS, puis MicroDMS ont été utilisés pour caractériser la dynamique du modèle DMS multi sectoriel de l'INSEE).

Il existe deux catégories de méthodes:

- les méthodes «externes» qui utilisent des simulations du modèle pour observer ses propriétés quantitatives, et en déduire un commentaire descriptif, statistique et économique ;
- les méthodes «internes» qui cherchent à expliquer les propriétés du modèle par ses caractéristiques structurelles, à l'aide d'outils mathématiques. Et cela sans faire nécessairement appel à de véritables simulations.

CHAPITRE 3 : LES TYPES DE MODELES

Nous allons maintenant chercher à établir une classification des modèles, en nous concentrant sur le lien entre les caractéristiques du modèle et l'objectif pour lequel il a été construit.

3.1 LE CHAMP

Le champ décrit par un modèle est caractérisé par les variables qu'il calcule, mais aussi par les hypothèses qu'il prend en compte.

Dans le sous-ensemble des modèles économiques, on peut considérer:

- un champ géographique: modèles nationaux, modèles multinationaux, modèles mondiaux. Ces derniers peuvent être construits de deux manières: en rassemblant des modèles nationaux préexistants, avec des structures potentiellement différentes, ou en construisant simultanément des modèles-pays de structure identique, peut-être avec une seule équipe de modélisateurs. Nous traiterons cet aspect plus tard.
- Un champ théorique: la théorie utilisée pour la formalisation du modèle peut approcher ou non certains aspects économiques spécifiques, ou les traiter de façon originale. Un modèle keynésien pourra limiter le traitement des aspects monétaires. Un modèle de court terme ne formalisera pas les évolutions démographiques.
- Un champ d'unités: un modèle peut présenter les seules variables à prix constants, ou des quantités physiques comme les barils de pétrole ou le nombre de porcs.
- Un champ d'agents: un modèle peut décrire le d'un agent unique: les ménages, les entreprises, l'État. Il peut détailler ce dernier entre État central et collectivités locales.
- Un champ de biens: un modèle peut ne considérer que la production et la consommation d'un bien, par exemple l'énergie. Un modèle énergétique peut utiliser des unités physiques.

Il existe d'autres types de champs. Cependant, la distinction n'est pas toujours facile: certains modèles décrivent sommairement un champ global, mais détaillent un certain aspect sur lequel ils se concentreront. Un modèle énergétique prendra en compte les interactions avec le reste de l'économie, qu'il faudra aussi modéliser, mais pas dans le même détail. Et il pourra mélanger des unités physiques (barils de pétrole ou gigawatts) avec des éléments de comptabilité nationale.

D'autre part, il sera toujours possible, et rendu plus facile par certains logiciels de modélisation, de changer (en fait, de restreindre) au moment de la simulation le champ d'application du modèle. La distinction n'est alors plus permanente: un modèle multinational peut être utilisé pour simuler l'évolution complète de l'économie mondiale, mais son utilisateur peut également limiter les calculs à l'évolution d'un groupe de pays, voire un seul, les autres éléments étant fixés. On peut simuler un modèle de

l'économie réelle avec ou sans caractéristiques monétaires. Ou un modèle utilisant normalement des éléments d'anticipations rationnelles pourra les mettre de côté et ne tenir compte que du passé dans les comportements des agents.

3.2 LA TAILLE

L'histoire de la modélisation montre que, pendant une longue période, les modèles nouveaux ont généralement vu leur taille augmenter, pour les raisons citées plus haut: les progrès de la construction de modèles techniques, la disponibilité accrue des données, les calculs informatiques plus rapides. De plus, pour un modèle donné, la taille augmente régulièrement au cours de son existence, tant que les nouveaux membres de l'équipe souhaitent y apporter leur contribution.

Cependant, les dernières décennies ont vu la tendance s'inverser, en faveur d'un retour à des modèles de taille limitée. L'amélioration de la productivité, demandée aux équipes de constructeurs de modèles, ainsi que le contrôle des valeurs des solutions, sont de moins en moins compatibles avec l'utilisation d'un grand modèle. Malgré les progrès des techniques de construction de modèle, le désir de réduire les coûts et les délais est entré en conflit avec la taille, en particulier (mais pas seulement) en ce qui concerne les opérations humaines: élaboration d'hypothèses, contrôle des projections et interprétation des résultats.

En outre, l'utilisation d'un modèle très détaillé peut faire apparaître trop élevé le coût des estimations et des contrôles de cohérence. L'attractivité d'un modèle EGC calibré et gémellaire va alors augmenter.

Enfin, la volonté de répondre aux critiques comparant les modèles à des «boîtes noires» a conduit les modélisateurs à trouver des instruments plus explicites et plus facilement gérables.

3.2.1 LES DÉTERMINANTS DE LA TAILLE

Les déterminants de la taille du modèle sont les suivants:

- la taille du champ couvert (voir ci-dessus) ;
- le degré d'agrégation, qui peut être
 - verticale: le nombre d'opérations prises en compte (par exemple, on peut distinguer plusieurs types de subventions, ou de prestations sociales),
 - ou horizontale: le nombre d'agents pris en compte ; on peut distinguer plus ou moins de secteurs d'entreprises ou de types de ménages.

Le degré d'agrégation ne sera pas forcément uniforme: un modèle énergétique utilisera un détail particulièrement élevé pour les produits de ce type.

En fait, le même modèle peut apparaître sous plusieurs versions de taille différente, en fonction notamment du degré d'agrégation. Chaque version a son champ d'utilisation: prévisions détaillées, simulations rapides, analyse mathématique et utilisations pédagogiques.

Ainsi, à la fin des années 1980, le modèle à 3 000 équations DMS (dynamique multi sectoriel) utilisé par l'INSEE pour ses prévisions à moyen terme était accompagné par deux versions de taille réduite: Mini-DMS (200 équations), utilisé pour les projections opérationnelles et les analyses qui ne réclamaient pas de détail par produits, et MicroDMS (45 équations), avec un but essentiellement éducatif.

Cette distinction a perdu de sa validité, cependant, suite à la réduction de la taille des modèles opérationnels.

3.2.2 UNE CLASSIFICATION

Nous proposons la classification suivante, forcément subjective:

- les petits modèles: 1 à 50 équations

Exemples: Klein-Goldberger, MicroDMS (INSEE), DMM (CEPREMAP), le modèle FAIR ;

- les modèles de moyenne: 150 à 400 équations

Exemples: MULTIMOD (FMI), TESTUS (Federal Reserve Board), AMADEUS (INSEE), HERMES (École Centrale de Paris), MESANGE (INSEE et ministère français des Finances), MZE (INSEE et ministère français des Finances) ;

- les grands modèles: 800 à 2 000 équations

Exemple: TESTMCM (Federal Reserve Board), METRIC (Direction de la Prévision), MOSAIQUE (OFCE), MASCOTTE (Banque de France), MULTIMOD (FMI) ;

- les très grands modèles: plus de 4 000 équations

Exemple: NiGEM (NIESR), INTERLINK (OCDE), le modèle LINK (Organisation des Nations Unies + Université de Toronto + partenaires), MacSim, MIMOSA (CEPII-OFCE). Aujourd'hui, cette dernière catégorie devrait s'appliquer uniquement à la modélisation internationale.

Des intervalles ont été laissés pour les cas intermédiaires.

En outre, un certain nombre de modèles construits par l'auteur dans le cadre de la coopération économique sont répertoriés dans l'étude sur l'efficacité des algorithmes, présentée plus tard. Les projets en cours incluent la Chine, le Vietnam, l'Algérie, le Maroc. Des projets passés ont concerné la Pologne, la

Slovaquie, la République tchèque, la Lituanie, l'Ukraine, la Tunisie, la Communauté Andine, l'Argentine, le Kazakhstan, le Tadjikistan.

3.3 L'HORIZON

3.3.1 POUR LA PREVISION¹

Si un modèle est conçu pour la prévision, son horizon sera défini lors de la construction du modèle. Il sera fortement lié à sa philosophie générale et à l'ensemble des mécanismes qu'il met en œuvre. Un modèle de long terme sera peu intéressé à des phénomènes conjoncturels (tels que les retards dans l'ajustement des salaires sur les prix), tandis qu'un modèle de court terme ne prendra pas en compte les tendances les plus longues (telles que l'influence de la conjoncture économique sur la démographie).

Ces différences semblent ignorer l'élaboration d'un modèle qui peut être utilisé à la fois à court et long terme. Mais nous allons voir que de bonnes raisons, en particulier économétriques, font que cette option apparaît maintenant comme la plus naturelle. Nous les développerons lorsque nous aborderons la périodicité, dans le paragraphe 3.4.

Dans tous les cas, on peut trouver une certaine asymétrie dans la pertinence de cette observation. Si les modèles de long terme peuvent négliger les périodes intermédiaires lorsqu'elles ne présentent pas de fluctuations importantes, la simulation des périodes au-delà de l'horizon opérationnel peut mettre en évidence des problèmes futurs, déjà présents mais non apparents à court terme

3.3.2 POUR L'ANALYSE

Ici, l'horizon dépend du type de l'analyse qu'on veut produire. Souvent, pour tester un modèle construit avec un horizon de prévision donné, une simulation sur une plus longue période est nécessaire. Encore plus que pour les prévisions, les chocs d'hypothèses, par exemple, pourront alors montrer et expliquer les anomalies qui n'étaient pas vraiment apparentes au cours de la période de projection normale, mais qui avaient déjà une influence négative non négligeable. Nous allons insister sur ces questions plus tard.

¹ On remarquera que l'on peut utiliser plusieurs mots pour caractériser ces exercices: prévisions, projections, scénarios, simulations. Tout dépend de la raison pour laquelle le test a été fait, et peut-être de la confiance donnée à ces résultats. Nous privilégions le dernier terme, qui doit malheureusement être complété en: «simulation sur le futur».

3.3.3 UNE CLASSIFICATION

On pourrait utiliser la classification suivante pour les différents modèles:

- à court terme: 1 trimestre² à 2 ans ;
- à moyen terme: 4 à 7 ans ;
- à long terme: 10 ans et plus.

Evidemment, en particulier pour une simulation dynamique, la trajectoire complète, y compris les valeurs intermédiaires, est intéressante.

3.4 LA PERIODICITE

La périodicité d'un modèle est liée aux mécanismes qu'il cherche à étudier et donc à son horizon.

- Les modèles de court terme exigent une périodicité élevée, afin de prendre en compte des phénomènes conjoncturels: les retards de l'indexation des salaires sur les prix, l'ajustement progressif du niveau de consommation à une augmentation des revenus.
- Les modèles de long terme peuvent utiliser une périodicité plus basse, moins pour des raisons théoriques (le comportement de long terme peut être décrit par un modèle à périodicité courte), que techniques: ce choix va réduire les contraintes sur la disponibilité des séries, faciliter la production des hypothèses, et limiter les coûts de simulation.

Cependant, nous verrons que l'utilisation de méthodes économétriques «modernes» appelle à une périodicité courte, pour tous les types de modèles, dès lors que des estimations sont prises en considération.

Cela signifie que le principal déterminant de la périodicité du modèle est fourni par les données disponibles. Les pays qui produisent des comptes nationaux trimestriels utilisent des modèles trimestriels, qui leur permettent d'appliquer des techniques modernes avec un certain confort, et de traiter simultanément les deux approches à court et à long terme. Lorsque seuls des comptes annuels sont disponibles, les techniques deviennent plus simplistes, et de véritables applications à court terme ne sont

² On peut envisager des modèles à périodicité inférieure (1 mois ou moins). Mais il faut pour cela disposer d'un ensemble de séries mensuelles recouvrant complètement le champ du modèle, ce qui est rarement le cas. C'est bien sûr moins vrai pour une monographie économétrique.

pas possibles. Malheureusement, ceci s'applique le plus souvent aux pays ayant une histoire courte en statistiques, rendant le problème encore plus difficile à traiter.

3.5 UNE DIMENSION SUPPLEMENTAIRE ?

Les variables des modèles que nous traiterons ici évoluent en général avec le temps. Mais une dimension supplémentaire peut apparaître: ils peuvent par exemple traiter des échanges entre plusieurs pays, ou séparer les biens en plusieurs produits, ou même les deux à la fois.

Nous traiterons de ce problème dans le chapitre 10 (10.3 à 10.6). Mais on comprend déjà qu'un choix devra être fait pour les valeurs des paramètres: communes et individualisées.

3.6 LES AUTRES MODELES

Nous nous sommes essentiellement concentrés sur le cas des modèles macroéconomiques. On peut également trouver des modèles:

- microéconomiques: ils décrivent le comportement des entreprises, des ménages.

Ces modèles seront parfois plus théoriques, faisant appel à des calculs d'optimisation (minimisation des coûts) ou à des éléments de stratégie (théorie des jeux). Ils vont souvent être estimés sur données d'enquête ;

- non-économiques: les modèles peuvent s'appliquer à la biologie, la physique, la chimie, l'astronomie, la météorologie, l'écologie, le contrôle de processus, et ainsi de suite Ils seront utilisés pour évaluer les conséquences de la construction d'un barrage, contrôler le fonctionnement d'un processus de fabrication, rechercher la meilleure organisation d'un projet, décrire un processus biologique. Ces modèles seront souvent conçus non pas comme un système d'équations formalisé, mais comme la maximisation d'un critère sous certaines contraintes, ou comme un système de propositions reliées par des opérateurs logiques.

CHAPITRE 4: LE PROCESSUS DE MODELISATION: PRESENTATION ET GENERALITES

À partir de maintenant, nous allons décrire le processus de développement, d'utilisation et de mise à jour d'un modèle, en attachant un intérêt particulier aux aspects techniques et notamment informatiques. Les applications à EViews seront présentées en détail, mais la plupart des enseignements proposés peuvent être appliqués aux autres logiciels du même type, parfois aussi aux produits qui ne sont pas dédiés à la modélisation (par exemple Excel).

Tout d'abord, établissons une brève description de l'organisation du processus de construction d'un modèle.

4.1 LES ETAPES DU PROCESSUS

4.1.1 LA PREPARATION DU MODELE

La première étape dans la construction d'un modèle est la production d'un projet qui assure une certaine compatibilité entre les données disponibles (d'où qu'elles viennent) et le type de modèle que son constructeur a en tête (objectif, champ, nature des variables, théorie sous-jacente).

Une fois connu le champ des données disponibles, l'économiste définira un cadre de modèle pour lequel des valeurs peuvent être attribuées à toutes les variables, soit en utilisant des éléments existants ou par le calcul. Cela signifie qu'une première décision doit être prise quant au domaine décrit par le modèle, les variables utilisées comme hypothèses et les variables qu'il doit calculer. En outre, il va falloir diviser les équations en identités, qui fixent des liens incontestables entre les variables, et équations décrivant le comportement des agents, dont la formulation finale sera basée sur les évolutions passées des éléments associés.

Sa première tâche sera de rassembler, par la lecture de fichiers et de transformation des données, l'ensemble complet des variables requises par le modèle, de définir la forme des identités, et d'annoncer ses intentions en termes de comportements. Il faudra vérifier sur quelles périodes les données nécessaires sont connues, et si sur ces périodes les identités sont bien vérifiées. Si certains éléments ne sont pas disponibles, il devra utiliser les meilleures approximations disponibles. Et si cela échoue aussi, il devra utiliser son imagination.

Il pourra également établir une première analyse économique du cadre impliqué par les spécifications du modèle (grandement aidé par EViews).

4.1.2 L'ESTIMATION

La deuxième phase va chercher à obtenir une description satisfaisante du comportement des agents, en vérifiant la compatibilité de la théorie économique avec les données disponibles. Le modélisateur doit définir un ensemble de formulations contenant des paramètres inconnus, calculer pour chaque formulation les valeurs qui donnent la meilleure explication des évolutions passées, et faire son choix, en utilisant comme critère à la fois les tests statistiques et la conformité à la théorie économique. Ce processus peut réclamer l'introduction de nouvelles variables, ou des changements dans certaines définitions, qui se traduiront par la reformulation des identités.

4.1.3 LA RESOLUTION ET LES TESTS SUR LE PASSE

Une fois le modèle complet défini, on peut chercher à le résoudre.

Il faut d'abord vérifier la cohérence de l'ensemble d'équations, données et paramètres, en appliquant chaque formule séparément sur la période d'estimation. Si les résidus d'estimation ont été introduits comme éléments supplémentaires, le processus devrait donner les valeurs historiques dans tous les cas.

On peut alors simuler le modèle complet sur la même période, en mettant (temporairement) les résidus à zéro. Cela montrera si la prise en compte des interactions instantanées et décalées n'amplifie pas trop les erreurs d'estimation.

Enfin, on mesurera les réactions de l'équilibre à un changement d'hypothèses, comme par exemple la composante exogène de la demande. Les résultats seront comparés avec les enseignements de la théorie économique, et ce qui est connu des valeurs données par d'autres modèles. Cependant, il ne faut pas consacrer trop de temps à cette tâche, car les simulations sur l'avenir fourniront un bien meilleur contexte.

Découvrir des écarts pourra conduire à des changements dans certains éléments du modèle, y compris son ensemble de variables. Cela signifie qu'il faudra revenir à l'étape 1 ou 2.

4.1.4 LA RESOLUTION ET LES TESTS SUR LE FUTUR

Une fois que le modèle a passé tous les tests sur le passé, d'autres tests seront effectués, dans des conditions plus représentatives de l'utilisation réelle: sur l'avenir. Pour cela, des valeurs devront être établies quant aux hypothèses futures. Encore une fois, la sensibilité du modèle aux chocs sera étudiée, cette fois sur une base plus longue et plus régulière. Quant à la fiabilité des résultats de simulation, on ne pourra compter cette fois que sur des simulations stochastiques, en l'absence d'éléments de comparaison.

4.1.5 L'UTILISATION DU MODELE POUR LES PREVISIONS ET LA POLITIQUE ECONOMIQUE

Enfin, le modèle sera considéré comme opérationnel pour les études économiques: prévisions et analyses de politique économique.

4.2 LE CHOIX DU LOGICIEL

Le premier choix que le modélisateur doit faire concerne le logiciel qu'il va utiliser au cours du processus de modélisation, et donc pour la réalisation de chacune des tâches décrites ci-dessus. L'option la plus simple, et celle que nous développerons dans nos exemples, est d'utiliser le plus largement possible un seul produit, proposant toutes les fonctions nécessaires à la construction de modèles, à leur gestion et à l'utilisation des résultats. Comme annoncé, nous allons baser notre présentation sur le logiciel EViews.

D'autres options peuvent être envisagées.

4.2.1 COMBINER UN LOGICIEL DE MODELISATION AVEC D'AUTRES ELEMENTS

On peut limiter le logiciel de modélisation à des tâches spécifiques, et utiliser un produit plus général pour les tâches communes.

De toute évidence, le logiciel de modélisation est plus adapté au cœur du processus: les estimations, l'établissement du modèle et la production de simulations, mais on peut envisager d'autres options pour les phases initiale et finale. Cela évite aussi des allers retours entre les produits, qui se montreront coûteux et source d'erreurs.

Cela signifie que l'on peut:

4.2.1.1 Pour la gestion des données

- Confier tout ou partie de la gestion des données à un logiciel spécialisé, comme Access ou Dbase (ou même Excel). Ceci est facilité par le développement de logiciels récents. Par exemple, à partir de la version 5, EViews permet de lire directement à partir d'un grand nombre de formats (EViews 7 rendra les choses encore plus faciles). Bien sûr, le modélisateur doit y avoir accès et se familiariser avec son utilisation.
- Utiliser un autre logiciel scientifique si l'on procède à plusieurs types de recherches sur le même ensemble de données. Encore une fois, les fonctions de transfert fournies par EViews facilitent cette option.

- Les avantages: faire appel à un produit de gestion des données plus efficace et plus polyvalent, en utilisant des procédures spécifiques, peut-être déjà développées avant le début du projet de modélisation.
- Les inconvénients: le coût et le risque des transferts, peut-être la nécessité d'acheter et d'apprendre l'usage d'un nouveau produit.

En tout cas, certains des calculs («à la volée») devront être exécutés à l'intérieur du logiciel de modélisation. Et avec ses versions successives, EViews est en progression constante pour la gestion des données, et le champ des applications scientifiques.

4.2.1.2 Pour les estimations

On peut aussi considérer l'estimation des équations à l'extérieur du logiciel. Ceci est bien sûr beaucoup moins naturel, bien plus dangereux et coûteux. Le seul cas que nous pouvons considérer est celui où la méthode demandée n'est pas disponible.

Pour EViews ceci est rare, car les options proposées sont assez complètes et à jour. Mais comme tout logiciel commercial, ses mises à jour ne suivent pas de trop près les développements récents de l'économétrie. Ses concurrents non plus, mais certains d'entre eux (comme Gauss) permettent à l'utilisateur d'aller plus loin dans la programmation des méthodes, ce qui permet une adaptation plus rapide au prix d'un effort personnel. En outre, cela permet de réaliser différentes variantes de méthodes, mieux adaptées aux cas particuliers si EViews ne fournit pas toutes les options nécessaires. Plus récemment, les versions 6, 7 et 8 donnent plus en plus de possibilités d'application de méthodes définies par l'utilisateur et d'appel à des sous programmes externes.

- Les avantages: l'application des dernières méthodes développées, en utilisant plus d'options.
- Les inconvénients: d'avoir à gérer (+ apprendre et acheter) deux logiciels distincts traitant les séries chronologiques, à transférer des données et des éléments tels que les formules de l'un à l'autre (avec un grand risque d'erreurs et d'incohérences). Le transfert d'éléments avec des partenaires ou d'autres utilisateurs risque également d'être plus difficile.

Le dernier élément est très important dans le cas d'EViews: la probabilité pour un nouveau partenaire d'utiliser ce produit est extrêmement élevée, en particulier dans le domaine des modélisateurs universitaires (qui ne peuvent pas se permettre l'acquisition des produits supposés «haut de gamme» comme Troll ou Aremos).

4.2.1.3 Pour la résolution

L'estimation et la résolution du modèle en deux produits séparés semble étrange et les cas où cela se révèle efficace sont rares. On peut penser à:

- La résolution de très grands modèles. Dans ce cas, on pourrait envisager d'utiliser un logiciel mathématique (MatLab ?), un programme spécialisé dans la résolution de modèles (Win Solve ?), ou d'écrire un programme spécifique tenant compte des spécificités du modèle, afin d'accélérer les calculs. L'importance de cette question diminue à partir d'EViews 6, environ 3 à 5 fois plus rapide qu'EViews 5 en temps de calcul, et capable de résoudre les grands modèles (notre plus grand modèle comprend 3 500 équations, et se résout assez vite comme nous le verrons plus tard).
- Construire un modèle dont les spécifications ne sont pas gérées par le logiciel. Deux cas viennent à l'esprit: les anticipations rationnelles stochastiques (non fournies par une version quelconque d'EViews) ou l'introduction d'un processus de maximisation.

4.2.1.4 Pour la présentation des résultats

C'est bien sûr le cas le plus acceptable. On peut considérer que les caractéristiques des présentations (tableaux et graphiques) proposées par le produit ne présentent pas une qualité suffisante ou suffisamment d'options. Dans ce cas, le transfert vers Excel ou un programme similaire vient à l'esprit. Ce transfert est de plus en plus facile, allant même jusqu'à l'introduction dans le logiciel externe de fonctionnalités de lecture des fichiers EViews. En outre, les partenaires non-modélisateurs n'auront pas besoin du logiciel de modélisation.

Toutefois, la qualité des graphiques (à partir d'EViews 4) et des tableaux (à partir d'EViews 5) s'est beaucoup améliorée, ce qui réduit le gain d'une telle technique. Cela est encore plus vrai pour les versions suivantes.

- Les avantages: une meilleure qualité de graphiques et de tableaux, une adaptation plus facile aux données, un transfert facilité vers des partenaires (tout le monde a accès à Excel ou si besoin à OpenOffice).
- Les inconvénients: le coût du transfert lui-même, une organisation plus complexe.

4.2.2 UTILISER UN LOGICIEL SPECIALISE DANS LA MODELISATION

On peut trouver sur le marché des logiciels scientifiques un ensemble de produits dédiés à la construction de modèles et à leur gestion. Bien qu'ils autorisent une certaine économétrie, celle-ci n'apparaît que comme un complément nécessaire, et ils ne peuvent être considérés comme des outils efficaces pour des chercheurs

Actuellement, les exemples les plus courants sont TROLL et AREMOS.

Ces deux produits ne sont pas conçus pour le grand public: ils sont assez chers, et disponibles uniquement sur la base d'une location. En échange, la société de logiciels fournit des services supplémentaires, tels qu'une assistance téléphonique et des séminaires.

Ceci n'est possible qu'avec un nombre limité de clients. De toute évidence, la création d'une hotline EViews coûterait très cher, et appellerait à une forte augmentation du coût du produit, réduisant ainsi son attractivité pour les utilisateurs universitaires.

Mais le forum EViews est maintenant tout à fait efficace, reliant les utilisateurs et permettant d'obtenir (assez rapidement) des réponses de la part des modérateurs EViews.

Bien sûr, les manuels comme celui-ci peuvent s'avérer utiles, et l'aide EViews est bien conçue et facile à gérer. En particulier, la description des éléments économétriques se fait par étapes, et ne nécessite pas de connaissances préalables trop pointues.

4.2.3 UTILISER UN LOGICIEL D'USAGE GENERAL

Le choix naturel dans ce cas est un tableur comme Excel ou OpenOffice.

En fait, Excel propose l'estimation d'équations et la résolution de systèmes. Cependant, les méthodes économétriques sont assez limitées et quelque peu difficiles à utiliser, en particulier pour la définition des éléments explicatifs, la construction des statistiques d'estimation et l'observation des résultats à la fois dans des tableaux et des graphiques.

En outre, le processus de résolution utilise uniquement une version limitée de la méthode de Gauss-Seidel, dans laquelle l'ordre des équations est le même que celui des variables dans la feuille de calcul. Comme cet ordre a des conséquences énormes pour la convergence du modèle, ceci exige l'utilisation d'un ensemble de deux feuilles liées, l'une pour la gestion des données et l'autre pour les variables du modèle. Les options pour le contrôle de la convergence et la localisation des erreurs ne sont pas présentes (comme l'exogénéisation d'éléments ou l'utilisation de scénarios alternatifs).

- Les avantages: le coût est négligeable¹, la gestion des données, la gestion et la présentation des résultats du modèle sont faites dans le même produit, ce qui réduit les transferts et les risques d'erreurs.
- Les inconvénients: l'économétrie est limitée, les méthodes sont difficiles à mettre en œuvre, et les résultats difficiles d'accès. Le contrôle de l'algorithme de résolution (unique) est limité, les

¹ On peut envisager des modèles à périodicité inférieure (1 mois ou moins). Mais il faut pour cela disposer d'un ensemble de séries mensuelles recouvrant complètement le champ du modèle.

options sont peu nombreuses. Trouver la source d'erreurs, les interpréter et les résoudre se révèle assez difficile.

4.2.4 CONCLUSION

À notre avis, la meilleure option est la suivante:

- Gérer les données officielles dans un logiciel spécifique, la plupart du temps.
- Si elles sont fournies par une source extérieure, transférer immédiatement les éléments vers le logiciel de modélisation, après une certaine réorganisation des fichiers (y compris l'introduction de noms de variables).
- Transformer les données dans les concepts du modèle.
- Construire le cadre du modèle, estimer les équations, résoudre le modèle dans le logiciel de modélisation.
- Gérer la présentation des résultats dans le logiciel pour la plupart des interprétations.
- Transférer les résultats dans un logiciel spécifique (Excel ?) pour les présentations les plus sophistiquées (même si cela est de moins en moins évident)².

4.3 COMMENT ORGANISER LE DEVELOPPEMENT DU MODELE

Considérons maintenant l'organisation de la tâche de production du modèle.

Pour créer un modèle, deux types d'organisation extrêmes peuvent être envisagés:

- Option méthodique

Le constructeur de modèles

- spécifie complètement un modèle cohérent (y compris les équations comptables), précisant la séparation des hypothèses et des résultats,
 - recherche les séries nécessaires,
 - estime les équations de comportement,
 - utilise le modèle qui en résulte.
-

L'application d'un tel cadre est évidemment illusoire, comme de nombreux retours en arrière seront nécessaires dans la pratique:

- Certaines séries apparaîtront comme indisponibles, et il sera nécessaire de les remplacer ou de les éliminer de formulations. Ainsi, en l'absence de séries pour les intérêts versés par les entreprises on devra se contenter des bénéfices avant intérêts.
- Certaines estimations donneront des résultats insatisfaisants: il sera nécessaire de modifier les formulations, d'utiliser des séries supplémentaires ou de remplacement. Ainsi, une formulation en niveaux pourra être remplacée par une formulation sous forme de logarithmes (élasticités constantes), ou en taux de croissance. Ou on sera amené à expliquer le salaire mensuel moyen au lieu du salaire horaire, et à introduire dans cette dernière explication l'évolution du salaire minimum. Pour un pays producteur de pétrole, il apparaîtra nécessaire d'identifier celui-ci (et les produits non pétroliers) à la fois dans la production et les exportations³.
- De nouvelles idées apparaîtront lors de l'estimation. Par exemple, un article récent sur le rôle de l'investissement direct étranger pourra conduire à tester une formulation originale.
- Des erreurs formelles vont être identifiées. Ainsi, un élément (un type de retraite) pourra avoir été oublié dans la définition du revenu des ménages.
- Certaines variables définies comme des hypothèses vont apparaître suffisamment influencées par les résultats pour voir leur statut modifié.
- Improvisation

Au contraire, un constructeur de modèle peut

- définir les options générales pour la structure du modèle et le cadre théorique,
- produire des formulations indépendantes les unes des autres,
- les estimer en accédant aux séries individuelles,
- et peu à peu connecter les éléments sélectionnés en complétant le modèle par les identités nécessaires, et la base de données par les éléments manquants.

Ce cadre sera encore moins efficace, si ce n'est que par le nombre d'opérations simples mais séparées sur des équations et des séries, qui présentera un coût prohibitif. En outre, l'obtention de la cohérence théorique et comptable du modèle pourra s'avérer difficile, et le processus de modélisation pourra ne jamais converger vers une version satisfaisante.

La solution optimale est évidemment intermédiaire:

- définir le plus précisément possible le champ et la nomenclature du modèle,
- définir ses options théoriques générales et son but,
- obtenir, créer et stocker l'ensemble total des séries présumées utiles,
- définir des domaines d'estimation, spécifier des variables associées et établir des connexions formelles, en particulier les équations comptables,
- procéder à des estimations,
- et passer par une suite de modifications (en nombre aussi limité que possible) jusqu'à ce qu'une forme acceptable soit obtenue.

Il est clair que ce type d'organisation est d'autant plus facile à mettre en œuvre que:

- La taille du modèle est faible: il est possible de mémoriser l'ensemble total des noms de variables pour un modèle de 30 équations, mais pour un grand modèle une documentation formelle sera nécessaire, si possible produite dès le début (pourquoi attendre ?) et régulièrement mise à jour.
- Le nombre de personnes concernées est faible (la distinction va essentiellement se faire entre un et plusieurs): pour un projet d'équipe, le rôle de chaque participant et son domaine de responsabilité doivent être clairement définis. Surtout, les changements physiques (sur les données et les spécifications du modèle) doivent être la responsabilité d'une seule personne, ce qui permettra de centraliser les demandes et de les appliquer de façon cohérente. Et les modifications doivent être clairement annoncées et documentées.

Les modifications individuelles du modèle peuvent être autorisées, cependant, à condition qu'une version de base soit préservée. Ainsi, plusieurs membres d'une équipe de modélisateurs pourront tester, l'un une nouvelle fonction de production, l'autre une description détaillée du secteur financier. Mais même dans ce cas, les mises à jour interfèrent souvent, au moment où des modifications générées dans les versions d'essai distinctes sont appliquées au modèle de base. Par exemple, une nouvelle définition du coût de salaires et de l'investissement, qui définissent les rôles optimaux du travail et du capital dans le processus de production, va influencer la cible de marges dans l'équation de prix.

CHAPITRE 5: LA PREPARATION DU MODELE

Nous allons maintenant décrire la première tâche (et probablement la plus importante): la préparation de la production du modèle.

On pourrait être tenté de commencer la production du modèle le plus rapidement possible. Mais il est extrêmement important de consacrer au début suffisamment de temps à évaluer les options et le choix d'une stratégie. Se rendant compte beaucoup plus tard qu'il a choisi les mauvaises options, le modélisateur devra faire face à deux mauvaises solutions: poursuivre un processus conduisant à un modèle de qualité limitée, ou revenir en arrière jusqu'au point où le choix a été fait.

Cela peut concerner:

- l'organisation des tâches, comme la production de premiers modèles mono pays, pour un projet de modélisation mondiale ;
- les questions économiques, comme le choix de la complexité de la fonction de production, ou de la décomposition des produits ;
- les éléments techniques, comme le nombre de lettres identifiant le pays dans un modèle mondial.

5.1 LA PREPARATION DU MODEE: LE CADRE

Au début du processus d'élaboration du modèle, le modélisateur (ou l'équipe) a au moins:

- des idées générales sur la logique du modèle qu'il veut construire,
- des informations sur l'ensemble des données disponibles.

En fait, les choses peuvent être plus avancées:

- Les données peuvent être directement disponibles, presque toujours sous forme de fichier informatique, mais pas nécessairement dans le format requis par le logiciel de modélisation.
- Les équations ont peut-être déjà été estimées, que ce soit sous forme de formules ou même d'articles, si la modélisation est le prolongement d'une étude économétrique.

Dans tous les cas, la première étape dans le processus devrait aboutir à:

- un ensemble d'équations entièrement déterminé, à l'exception des formules à estimer,
- l'ensemble correspondant de données.

Bien évidemment, ces deux tâches sont liées, car les équations sont établies sur la base des données disponibles, et les données sont produites en fonction des équations du modèle. Cela signifie qu'elles sont normalement traitées en parallèle. Cependant, il est tout à fait possible :

- De produire la plupart des données avant que les équations ne soient définies. Certains concepts (l'équilibre offre-demande aux prix constants et courants, l'emploi, les taux d'intérêt) apparaîtront certainement dans le modèle. Mais certaines variables spécifiques au modèle devront attendre.
- De produire la spécification du modèle avant que les données soient disponibles. De toute évidence, l'écriture d'une identité, ou l'indication d'une équation à estimer, ne requiert pas de données. Ce n'est que l'application (la vérification que l'identité est conforme avec les données, ou l'estimation de l'équation) qui le fait. Mais pour établir une équation il faut être raisonnablement sûr que les données seront disponibles, ou que l'on disposera d'une technique raisonnable pour l'estimer.

On peut même produire une première version du programme qui transformera en concepts du futur modèle les données d'origine, une fois que ces concepts sont complètement définis, ainsi que la définition des secondes.

On peut comparer la situation avec la construction d'une maison : on peut dessiner les plans avant que les matériaux ne soient achetés, mais leur disponibilité éventuelle (au bon moment) doit être certaine. Et les matériaux peuvent être achetés avant que les plans ne soient complètement établis (mais la probabilité d'avoir à les utiliser doit être assez élevée).

Ces options ne sont pas optimales dans le cas général, mais elles peuvent faire gagner du temps. La plupart des projets de modélisation sont confrontés à une échéance et, une fois que la force de travail est disponible, les tâches doivent être traitées dès que possible, si l'on veut avoir la meilleure chance de la respecter.

On peut s'interroger sur la possibilité de produire un jeu complet d'équations avant toute estimation. La technique que nous proposons consiste à remplacer les formulations futures par une « déclaration d'intention », qui stipule la variable à expliquer, et les éléments qui l'expliquent. Pour chaque équation, le format doit être aussi proche que possible de :

variable = f (liste des variables)

Les avantages de la définition d'un modèle complet sont nombreux :

- Le modélisateur sera en mesure de vérifier visuellement la logique de son modèle.

- Le texte peut être fourni à d'autres économistes pour obtenir des conseils.
- La liste complète des variables demandées peut être établie, permettant de produire un programme de transfert complet.
- Le traitement des équations par EViews donnera des informations intéressantes sur plusieurs éléments:
 - Les équations:

l'acceptabilité grammaticale d'équations sera vérifiée: par exemple, l'égalité du nombre de parenthèses gauches et droites ;

De plus, le fait que chaque variable endogène n'est calculée qu'une seule fois.

- Les variables:

l'information la plus importante viendra de la liste des exogènes: on peut y trouver des éléments qui auraient dû être déterminés par le modèle, selon sa logique. En général, cela signifie qu'on a oublié de préciser l'équation associée. En outre, certains éléments peuvent apparaître, qui ne devraient pas faire partie du modèle. Normalement, cela correspond à des erreurs de frappe.

- La structure par blocs:

EViews décompose l'ensemble des équations en une séquence de blocs, soit récursive (chaque variable du bloc ne dépend que des éléments précédents) ou simultanée (certaines variables sont utilisées avant d'avoir été calculées). Si les équations que l'on va réussir à estimer suivent la même logique, la structure par blocs décrite à ce stade sera déjà pleinement représentative de la future. On pourra donc détecter:

- Des simultanités anormales: une boucle causale peut apparaître, qui n'est pas acceptée par la théorie économique associée au modèle.
- Des liens récursifs anormaux: un bloc d'équations contenant une boucle théorique (la boucle prix salaire, la boucle keynésienne) peut apparaître comme récursif. Cela peut provenir d'une équation oubliée, d'une faute de frappe ...
- Dans tous les cas, l'observation de la structure causale du modèle donnera des informations préliminaires sur sa logique générale, et ses propriétés potentielles.

Toutes ces informations vont permettre d'éliminer à la source des erreurs dont la détection ultérieure conduira souvent à un coût non négligeable.

5.2 LA PREPARATION DU MODELE: ASPECTS SPECIFIQUES DES DONNEES

Détaillons le processus.

5.2.1 TYPES DE DONNEES

Dans le cas d'un modèle macroéconomique national, les données nécessaires peuvent représenter:

- Des éléments des comptes nationaux: opérations sur biens et services, transferts entre agents, mesurés en valeur, à prix constants, ou aux prix de l'année précédente. Le producteur sera généralement l'Office national de statistique. Pour la France, ce sera l'INSEE (Institut national de la statistique et des études économiques).
- Les déflateurs correspondants.
- Leurs équivalents étrangers, en utilisant le système de comptabilité et l'année de référence correspondante des pays particuliers, ou mieux une synthèse produite par un organisme international (OCDE, Fonds Monétaire International, Eurostat ...).
- Des variables présentant un plus grand détail, peut-être mesurées en quantités physiques (barils de pétrole, tonnes de riz). Elles peuvent provenir d'un organisme public ou privé, ou des producteurs eux-mêmes. Pour la France les éléments énergétiques viendront de l'Observatoire de l'Energie.
- Des données monétaires et financières, provenant principalement de la Banque nationale locale (en France, la Banque de France ou la Banque Centrale Européenne ...), d'une banque internationale (Banque Européenne pour la Reconstruction et le Développement, Banque Asiatique du Développement), ou du Fonds Monétaire International.
- Des données sur l'emploi ou le chômage. On peut obtenir des statistiques sur le travail détaillées (par âge, qualification, sexe ...) de l'US Bureau of Labor ou du ministère français du Travail.
- Des données démographiques: la population, la population en âge de travail, les classes d'âge (INSEE en France).
- Des données d'enquête: la croissance et les perspectives d'investissement selon les chefs d'entreprise, la capacité de production, les conditions de vie des ménages (en provenance d'instituts publics ou privés).
- Des éléments qualitatifs: le fait d'appartenir à un ensemble, répondant à une contrainte spécifique.

Les modèles microéconomiques utilisent généralement des données d'enquête (ménages, entreprises) avec parfois une dimension temporelle (panneaux, cohortes) et peuvent inclure certains des éléments ci-dessus ainsi que des indicateurs mondiaux.

Comme le champ d'application de modèles est quasiment illimité, le champ de données potentiellement pertinentes l'est également. Un modèle sur l'économie des transports comprendra des données techniques sur le système ferroviaire et sur les distances entre les villes, un modèle agricole des données météorologiques et des informations sur les variétés de plantes.

5.2.2 L'ACCES AUX DONNEES

Le moyen par lequel les données peuvent être obtenues jouera un rôle important. L'accès aux données nécessite de prendre en compte plusieurs caractéristiques:

- le mode de transmission,
- le format utilisé,
- les aspects institutionnels.

5.2.2.1 Le mode de transmission

Plusieurs options sont disponibles quant au mode de transmission, parmi lesquelles on peut citer :

5.2.2.1.1 La transmission physique

Les données peuvent être obtenues à partir d'un support physique, qu'il soit commercial ou créé à cet effet. Il peut s'agir d'un Compact Disk ou d'un DVD-ROM, ou de tout autre support réinscriptible comme une clé USB ou une carte mémoire. Par exemple, l'INSEE fournit des CD-ROM contenant les Comptes Nationaux complets.

5.2.2.1.2 La transmission par e-mail

Les fichiers peuvent être transférés d'un utilisateur à un autre par e-mail, en pièce jointe à un message.

5.2.2.1.3 Transmission par Internet

Les fichiers disponibles peuvent être téléchargés depuis un site web, commercial ou non. Le site de l'INSEE www.insee.fr permet l'accès à un ensemble de séries des Comptes Nationaux, dans des formats Excel et HTML.

La plupart du temps, ces fichiers sont fournis en format Excel, le plus fréquemment utilisé pour le traitement des données, et pour lequel chaque logiciel présent sur le marché fournit une interface simple.

Quant à la transmission entre utilisateurs, elle peut utiliser le transfert direct par FTP (File Transfer Protocol) utilisant un logiciel comme Filezilla, ou l'accès à un site spécialisé comme DropBox.

5.2.2.1.4 Autres media

Dans des cas de moins en moins fréquents, certaines données ne sont pas disponibles sous forme magnétique: séries obtenues dans les documents imprimés ou par fax, obtenues directement à partir d'autres experts, ou fixées par l'utilisateur (qui joue alors le rôle d'expert). Ces données devront généralement être saisies à la main, même si une interprétation directe par l'ordinateur grâce à la reconnaissance optique de caractères (OCR) est tout à fait opérationnelle (mais cette technique réclame des documents de bonne qualité).

Dans ce cas, il est essentiel de ne pas transférer les chiffres directement vers le fichier du modèle, mais de créer un fichier provisoire (comme une feuille de calcul Excel ou même un fichier ASCII) à partir duquel les informations seront lues.

5.2.2.2 Changement de format

Comme indiqué plus haut, le format des données d'origine est généralement différent de celui utilisé par le logiciel de création de modèle.

Dans le pire des cas, passer d'un logiciel à un autre va nécessiter la création d'un fichier intermédiaire dans un format donné que le logiciel de création de modèle peut interpréter. Le format Excel est l'intermédiaire le plus naturel, car il est lu et produit par tous les logiciels. Dans ce cas, il n'est pas nécessaire de posséder un exemplaire du logiciel pour utiliser son format.

Mais de toute façon, il est toujours possible de demander au premier programme de produire un fichier de caractères (dans la norme ASCII) qui peut, avec un minimum de corrections, être interprété par le deuxième programme comme la séquence d'instructions permettant de créer la série transférée, avec ses données et ses définitions¹.

¹ Par exemple, la séquence:

utiliser de 1970 à 2007

lire x

----- Valeurs

fin

Cependant, la situation s'est améliorée ces dernières années, car les logiciels fournissent de plus en plus souvent un accès direct aux formats utilisés par les logiciels les plus courants. Par exemple, EViews créera automatiquement un fichier de travail à partir d'une liste de 12 formats :

Accès. Mdb

Aremos. Tds

Dbase.dbf

Xls Excel.

Gauss. Dat

Givewin.im7.

HTML. Htm,. Html

Lotus 123. Semaines,. Wk1,. Wk3

MicroTSP. Wf

ODBC. Dsn,. Dqy

Rats. Rat,. Trl

SAS (divers)

SPSS. Sav,. Por

Stata. Dta

peut être traduite facilement par un traitement de texte en

smpl 1970 2007

x série

----- Valeurs;

Texte. Txt., Csv., Prn., Dat

TSP. C

5.2.2.3 Aspects institutionnels

Bien sûr, il faut aussi considérer la relation entre les producteurs de données et les instituts de modélisation. Les transferts les plus techniquement complexes ne vont pas nécessairement se produire entre des institutions distinctes. Un contrat commercial pourrait donner à l'institution de modélisation l'accès direct (via un modem ou l'accès à un réseau global) à l'information gérée par une société de production de données, sous le même logiciel, tandis qu'une grande institution pourra encore utiliser des CD-Roms comme un intermédiaire entre des unités séparées.

Cependant, il faut aussi considérer le coût de conclusion des marchés, y compris peut-être certains le troc entre les producteurs de données et les instituts d'étude.

5.2.2.4 Comment traiter des sources disparates

En règle générale, il convient de privilégier l'utilisation d'une seule source. Mais ce n'est pas toujours possible. Dans ce cas, il convient de définir une source primaire, et de tirer de la source alternative les seules séries supplémentaires. Les principaux problèmes peuvent provenir de:

- déflateurs et valeurs à prix constants utilisant une année de référence différente,
- données financières et d'emploi venant d'autres institutions que les comptes nationaux,
- variables mesurées en unités physiques (tonnes, mètres carrés) ayant dans le modèle des homologues en valeurs.

Dans tous ces cas, la priorité est la cohérence des équations du modèle, basée sur les données de la source primaire. Des résidus peuvent être introduits pour établir cette cohérence. Ceci s'applique en particulier au respect des équations d'équilibre (offre = demande), ou aux sommes (demande totale = somme de ses composants).

5.2.3 LA PREPARATION DES DONNEES POUR LE TRANSFERT

Nous allons maintenant définir la meilleure organisation (à notre sens) permettant de transférer des données de la source originale au logiciel (nous utiliserons Excel et EViews à titre d'exemple).

Nous devons garantir plusieurs choses:

- les données originales doivent rester disponibles en l'état;
- la base doit pouvoir être facilement mise à jour ;

- le transfert doit être aussi facile que possible.

Pour atteindre ces objectifs, la meilleure organisation doit être:

- copier le fichier original sous un autre nom,
- dans ce fichier, créer une nouvelle page,
- copier l'ensemble des séries originales dans cette page, en utilisant un lien (« copier avec liaison » sous Excel).

Nous supposons que les données d'origine sont organisées comme une matrice (ou un ensemble de matrices), avec des séries soit en lignes soit en colonnes. Sinon, une phase intermédiaire supplémentaire peut être nécessaire² ;

- insérer une ligne de noms de séries au-dessus des données de la première période (ou une colonne à gauche de la première colonne). La matrice peut ne pas démarrer dans la cellule B2,
- lire la feuille dans EViews (en utilisant l'importation ou la copie).

Cela garantit que:

- Les données d'origines ne sont pas modifiées.
- Les mises à jour sont simples: il suffit de copier la nouvelle page dans l'originale (et créer une nouvelle colonne dans la deuxième page si de nouvelles observations ont été publiées).

Le seul changement dans les programmes de transfert EViews concernera la période.

5.2.4 LE TRAITEMENT PRELIMINAIRE DES SERIES

Très souvent, la nature des séries disponibles n'est pas vraiment adaptée aux besoins du modèle. Un traitement préliminaire est alors nécessaire. Cela peut s'appliquer à plusieurs fonctions.

² Par exemple, les données trimestrielles peuvent apparaître dans des lignes annuelles de quatre colonnes.

5.2.4.1 Transformations temporelles

La plupart du temps le constructeur du modèle accédera à des séries ayant la bonne périodicité. Mais des exceptions individuelles peuvent se produire. La nouvelle série devra être calculée (à l'intérieur du logiciel de modélisation).

Le changement peut être entrepris dans deux directions: l'agrégation et la désagrégation.

5.2.4.1.1 Agrégation

Le cas le plus simple intervient si la périodicité disponible est trop courte. La nature de la variable conduira naturellement à une méthode d'agrégation, ce qui donne la valeur exacte de la série.

Si nous notons t la variable agrégée en t , et t, i la variable de la sous-période i en t , on peut considérer les techniques suivantes:

- somme, pour un flux (comme la production d'une branche).

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

- Moyenne, pour un niveau (comme le chômage pour une période donnée).

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

Première ou dernière valeur, pour un niveau à une date donnée (par exemple le capital productif, le premier jour d'une année ou d'un trimestre). Cette mesure s'appliquera aux variables «de stock».

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

or

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

5.2.4.1.2 Désagrégation

Lors du passage à une périodicité plus courte, le calcul nécessitera une approximation. On peut imaginer de:

- Diviser la valeur totale par le nombre de sous-périodes (pour un flux).

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

- Répartir sur chaque sous-période la variation totale observée sur la période (pour un niveau).

On pourra raisonner de façon additive

$$x_{t,i} = X_{t-1} + i/n \cdot (X_t - X_{t-1})$$

ou logarithmique

$$x_{t,i}/x_{t-1,n} = (X_t / X_{t-1})^{i/n}$$

- Filtrer ~~de~~ la série au cours de la période, de sorte que la série transformée soit une moyenne mobile de l'original.

Cependant, il faut être conscient que les deux dernières méthodes ne garantissent pas que la somme des valeurs détaillées soit égale à la valeur globale. La plupart des logiciels (y compris EViews) peuvent respecter cette contrainte sur demande (voir le manuel de l'utilisateur pour connaître les options disponibles).

5.2.4.1.3 Lissage

Le lissage représente un cas particulier: la préservation de la même périodicité que la série originale, mais avec la contrainte d'une évolution régulière, par exemple un taux de croissance constant. Au lieu de n valeurs libres, le choix est réduit à la valeur d'un (ou peut-être deux) paramètres.

Nous serons plus ou moins ramenés à la désagrégation, sauf que nous serons en mesure d'utiliser l'information supplémentaire. Par exemple, il est maintenant possible d'estimer la formulation à taux constant qui présente la plus petite distance à la série observée. Le respect des contraintes de cohérence, par exemple sur la moyenne des variables lissée et originale, conduira souvent à des méthodes d'estimation spécifiques (une régression simple établirait seulement cette contrainte sur les variables transformées, les taux de croissance par exemple).

Des techniques de lissage plus complexes peuvent être appliquées, les plus populaires étant les filtres d'Hodrick Prescott (pour les cas simples) et de Kalman (plus complexe).

5.2.4.1.4 Désaisonnalisation

Comme nous l'avons expliqué précédemment, une méthode permettant de traiter les variables présentant un caractère saisonnier est d'éliminer celui-ci, et de travailler avec des séries corrigées des variations saisonnières (ou CVS).

Plusieurs algorithmes peuvent être pris en considération. Les plus connus sont sans doute X-13ARIMA-SEATS géré par le Census Bureau des USA, et Tramo-Seats, géré par la Banque d'Espagne, tous deux disponibles sous EViews.

Évidemment, il convient de ne pas mélanger les séries originales et ajustées dans le même ensemble de variables du modèle.

5.2.4.2 Changement de nomenclature

Nous avons déjà abordé ce problème lorsque nous avons abordé les domaines de modèles.

La modification des catégories correspond habituellement à une agrégation. Dans le cas des modèles économiques, cette règle s'appliquera essentiellement:

- aux agents économiques: on peut séparer plus ou moins précisément les catégories des ménages (leur revenu, leur profession, leur taille, ...), les types d'entreprises (en fonction de leur taille, de la nature de leur production ...), les institutions gouvernementales (centrale et locale, la sécurité sociale, les agences spécifiques ...);
- aux produits (la production peut être décrite en plus ou moins de détail);
- aux opérations (on peut séparer les prestations sociales en fonction du risque, de l'agence qui les contrôle, ou ne considérer que la valeur globale);
- aux unités géographiques (un modèle mondial peut regrouper les pays en zones).

5.2.4.3 Transformations formelles

Certaines variables requises par le modèle ne seront pas disponibles en tant que telles, mais devront être calculées à partir de séries existantes, par une formule mathématique. Par exemple, le taux d'utilisation des capacités de production sera défini comme le rapport entre production effective et capacité, provenant éventuellement de sources différentes. Ou encore le coût relatif des salaires et du capital (utilisé pour définir le processus de production optimal) prendra en compte le prix de ces deux facteurs, mais aussi le taux d'intérêt, le taux d'amortissement, l'évolution attendue des salaires et des taux d'imposition.

5.2.5 MISES A JOUR

Une fois adaptées aux besoins du constructeur de modèles, les séries devront souvent être modifiées.

La modification des valeurs des séries existantes peut avoir plusieurs objectifs:

- la correction d'une erreur formelle, faite par le constructeur du modèle ou le producteur de la série: fautes de frappe ou erreurs de conception ;
- l'allongement de l'échantillon disponible: de nouvelles observations sont apparues pour les périodes les plus récentes ;
- l'amélioration de l'information: pour les dernières années connues, les dernières valeurs des séries de la comptabilité nationale française apparaissent successivement comme provisoires, semi-définitives et définitives ;
- la modification de la définition de certaines variables.

On peut aussi ajouter à la banque une toute nouvelle série

- qui est apparue récemment comme utile pour le modèle,
- qui a été mise à disposition par l'accès à une nouvelle source d'information, ou la création par les constructeurs de données d'un nouveau concept, plus intéressant.

Cette multiplicité des changements possibles interdit à l'ensemble global des séries utilisées par le modèle de rester stable, y compris pour une courte période. Adapter en permanence les spécifications du modèle (en particulier les équations estimées) à cette évolution demanderait beaucoup du modélisateur, au détriment de tâches plus productives. Cela signifie qu'il faut limiter la fréquence des mises à jour, pour l'ensemble des éléments opérationnels (par exemple une ou deux fois par an pour un modèle annuel, ou tous les trimestres pour un trimestriel), à quelques exceptions: la correction des fautes graves ou l'introduction d'informations vraiment importantes.

Sans aucun doute, la meilleure solution est de gérer en fait deux ensembles de données, une mise à jour assez fréquente intégrant les dernières valeurs, et une autre construite à intervalles plus longs (la périodicité du modèle par exemple). Cette solution permet d'étudier à l'avance, par des estimations basées sur le premier ensemble, les conséquences de l'intégration de nouvelles valeurs sur les spécifications et les propriétés de la version suivante du modèle.

5.2.6 SUPPRESSIONS

Il est bénéfique de supprimer dans la base les séries qui sont devenues inutiles:

- **cela permet de gagner de l'espace,**
- les recherches seront plus rapides,
- la banque sera plus cohérente avec le modèle,
- le constructeur du modèle aura moins d'informations à mémoriser, et l'architecture de la banque sera plus facile à maîtriser (on aura deviné que c'est l'élément le plus important, à notre sens).

Les séries inutiles conservées trop longtemps vont conduire à l'oubli de ce qu'elles représentent, et leur destruction réclamera un processus d'identification fastidieux.

Pour EViews, la suppression présente un intérêt supplémentaire: les éléments du fichier de travail étant affichés dans une fenêtre unique de taille limitée, il est essentiel que cette fenêtre concentre autant d'éléments intéressants que possible.

5.2.7 LA DOCUMENTATION

De même, l'investissement dans la documentation des séries est rapidement amorti. Il peut concerner:

- la définition, peut-être à deux niveaux: un court pour afficher les titres dans les tableaux ou les graphiques, et un plus long pour décrire complètement le concept utilisé ;
- la source: fichier (et feuille) d'origine, l'institution productrice et peut-être la façon de contacter l'auteur ;
- les unités dans lesquelles la série est mesurée ;
- des remarques supplémentaires, telles que la qualité et le statut (final, provisoire, estimé) de chaque observation ;
- la date de production et la dernière mise à jour (heures voire minutes, ce qui peut servir à déterminer exactement quel ensemble de valeurs une application a utilisées). Cette information est souvent enregistrée automatiquement par le logiciel ;
- si cela est pertinent, la formule utilisée pour le calculer.

Exemple: Taux de salaire = salaire versé / (emploi x nombre de semaines de travail x durée du travail hebdomadaire).

EViews permet de spécifier les quatre premiers types, en utilisant la commande «label», et produit automatiquement les deux derniers.

Par exemple, une série appelée PIB peut être définie par la séquence:

```
GDP.label (c)
GDP.label (d) Produit intérieur brut à prix constants
GDP.label (u) En Euros 2005
GDP.label (s) à partir des fichiers Excel accounts.xls produits par l'Office statistique
GDP.label (r) 2011T4 est provisoire
```

qui efface le contenu précédent, donne la définition, décrit les unités, la source, et ajoute des remarques.

EViews 8 permet à l'utilisateur d'introduire ses propres labels. Ceci peut être utilisé pour caractériser les éléments du modèle, par exemple le pays pour un modèle multinational, l'agent en comptabilité nationale, le type (endogène / exogène), ou tout simplement le fait que la série est utilisée par un modèle particulier. Vous avez simplement à appliquer spécifier une déclaration ????? ordre «label» appliqué à une nouvelle catégorie. Aucune déclaration préalable n'est nécessaire. .

Par exemple, vous pouvez utiliser:

```
HI.label (agent) Ménages
MARG.label (agent) Entreprises
```

Une fois qu'un type a été créé, il doit être déclaré dans son intégralité (vous ne pouvez pas utiliser «a» pour «agent»).

Si l'écran de la fenêtre workfile est en mode «Display +», vous pouvez maintenant trier les éléments en fonction de leurs caractéristiques. Outre le nom, le type et la date de dernière modification (ou création) vous avez accès à la description.

En définissant les éléments selon un certain code, vous pouvez obtenir un affichage dans lequel les éléments liés apparaîtront ensemble, dans l'ordre alphabétique.

Par exemple en utilisant:

```
HI.label(d) Modèle Fra / Ménages: revenu disponible  
MARG.label(d) Modèle Fra / Entreprises: Marges
```

Lors d'un tri sur la description, les variables du modèle seront identifiées et réparties en fonction de l'agent.

Par défaut, l'écran présentant les éléments ligne à ligne (obtenu par appel à «Details+» dans la barre supérieure) se limite aux seules descriptions. Pour accéder aux autres éléments (et donc au tri) sur les autres éléments, il convient de cliquer avec le bouton droit sur la barre supérieure (par exemple sur «Description»). Ceci fait apparaître un menu contextuel où l'on sélectionnera l'élément «Edit Columns...» (le dernier). On peut alors choisir les éléments affichés sur la ligne (y compris ceux créés par l'utilisateur).

5.2.8 LA COMPARAISON DES DONNEES SOUS EVIEWS 8

Vous pouvez comparer les éléments entre fichiers de travail et pages à l'intérieur du même fichier de travail. EVIEWS affiche une ligne par élément, dans lequel sera indiqué son statut, entre: inchangé, modifié (numériquement), ajouté, supprimé, remplacé (logiquement, ce dernier cas s'applique par exemple à une variable liée dont le lien a été modifié). Un filtre peut être appliqué.

Pour les séries, un niveau de tolérance peut être défini, en deçà duquel la série ne sera pas considérée comme modifiée. L'affichage indiquera le nombre de périodes montrant une différence supérieure.

Par défaut, tous les éléments seront affichés, mais on peut limiter les cas (par exemple, pour toutes les variables présentes dans les deux pages avec une différence supérieure au critère).

Les équations et les modèles ne sont pas comparés mais apparaissent dans la liste.

La commande associée est **wfcompare**.

```
wfcompare (tol = critère, list = type-de-comparaison) liste-des series-de-référence liste-des-  
séries- comparées
```

Pour plus de détails veuillez vous reporter à l'Aide EVIEWS.

Par exemple, si vous voulez comparer toutes les séries françaises (à commencer par «FRA_») entre les pages «base» et «updated», pour un niveau de tolérance de 0,00001, vous indiquerez:

```
wfcompare (tol = 1E-5, list = m) updated\fra_* base\ fra_*
```

Cette commande peut être particulièrement utile pour

- contrôler l'évolution des valeurs historiques pour un ensemble, identifiant les équations qui devront être estimées à nouveau avec les données du modèle ;
- résumer les résultats d'une vérification résiduelle, montrant pour quelles équations le terme de droite (en utilisant les valeurs historiques de la variable expliquée) est différent du terme de gauche (le résultat du calcul). En fixant un seuil de tolérance légèrement supérieur à zéro (par exemple 0,0001) on peut limiter l'affichage aux erreurs jugées significatives.

5.2.9 CONSEQUENCES SUR L'ORGANISATION DU TRAVAIL

Donnons maintenant quelques considérations spécifiques sur la gestion des données.

Dans le cas général, les constructeurs de modèles seront confrontés à un grand nombre de séries de plus ou moins diverses origines. La stratégie de gestion optimale pourrait sembler varier en fonction des cas, mais en fait elle est unique dans sa caractéristique principale: il faut produire un fichier unique, dans le standard du logiciel de modélisation, et contenant toutes les séries ayant une chance d'être utiles.

Cela est vrai même si l'ensemble des séries nécessaires est produit et géré sur le même ordinateur ou un réseau informatique, en utilisant le même logiciel (la tâche de transfert sera tout simplement facilitée): il est essentiel que le constructeur du modèle contrôle les séries qu'il utilise, et surtout qu'il gère leurs changements (en particulier les mises à jour des séries d'usage courant). Lors de l'interprétation d'un changement dans les propriétés du modèle (simulations, estimations), il doit être capable de rejeter la modification des données source, sauf si celle-ci a été introduite en connaissance de cause par le constructeur du modèle lui-même³.

³ Cette remarque est une application particulière du principe général « évitons les problèmes potentiels qui peuvent se révéler coûteux en temps de réflexion ».

Une telle organisation facilite également la gestion des séries. En particulier, la limitation du nombre de séries dans la banque, en dehors du fait qu'elle permettra de gagner du temps et de l'espace informatique, rendra l'ensemble plus facile à manipuler intellectuellement.

En ce qui concerne le champ des séries, deux options extrêmes peuvent toutefois être prises en considération⁴:

- transférer dans la banque du modèle l'ensemble des séries qui ont une chance (même petite) de devenir utiles à l'élaboration du modèle ;
- transférer le minimum, puis compléter l'ensemble en fonction des besoins.

Si une solution médiane peut être envisagée, le choix penche fortement en faveur de la première solution. Elle sera peut-être plus chère au départ, en temps humain, et en taille des fichiers, mais elle se révélera généralement un bon investissement, car elle évite souvent un grand nombre de transferts limités coûteux, et donne une certaine stabilité à la banque ainsi qu'à ses procédures de gestion.

5.2.10 LES OPTIONS PRATIQUES

Pour les modèles gérés par des institutions, l'organisation la plus fréquemment rencontrée est une équipe qui travaille sur un réseau local ou des ordinateurs connectés, où l'on peut transférer une sélection de données provenant de sources éloignées. On pourra, dans certains cas, accéder directement au site du fournisseur, de l'intérieur d'une session de création de modèle. Les producteurs de logiciels de modélisation attribuent une priorité importante à ce type d'option.

Pour des équipes importantes ou travaillant dans des endroits différents, la communication n'est plus un problème, si l'accès Internet est disponible. Une connexion Internet ou Intranet est même généralement utilisée par les économistes travaillant dans le même bâtiment.

Il faut cependant faire attention aux incompatibilités de format, en particulier si le système d'exploitation est différent (Windows et ses versions, Linux, UNIX, Macintosh ...) ⁵.

⁴ Même si elles ne sont pas considérées comme de véritables variables du modèle. Par exemple, on peut être intéressé par la comparaison du ratio capital-travail du pays modélisé avec celui d'autres pays.

⁵ En fait, ce problème apparaît assez rarement, comme la plupart des logiciels de modélisation (sauf peut-être Troll) ne fonctionnent que sous Windows.

5.2.11 RETOUR A NOTRE EXEMPLE

Maintenant que nous avons établi les principes, nous allons voir comment les appliquer au cas que nous avons défini plus tôt. Pour éviter des allers retours entre des pages lointaines, nous allons répéter sa présentation.

- Sur la base de leurs prévisions de production et de la productivité des facteurs, les entreprises investissent et embauchent des travailleurs afin d'adapter leur capacité de production. Toutefois, elles exercent une certaine prudence dans ce processus, car elles ne veulent pas se retrouver avec des éléments inutilisés.
- Les niveaux atteints dans la pratique permettent de définir le potentiel de production.
- Les entreprises constituent également des stocks.
- Les ménages obtiennent des salaires, sur la base de l'emploi total (comprenant les fonctionnaires), mais aussi une part du produit intérieur brut. Ils consomment une partie de ce revenu.
- La demande finale est définie comme la somme de ses composantes: la consommation, l'investissement productif, l'investissement logement, les variations de stocks et la demande publique.
- Les importations représentent une part de la demande locale («la demande intérieure»). Mais moins il reste de capacités disponibles, plus une augmentation de la demande fera appel aux importations.
- Les exportations suivent la demande mondiale, mais les producteurs sont limités par les capacités disponibles, et leur priorité est de satisfaire la demande locale.
- L'offre est égale à la demande.
- Le capital productif augmente avec l'investissement, mais il est soumis à une certaine dépréciation.

Nous avons volontairement gardé un cadre simple (peut-être pas assez), car notre but actuel est seulement explicatif. Toutefois, le modèle que nous construisons a une certaine cohérence économique, et peut réellement représenter un noyau pour d'autres extensions que nous présenterons plus tard.

Nous supposons également que les données suivantes sont disponibles dans un fichier Excel appelé FRA.XLS, choisies parmi les données de l'OCDE «Perspectives économiques d'ensemble». Ces séries sont disponibles à partir du premier semestre de 1970 jusqu'à la fin 2004.

La raison du préfixe «FRA» est d'identifier les séries relatives à la France parmi un grand nombre de pays, représentant tous les membres de l'OCDE ainsi que certains regroupements.

Ils utilisent les unités suivantes:

Valeurs: Euros

Déflateurs: base 100 en 1995.

Volumes (ou quantités): Millions d'Euros de 1995

Populations: personnes

FRA_CGV	Consommation du gouvernement
FRA_CPV	Consommation privée
FRA_ET	Emploi total
FRA_CPV	Consommation privée
FRA_EG	Emploi, Gouvernement
FRA_FDDV	Demande intérieure finale
FRA_GAP	Écart entre production effective et potentielle
FRA_GDPTR	Production potentielle, de l'économie totale
FRA_GDPV	Produit intérieur brut (prix du marché)
FRA_IBV	Formation brute de capital fixe
FRA_ICV	Consommation intermédiaire
FRA_IGV	Investissement du gouvernement
FRA_IHV	Investissement en logement
FRA_ISKV	Variations de stocks
FRA_KBV	Stock de capital des entreprises
FRA_MGSV	Importations de biens et services
FRA_PCP	Déflateur de la consommation privée
FRA_TDDV	Demande intérieure totale
FRA_WSSS	Rémunération des employés, à prix courants
FRA_XGSV	Exportations de biens et services
FRA_XGVMKT	Exportations de biens et services, marché potentiel
FRA_YDRH	Revenu disponible réel des ménages, prix courants

Toutes les opérations sont mesurées en volume, sauf indication contraire.

En appliquant les principes que nous avons définis ci-dessus, nous allons:

- définir le cadre du modèle,
- identifier ses variables,
- les séparer entre endogènes et exogènes,
- spécifier complètement les identités,
- établir chaque équation de comportement comme une identité, présentant de la façon la plus simple la variable qu'elle définit, et ses éléments explicatifs,
- créer les séries associées, à partir des données disponibles,

- transférer les éléments déjà disponibles dans des séries du modèle en utilisant les noms que nous leur avons attribués,
- spécifier les formules du calcul des éléments restants.

Maintenant que nous avons obtenu les données, nous pouvons passer aux deux tâches suivantes: les transformer pour répondre aux besoins du modèle, et commencer à préciser les équations.

Il devrait être clair que cela va se faire, non pas «à la main», mais grâce à un ensemble d'instructions enregistrées dans un langage lisible (un programme). Cette option permettra:

- d'établir un ensemble apparemment cohérent d'instructions, qui peuvent être contrôlées visuellement ;
- de localiser les erreurs et introduire les corrections le plus simplement et le plus clairement possible ;
- de conserver des versions successives, dont la dernière est la plus correcte, jusqu'à ce qu'une version satisfaisante soit établie ;
- de répéter le processus au moindre coût ;
- d'afficher les étapes du processus aussi clairement que possible, par l'introduction de commentaires ;
- une fois qu'un stade satisfaisant a été atteint, de mémoriser les actions pour une utilisation ultérieure (surtout si le projet de modélisation doit faire face à des interruptions, plus ou moins longues) ;
- de permettre aux utilisateurs extérieurs de maîtriser le stade actuel des opérations, afin d'évaluer l'état actuel du développement du projet.

Des commentaires peuvent être insérés dans le programme, ce qui rend la séquence des tâches et le rôle des différentes commandes plus clairs, et permet d'avertir de la présence de problèmes locaux et la façon dont ils ont été traités. Ceci est particulièrement utile pour un projet d'équipe. Le nom de l'auteur doit alors également être inclus.

Sous EViews, deux autres méthodes sont disponibles:

- utiliser une séquence de menus et sous-menus,
- saisir les commandes sans les enregistrer, directement à partir de la fenêtre de commande.

Ces deux méthodes échouent sur tous les critères. La séquence des tâches n'est pas disponible, ce qui signifie que les erreurs sont difficiles à détecter. Reproduire la tâche, que ce soit pour corriger des erreurs

ou mettre à jour les spécifications ou les données, doit faire appel à une nouvelle séquence de menus ou une nouvelle saisie⁶.

En fait, la seule option raisonnable est celle que nous avons proposée ci-dessus: la définition d'un programme produisant toutes les informations nécessaires et le cadre des équations qui vont les utiliser. Mais l'ordre des tâches peut être remis en question, comme nous avons commencé à l'expliquer plus tôt. Jusqu'à ce que les deux tâches soient terminées, elles sont techniquement indépendantes: on n'a pas besoin du modèle physique pour créer les données, ou des valeurs des séries pour spécifier les équations. Cela signifie que l'on peut envisager deux méthodes extrêmes:

- création de toutes les données nécessaires par le modèle, puis spécification du modèle ;
- spécification de toutes les équations du modèle, puis production des données associées.

Le critère est la possibilité technique.

Clairement la première option n'est pas réaliste, car l'écriture des équations révélera certainement le besoin d'éléments supplémentaires. La seconde est plus réalisable, car on n'a pas besoin des valeurs des séries pour écrire une équation. Mais au fur et à mesure que les équations sont définies, il faut vérifier que tous les éléments qu'elles contiennent sont ou seront disponibles sous la forme requise, que ce soit en tant que concepts réels (PIB) ou comme des transformations des concepts (le déficit budgétaire en points de PIB réclame les séries de déficit et de PIB). Si un concept semble faire défaut, on devra utiliser un élément alternatif disponible («proxy»), établir une hypothèse, chercher un remplacement dans les séries disponibles, ou éliminer l'élément du modèle.

Cela montre que si la production des séries peut être faite dans n'importe quel ordre, il vaut mieux commencer par la spécification des équations. Si les données ne sont pas prêtes, mais leur définition est déjà connue, il est possible d'écrire les équations et de demander au logiciel d'interpréter le texte. L'utilisateur sera prévenu des erreurs de syntaxe possibles, de la nature des variables (endogènes / exogènes), et de l'architecture de son modèle. Cela permettra les premières corrections du modèle, faisant gagner du temps et évitant de prendre déjà de mauvaises directions. Et si certaines parties du modèle sont encore en chantier, il est possible de construire une première version de l'ensemble de données déjà établi, qui sera mise à jour lorsque le modèle sera complet.

⁶ Toutefois, à partir d'EViews 7, on peut transférer la séquence des instructions spécifiées dans la fenêtre de commande vers un fichier programme.

Dans la pratique, en particulier dans les cas les plus simples, on peut aussi commencer à définir le programme en constituant deux paragraphes vides et en les remplissant progressivement avec des données et des équations, jusqu'à ce que les deux tâches soient terminées. Les huit paragraphes originaux dans nos spécifications du modèle peuvent être traités un par un de cette façon (pas nécessairement dans l'ordre présenté).

En définitive, parmi les propositions ci-dessus nous privilégions deux techniques alternatives:

- **Modèle puis données:** spécifier d'abord le modèle complet, en vérifiant que tous les éléments utilisés pourront être produits soit directement soit par l'intermédiaire d'une formule. Puis produire l'ensemble de données, de préférence par le biais d'un transfert direct ou une transformation.
- **Modèle et données:** produire les équations dans l'ordre, ou bloc par bloc d'éléments liés, et établir simultanément les déclarations qui créent toutes les séries dont ils ont besoin.

5.2.12 APPLICATION A NOTRE EXEMPLE

Montrons maintenant sur notre exemple comment le processus peut être effectué en utilisant la seconde méthode, sans doute la plus adaptée à un si petit modèle.

Nous allons d'abord présenter le processus en termes généraux (non EViews), en traitant chaque cas en séquence, et présentant à la fois les équations et les instructions générant les variables associées. Pour que les choses soient plus claires, les équations seront numérotées et les instructions de création commenceront par le symbole «>>»

En outre, les variables endogènes utiliseront des caractères majuscules, les exogènes des minuscules⁷.

- Sur la base de leurs prévisions de production et de la productivité des facteurs, les entreprises investissent et embauchent des travailleurs.

Ceci définit deux équations de comportement définissant la demande des facteurs, l'emploi (que nous appellerons LE) et l'investissement (appelé I) dépendant du PIB, appelé Q.

$$\begin{aligned}LE &= f(Q) \\ I &= f(Q)\end{aligned}$$

⁷ Ceci n'a aucun effet sur le résultat des instructions, il s'agit seulement de clarifier le processus pour l'utilisateur et ses partenaires.

Nous avons besoin de:

```
>> IP = FRA_IBV
>> Q = FRA_GDPV
```

Mais pour LE, nous sommes confrontés à notre premier problème. L'emploi privé n'est pas directement disponible. Cependant, nous avons supposé que les seuls employeurs étaient le secteur public (gouvernement) et les entreprises. Cela signifie que nous pouvons utiliser:

```
= >> LE=FRA_ET-FRA_EG
```

- Le capital productif augmente avec l'investissement, mais il est soumis à l'amortissement.

Le capital K , mesuré à la fin de la période, est défini par une identité. Au niveau initial, on applique un taux d'amortissement (appelé dr) et on ajoute l'investissement. L'équation s'écrit:

$$K(t) = K(t-1) \cdot (1-dr(t)) + I(t)$$

Le définir au début de la période ne changerait quasiment pas les notations.

Nous avons besoin des données pour K

```
>> K = FRA_KBV
```

et nous obtenons dr en inversant la formule:

```
>> dr = ((K (-1) + IP) - K) / K (-1)
```

En d'autres termes, dr sera la division, par rapport au niveau de capital initial, de la différence entre deux niveaux de capital: la valeur que nous aurions obtenue sans dépréciation, et la valeur observée.

Les niveaux effectivement atteints définissent le potentiel de production.

- La capacité (appelé CAP) dépend des facteurs LE et K.

$$\text{CAP}(t) = f(\text{LE}(t), \text{K}(t))$$

Elle peut être calculée directement comme:

$$\gg \text{CAP} = \text{FRA_GDPVTR}$$

ce qui représente plutôt une valeur « normale » du PIB compte tenu du niveau actuel des facteurs.

La disponibilité directe de ce concept comme une série représente le meilleur des cas, mais n'est pas souvent observée dans la pratique. Plus tard dans le texte nous allons aborder les techniques alternatives disponibles dans des situations moins favorables.

- Les entreprises ont besoin d'intrants, et aussi de constituer des stocks.

La consommation intermédiaire peut être définie comme proportionnelle au PIB, en utilisant la valeur à prix constants. Cela signifie que, pour n'importe quel niveau de la production, chaque unité produite aura besoin de la même quantité de produits intermédiaires.

Nous appliquerons donc un facteur exogène à Q.

$$\text{IC} = r_{\text{icq}} \cdot \text{Q}$$

Pour les stocks, nous allons estimer leur évolution:

$$\text{IC} = f(\text{Q})$$

Pour cela, nous devons calculer:

$$\gg \text{IC} = \text{FRA_ISKV}$$

```
>> r_icq = IC / Q (ou FRA_ISKV / FRA_GDP)
>> CI = FRA_CIV
```

- Les ménages obtiennent des salaires, sur la base de l'emploi total (comprenant les fonctionnaires) et une part du produit intérieur brut. Ils consomment une partie de ce revenu.

Maintenant nous avons besoin de définir l'emploi total, en ajoutant l'emploi du gouvernement (appelé ig) à LE .

$$LT = LE + ig$$

Les nouvelles séries sont obtenues par:

```
>> LT = FRA_ET
>> Ig = FRA_EG
>> LE = LT - Ig
```

Nous devons maintenant calculer les revenus des ménages, que nous appellerons R_{HI} . Nous supposons que le salaire est le même pour tous les travailleurs, et que le revenu non salarial des ménages représente une part donnée du PIB, une série appelée r_{rhiq} . Ceci donne:

$$R_{HI} = wr. LT + r_{rhiq}. Q$$

En fait, l'hypothèse ci-dessus, bien que simpliste, n'est probablement pas très éloignée de la vérité. La sensibilité au PIB des éléments inclus dans ce concept hétérogène peut être faible (comme les retraites ou les intérêts des obligations à long terme), élevée (les revenus des propriétaires de petites entreprises, avec des coûts fixes et donc des marges très variables) ou moyenne (travailleurs individuels travaillant comme des salariés).

- La consommation des ménages est donnée par l'application à R_{HI} d'un taux d'épargne que nous appellerons sr . Pour le moment, le taux d'épargne est exogène.

$$CO = RHI. (1 - sr)$$

- L'investissement en logements est également une part de RHI, que nous appellerons r_{ih} .

$$IH = r_{ih}. RHI$$

Les nouvelles variables sont RHI, wr , r_{rhiq} , sr , IH et r_{ih} .

RHI est donné simplement par:

$$\gg RHI = FRA_YDRH$$

Nous devons maintenant calculer le taux de salaire réel wr . Cela se fait par le calcul suivant.

Diviser FRA_WSSS par FRA_ET donne la valeur nominale individuelle, que nous divisons à nouveau par $FRA_CPI/100$ ⁸ pour obtenir la valeur réelle⁹.

$$\gg wr = (FRA_WSS / FRA_ET) / (FRA_CPI/100)$$

r_{rhi} sera obtenu comme le rapport au PIB du revenu des ménages diminué des salaires

$$\gg r_{rhi} = (RHI - Wr LT) / Q$$

⁸ Les déflateurs produits par l'OCDE prennent la valeur 100 en 1995, l'année de base.

⁹ Compte tenu de la liste ci-dessus des séries disponibles, on peut observer que d'autres options sont possibles.

La consommation et l'investissement en logements seront obtenus directement:

```
>> CO = FRA_CPV  
>> IH = FRA_IHV
```

Le calcul du taux d'épargne et r_{ih} utilisera l'inversion de l'équation associée:

```
>>sr = (RHI-CO) / RHI
```

ou

```
>>sr = (FRA_YDRH-FRA_CPV) / FRA_YDRH
```

(l'épargne divisée par le revenu)

```
>> r_{ih} = IH / RHI
```

ou

```
>> FRA_IHV / FRA_YDRH
```

La demande finale est la somme de ses composantes: la consommation, l'investissement productif, l'investissement immobilier, les stocks et la demande du gouvernement. La demande totale comprend les consommations intermédiaires.

```
FD = IP + CO + H + gd + CI  
TD = FD + r_{ic}. Q
```

Nous avons besoin de calculer gd comme la somme de FRA_IGV et FRA_CGV .

```
>> gd = FRA_IGV + FRA_CGV
>> FD = FRA_TDDV
>> r_icq = FRA_ICV / FRA_GDPV
```

Les importations représentent une part de la demande locale (la «demande intérieure»). Mais moins il reste de capacités disponibles, plus une demande supplémentaire devra être importée.

Cela nécessite:

```
UR = Q / CAP
M = f (FD + IC, UR)
```

Nous avons besoin de calculer:

```
>> UR = Q / CAP (sa définition)
>> M= FRA_MGSV
```

Les exportations dépendront essentiellement de la demande mondiale. Mais nous allons aussi supposer que si les tensions apparaissent (par UR) les entreprises locales vont transférer une partie de leur production vers la demande locale, et être moins dynamiques dans leur recherche de contrats à l'étranger.

```
X = f (WD, UR)
```

Nous avons besoin de:

```
>> X = FRA_XGSV
>> WD = FRA_XGVTR
```

L'offre est égale à la demande.

L'équation offre et demande va pour le moment utiliser la formule implicite ci-dessous:

$$Q + M = X + FD$$

(Toutes les valeurs des variables ont été obtenues plus tôt)

Nous pouvons maintenant réorganiser le cadre de notre modèle selon le schéma suivant:

$$[1] LE = f(Q)$$

$$[2] IP = f(Q)$$

$$[3] K = K-1 (1-DEPR) + IP$$

$$[4] CAP = f(LE, K-1)$$

$$[5] IC = r_{icq} \cdot Q$$

$$[6] CI = f(Q)$$

$$[7] LT = LE + I_g$$

$$[8] RHI = w_r RHI \cdot LT + r_{rhiq} \cdot Q$$

$$[9] CO = (1-r_s) \cdot RHI$$

$$[10] = I_H r_{ih} \cdot RHI$$

$$[11] FD = CO + H + IP + CI + gd$$

$$[12] TD = FD + r_{ic} \cdot Q$$

$$[13] UR = Q / CAP$$

$$[14] M = f(TD, UR)$$

$$[15] X = f(w_d, UR)$$

$$[16] Q = M + X + FD$$

Les variables endogènes

I	Investissement des entreprises
LE	Emploi des entreprises
K	Capital productif des entreprises

CAP	Capacités de production des entreprises
LT	Emploi total
CI	Variation des stocks
IC	Consommations intermédiaires
IH	Investissement en logements
CO	Consommation des ménages
FD	Demande finale locale
TD	Demande totale locale
M	Importations
RHI	Revenu des ménages réel
UR	Taux d'utilisation des capacités
X	Exportations françaises
Q	Produit intérieur brut

Les variables exogènes

dr	taux d'amortissement du capital
gd	consommation et l'investissement de l'État
lg	emploi public
r_ih	ratio de l'investissement logement des ménages au revenu
r_rhiq	part du PIB transférée aux ménages, en plus du salaire
wd	demande mondiale adressée à la France
r_icq	ratio de la consommation intermédiaire au PIB
wr	taux de salaire réel moyen

On observe que:

- le nombre d'équations est aussi celui des variables à calculer,
- nous avons séparé les équations de comportement et les identités,

- les identités comptables sont complètement définies,
- la forme des équations de comportement est encore indéterminée, bien que leurs éléments explicatifs soient connus (au moins comme une première version).

Cette distinction est normale. Comme nous l'avons déjà indiqué, les identités représentent généralement un lien formel obligatoire, tandis que le lien entre équations de comportement et théorie économique n'est pas aussi restrictif.

5.2.12.1 Les formules de calcul

En observant les formules que nous avons obtenues, nous pouvons voir que la plupart des données nécessaires sont disponibles directement, donc un simple transfert devrait être suffisant. On aurait même pu envisager d'utiliser les noms d'origine. Mais comme notre modèle s'applique uniquement à la France, il n'y a aucune raison de garder le préfixe, qui a permis d'identifier les données françaises au sein d'un fichier multi pays beaucoup plus important. Et on peut décider (à juste titre à notre sens) que les noms que nous avons choisis sont plus clairs.

Les correspondances sont les suivantes :

```
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
```


Seulement huit éléments font défaut, sept d'entre eux correspondant à des variables exogènes:

$gd = FRA_CGV + FRA_IGV$	Demande de l'Etat
$UR = Q / CAP$	Taux d'utilisation des capacités
$dr = ((K(t-1) + IP) - K(t))$	Taux d'amortissement du capital
$r_ic = IC / Q$	Ratio de la consommation intermédiaire par rapport au PIB
$r_ih = IH / RHI$	Part de l'investissement logement dans le revenu
$r_rhiq = (. RHI - wr LT) / Q$	Part non salariale du PIB salarié dans le revenu des ménages
$sr = (RHI-CO) / RHI$	Taux d'épargne
$wr = (FRA_WSS / FRA_ET) / (FRA_CPI/100)$	Taux de salaire réel moyen

Dans les cas réels, ce type de calcul sera souvent utilisé. Il faut être conscient **d'un élément important**:

L'utilisation de ces formules est distincte de la définition des équations du modèle. La seule raison pour laquelle nous en avons besoin est de produire les valeurs **historiques** des séries, quand celles-ci ne sont pas encore disponibles. Si les statisticiens avaient fait un travail complet (et s'ils connaissaient les exigences du modèle) ils en auraient fourni l'ensemble, et aucun calcul n'aurait été nécessaire (peut-être seulement quelques changements de noms).

Ainsi, ces deux types de formules ont des objectifs totalement différents

- Appliquer les **instructions de calcul** garantit que toutes les données requises seront disponibles. En associant des formules aux éléments manquants, elles permettent de produire l'ensemble nécessaire à la simulation et l'estimation. Si les données étaient déjà disponibles dans le bon format, et les noms donnés aux variables étaient acceptables, aucune déclaration ne serait nécessaire. Et on peut vérifier que, dans notre cas, la plupart des calculs sont en fait des transferts directs, qui permettent de créer un élément de modèle tout en conservant la série originale.

En fait, on peut s'interroger un instant sur la nécessité de disposer d'un ensemble complet de valeurs historiques des variables endogènes. Celles-ci ~~seront-vont être~~ calculées par le modèle, qui est simulé sur l'avenir de toute façon. Les raisons ~~de la construction~~ **l'établissement** d'un ensemble complet sont les suivantes:

- L'estimation aura besoin de tous les éléments des équations associées.

- Le contrôle de la cohérence des identités avec les données est un préalable avant toute simulation, pour éviter de commencer avec un ensemble d'éléments défectueux¹⁰.
- Pour vérifier que le modèle donne des simulations précises sur le passé, on aura besoin de tous les éléments historiques.
- De nombreuses équations utilisent des valeurs retardées. Cela nécessite des valeurs pour les périodes précédant le début des prévisions.

Ces formules peuvent inclure des données originales, des données transformées calculées tôt dans le programme, ou tout simplement des hypothèses. Par exemple:

- Le PIB a été tiré directement de la série originale.
- Le taux de dépréciation s'obtient à partir du capital et de l'investissement.

En l'absence d'autres informations, la cible de l'inflation peut être fixée à 2 %.

Les **équations du modèle** établissent un lien logique entre les éléments, qui sera utilisé par le modèle pour produire un équilibre cohérent. Cela signifie que si la formule de calcul de la variable A contient la variable B, la variable A est supposée dépendre de B, en termes économiques.

C'est évidemment le cas pour les équations estimées. Par exemple, le taux de salaire peut dépendre de l'inflation, et les exportations de la demande mondiale. Mais cela est également vrai pour les identités.

Le revenu des ménages est la somme de ses éléments. Si l'un change, le revenu va changer de la même manière (*ex ante*, bien sûr). Fondamentalement, nous supposons que les comportements s'appliquent de la même manière à tous les éléments du revenu, quelle que soit leur source.

Si la consommation des ménages est estimée, l'épargne est la différence entre les recettes et la consommation.

Il est extrêmement important d'assimiler cet aspect des choses, au début de tout projet de modélisation.

Il est tout à fait possible, cependant, que la même formule soit utilisée dans les deux situations. Par exemple nous pourrions ne pas disposer de valeurs pour FD, et nous pensons que CO, I, IH et gd représentent la totalité de ses composantes. Dans ce cas la formule:

¹⁰ En particulier, l'incohérence des identités peut permettre de détecter des erreurs dans les variables intervenant dans les équations estimées. Poursuivre le processus sans corriger ces erreurs peut conduire à un modèle économiquement faussé.

$$FD = CO + H + IP + gd$$

sera utilisée à la fois pour calculer les valeurs historiques de FD et définir FD dans le modèle. Cela introduit un problème évident: si nous faisons une erreur dans la formule, ou nous utilisons des données erronées, il n'y a aucun moyen de le détecter.

5.2.12.2 le programme EViews

Voyons maintenant comment la tâche ci-dessus peut être réalisée. Nous voulons créer:

- un fichier de travail pour tous les éléments du modèle,
- une image du modèle, avec des identités bien définies, et des indications sur les équations estimées envisagées,
- les données associées.
- Le fichier de travail («Workfile»)

Tout d'abord, nous avons besoin d'un fichier de travail. Sous EViews, toutes les tâches sont effectuées en mémoire, mais elles s'appliquent à l'image d'un fichier physique qui contiendra tous les éléments gérés à un moment donné.

Nous pouvons créer le fichier (comme une image de la mémoire) ou partir d'un fichier préexistant, auquel cas le dossier sera transféré de son emplacement physique vers la mémoire.

Certaines précautions doivent être prises.

Tout d'abord, une seule version du fichier doit être ouverte dans la mémoire. Comme nous l'indiquerons ailleurs, EViews permet à l'utilisateur d'ouvrir une deuxième version (ou même une troisième, quatrième ...) d'un fichier déjà ouvert. Ainsi, des changements peuvent être appliqués séparément aux deux versions présentes en mémoire, ~~tels que comme (ou telles que si cela se rapporte à versions ???)~~ la production de séries pour l'une et ~~les-des~~ estimations pour l'autre.

Ceci nous paraît très dangereux¹¹. À tout le moins, nous perdrons l'un des ensembles de modifications, car il n'y a aucun moyen de transférer des éléments d'une image mémoire à l'autre. Bien sûr, chaque

¹¹ Cette opinion n'est pas partagée par les gestionnaires d'EViews, sans doute pour de bonnes raisons, dont nous sommes prêts à discuter.

fichier peut être sauvegardé sous un nom différent, mais cela ne permet pas de fusionner les modifications¹². Au pire, on oublie la répartition des modifications apportées aux fichiers, et l'un ou les deux deviendront incohérents, la meilleure option étant d'éviter d'enregistrer l'un d'eux, et de repartir à zéro.

Cela signifie qu'il faut:

- en mode commande, vérifier qu'aucun fichier du même nom n'est ouvert, et le fermer si nécessaire ;
- en mode programme (le cas ici), s'assurer dans un premier temps qu'aucun fichier de même nom n'est ouvert. Ceci peut être réalisé par l'ordre «CLOSE», qui n'aura aucun effet la plupart du temps¹³, mais garantira que nous sommes dans la situation requise ;
- Deuxièmement, un nouveau projet doit commencer par un nettoyage du fichier de travail. Pour un fichier initial, contenir des éléments est au mieux source de confusion, au pire dangereux. Par exemple, des séries ayant le même nom que les éléments de notre projet peuvent déjà être présentes avec un sens différent (PIB pour un pays différent ?), et disponibles pour une période plus longue. Permettre à EViews d'estimer des équations sur la plus longue période disponible introduira dans l'échantillon des valeurs non pertinentes.

Une façon simple de résoudre le problème consiste à supprimer tout élément existant, grâce à la déclaration:

```
DELETE *
```

qui va détruire tout élément déjà existant, à l'exception de C (vecteur de coefficients générique) et RESID (série générique de résidus) qui sont créés automatiquement avec le fichier de travail, et ne peuvent pas être supprimés.

Nous voyons un seul cas acceptable pour garder des éléments préexistants: si le fichier de travail contient des informations d'origine, fournies à l'utilisateur par une source externe. Mais même dans ce

¹² Fournir cette option ne semble pas impossible.

¹³ Malheureusement sans fournir de message d'erreur.

cas, le fichier doit être enregistré tel quel, pour permettre de remonter les étapes jusqu'au tout début où seule cette information originale était présente, dans sa forme originale.

Dans tous les cas, à partir d'EViews 5, la possibilité de définir des feuilles séparées à l'intérieur du fichier de travail permet de résoudre le problème. Comme nous l'avons vu précédemment, on peut simplement stocker les données d'origine dans une feuille, et commencer à construire les données transformées dans une feuille vierge, qui vont contenir des séries liées logiquement aux originales.

Principe de modélisation: toujours organiser votre travail de telle sorte que si l'étape n échoue, vous pouvez toujours revenir au résultat de l'étape n-1.

Principe de modélisation (version alternative): toujours organiser vos programmes de telle sorte que vous pouvez produire à nouveau tous les éléments liés à la situation actuelle.

Ce discours (long) conduit à des instructions suivantes:

```
CLOSE SMALL
WFCREATE (page = model) Q 1970Q1 2005Q4 small
DELETE *
```

Les appliquer garantit:

- que le fichier small.wf1 est ouvert en mémoire avec les caractéristiques nécessaires pour une page appelée « modèle » ;
- qu'une seule version du fichier est ouverte (nous supposons que vous allez suivre nos suggestions) ;
- que la page est vide (en fait, qu'elle ne contient que C et RESID).

5.2.12.3 Les données

Maintenant que nous avons un fichier de travail, nous pouvons le remplir avec les informations nécessaires.

L'information d'origine est représentée par 20 séries, dans le fichier Excel FRA.XLS¹⁴. Nous les importerons en utilisant l'instruction IMPORT. Cette instruction est très simple (voir le manuel de l'utilisateur pour les options détaillées):

```
READ fra.xls 20
```

Mais attention: même si le fichier Excel contient des dates (dans la première colonne ou ligne) cette information n'est pas prise en compte. Ce qui est utilisé est plutôt la période de travail courante, définie par la dernière instruction SMPL. Heureusement, dans notre cas, l'échantillon actuel, défini lors de la création du fichier de travail, est le même que celui du fichier Excel. Mais ce ne sera pas toujours le cas: il vaut mieux préciser le SMPL avant la lecture.

```
SMPL 1970Q1 2004Q4  
READ fra.xls 20
```

Principe de modélisation: si l'introduction d'une déclaration (à faible coût) peut être utile, même très rarement, le faire.

Il faut aussi faire attention à l'orientation de la série: par défaut, elles apparaissent comme des colonnes, et les données commencent à partir de la cellule B2 (deuxième ligne, deuxième colonne). Tous les autres cas doivent être précisés, ainsi que le nom de la feuille pour un fichier de plusieurs feuilles.

Une option supplémentaire (sans doute meilleure)

Si l'on suit la méthode ci-dessus, toutes les données seront transférées dans la page «modèle». Cela facilite les choses d'une certaine manière, car toutes les informations seront immédiatement disponibles. Mais

- la séparation entre les données originales et celles du modèle ne sera pas claire ;

¹⁴ À partir d'EViews 7 on peut également lire des fichiers Excel 2010 de type xlsx (et en produire à partir d'EViews 8).

- la stabilité des données originales n'est pas garantie, car la construction du modèle peut les modifier par inadvertance ;
- comme les séries originales sont probablement plus nombreuses, la plus grande partie de l'écran sera occupée par des éléments devenus inutiles.

Bien sûr, on peut séparer les données originales de celles du modèle en utilisant un préfixe pour le premier type. Mais il est encore préférable de séparer les deux ensembles physiquement. Cela peut être fait par l'intermédiaire de la fonction EViews «LINK».

Au lieu de charger la série originale dans la page modèle, une page spécifique est créée (son nom peut être «OCDE») dans laquelle les données sont importées.

Ensuite, dans la page modèle, les variables du modèle sont déclarés comme «liées», et un lien est défini avec la série originale de la page «OCDE».

La syntaxe associée sera présentée plus tard.

5.2.12.4 Le modèle

Maintenant nous devons définir le modèle sur lequel nous allons travailler.

La production d'un modèle commence par la phrase:

```
Nom-du-modèle MODEL
```

Appelons notre modèle `_fra_1`.

Une astuce: commencer le nom des éléments importants par un «blanc souligné» leur permet d'être affichés au début de l'écran du fichier de travail, évitant ainsi un défilement fastidieux si le nombre d'éléments est important. Pour des éléments très importants (comme le modèle lui-même), vous pouvez même utiliser un double soulignement.

La déclaration

```
MODEL _fra_1
```

définit `_fra_1` comme modèle «actif».

Deux cas peuvent être envisagés:

- le modèle n'existe pas. Il est créé sans équations (pour l'instant) ;
- le modèle existe. Il est ouvert, avec ses équations actuelles.

La seconde option est dangereuse dans notre cas, car nous voulons commencer à partir de zéro. Pour s'assurer de cela, la technique la plus efficace (et brutale) consiste à supprimer le premier modèle, ce qui nous place dans le premier cas.

```
DELETE _fra_1  
MODEL _fra_1
```

Cela introduit un léger problème, cependant. Dans la plupart des cas (y compris actuellement), le modèle n'existe pas, et l'instruction DELETE échoue. Pas de problème, ce que nous voulions est de s'assurer qu'aucun modèle ne préexistait, et c'est bien là la situation que nous obtenons. Mais EViews va se plaindre, car il n'a pas pu effectuer la tâche requise. Et si le nombre maximum d'erreurs acceptées est 1 (option par défaut), le programme s'arrêtera.

Nous pouvons résoudre le problème en appliquant la fonction @isobject qui donne «vrai» (ou 1) si l'élément existe, et «faux» (ou 0) sinon:

```
if @isobject ("_fra_1"), then  
DELETE _fra_1  
endif  
MODEL _fra_1
```

Note: dans ce cas bien localisé il est en fait plus simple de débrancher la détection des erreurs, par l'option «noerr»:


```
DELETE(noerr) _fra_1
MODEL _fra_1
```

Mais il ne saurait être question d'appliquer cette option de manière générale.

Nous pouvons aussi modifier le nombre par défaut d'erreurs acceptées.

Une autre façon d'éviter cette situation est évidemment de ~~définir le~~donner au nombre maximum d'erreurs une valeur ~~???~~ supérieure à 1. Cela se fait en changeant le nombre ~~dans de~~ la case «maximum errors before halting» (maximum d'erreurs ~~maximales~~ avant l'arrêt) ~~de dans~~ la fenêtre File > Run du menu.

En fait, si vous avez suivi le principe ci-dessus, vous ne courez aucun risque en exécutant un programme qui produit des messages d'erreur, même valables. Vous avez enregistré les éléments liés à la situation initiale, et même si vous avez oublié de le faire, vous pouvez toujours répéter les étapes qui ont y ont conduit.

L'avantage de cette option:

- le programme se poursuivra après les messages d'erreur inutiles ;
- vous pouvez produire des erreurs artificielles, qui peuvent être très utiles pour localiser les vraies erreurs¹⁵ ;
- les messages peuvent être associés à plusieurs erreurs logiquement indépendantes, qui peuvent être corrigées simultanément, ce qui conduit plus rapidement à une version correcte.

Maintenant, quel nombre faut-il indiquer ? Pour l'auteur, en fonction de la taille du modèle, de 1 000 à 10 000. Le nombre doit être supérieur au nombre d'erreurs potentielles, car vous voulez que le programme s'effectue complètement. Bien sûr, vous ne ferez jamais 10 000 erreurs logiques. Mais le comptage est effectué sur le nombre de messages d'erreur. Et dans un modèle à 2 000 équations, si vous avez mis toutes les endogènes à zéro et que vous calculez leurs taux de croissance, cette seule erreur va générer 2 000 messages.

¹⁵ La place du message associé à une vraie erreur permettra de la localiser entre les deux erreurs artificielles précédente et suivante.

Le seul véritable inconvénient est que si votre programme utilise une boucle sur le nombre d'éléments d'un groupe, et que ce groupe n'a pas pu être créé, une erreur est répétée indéfiniment avec le message suivant:

```
Syntax error in "FOR! I = 1 TO G.@COUNT"
```

Vous devrez attendre que le nombre maximum soit atteint.

5.2.12.4.1 Introduction des équations.

Maintenant que nous avons un modèle vierge, nous pouvons y introduire les équations une par une. Le texte de ces équations a déjà été défini, il nous reste simplement à présenter les commandes EViews.

Cela se fait par la commande APPEND.

La première va définir l'investissement:

```
_fra_1.append IP = f*(Q)
```

De toute évidence la syntaxe

- contient la déclaration «append»,
- ajoute le nom du modèle sur la gauche, avec un point,
- ajoute le texte de l'équation sur la droite, en commençant par un blanc séparateur.

Nous devons maintenant expliquer la syntaxe étrange de notre équation.

À l'heure actuelle, nous prévoyons que le modèle va expliquer la décision d'investissement par l'évolution du PIB. Cela semble tout à fait logique, mais nous n'avons pas établi les formes possibles de l'équation théorique, et nous n'avons pas vérifié qu'au moins une de ces équations était validée par tous les tests économétriques requis.

Mais en même temps, nous voulons qu'EViews nous donne autant d'informations que possible sur la structure de notre modèle: simultanités, exogènes ...

Le meilleur compromis est clairement de produire un modèle qui, bien que dépourvu de toute équation estimée, présente néanmoins les mêmes relations de causalité que le modèle (futur) que nous considérons.

Le choix le plus simple devrait être, comme si nous étions en train de décrire les spécifications du modèle dans un document ou sur un tableau noir, de spécifier:

$$IP = f(Q)$$

Malheureusement, EViews n'accepte pas une équation écrite de cette façon. Il supposera que nous utilisons une fonction appelée f , avec l'argument Q . Comme cette fonction n'existe pas, l'équation sera rejetée.

L'astuce que nous proposons est de placer un astérisque entre « f » et la première parenthèse, ce qui donne

$$IP = f^*(Q).$$

Et de définir f comme un scalaire (pour éviter l'apparition d'une exogène supplémentaire).

Si plus d'une variable explicative est utilisée, comme dans l'équation donnant la capacité de production, nous aimerions écrire:

$$CAP = f (LE, K)$$

Encore une fois, cette forme n'est pas acceptée par EViews, et on peut écrire à la place:

$$CAP = f^*(LE+K)$$

Pour que les choses soient claires (y compris pour des lecteurs extérieurs) il suffit d'indiquer ses conventions, et vous êtes invités à utiliser les vôtres si vous le désirez.

Cependant, l'abandon pur du paramètre f est dangereux, comme dans:

$$M = FD + TD$$

Cette méthode fonctionnera aussi, mais l'équation peut être confondue avec une véritable identité, tout à fait trompeuse dans ce cas.

Le jeu complet des équations devient le suivant:

$$\text{_fra_1.append } LE = f * (Q)$$

$$\text{_fra_1.append } I = f * (Q)$$

$$\text{_fra_1.append } K = K (-1) * (1-DR) + I$$

$$\text{_fra_1.append } CAP = f * (LE + K)$$

$$\text{_fra_1.append } IC = r_icq * Q$$

$$\text{_fra_1.append } IC = f * (Q)$$

$$\text{_fra_1.append } LT = LE + I_g$$

$$\text{_fra_1.append } RHI = wr * LT + r_rhiq * Q$$

$$\text{_fra_1.append } CO = (1-sr) * RHI$$

$$\text{_fra_1.append }] = IH r_ih * RHI$$

$$\text{_fra_1.append } FD = CO + H + IP + CI + gd$$

$$\text{_fra_1.append } TD = FD + r_ic * Q$$

$$\text{_fra_1.append } UR = Q / CAP$$

$$\text{_fra_1.append } M = f * (TD + UR)$$

```
_fra_1.append X = f * (wd + UR)
```

```
_fra_1.append Q + M = FD + X
```

Nous avons produit un modèle à 16 équations appelé `_fra_1`. Après avoir exécuté les instructions, un élément sera créé dans le fichier de travail, avec le nom «`_fra_1`» et le symbole «`M`» (en bleu).

Double-cliquer sur cet élément va ouvrir une fenêtre spéciale, avec la liste des équations:

- son texte (avec l'icône «TXT» sur la gauche),
- son numéro d'équation (dans l'ordre d'introduction dans le modèle),
- ses causalités économiques: la variable expliquée à gauche, les éléments explicatifs à droite, avec la syntaxe que nous n'avons pas pu utiliser plus tôt. Les retards ne sont pas spécifiés, comme nous le verrons plus tard. Par exemple, `K` est présenté comme dépendant de `K`.

En fait, trois autres modes d'affichage sont disponibles, en utilisant le bouton «View»:

- Variables: affiche les variables endogènes (en bleu avec «En»), et exogènes (en jaune avec «X»). Pour les endogènes, le numéro de l'équation est donné. Cela permet de localiser l'équation dans le texte du modèle, ce qui est utile pour les grands modèles.

Le bouton «Dependencies» donne accès à un sous-menu qui permet d'associer à la variable celles qui en dépendent (Up) et qu'elle influence (Down).

Par exemple, pour `FD`, «Up» donnera `TD` et `Q`, «Down» donnera `CO`, `I`, `G`, et `IH`.

Bien sûr, «Up» ne sera pas disponible pour les exogènes.

L'option «Filter» permet la sélection des variables en utilisant un «masque» spécifique. Par exemple, dans un modèle multi pays les variables françaises peuvent être identifiées par `FRA_*`, à condition que l'on ait utilisé une telle convention.

- Source text: ceci est de fait le texte du code modèle. Nous verrons que cela change avec des équations estimées.

Block structure: ceci donne des informations sur la structure logique du modèle (cet aspect est détaillé plus loin dans le chapitre 7).

Nous obtenons:

- le nombre d'équations,

- le nombre de blocs, en distinguant simultanés et récur­sifs,
- les variables contenues dans chaque bloc.

Pour le moment, disons seulement qu'un bloc **simultané** contient des éléments interdépendants. Pour tout couple de variables du bloc, un chemin peut conduire de la première à la seconde, et vice versa. Bien sûr, cette propriété ne dépend pas de l'ordre des équations à l'intérieur du bloc.

EViews 6 et 7 donnent également le nombre de variables de bouclage (ceci sera aussi expliqué plus tard).

Au contraire, un bloc **récur­sif** peut être organisé (EViews le fait lui-même) de telle sorte que chaque variable dépende seulement (pour la période actuelle) des variables définies précédemment.

Cette information est utile pour améliorer la compréhension du modèle, localiser les incohérences et corriger des problèmes techniques.

EViews peut détecter des erreurs si:

- une variable est définie deux fois,
- la syntaxe d'une équation est erronée (une parenthèse qui manque par exemple)

et permettre à l'utilisateur d'observer les erreurs lui-même si:

- des éléments normalement endogènes apparaissent comme exogènes: l'équation définissant la variable a été oubliée, ou mal écrite ;
- des éléments étrangers au modèle apparaissent: les variables ont été mal orthographiées ;
- une boucle apparaît là où il ne devrait pas y en avoir ;
- ou (plus probablement) une boucle qui devrait apparaître ne s'affiche pas: par exemple, un modèle keynésien est décrit comme récur­sif, ou un modèle à deux pays qui commercent entre eux apparaît comme constitué de deux blocs indépendants.

Toutes ces erreurs peuvent être détectées (et corrigées), sans faire appel aux données. Cela permet d'accélérer le processus de construction, en particulier si les données ne sont pas encore disponibles.

Pour la production des séries, on peut considérer deux options.

- Si les séries originales et celles du modèle sont dans la même page, on utilisera simplement l'instruction «GENR» selon la séquence.

```
genr Q = FRA_GDPV
```

```

genr PAC = 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 (ou 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)

```

- Si les séries originales sont gérées dans leur propre page (une meilleure option à notre avis), on utilisera:

FOR %1 Q CAP CI CI LT LG FD CO RHI je WD IH X M

LINK {%1}

NEXT

Q.linkto OCDE\FRA_GDPV

CAP.linkto OCDE\FRA_GDPVTR

CI.linkto OCDE\FRA_ISKV

IC.linkto OCDE\FRA_ICV

LT.linkto OCDE\FRA_ET

LG.linkto OCDE\FRA_EG

FD.linkto OCDE\FRA_TDDV

CO.linkto OCDE\FRA_CPV

RHI. OCDE\FRA_YDRH

I.linkto OCDE\FRA_IBV

IH.linkto OCDE\FRA_IHV

WD.linkto OCDE\FRA_XGVMKT

X.linkto OCDE\FRA_XGSV

M.linkto OCDE\FRA_MGSV

GD.Linkto OCDE\FRA_IGV+FRA_CGV

Cependant, un problème demeure pour GD, la somme des deux variables originales FRA_IGV et FRA_CGV. La fonction LINK permet de se référer à des variables individuelles et non des fonctions (comme Excel le fait). Il y a deux options, aucune d'entre elles n'étant vraiment satisfaisante.

Créer des liens vers les éléments originaux dans la page modèle.


```
LINK FRA_IGV
LINK FRA_CGV
FRA_IGV.linkto OCDE\FRA_IGV
FRA_CGV.linkto OCDE\FRA_CGV
genr gd = FRA_IGV + FRA_CGV
```

ou calculer une variable FRA_GDV dans la page d'origine.

```
genr FRA_GDV = FRA_IGV + FRA_CGV
```

et la relier

```
LINK GD
GD.linkto OCDE\FRA_GDV
```

Les instructions qui impliquent seulement des éléments du modèle restent inchangées.

Nous avons maintenant réalisé une première version du modèle, et généré les données associées. Comme les comportements n'ont pas été établis, nous ne pouvons évidemment pas encore le résoudre. Mais nous pouvons vérifier deux choses importantes:

- les données nécessaires à l'estimation sont présentes,
- les données sont compatibles avec les identités.

Ces conditions sont nécessaires pour commencer l'estimation, la prochaine étape dans le processus. La première est évidente, la seconde un peu moins. Mais des incohérences dans les identités peuvent provenir de l'utilisation d'un mauvais concept pour une variable, ou d'un mauvais calcul. Si cette variable va être utilisée dans les estimations, qu'elle soit expliquée ou explicative, l'ensemble du processus sera basé sur des éléments erronés.

- Le temps passé à l'estimation sera perdu.
- Celui-ci sera probablement plus long que d'habitude, car il est généralement plus difficile (voire impossible) de trouver un bon ajustement basé sur des données erronées (heureusement ?).

Si un ajustement est jugé satisfaisant, l'équation associée peut rester longtemps dans le modèle (peut-être indéfiniment), et tous les résultats ultérieurs seront invalidés. Si l'on est honnête, une découverte ultérieure de l'erreur signifiera que beaucoup de travail devra être recommencé, portant éventuellement sur des résultats publiés.

Ce test peut être réalisé grâce à une technique très simple: le contrôle résiduel.

5.2.13 UN PREMIER TEST: LA VERIFICATION RESIDUELLE DES IDENTITES

À ce stade, une solution complète du modèle ne peut pas être prise en considération. Toutefois, certains contrôles peuvent être effectués, qui demandent une «simulation» très spécifique. Cette technique est appelée «contrôle résiduel».

Cette méthode va appliquer chaque formule du modèle séparément, en utilisant les valeurs historiques des variables. Cela peut être fait en créant pour chaque équation une formule donnant la valeur de l'expression de droite (avec l'instruction «genr» dans EViews). Cependant, il existe une méthode beaucoup plus simple, proposée par EViews.

Si nous considérons un modèle écrit sous la forme:

$$y_t = f(y_t, y_{t-1}, x_t, \hat{\alpha})$$

avec y et x vecteurs de variables endogènes et exogènes.

On peut effectuer une «simulation» très spécifique, dans laquelle chaque équation est calculée séparément en utilisant des valeurs historiques.

Techniquement, cela signifie:

- découper le modèle en modèles à équation unique, autant qu'il existe d'équations ;
- résoudre chacun de ces modèles séparément, en utilisant comme valeurs explicatives les données historiques.

Cela signifie que nous allons calculer:

$$y_t = f(y_t^0, y_{t-1}^0, x_t, \hat{\alpha}) + e_t$$

Cette méthode va contrôler:

- pour les identités, la cohérence entre les données et la formulation ;
- pour les équations de comportement, la disponibilité des variables demandées par les estimations envisagées. Mais on ne s'intéressera pas aux valeurs obtenues (en fait la méthode que nous proposons donnera une valeur nulle).

En fait EViews permet l'utilisation d'une expression sur le côté gauche. Ceci est également valable ici, la comparaison étant faite entre les expressions gauche et droite.

L'intérêt de cette méthode est évident: si le résidu de l'équation n'est pas nul, cela signifie qu'il y a au moins une erreur dans cette équation particulière. Bien sûr, le problème n'est pas résolu, mais son emplacement est identifié. Nous verrons plus loin que cette méthode est encore plus efficace pour un modèle complètement estimé, et nous allons étendre notre discussion à ce moment-là.

Il serait toutefois illusoire d'espérer obtenir un modèle correct immédiatement: certains diagnostics d'erreur auront pu être mal interprétés et des corrections mal effectuées. Mais même si l'erreur a été corrigée:

- il pouvait y avoir plusieurs erreurs dans la même équation,
- le processus de correction peut introduire une erreur dans une autre équation qui semblait auparavant exacte, mais contenait en fait deux erreurs se compensant. Nous allons préciser ce cas.

Prenons notre exemple. Si nous avons utilisé pour l'investissement en logements de la valeur à prix courants:

```
genr IH = FRA_IH
genr r_ih = IH/RHI
```

Alors l'équation donnant FD

```
_fra_1.append FD = CO + IH + IP + CI + gd
```

ne tomberait pas juste, mais celle donnant IH

```
_fra_1.append IH = r_ih. RHI
```

oui, car le calcul de r_{ih} comme le rapport de IH à RHI va compenser l'erreur par une autre erreur.

Si nous corrigeons l'erreur sur IH sans corriger r_{ih} , l'équation donnant IH apparaît désormais comme erronée, alors que son nombre d'erreurs a diminué, passant de 2 à 1.

Cela signifie que l'obtention d'un ensemble complet de résidus nuls peut prendre un peu de temps, et quelques itérations, mais la méthode devrait converger régulièrement jusqu'à ce que toutes les erreurs aient disparu¹⁶.

5.2.13.1 Les types d'erreurs rencontrés

Le contrôle résiduel permet de diagnostiquer les erreurs suivantes

Absence de solution

- erreur de syntaxe (appel à une fonction inexistante, parenthèses déséquilibrées) ;
- série avec le bon nom, mais indisponible, soit complètement (elle n'a pas été obtenue), ou partiellement (certaines périodes manquent) ;
- mauvaise orthographe (appel à une série inexistante).

Résidus non nuls

- mauvaise orthographe (appel à la mauvaise série) ;
- erreurs de logique. Cela peut être plus ou moins grave, car cela peut provenir d'une erreur purement technique: oubli d'un terme, par exemple, ou d'une erreur conceptuelle: assumer une identité non vérifiée théoriquement ;

¹⁶ À moins que le modélisateur n'autorise certaines identités à n'être vérifiées qu'approximativement.

- erreur de données: informations mal introduites, séries mal calculées, informations provenant de sources non cohérentes ou de versions différentes d'une même banque.

Quant aux équations de comportement non vérifiées (ou à résidus erronés), cette question sera applicable (et traitée) plus tard.

Observer les valeurs des erreurs peut donner des indices quant à leur origine:

- Si certaines périodes donnent un résultat correct:
 - À l'année de base (où les éléments à prix constants et prix courants sont identiques): les indices de prix pourraient avoir été confondus les uns avec les autres, ou les valeurs confondues avec des volumes.
 - Si une variable dans la formule est nulle pour ces périodes, elle pourrait être responsable.
 - Sinon, l'erreur pourrait provenir d'une faute de frappe (faite par l'utilisateur ou le producteur de données).
 - Ou si elle apparaît dans les dernières périodes, les éléments provisoires pourraient être incompatibles.
- L'observation de l'ampleur de l'erreur peut également être utile: un dépassement d'une magnitude économique aberrante (100 000 % par exemple) devrait provenir d'une erreur de spécification: mauvais opérateur, inversion des coefficients, confusion entre valeurs et valeurs par habitant. Une erreur faible vient souvent de la confusion entre deux concepts proches (le déflateur du prix à la consommation excluant ou incluant la TVA).
- Pour les équations additives, l'absence d'un élément peut être déterminée en comparant la valeur du résidu aux valeurs effectives des variables. Par exemple, si l'erreur sur la demande finale est 56734 en 2000Q1, qui est justement la valeur de l'investissement en logements.
- Si le signe de l'erreur est constant (et surtout si l'ordre de grandeur est le même pour toutes les périodes), l'erreur pourra provenir de l'absence d'un élément, une multiplication par un mauvais facteur ou l'absence d'une influence positive.
- Si les erreurs sur plusieurs variables ont des valeurs identiques, ils doivent avoir la même origine.
- Si au contraire deux variables montrent des erreurs à peu près identiques avec un signe opposé, cela peut provenir du fait que l'une d'elles a des valeurs erronées et explique l'autre.

Par exemple, si les valeurs historiques pour Q sont surestimées, les erreurs relatives sur UR et Q seront similaires avec des signes différents.

$$UR = Q / CAP$$

$$Q + M = FD + X$$

5.2.13.2 Le traitement des erreurs

Le diagnostic des erreurs dans la phase de vérification résiduelle peut faire remonter à différentes phases du processus de modélisation:

- la gestion des données: les données obtenues auprès des producteurs externes ne sont pas cohérentes, pour une série complète ou des observations particulières (cela arrive !);
- la production des données du modèle: on utilise la mauvaise série originale, ou effectue un calcul erroné,

exemple: utilisation d'une variable à prix courants au lieu de constants, ou oubli d'un élément dans une somme;

- la spécification du modèle (équations mal écrites),

exemple: oublier le terme d'investissement logement dans la définition de la demande;

- l'estimation (série modifiée depuis l'estimation, mauvais coefficients),

exemple: une erreur dans l'équation des importations montre que la série de demande intérieure a été changée depuis l'estimation.

L'application de ce processus un certain nombre de fois sera nécessaire pour produire un modèle cohérent.

5.2.13.3 Retour à l'exemple

Effectuer une vérification résiduelle est facile dans EViews: il suffit de spécifier l'option «d = f» dans l'ordre de SOLVE:

```
_fra_1.solve(d=f)
```

Bien sûr, comme toutes les équations seront calculées séparément, toutes les informations doivent être disponibles sur la période en cours, y compris les variables endogènes (qui deviennent localement exogènes). Contrairement aux calculs et estimations, EViews n'adaptera pas le processus de simulation à la période utilisable (ce qui semble assez logique).

Comme le modèle prend une forme récursive (super-récursive ?) le calcul donne le résultat directement, et aucun élément décrivant la méthode de résolution n'est nécessaire (nous les verrons plus tard).

Toutefois:

- Il faut spécifier le nom attribué aux variables calculées.

Chaque fois que EViews doit résoudre un modèle, le nom donné aux résultats sera construit à partir du nom d'origine de la variable, auquel sera ajouté un suffixe (un préfixe est également possible, mais moins facile à gérer, à notre avis). Cela permet d'éviter la disparition des éléments originaux, et permet de comparer des solutions alternatives.

Le préfixe est spécifié en utilisant l'instruction:

```
nom-du-modèle @all suffixe
```

(Rappelez-vous: append ajoute du texte au modèle, une équation d'identité n'est qu'un cas particulier de texte).

Dans notre cas, en appliquant le suffixe «_C» selon:

```
_fra_1.append @all _C
```

Le calcul de l'équation pour FD donnera FD_C, que nous pouvons comparer avec les valeurs observées de FD.

Calculer les différences entre les valeurs observées et calculées peut être fait dans une boucle, en utilisant la syntaxe décrite plus loin. Les éléments de la boucle peuvent être définis «à la main» mais il est plus efficace d'utiliser l'instruction «makegroup».

```
nom-du-modèle.makegroup (a,n) nom-du-groupe @endog
```

```
nom-du-modèle.makegroup (a,n) nom-du-groupe @exog
```

Dans notre cas:

```
_fra_1.makegroup (a,n) g_vendo @endog
```

```
_fra_1.makegroup (a,n) g_vexo @exog
```

Deux remarques:

- Vous avez sûrement dû vous interroger sur la raison de l'ajout de «(a,n)». Ceci modifie les options de la commande «makegroup», qui par défaut produirait un groupe avec les noms de base (dans notre cas avec l'ajout de _C) et laisserait de côté les noms réels.

Indiquer (a, n):

- introduit les noms réels (par «a» pour « actual »),
- élimine la référence aux solutions («n» pour non).

Il est clairement préférable de limiter les calculs aux identités. Les résidus sur les équations estimées n'ont pas de sens: comme le «f» scalaire est nul, le côté droit sera calculé comme égal à zéro, et le pourcentage d'erreur sera de 100 % comme $100 \times (\text{valeur} - 0) / \text{valeur}$. Mais être capable de calculer des résidus sur l'ensemble du modèle prouve que les estimations pourront être réalisées sur cette période.

On peut créer deux sous-groupes par

```
group g_vbeha CI I LE M X  
group g_viden CAP CO FD IC IH K LT Q RHI TD UR
```

ou

```
group g_vbeha CI I LE M X  
_fra_1.makegroup(a,n) g_viden @endo  
g_viden.drop CI I LE M X
```

La deuxième option crée d'abord un groupe *g_viden* complet, puis élimine de celui-ci les variables estimées.

Cette dernière technique est clairement inefficace ici, mais le sera beaucoup plus avec un modèle de 500 équations dont 50 estimées (une situation plus habituelle).

Cependant, les deux techniques font appel à une liste définie par l'utilisateur, qui devra être mise à jour à chaque fois qu'elle évolue, une chose que nous voulons éviter: nous proposons d'utiliser une méthode plus complexe, mais automatique.

Un conseil: un contrôle visuel est rendu difficile par l'imprécision d'EViews, ce qui produit souvent des petits résidus pour des équations exactes. En format scientifique (exposant + mantisse), ces résidus apparaissent sous forme de valeurs aux exposants fortement négatifs, qui sont difficiles à identifier visuellement. Une solution est de passer à une présentation en virgule fixe, en sélectionnant une zone (dans la présentation «Spreadsheet»), puis en utilisant le bouton droit de la souris pour accéder à «Display Format» puis «Fixed Decimal».

Une solution plus simple pour observer s'il n'y a aucune erreur est d'afficher tous les résidus en un seul graphique, et d'observer, non les séries (elles doivent suivre un mouvement brownien), mais l'échelle: maximum et minimum doivent être très faibles.

Une autre idée est de transférer les résidus vers Excel et de trier la feuille (malheureusement EViews ne peut pas trier une feuille avec comme critère les valeurs à une période donnée des séries). Les éléments non négligeables vont apparaître en haut et en bas de celle-ci en fonction de leur signe et de l'ordre de tri. Cette technique prend plus de temps mais permet d'identifier immédiatement et complètement les éléments en cause.

5.2.13.4 Une astuce: comment établir les groupes d'identités et de relations de comportement

Vous avez certainement réalisé maintenant (et vous le saviez probablement déjà) que l'on doit éviter autant que possible d'avoir à modifier le texte des programmes de modélisation, chaque fois qu'un changement a été fait à un autre endroit du processus. Cela représente au mieux du travail supplémentaire, au pire une source d'erreur. Nous venons de violer ce principe, en séparant nous-mêmes les endogènes en variables de comportement et d'identité.

Cela introduira des problèmes, en particulier dans les grands modèles: le processus sera fastidieux et source d'erreurs, et il faudra se rappeler de mettre à jour la liste à chaque modification de la structure du modèle.

Nous proposons une technique simple pour éviter ce problème, permettant d'établir la séparation initiale et sa mise à jour automatique. Elle est basée sur la présence du scalaire «f» dans les équations de comportement.

Nous devons simplement:

- simuler le modèle avec l'option «d = f» et $f = 1$, et enregistrer les résultats sous un suffixe donné ;

- faire passer f à 2 et mettre à jour le modèle, par «UPDATE» (ceci est nécessaire pour que le changement soit pris en compte) ;
 - simuler le modèle de nouveau avec $f = 2$ et un autre suffixe ;
 - créer des groupes vides pour les variables estimées et les identités,
 - produire une boucle sur l'ensemble du groupe d'endogènes, et tester à chaque fois si les résultats des deux simulations sont différents.
- s'ils le sont, ajouter la variable à la liste des éléments estimés,
 - sinon, à la liste des éléments d'identité.

Nous pouvons utiliser le programme suivant (pour la période 2000-2002). Nous supposons que tout écart en pourcentage plus élevé que 0,00001 indique une erreur.

```

_fra_1.makegroup(a,n) g_vendo @endog
_fra_1.makegroup(a,n) 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 2000S1 2002S2
_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

```

Cette séquence appelle quelques explications.

- La boucle (de «for» à «next») sera reproduite pour chaque variable dans la liste `g_vendo`. Le nombre de ces variables est `g_vendo.@count` (Pour EViews, `x.@count` représente un scalaire entier contenant le nombre d'éléments du groupe `x`).
- `!i` est le rang de la variable dans le groupe `g_vendo` (de 1 à `g_vendo.@count`).

- %1 reçoit comme une chaîne de caractères le contenu de `g_vendo.@seriesname(!i)`, c'est-à-dire le nom de la variable dans le groupe `g_vendo`, avec le rang `!i`.
- Les formules suivantes remplacent %1 par la valeur de la chaîne, et les crochets sont ensuite éliminés, en laissant les caractères dans la déclaration.

Pour les utilisateurs réguliers d'EViews ou des personnes familières avec la programmation, ce qui précède est probablement clair. Pour les autres, c'est sans doute le moment de donner des informations très élémentaires sur la programmation en EViews (même si ce n'est pas le but de ce livre).

5.3 ÉLÉMENTS DE PROGRAMMATION SOUS EViews

Dans les programmes que nous allons présenter, une utilisation intensive sera faite de deux éléments: les groupes et les boucles.

Comme c'est la première fois que nous abordons les applications sous EViews, nous allons maintenant donner quelques éléments de programmation.

Nous commencerons par traiter les groupes et les boucles, puis nous décrirons comment EViews 8 facilite la gestion de l'exécution.

5.3.1 LES GROUPES 6

Les groupes sont des éléments nommés qui font référence à un ensemble d'objets (seulement des séries ou expressions de séries jusqu'à EViews 7), ce qui permet de les traiter soit globalement soit en séquence.

La déclaration la création d'un groupe se fait par l'ordre `GROUP`

```
group nom-du-groupe liste-de-séries-du-groupe
```

par exemple

```
group g x y z x/y
```

créera un groupe nommé `g` contenant les trois séries `x`, `y` et `z`, et le rapport de `x` et `y`.

Jusqu'à la version 6 incluse, l'élément doit être une série ou une formule appliquée à des séries, mais on peut tricher en créant des séries artificielles contenant le nom de l'élément demandé.

On peut:

- grouper des groupes
- ajouter et retirer des éléments des groupes:

```
g.add x
```

ajoutera la série au groupe g

tandis que

```
g.drop x
```

enlèvera la série x du groupe g.

Deux éléments utiles peuvent être associés avec le groupe:

- [g.@count](#) est un scalaire entier qui contient le nombre d'éléments du groupe g.
- [g.@seriesname](#) est un vecteur de chaînes de caractères qui contient le nom des séries du groupe g.

Enfin, les groupes peuvent être créés à par un masque:

group [g_fra fra_*](#) va créer un groupe contenant tous les éléments commençant par fra suivi d'un blanc souligné.

group [g_GDP ???_GDP](#) va créer un groupe contenant tous les PIB des pays de l'OCDE (en utilisant trois lettres comme étiquette).

group [g_3 ???_*](#) va créer un groupe de tous les éléments dont le nom commence par trois caractères, suivis d'un blanc souligné.

Les groupes peuvent être utilisés pour afficher une liste des séries, sous forme de feuille de calcul ou de graphique, en double-cliquant sur son nom dans la fenêtre du fichier de travail associé (où ils apparaissent avec le symbole «G» en bleu), ou en l'appelant.

L'affichage par défaut est un format «feuille de calcul», mais on peut passer à des graphiques en utilisant le bouton «View» puis «Graph», ou même éditer la liste des éléments par «Show» puis «Group members».

5.3.2 LES BOUCLES

EViews autorise deux types de boucles:

- Sur des éléments (une liste ou un groupe)

La syntaxe est la suivante:

```
FOR %paramètre liste-de-variables ou groupe-de-noms-de-variables  
  bloc d'instructions incluant {%paramètre} ou %paramètre  
NEXT
```

Le bloc d'instructions sera répété en séquence pour chaque élément de la liste, qui remplacera tour à tour le paramètre.

La présence de crochets autour du paramètre change son statut. Avec des crochets les éléments associés sont inclus dans les instructions sous forme de chaîne de caractères, puis les crochets sont supprimés et l'instruction appliquée. Sans parenthèses, le paramètre est considéré comme une variable chaîne de caractères.

Par exemple, avec

```
%1 = "111"
```

la déclaration

```
genr xxx = {%1}
```

donnera à la série xxx la valeur 111,

alors que

```
xxx = %1
```

permettra de créer une chaîne de caractères avec la valeur "111"

La déclaration

```
genr xxx = %1
```

sera illégale car elle tente de transférer une chaîne de caractères dans une série.

On obtient le message:

```
can not assign string expression to numeric variable in "GENR XXX="111"
```

D'autre part, la déclaration suivante:

```
%2 = %1 + "333"
```

va créer une chaîne "111333", tandis que

```
%2 = {%1} + "333"
```

sera illégale car elle mélange une chaîne et des valeurs:

```
Scalar assigned to string in "%2=111+"333"
```

- Sur des nombres entiers

La syntaxe est la suivante:

```
for !entier=premier-entier to second-entier [by troisième-entier]  
bloc d'instructions incluant {!entier}  
next
```

Le bloc d'instructions est répété dans l'ordre à partir du premier nombre entier jusqu'au deuxième entier, le cas échéant en incrémentant la valeur du paramètre courant par le troisième nombre entier.

Les entiers peuvent être négatifs. Si le troisième entier est omis, l'incrément sera de 1.

Ce type de boucle peut également être appliqué à un groupe

```
for !entier=1 to nom-du-groupe.@count  
%1=nom-du-groupe.@seriesname(!entier)  
bloc d'instructions incluant !entier, %1, et {%1}  
next
```

nom-du-groupe.@count est le nombre d'éléments du groupe *nom-du-groupe*

% 1 reçoit sous forme de chaîne de caractères le contenu de *nom-du-groupe*.@seriesname(!entier), le nom de la variable dans le groupe *nom-du-groupe*, avec le rang !entier.

5.3.3 LES PROGRAMMES: NOUVEAUTES EViews 8

EViews 8 améliore la façon dont les programmes peuvent être exécutés.

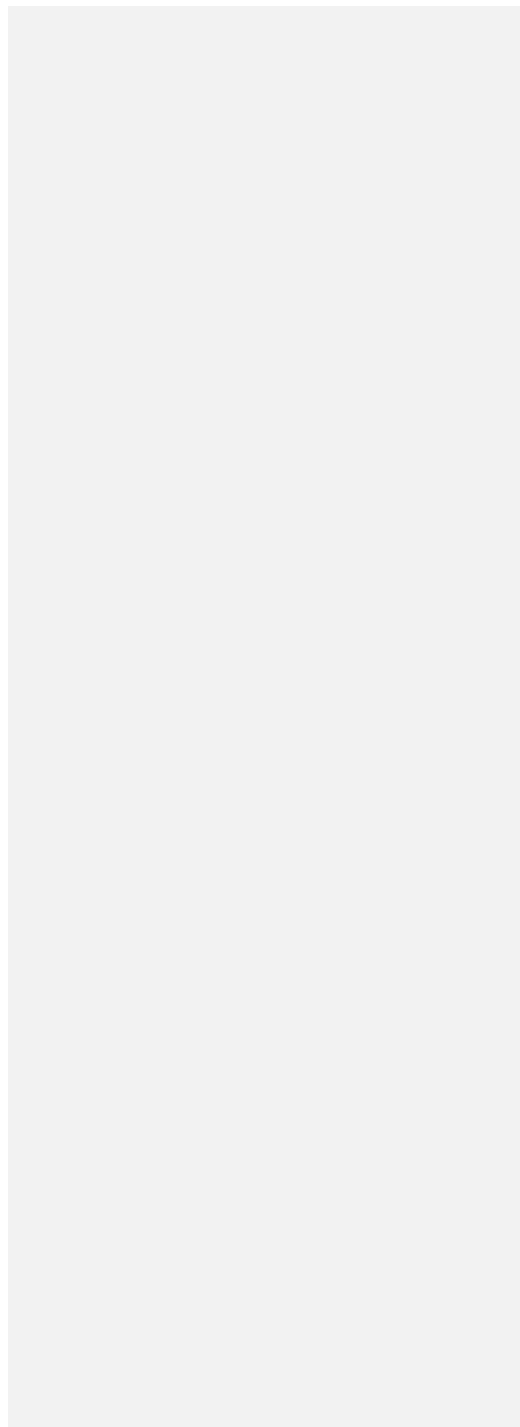
- Vous pouvez exécuter une partie d'un programme, en sélectionnant avec la souris (selon la méthode habituelle), en cliquant sur le bouton droit et en choisissant «Run selected».

Cela est généralement plus efficace que la méthode ancienne consistant à copier la partie sélectionnée dans un programme vide, et en l'exécutant. Toutefois, cette nouvelle méthode ne permet pas la modification.

- Symétriquement, on peut exclure temporairement une partie de l'exécution d'un programme, en la transformant en commentaire. Pour ce faire, il faut sélectionner la partie concernée, cliquer sur le bouton droit et choisir «Comment Selection». Pour réactiver les instructions, il faut les sélectionner à nouveau et utiliser «Uncomment selection».

Cela peut être un peu dangereux, surtout si vous (comme moi) avez le réflexe de sauvegarder le programme avant chaque exécution. Pour éviter de détruire l'original, on peut sauvegarder d'abord le programme à modifier sous un autre nom.

Enfin, on peut demander d'afficher une colonne de chiffres à gauche des lignes de programme. Ceci est particulièrement efficace si vous utilisez la fonction «Go To Line».



CHAPITRE 6: L'ESTIMATION DES EQUATIONS

Nous avons maintenant

- une description complète du cadre du modèle, dans laquelle toutes les identités sont complètement spécifiées, et les intentions en termes de comportements sont également décrites aussi clairement que possible ;
- une base de données complète contenant toutes les séries nécessaires au modèle actuel, endogènes et exogènes, avec leur description.

Nous avons également vérifié que:

- la spécification des identités est compatible avec les données disponibles ;
- les informations obtenues sur la structure du modèle (causalités, interdépendances) sont conformes à nos idées économiques.

Tant la liste des variables que celle des équations sont disponibles sous forme de documents imprimables.

La prochaine étape est évidemment le remplacement de chacun des comportements indicatifs par des comportements effectifs, validés à la fois par la théorie économique et les critères statistiques.

6.1 LE PROCESSUS D'ESTIMATION

Notre but ici n'est pas de donner un cours d'économétrie ou de statistiques¹⁷. Par conséquent, notre approche sera essentiellement pratique, et nous nous contenterons d'éléments de base. Nos raccourcis pourront même introduire quelques légères approximations, qui à notre avis ne sont pas dommageables.

Nous commencerons par décrire les méthodes d'estimation les plus élémentaires.

Ensuite, nous décrirons les tests de base utilisés dans la pratique pour juger d'une formulation, estimée par une méthode quelconque, en observant la qualité de sa description des évolutions passées.

¹⁷ Ce sera également vrai lorsque nous considérerons l'économie. Nous considérons que notre expertise principale n'est pas dans un de ces domaines, mais dans la capacité à les rassembler et à formaliser la méthodologie associée.

Enfin, nous aborderons les techniques plus sophistiquées (mais en fait assez simples): les modèles à correction d'erreur et la cointégration.

6.1.1 LES METHODES D'ESTIMATION

Encore une fois, notre but n'est pas de décrire en détail les différentes méthodes d'estimation, mais de présenter les plus couramment utilisées et donner des indications pour le choix d'une méthode particulière.

6.1.1.1 Moindres carrés ordinaires

Cette méthode s'applique à l'estimation d'une équation linéaire par rapport à des coefficients:

$$y_t = X_t \cdot a + e_t$$

Avec t la dimension du temps (de 1 à T)

y_t le vecteur des observations de la variable calculée (dimension $T \times 1$)

X_t une matrice (dimension $T \times n$, où n est le nombre de variables explicatives)

Aussi bien y et X peuvent représenter des expressions de variables.

Comme son nom l'indique, la méthode consiste à minimiser la somme au cours du temps des carrés des différences entre les valeurs observées et les valeurs obtenues par la formule estimée.

Un calcul simple montre que la matrice des valeurs optimales est

$$\hat{a} = (X'X)^{-1} X'y$$

(avec X' transposée de X)

Pour utiliser véritablement les moindres carrés «ordinaires», l'équation doit inclure un terme constant. Il s'agit d'une condition nécessaire pour que la somme des résidus soit nulle.

Ceci peut être obtenu en introduisant comme dernière variable une série comportant un 1 dans chaque position.

On peut montrer que, dans ce cas, le coefficient de la constante est le suivant:

$$\hat{a}_n = \bar{y} - \sum_{i=1, n-1} \hat{a}_i \cdot \bar{x}_i$$

avec \bar{x} = moyenne de x dans l'échantillon.

et l'estimation des coefficients est effectuée sur la différence des éléments à leur moyenne.

Par exemple l'estimation de la formule

$$y_t = a x_{1t} + b x_{2t} + c$$

conduit à

$$(y_t - \bar{y}) = a (x_{1t} - \bar{x}_1) + b (x_{2t} - \bar{x}_2)$$

$$c = \bar{y} - a \bar{x}_1 - b \bar{x}_2$$

Cette méthode sera souvent utilisée en pratique par les modèles opérationnels. Chaque équation sera estimée de façon indépendante.

6.1.1.2 La prise en compte de l'autocorrélation: une première technique

Même si elle est assez simple, l'application de la méthode précédente exige que les termes aléatoires contenus dans la formulation ne soient pas autocorrélés, ce qui signifie que pour une période donnée le terme est indépendant de tout terme associé aux périodes précédentes. Si ce n'est pas le cas, et surtout si l'autocorrélation est de premier ordre (le diagnostic peut être obtenu par les tests de Durbin-Watson, voir ci-dessous, ou également Breush-Godfrey et Q de Ljung-Box), les résultats d'estimation ne sont pas acceptables, et ne peuvent être utilisés même si les autres statistiques semblent très satisfaisantes. En d'autres termes, la qualité des autres statistiques comme le R2 ou les statistiques de Student ne peut pas

compenser la présence d'autocorrélation, de même que la qualité d'un article ne peut pas (normalement) compenser le fait qu'il ait été envoyé après la date limite.

Cependant, il existe plusieurs façons d'éliminer l'autocorrélation. La plus ancienne et la plus simple est la suivante (la stationnarité et les modèles à correction d'erreur seront traités plus loin).

Supposons que l'équation

$$y_t = X_t \cdot a + e_t$$

est sujette à une autocorrélation du premier ordre, ce qui signifie que les perturbations successives se décomposent en

$$e_t = u_t + r \cdot e_{t-1}$$

où u représente un terme aléatoire non autocorrélé, et r le coefficient de corrélation entre les valeurs consécutives de e .

Il suffit de transformer l'équation en soustrayant des deux côtés son expression retardée, multipliée par r pour obtenir:

$$y_t - r \cdot y_{t-1} = (X_t - r \cdot X_{t-1}) \cdot a + u_t$$

une équation sans autocorrélation, mais incluant un coefficient supplémentaire.

Malheureusement, l'équation n'est plus linéaire par rapport aux coefficients. Ce n'est pas un problème pour EViews, qui ne demande pas de choisir entre moindres carrés ordinaires et non linéaires. La non-linéarité sera détectée automatiquement et la méthode associée appliquée, faisant appel ici à un processus itératif, avec r comme coefficient supplémentaire.

Cette technique est appelée estimation autorégressive (AR) d'ordre 1. L'ordre peut être plus élevé (par exemple, si une erreur trimestrielle dépend de la valeur d'il y a quatre trimestres). Elle peut être combinée avec la moyenne mobile (MA) d'erreurs:

$$e_t = u_t + r \cdot u_{t-1}$$

La combinaison des deux représente un processus ARMA (AR + MA).

6.1.1.3 L'homoscédasticité

Une autre condition nécessaire pour des résultats acceptables est l'homoscédasticité, ce qui signifie que l'écart-type de l'erreur doit être indépendant du temps. Ce ne sera pas le cas, par exemple, pour une variable à prix constants, ou un déflateur, qui suivent généralement l'un la croissance économique, l'autre l'inflation.

Encore une fois, la mise en évidence de l'hétéroscédasticité (à l'opposé de l'homoscédasticité) interdit l'utilisation des résultats de l'estimation, la méthode s'appuyant sur son absence.

Il n'existe pas de méthode générale pour éliminer l'homoscédasticité. En fait, comme nous allons le voir, la plupart des équations estimées le feront naturellement, en utilisant des éléments sans dimension, comme les logarithmes, les taux de croissance ou les ratios. Si ce n'est pas le cas, une méthode simple consiste à diviser tous les éléments par la même variable, par exemple le PIB du pays si l'on veut expliquer une variable à prix constants, ou le déflateur du PIB si l'on estime un indice de prix. Pour éviter tout problème de simultanéité, il est peut être préférable de retarder cette variable d'une période, ou d'utiliser un élément exogène.

6.1.1.4 Les retards échelonnés

Il arrive souvent, surtout dans les modèles à périodicité courte, que plusieurs valeurs décalées de la même variable soient utilisées dans une équation. Dans le cas d'un modèle mensuel, l'évolution du taux de salaire peut être expliquée par les variations des prix à la consommation des 6 mois précédents.

Laisser l'estimation de ces coefficients complètement libre est une source de problèmes :

- Statistiquement, leur signification individuelle peut devenir faible.
- Sur le plan économique, leur profil peut être contraire à la logique: si l'on explique la formation des anticipations d'inflation par les dernières valeurs observées, on s'attend généralement à voir la plus forte influence attribuée aux plus récentes, et chaque influence à être positive, peut-être avec la diminution régulière de sa valeur.

En termes de propriétés du modèle, même si aucune contrainte théorique n'est mise sur les coefficients, une formulation aux coefficients libres pourrait donner à l'équation et au modèle une sensibilité dynamique aux chocs erratique, difficile à interpréter et à justifier.

On peut alors être amené à réduire le nombre de coefficients estimés, en limitant ces coefficients à une relation polynomiale, fonction du décalage de la variable:

$$y_t = \sum_{i=0}^n a_i x_{t-i}$$

va devenir

$$y_t = \sum_{j=0}^m p(i) x_{t-i}$$

où $p(i)$ est un polynôme en i , par exemple $p(i) = a + b \cdot i + c \cdot i^2$

Cette technique est aussi appelée la méthode des polynômes d'Almon¹⁸. On observe qu'elle applique dans la pratique les moindres carrés ordinaires à des variables transformées (s'il n'y a pas de contrainte supplémentaire sur les coefficients¹⁹), pour obtenir les termes des polynômes. Puis, à partir des valeurs numériques obtenues, on peut reconstruire les coefficients des variables réelles.

6.1.1.5 Les moindres carrés non linéaires

Il n'existe aucune obligation de supposer une relation linéaire par rapport aux coefficients. Toutefois, dans le cas contraire, l'expression des coefficients ne peut plus être obtenue analytiquement comme ci-dessus.

¹⁸ Almon, Shirley, "The distributed lag between capital appropriations and net expenditures," *Econometrica* 33, 1965, 178-196.

¹⁹ On peut par exemple obtenir un ensemble de valeurs diminuant progressivement, en fixant le dernier coefficient de décalage à zéro.

On devra utiliser une méthode d'estimation plus complexe, fonctionnant par une suite d'approximations (une technique proche de la méthode de Newton que nous allons aborder lors de l'examen des simulations de modèle).

Parfois, il est possible, grâce à une transformation formelle, de transformer une équation non linéaire en une relation linéaire, avec deux ensembles de coefficients en correspondance biunivoque, ce qui signifie que les ensembles peuvent être déterminés à partir de l'autre d'une manière unique (le nombre de coefficients doit évidemment être le même, mais ce n'est pas suffisant). L'application de la méthode des moindres carrés à l'équation transformée donnera le résultat plus rapidement et sans problèmes de convergence, et les vraies valeurs peuvent être obtenues par une transformation inverse. Mais pas les statistiques de Student associées, ce qui signifie que l'on ne peut pas évaluer la qualité des coefficients originaux. Une solution très simple consiste à recommencer l'estimation non-linéaire à partir des valeurs transformées. Elle devrait converger en un seul passage donnant les «vraies» T-statistiques.

Par exemple, si l'agent est censé adapter sa décision à une cible donnée avec une certaine inertie:

$$y_t = ay_{t-1} + (1-a).(bx_t + c) + e_t$$

On peut estimer avec les MCO:

$$y_t = ay_{t-1} + b'x_t + c'$$

et calculer

b comme $b' / (1-a)$

c comme $c' / (1-a)$

Une estimation non-linéaire à partir de ces valeurs va converger en un seul passage, et donner les bonnes T-statistiques.

L'avantage évident de la forme non linéaire est qu'elle donne la valeur et la significativité de b, l'impact à long terme de x sur y.

6.1.1.6 Les estimations simultanées

Souvent, on sera amené à estimer simultanément plusieurs équations:

- si le même coefficient théorique apparaît dans plusieurs équations. Cela peut se produire par exemple si les équations représentent des dérivées de la même expression par rapport à différents éléments.

L'exemple suivant sera développé plus loin lorsque nous aborderons les modèles opérationnels. Ici, nous allons seulement donner les bases.

Supposons qu'une entreprise veut produire une quantité Q_t , au moindre coût ; son processus de production est décrit par une fonction Cobb-Douglas:

$$Q_t = c \cdot e^{bt} L_t^\alpha \cdot K_t^{(1-\alpha)}$$

ou

$$\text{Log}(Q_t) = c' + b \cdot t + \alpha \cdot \text{Log}(L_t) + (1-\alpha) \cdot \text{Log}(K_t)$$

où L représente l'emploi, et K le capital.

Les prix du travail et du capital sont w et p .

La minimisation du coût, sous la contrainte de la fonction de production, donne pour la combinaison des deux facteurs une élasticité unitaire aux coûts relatifs:

$$w_t / p_t = (1-\alpha) / \alpha \cdot (K_t / L_t)$$

En divisant dans (1) Q par K et L successivement on obtient:

$$\text{Log}(K_t / Q_t) = \alpha \cdot \text{Log}(L_t / K_t) - \text{Log}(c) - b \cdot t$$

$$\text{Log}(L_t / Q_t) = (\alpha - 1) \cdot \text{Log}(L_t / K_t) - \text{Log}(c) - b \cdot t$$

Ce qui avec l'équation précédente donne:

$$\text{Log}(K_t / Q_t) = \alpha \cdot \text{Log}((1 - \alpha) / \alpha \cdot (w_t / p_t)) - \text{Log}(c) - b \cdot t$$

$$\text{Log}(L_t / Q_t) = (\alpha - 1) \cdot \text{Log}((1 - \alpha) / \alpha \cdot (w_t / p_t)) - \text{Log}(c) - b \cdot t$$

dont l'estimation simultanée de coefficients donne α , b et c

- soit parce qu'une variable expliquée par une équation est utilisée comme élément d'explication dans une autre. Cela interdit une estimation séparée (qui donne des estimateurs biaisés et incohérents).

Exemple: les prix à l'exportation, calculés à partir des prix de production et des prix à l'étranger, déterminent la compétitivité des exportations dans l'équation associée.

Deux types de méthodes peuvent être envisagés pour des estimations simultanées:

- une information limitée: chaque équation est estimée séparément, mais prend en compte les informations fournies par les autres (exemple: doubles moindres carrés).
- une information complète: on estime l'ensemble des équations simultanément, y compris les identités si nécessaire (par exemple: maximum de vraisemblance à information complète).

On peut remarquer que ces méthodes:

- demandent un plus grand nombre d'observations pour donner des résultats significatifs (ou même théoriquement valides) ;
- doivent être appliquées à une période où toutes les variables sont connues (donc à l'intersection de chaque échantillon) ;
- doivent être appliquées à nouveau quand une seule valeur est modifiée ;
- sont beaucoup plus coûteuses que d'autres en temps de calcul, compte tenu de la complexité des algorithmes et leur caractère itératif (mais ce critère n'est plus vraiment pertinent) ;

- réclament un processus itératif qui peut rencontrer certaines difficultés dans l'obtention de la solution (et échoue souvent, surtout si l'on n'a pas de point de départ satisfaisant pour les coefficients).

Pour toutes ces raisons, ce type de méthode est généralement limité à de petits modèles. Pour un grand modèle, cette méthode sera limitée à un sous-ensemble choisi pour la densité de ses interactions, comme l'ensemble des prix et des salaires, ou les éléments du commerce extérieur.

Et en fait, comme nous le verrons, les tests de cointégration entre un sous-ensemble de ces variables peuvent libérer de cette contrainte de simultanéité globale.

6.1.1.7 Les estimations en données de panel

Nous avons déjà évoqué l'existence de modèles temporels présentant une dimension supplémentaire: par exemple le pays ou le produit. Dans le cas où ces modèles présentent des équations de comportement similaires, se pose le problème d'estimer des coefficients communs ou individualisés.

À partir du moment où un ensemble d'équations contient un coefficient à valeur commune, l'estimation doit se faire de façon simultanée²⁰.

Il faut simplement distinguer les coefficients communs et individuels.

EViews propose deux méthodes.

6.1.1.7.1 Les «pools»

On commence par définir un ensemble de préfixes (ou suffixes) par l'instruction:4

```
POOL nom_de_l'ensemble liste_des_éléments.
```

Puis on applique une méthode d'estimation à l'ensemble, en faisant apparaître dans le liste des éléments (expliqués et explicatifs) le symbole «?». L'équation sera développée en autant de formules que le «pool».

²⁰ Même si on peut sélectionner et estimer une première équation, puis le reste de l'ensemble en utilisant la valeur trouvée comme un scalaire fixe.

Le mot-clé «@cxreg» permettra de séparer les coefficients à valeur commune et individualisée. Par exemple avec les moindres carrés:

```
Nom_du_pool.ls élément_expliqué éléments_explicatifs_à coefficient_commun @cxreg
éléments_explicatifs_à coefficient_individuel
```

Exemple d'application:

Avec un ensemble de pays: Allemagne, France, Grande Bretagne, Italie, Japon, USA associé aux préfixes: de_fr_gb_it_jp_us_

L'estimation de l'élasticité des importations M à la demande TD (avec prise en compte d'une tendance temporelle et d'une autocorrélation du premier ordre) se fera par

```
POOL pays de_fr_gb_it_jp_us_
pays.LS log(?m) log( ?td) @cxreg c t ar(1)
```

ce qui donnera:

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(?TD)	1.668930	0.090295	18.48307	0.0000
DE_--C	-86.18276	7.952445	-10.83727	0.0000
FR_--C	-36.74548	11.99882	-3.062424	0.0023
GB_--C	-22.47540	3.868119	-5.810422	0.0000
IT_--C	-29.42198	4.705607	-6.252536	0.0000
JP_--C	-56.84432	58.54691	-0.970919	0.3319
US_--C	-36.79270	6.926550	-5.311835	0.0000
DE_--T	0.037652	0.004099	9.184679	0.0000
FR_--T	0.012968	0.006147	2.109639	0.0352
GB_--T	0.005998	0.002487	2.411626	0.0161
IT_--T	0.009313	0.002704	3.444734	0.0006
JP_--T	0.020412	0.028904	0.706215	0.4803
US_--T	0.011963	0.003933	3.041377	0.0024
DE_--AR(1)	0.833152	0.082015	10.15859	0.0000
FR_--AR(1)	0.966441	0.032854	29.41611	0.0000
GB_--AR(1)	0.735066	0.052790	13.92434	0.0000
IT_--AR(1)	0.908306	0.027937	32.51251	0.0000
JP_--AR(1)	0.986891	0.013911	70.94257	0.0000
US_--AR(1)	0.929763	0.027704	33.56102	0.0000
R-squared	0.999830	Mean dependent var	11.94817	
Adjusted R-squared	0.999827	S.D. dependent var	1.969400	
S.E. of regression	0.025930	Akaike info criterion	-4.444664	
Sum squared resid	0.556027	Schwarz criterion	-4.338198	
Log likelihood	1899.093	Hannan-Quinn criter.	-4.403873	
F-statistic	270762.3	Durbin-Watson stat	2.149281	
Prob(F-statistic)	0.000000			

On constatera que l'on peut introduire des expressions. Par contre l'équation doit être linéaire par rapport aux coefficients.

6.1.1.7.2 Les systèmes

Cette méthode consiste simplement à définir un groupe d'équations (éventuellement par une boucle) en associant chaque coefficient soit à l'identifiant du pays soit à un identifiant global.

```
pool _pays_p de_fr_gb_it_jp_us_
delete(noerr) _pays _s
system _pays_s
coef(10) c_to
for %1 de fr gb it jp us
coef(10) c_{%1}
_pays_s.append log({%1}_m)=c_to(1)*log({%1}_td)+c_{%1}(2)+c_{%1}(2)*t+[ar(1)=c_{%1}(4)]
next
pays_s.sur(p)
```

On obtient:

System: _COUNTRY_S
 Estimation Method: Seemingly Unrelated Regression
 Date: 05/23/13 Time: 12:59
 Sample: 1970Q2 2008Q4
 Included observations: 156
 Total system (unbalanced) observations 846
 Iterate coefficients after one-step weighting matrix
 Convergence achieved after: 1 weight matrix, 14 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C_TO(1)	1.672614	0.087977	19.01194	0.0000
C_DE(2)	-86.01534	5.023653	-17.12207	0.0000
C_DE(3)	0.037543	0.002697	13.91914	0.0000
C_DE(4)	0.839766	0.049179	17.07586	0.0000
C_FR(2)	-35.04651	4.961937	-7.063069	0.0000
C_FR(3)	0.012095	0.002812	4.301701	0.0000
C_FR(4)	0.946157	0.019253	49.14282	0.0000
C_GB(2)	-22.30293	3.831344	-5.821178	0.0000
C_GB(3)	0.005888	0.002448	2.405362	0.0164
C_GB(4)	0.746599	0.053806	13.87587	0.0000
C_IT(2)	-29.08744	4.810399	-6.046783	0.0000
C_IT(3)	0.009122	0.002739	3.330232	0.0009
C_IT(4)	0.895795	0.032271	27.75855	0.0000
C_JP(2)	-31.72246	20.23412	-1.567771	0.1173
C_JP(3)	0.007904	0.010148	0.778800	0.4363
C_JP(4)	0.972514	0.015556	62.51750	0.0000
C_US(2)	-35.67751	5.641492	-6.324127	0.0000
C_US(3)	0.011378	0.003353	3.393792	0.0007
C_US(4)	0.902671	0.025650	35.19168	0.0000

$$\text{Equation: LOG(IT_M)=C_TO(1)*LOG(IT_TD)+C_IT(2)+C_IT(3)*T} \\ +[\text{AR}(1)=\text{C_IT}(4)]$$

Observations: 155

R-squared	0.996307	Mean dependent var	10.64985
Adjusted R-squared	0.996233	S.D. dependent var	0.503444
S.E. of regression	0.030898	Sum squared resid	0.144162
Durbin-Watson stat	2.293181		

$$\text{Equation: LOG(JP_M)=C_TO(1)*LOG(JP_TD)+C_JP(2)+C_JP(3)*T} \\ +[\text{AR}(1)=\text{C_JP}(4)]$$

Observations: 155

R-squared	0.996538	Mean dependent var	15.76677
Adjusted R-squared	0.996470	S.D. dependent var	0.513102
S.E. of regression	0.030487	Sum squared resid	0.140351
Durbin-Watson stat	1.807015		

$$\text{Equation: LOG(US_M)=C_TO(1)*LOG(US_TD)+C_US(2)+C_US(3)*T} \\ +[\text{AR}(1)=\text{C_US}(4)]$$

Observations: 155

R-squared	0.998734	Mean dependent var	11.94501
Adjusted R-squared	0.998709	S.D. dependent var	0.720805
S.E. of regression	0.025900	Sum squared resid	0.101295
Durbin-Watson stat	2.293713		

Determinant residual covariance 1.69E-20

$$\text{Equation: LOG(DE_M)=C_TO(1)*LOG(DE_TD)+C_DE(2)+C_DE(3)*T} \\ +[\text{AR}(1)=\text{C_DE}(4)]$$

Observations: 71

R-squared	0.997128	Mean dependent var	11.92017
Adjusted R-squared	0.996999	S.D. dependent var	0.295246
S.E. of regression	0.016174	Sum squared resid	0.017526
Durbin-Watson stat	2.175192		

$$\text{Equation: LOG(FR_M)=C_TO(1)*LOG(FR_TD)+C_FR(2)+C_FR(3)*T} \\ +[\text{AR}(1)=\text{C_FR}(4)]$$

Observations: 155

R-squared	0.999109	Mean dependent var	10.86283
Adjusted R-squared	0.999091	S.D. dependent var	0.542555
S.E. of regression	0.016354	Sum squared resid	0.040386
Durbin-Watson stat	1.867655		

$$\text{Equation: LOG(GB_M)=C_TO(1)*LOG(GB_TD)+C_GB(2)+C_GB(3)*T} \\ +[\text{AR}(1)=\text{C_GB}(4)]$$

Observations: 155

R-squared	0.997726	Mean dependent var	10.52924
Adjusted R-squared	0.997681	S.D. dependent var	0.570886
S.E. of regression	0.027492	Sum squared resid	0.114129
Durbin-Watson stat	2.151890		

On constate que les résultats sont légèrement différents, la méthode employée l'étant aussi.

Entre les deux, nous préférons logiquement la seconde, qui permet d'employer des formules non linéaires par rapport aux coefficients, et identifie clairement le pays d'appartenance de ces derniers (ou leur caractère global).

6.1.2 LES TESTS

Nous supposons que nous testons un lien explicite causal, dans lequel une expression donnée des variables du modèle sera expliquée par une formule incluant des paramètres inconnus.

L'équation va calculer la variable i du vecteur y endogène à l'aide de:

$$g(y_t^i, y_{t-1}, x_t) = f(y_t, y_{t-1}, x_t, \alpha)$$

Le fait que l'expression, et non une seule variable, apparaît sur le côté gauche ne rend pas le processus d'estimation plus complexe, car on peut calculer avant toute estimation les valeurs numériques de ce terme. C'est seulement lors de la simulation du modèle que des problèmes peuvent apparaître.

Par exemple, nous supposons que les importations M présentent une élasticité constante à la demande locale totale TD , avec des coefficients inconnus, et un résidu (évidemment avec une répartition différente):

$$\Delta M_t / M_t / (\Delta TD_t / TD_t) = a + u_t$$

ou

$$\text{Log}(M_t) = a \cdot \text{Log}(TD_t) + b + u_t$$

Dans la pratique, comme si une seule variable apparaît à la gauche du signe égale, on peut toujours formellement remplacer l'expression côté gauche par une nouvelle variable, définie par une identité supplémentaire.

Nous ouvrons écrire:

$$g(y_t, y_{t-1}, x_t) = z_t$$

La formulation estimée sera représentée par:

$$\hat{z}_t = f(y_t, y_{t-1}, x_t, \hat{a})$$

et l'erreur d'estimation par:

$$e_t = z_t - \hat{z}_t$$

L'erreur est l'élément qui doit être ajouté à la suite de la formule pour obtenir la valeur exacte de la variable.

6.1.2.1 Qualité globale

Ici, l'intérêt principal réside dans la qualité globale de l'estimation.

6.1.2.1.1 L'autocorrélation temporelle

Comme nous l'avons indiqué plus haut, la plupart des méthodes d'estimation nécessitent, pour que leurs résultats puissent être utilisés, l'absence d'autocorrélation temporelle, ou d'un lien entre les valeurs des termes aléatoires associées à des périodes différentes. La tradition propose le test de Durbin et Watson:

$$DW = T/(T-1) \sum_{t=k+1}^T (e_t - e_{t-k})^2 / \sum_{t=1}^T e_t^2$$

Ce test va chercher à mesurer si les termes aléatoires sont temporellement indépendants, dans la plupart des cas, d'une période à l'autre ($k = 1$).

$$E(e_t \cdot e_{t-k}) = 0$$

Ce test estime approximativement

$$2(1 - \rho)$$

Dans le cas où le terme aléatoire suit la loi:

$$e_t = \rho \cdot e_{t-k} + u_t$$

avec u_t non autocorrélé.

L'évolution du coefficient d'autocorrélation entre ses valeurs extrêmes -1 et +1 donne au test un intervalle compris entre 0 et +4²¹, la valeur 2 traduisant l'absence de corrélation²².

Bien que le test de Durbin-Watson soit le plus familier, il présente des lacunes: il est indécis pour certains intervalles, et échoue si l'explication contient la variable dépendante retardée. À l'heure actuelle la plupart des économistes préfèrent utiliser le test de Breusch-Godfrey ou le test Q de Ljung-Box. Les deux sont disponibles dans EViews mais ne s'affichent pas (encore) sur l'écran principal des résultats.

En général, on sera intéressé (sauf si l'équation présente un caractère saisonnier potentiel) par l'autocorrélation de premier ordre, associée à des valeurs consécutives des résidus.

Mais dans la pratique, une mauvaise statistique sera souvent le symptôme d'une mauvaise spécification, plutôt que de la présence d'autocorrélation dans la bonne formulation²³. Un exemple est d'estimation linéaire d'une formulation vraie en logarithmes. Notre équation (14):

²¹ En pratique, la valeur du test ne peut pas être égale à 0 ou 4 avec les moindres carrés ordinaires, car la somme des résidus étant nulle ils ne peuvent pas tous avoir le même signe.

²² En pratique, les choses sont un peu plus complexes et il faudra faire appel à des tables en tenant compte du nombre d'observations.

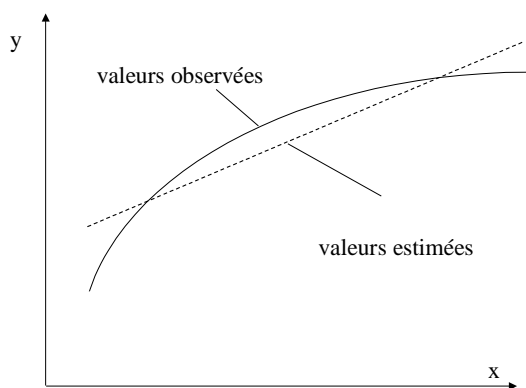
²³ En fait, un cas très particulier de mauvaise spécification.

$$M_t = a \cdot TD_t + b$$

peut présenter ce problème si la vraie formulation est:

$$\text{Log}(M_t) = a \cdot \text{Log}(TD_t) + b$$

Essayer d'adapter une droite à une courbe convexe mènera (voir graphique) à un ensemble de résidus ne changeant de signe que deux fois.



Des formulations mieux adaptées seront

$$\Delta x_t = a + b \cdot \Delta y_t + u_t$$

ou même

$$\Delta x_t / x_t = a + b \cdot \Delta y_t / y_t + u_t$$

qui se rapproche de la formulation en logarithmes (avec un terme constant qui s'associera à une tendance temporelle).

La bonne estimation donnera une bonne valeur du test, l'autre une très mauvaise.

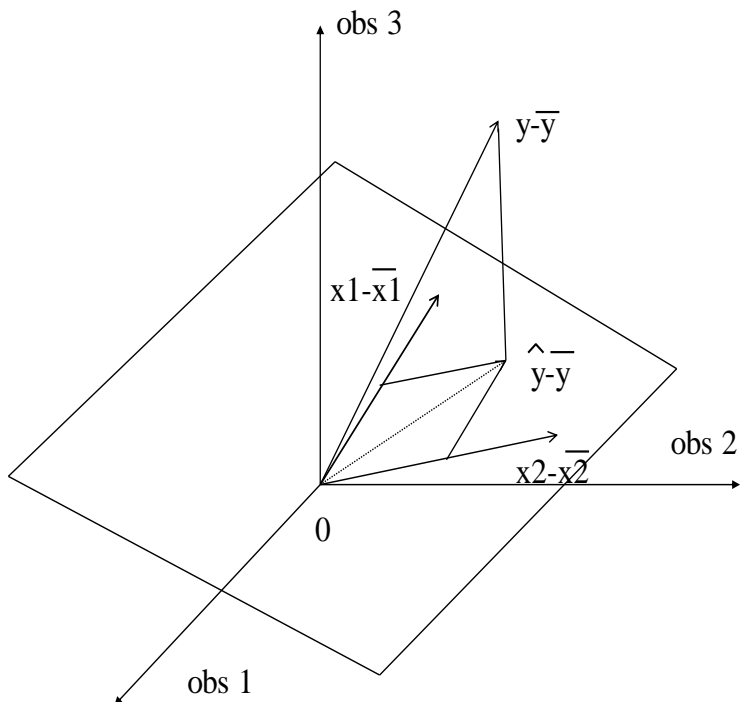
6.1.2.1.2 La statistique du R2 ou R-carré

La statistique appelée «R2» ou «R-carré» est la plus couramment utilisée pour juger de la qualité globale d'une estimation. Elle est définie par la formule suivante.

$$R^2 = \frac{\sum_{t=1}^T (\hat{x}_t - \bar{x})^2}{\sum_{t=1}^T (x_t - \bar{x})^2}$$

Cette statistique peut donc être interprétée comme la proportion de la variance de la variable observée x qui va être expliquée par la formule estimée.

Une explication géométrique peut également être utilisée: si l'on considère l'espace des variables (dimension T = nombre d'observations), la méthode d'estimation consistera à minimiser la distance entre la variable expliquée et l'espace (le plan ou l'hyper plan) généré par le vecteur des séries explicatives, en utilisant des combinaisons de valeurs des paramètres.



Et surtout, si la formule est linéaire par rapport aux paramètres estimés et contient un terme constant, on peut considérer que l'estimation est basée sur la différence des variables (expliquée et explicatives) à leur moyenne. Dans ce cas, réduire au minimum la distance euclidienne conduira (comme on peut le voir sur le graphique) le vecteur $(\hat{y}_t - y_t)$ à être orthogonal à l'espace et donc au vecteur $(\hat{y} - \bar{y})$.

Ces deux éléments représentent les parts non-expliquées et expliquées de la variance. Le R^2 peut être interprété comme le carré du cosinus de l'angle entre la série observée et ajustée: plus le R^2 est proche de 1, plus l'angle sera faible et plus la part de la variable estimée dans l'explication de la variance sera

élevée. L'explication sera parfaite si $y - \bar{y}$ appartient à l'espace, et nulle si la perpendiculaire rencontre l'espace à l'origine.

Si l'équation ne présente pas de terme constant, le même raisonnement peut être appliqué, mais cette fois, la moyenne n'est pas soustraite. Toutefois, le R2 n'a plus le même sens: au lieu de la variance, la somme directe des carrés sera utilisée.

Nous n'irons pas plus loin dans l'explication de ce test, préférant nous concentrer sur ses propriétés pratiques.

6.1.2.1.3 Le R2 en question

Il faut être très prudent lorsque vous utilisez la statistique du R2.

- Elle augmente avec les tendances sur les variables

La statistique du R2 sera d'autant plus élevée que la variable expliquée et au moins l'une des variables explicatives présente une tendance temporelle selon le rang de l'observation. Ainsi, les composantes de chacune de ces variables sur les axes des observations augmenteront dans le même sens ou en sens opposé (de très négatif à très positif, ou l'inverse), ce qui donnera aux vecteurs associés des orientations très proches. Sur le graphique ci-dessus, les composantes de variables sur les axes seront plus ou moins ordonnées selon la numérotation des axes eux-mêmes. Les premières observations seront les plus négatives, les valeurs augmenteront vers zéro et atteindront les plus positives à la fin. La même observation est valable si les directions sont opposées: l'estimation mettra alors en évidence un lien négatif.

Dans ce cas, même en l'absence d'une véritable relation de causalité entre les variables, l'orientation des vecteurs sera semblable, à un facteur multiplicatif près donné, et le test R2 semblera valider la formulation. Et comme la plupart des séries temporelles (comme les valeurs, les quantités ou les prix) présentent généralement une tendance positive, ce phénomène a une bonne chance de se produire. Par exemple, si nous appliquons l'équation précédente aux importations françaises:

$$\text{Log}(M_t) = a \cdot \text{Log}(TD_t) + b + u_t$$

Le remplacement de TD par une variable présentant une tendance croissante²⁴ va donner un « bon » R2, meilleur peut-être que la demande française elle-même.

En fait, il peut être montré que si l'on teste pour un pays donné de l'OCDE l'explication de ses importations par la demande de tous les pays en séquence, la « vraie » équation ne donne pas généralement le meilleur résultat, même si elle n'en est généralement pas loin.

- Elle donne des diagnostics trompeurs lorsque l'on compare des estimations expliquant des éléments différents.

Cela se produit en particulier lorsque nous expliquerons le même concept en utilisant une transformation différente.

Dans notre équation 14, faisons passer le terme retardé à droite.

$$\text{Log}(M_t) = \text{Log}(M_{t-1}) + a \cdot \Delta \text{Log}(TD_t) + b + v_t$$

Par rapport à la formule initiale, cette transformation ne change pas l'explication²⁵, comme la minimisation de la somme des carrés des résidus est évidemment associée au même processus. La seule statistique modifiée sera le R2, qui augmentera beaucoup, car des éléments identiques de variance élevée (comparée à celle de $\Delta \text{Log}(M_t)$) ont été ajoutés des deux côtés.

Le choix entre les deux formulations ne doit pas se baser sur le R2 mais sur l'autocorrélation du résidu: si u_t n'est pas corrélé on doit utiliser (1), s'il l'est, on doit essayer (2). Mais de toute façon, le débat sera tranché par les modèles de correction d'erreur et de cointégration, que nous allons aborder plus tard.

En conclusion:

²⁴ Par exemple la demande australienne, ou le prix d'un paquet de cigarettes en Ouzbékistan.

²⁵ Avant l'estimation EViews déplace le terme décalé vers la partie gauche.

Dans la pratique, une «bonne» variable explicative doit faire mieux que d'expliquer une simple tendance. Cela signifie qu'elle doit aussi suivre ses fluctuations autour de la tendance, ou présenter les mêmes ruptures de tendance, quand elles se produisent.

Nous verrons que si nous expliquons les importations par la demande, les périodes pour lesquelles les importations augmentent plus lentement que la moyenne correspondent à un ralentissement de la croissance de la demande. Et si la tendance de la demande se modifie, l'évolution des importations suit.

Malheureusement, ni l'étude R2 ni celle de la statistique de Student (expliquée plus loin) associé à l'élément explicatif ne vont nous donner cette information. Nous devons compter sur:

- un test d'autocorrélation comme condition de la cohérence des estimations,
- le graphique des valeurs observées et estimées,
- l'écart-type des estimations, surtout quand il peut être interprété directement (voir ci-dessous),
- la comparaison des statistiques entre les formulations possibles (y compris le R2 à variable expliquée inchangée).

6.1.2.1.4 Autres tests de qualité globale

Lorsque les séries présentent une tendance croissante, il est conseillé de compléter la statistique R2 par une autre, qui présentera l'avantage supplémentaire d'être plus facile à interpréter:

- L'écart type des résidus d'estimation:

$$see = \sqrt{1/(T-k) \sum_{t=1}^T (y_t - \hat{y}_t)^2}$$

où k est le nombre de variables explicatives, éventuellement divisé par la moyenne de la variable.

- L'erreur moyenne absolue en niveau

$$eam = 1/T \sum_{t=1}^T |y_t - \hat{y}_t|$$

- ou en pourcentage:

$$\text{eamp} = 100/T \sum_{t=1}^T |y_t - \hat{y}_t|$$

Pour certaines expressions estimées, l'interprétation de l'écart-type sera particulièrement facile. Pour une équation en élasticités, comme l'équation (14) de notre modèle, qui relie les importations à la demande:

$$\text{Log}(M_t) = a \cdot \text{Log}(TD_t) + b$$

L'écart-type estimé de l'expression peut être interprété comme une erreur relative: une erreur absolue moyenne de 0,05 sur $\text{Log}(M)$ représente une erreur de 5 % sur M^{26} .

De même, si l'on estime un taux de croissance, l'erreur absolue sera interprétée comme la variation de l'erreur relative introduite par la période. Si en estimant:

$$\Delta \text{Log}(M_t) = a \cdot \Delta \text{Log}(TD_t) + b$$

on obtient une erreur de 0,05, cette valeur sera interprétée, non plus comme le pourcentage d'erreur moyen sur M_t , mais comme l'erreur absolue sur la variation relative entre la période t et la période $t-1$.

²⁶ En effet en notant \hat{M}_t la valeur estimée de M , on obtient:

$$\begin{aligned} \text{Log}(M_t) - \text{Log}(\hat{M}_t) &= -\text{Log}(\hat{M}_t / M_t) \\ &= \text{Log}(1 + (M_t - \hat{M}_t) / M_t) \\ &\approx (M_t - \hat{M}_t) / M_t \end{aligned}$$

En d'autres termes, dans une formule dynamique, 5 points seront ajoutés à l'erreur relative précédente²⁷. Nous verrons comment comparer ces deux critères lorsque nous aborderons la simulation des modèles.

D'autres tests plus sophistiqués permettent de juger de la qualité globale de l'estimation. Nous ne les traiterons pas ici.

6.1.2.1.5 Le test de rupture de Chow

Un autre test important est utilisé pour vérifier que la formulation ne présente pas de rupture dans la période de référence, en d'autres termes que, étant donné une équation de comportement contenant des coefficients inconnus, leurs valeurs ne changent pas d'une sous-période à l'autre. Cela se fait en testant si la qualité de l'explication se détériore si l'on considère la période d'estimation toute entière, comparativement à des explications individuelles données par des estimations distinctes sur des sous-échantillons. Ce critère ne tient pas compte des coefficients individuels, ou des formulations alternatives. La question est simplement de savoir si le comportement logique décrit par la formule unique conserve le même jeu de coefficients au cours de la période, et si la qualité de la description globale n'est pas affectée par cette hypothèse.

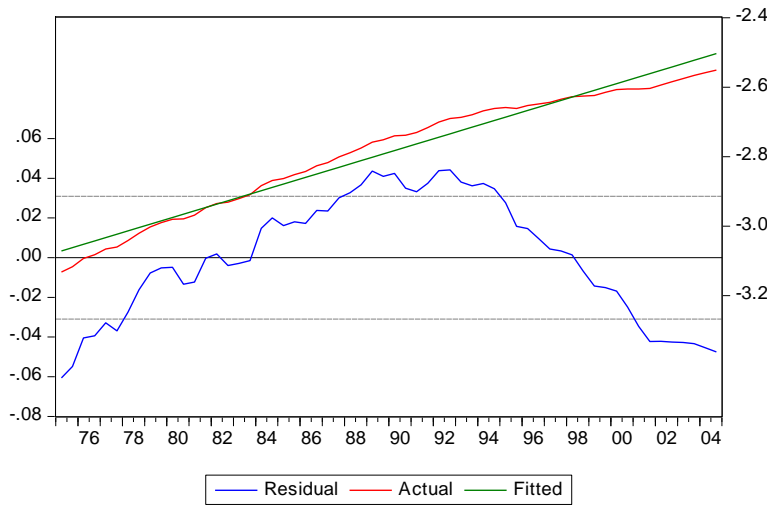
La base du test est assez simple.

- L'équation est estimée à trois reprises (en fait 2 plus le nombre de ruptures): sur l'échantillon complet et chacune des sous-périodes.
- Le rapport de la somme des variances des sous-échantillons à la variance globale est calculé.
- Il est comparé à une table de valeurs pour déterminer la probabilité de l'absence de rupture. Plus le ratio est faible, plus la probabilité d'une rupture est importante.

Un problème posé par ce test est qu'une rupture évidente sera acceptée, même pour des périodes assez éloignées de sa date réelle.

Prenons la productivité du travail en France, traduite par une série biannuelle de 1974 à 2004, et envisageons des ruptures dans son taux de croissance (ou des ruptures linéaires dans son logarithme). Comparer son évolution avec une tendance met en évidence une rupture, autour de 1992:

²⁷ Cela ne veut pas dire qu'il va augmenter, car sa valeur initiale peut être négative.



Le test de Chow accepte la présence d'une rupture.

Chow Breakpoint Test: 1992S1

Null Hypothesis: No breaks at specified breakpoints

Equation Sample: 1975S1 2004S2

F-statistic	594.9474	Prob. F(2,56)	0.0000
Log likelihood ratio	186.1355	Prob. Chi-Square(2)	0.0000
Wald Statistic	1189.895	Prob. Chi-Square(2)	0.0000

mais 1985 fonctionne aussi :

Chow Breakpoint Test: 1985S1

Null Hypothesis: No breaks at specified breakpoints

Equation Sample: 1975S1 2004S2

F-statistic	98.13185	Prob. F(2,56)	0.0000
Log likelihood ratio	90.30740	Prob. Chi-Square(2)	0.0000
Wald Statistic	196.2637	Prob. Chi-Square(2)	0.0000

Avec une probabilité beaucoup plus faible, toutefois.

On doit répéter le test pour connaître les dates potentielles, en choisissant celle qui donne la meilleure valeur.

6.1.2.2 La qualité des influences individuelles

On peut également tester la signification individuelle de chaque coefficient, en comparant leur valeur estimée avec leur écart-type estimé (la précision absolue avec laquelle ils sont évalués). Le rapport (appelé le T de Student) donne une évaluation de la précision relative. Si nous supposons que les coefficients suivent une loi normale, on peut associer le rapport avec une probabilité pour le coefficient d'être significativement différent de zéro (donc d'avoir le bon signe).

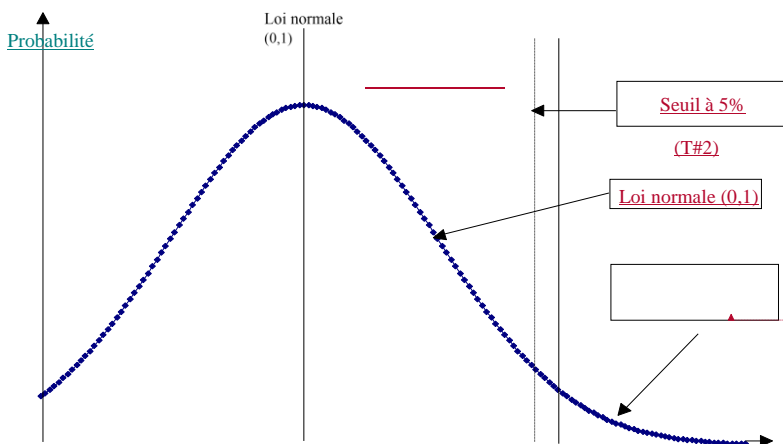
En supposant que le coefficient suit une loi normale avec un écart-type égal à la valeur estimée, le diviser par cette valeur donne un résultat suivant une loi normale (0,1). Une comparaison avec les tables associées donne la probabilité d'atteindre ce résultat ou une valeur supérieure.

Comme avec une loi normale (0,1) environ 95 % des valeurs se situent dans l'intervalle [-2, +2], atteindre une valeur supérieure à 2 (ou inférieure à -2) a une probabilité inférieure à 2.5 % si la loi normale s'applique bien.

La statistique:

$$T = \text{coefficient estimé} / \text{écart-type estimé du coefficient}$$

sera appelée T-statistique de Student.



Mis en forme : Français (France)

Bien sûr, l'équation doit être bien formulée par ailleurs, en particulier en ce qui concerne les autres coefficients.

L'expérience montre qu'il ne faut pas se concentrer sur la valeur 2. Tout d'abord, elle n'est pas vraiment associée à un seuil de 5 %, en particulier pour de petits échantillons. Deuxièmement, les probabilités diminuent très vite au-dessus de 2 (et augmentent lorsque nous revenons vers 0). Ainsi une valeur de 3 est associée à 0,5 %, mais une valeur de 1,30 avec 20 %. Troisièmement, l'objection «croissance similaire» vaut autant que pour le R2: si la variable expliquée et les explicatives croissent dans le même sens ou dans une direction opposée, une valeur élevée de T sera obtenue systématiquement, et il faut être beaucoup plus exigeant²⁸.

Le seuil dépend en fait du nombre d'observations et de variables dans l'équation. Plus la quantité d'information dont nous disposons est importante, et plus le nombre de coefficients est bas, meilleure sera la qualité de l'explication. Plus la différence entre les deux (le nombre de degrés de liberté) augmente, et plus la probabilité associée à une valeur donnée de T s'améliorera.

²⁸ Mais avec plus d'une variable explicative, la nature du problème change. Si au moins deux d'entre elles présentent des tendances, elles doivent fournir plus de l'explication que la tendance de la variable expliquée, pour que le processus d'estimation puisse déterminer leur contribution.

Il est nécessaire toutefois de ne pas surestimer la qualité du résultat. Une réussite proche du seuil ne garantit qu'un signe, et laisse une latitude importante aux valeurs que le coefficient peut prendre. Si l'on veut un modèle aux propriétés numériques fiables, une valeur beaucoup plus élevée de T sera nécessaire²⁹.

En fait, la statistique T peut être généralisée à la comparaison avec n'importe quelle valeur numérique. On peut calculer un intervalle «avec une probabilité de 95 %» autour de \hat{a} , comme

$$E = [\hat{a} - 2\hat{\sigma}, \hat{a} + 2\hat{\sigma}]$$

Par rapport à tout élément a , situé en dehors de cet intervalle, \hat{a} est en effet significativement plus grand si

$$a < \hat{a} - 2\hat{\sigma} \Leftrightarrow (\hat{a} - a) / \hat{\sigma} > 2$$

et plus faible si

$$a > \hat{a} + 2\hat{\sigma} \Leftrightarrow (\hat{a} - a) / \hat{\sigma} < -2$$

Par exemple, si l'on veut être sûr à 10% près de la valeur du coefficient, le T va diminuer de 90%. La valeur affichée par EViews devra être d'au moins 20.

En fait, beaucoup de coefficients ont déjà une valeur théorique, ou un intervalle théorique. Par exemple, l'indexation à long terme des salaires sur l'inflation peut être considérée comme unitaire, la sensibilité à court terme de l'emploi à la production doit être comprise entre 0 et 1 (et pas trop près de ces limites), la part du travail dans la fonction de production Cobb-Douglas est généralement censée être proche de la part du coût salarial dans le prix à la valeur ajoutée, ce qui représente généralement les deux tiers. La

²⁹ Par exemple, une valeur T de 4 signifie que le vrai coefficient doit être au moins supérieur à la moitié de sa valeur estimée.

procédure ci-dessus peut nous dire si le coefficient associé est significativement différent de cette valeur théorique, ou nettement en dehors de l'intervalle théorique.

Si ce n'est pas le cas, alors la fixation de la valeur à la valeur théorique, ou à la valeur la plus proche à l'intérieur des limites théoriques, peuvent être autorisées. Par exemple, si nous estimons un coefficient d'indexation des salaires sur les prix de 1,05 avec un écart type de .06, on peut considérer le mettre à 1, la valeur acceptable ayant la plus forte probabilité (car une sur indexation peut nuire aux propriétés du modèle, ou même mettre en cause sa convergence).

On peut également appliquer d'autres tests, sur des hypothèses plus ou moins complexes: la stabilité des coefficients individuels à l'intérieur de la période, la significativité globale d'un sous-ensemble, l'exogénéité, la cointégration. Nous ne les traiterons pas ici (seulement pour la dernière fonctionnalité), mais nous donnerons plutôt des références à la fin du livre.

6.1.2.3 Le terme constant

Lorsque l'on observe la validité des influences individuelles, un élément joue un rôle très spécifique: le terme constant.

Cet élément peut avoir deux objectifs:

- Gérer le fait que l'équation ne tient pas compte des éléments bruts, mais de leurs écarts par rapport à leur moyenne. Avec les moindres carrés ordinaires, même si le résultat final est une formulation linéaire des variables et un terme constant, le processus
 - calcule les écarts à la moyenne,
 - les utilise pour évaluer une formule sans constante³⁰,
 - recombine les coefficients estimés et les moyennes pour construire un terme constant.

Cette constante est donc une partie intégrante du processus. Elle devrait être incluse à chaque fois qu'au moins un des éléments explicatifs n'a pas de moyenne nulle.

- Décrire un mécanisme économique.

Donnons un exemple pour le premier cas: si les importations ont une élasticité constante par rapport à la demande, nous estimons:

³⁰ Comme tous les éléments de la formule ont une moyenne nulle, la somme des résidus également.

$$\Delta M_t / M_t = a \cdot \Delta TD_t / TD_t$$

ou

$$\text{Log}(M_t) = a \cdot \text{Log}(TD_t) + b$$

mais le processus d'estimation va d'abord utiliser la différence de la moyenne pour obtenir

$$\text{Log}(M_t) - \text{Log}(\bar{M}) = a \cdot (\text{Log}(TD_t) - \text{Log}(\bar{TD}))$$

et estimer

$$\text{Log}(M_t / \bar{M}) = a \cdot \text{Log}(TD_t / \bar{TD})$$

Puis la constante

$$b = \text{Log}(\bar{M}) - a \cdot \text{Log}(\bar{TD})$$

Nous pouvons voir en particulier les conséquences d'un changement d'unités (milliers, millions, milliards ...). Le terme constant va l'absorber, ce qui laissera «a» inchangé. En l'absence de «b», les «a» changeront de valeur, sans aucune justification économique.

Bien sûr, plus le «b» est important, plus son absence sera préjudiciable à la qualité de l'estimation (et plus les «a» seront touchés). Mais ce n'est pas une raison pour juger «b». Tout se passe comme si nous pesons un objet avec une balance: les deux plateaux ne pourront jamais avoir le même poids, et même si l'erreur augmente avec la différence, il sera toujours utile de la corriger. Et dans notre cas, c'est gratuit !

Il n'est pas fréquent que le terme constant ait un sens théorique. La majorité de ces cas provient d'une formule en taux de croissance ou en variations, où le terme constant a disparu avec la dérivation, et sa présence dans la nouvelle équation est associée à une tendance.

La seule justification de l'absence d'un terme constant apparaît quand cette constante est nulle en théorie. Dans ce cas, l'observation d'une valeur significative pose problème. Nous allons bientôt donner un exemple.

6.1.3 UN SUJET SPÉCIFIQUE: LA STATIONNARITE, LA COINTEGRATION ET LES MODELES A CORRECTION D'ERREUR

Depuis longtemps maintenant, la présence de ce qu'on appelle «cointégration» d'un ensemble de variables a été présentée comme un critère important, si ce n'est pas une condition préalable à l'établissement d'un lien entre des variables de comportement. Nous ne chercherons pas à décrire entièrement cette technique, mais à donner quelques éléments pour aider à comprendre ses fondements et son but. Le lecteur trouvera quelques références utiles à la fin du volume.

Nous allons commencer par les modèles à correction d'erreur. Bien plus tôt que la cointégration, ce type de modèle a été favorisé à la fois par les modélisateurs et les économètres, car il assure une structure plus cohérente et permet une meilleure interprétation des spécifications de l'équation. Il permettra également, si l'échantillon est suffisamment grand, de relier le diagnostic d'estimation à la cointégration.

Mais d'abord, nous devons aborder la question de la stationnarité.

6.1.3.1.1 La stationnarité et le test de Dickey-Fuller

En fait, la question fondamentale de la stationnarité ne devrait pas s'appliquer aux équations estimées, car elle considère une seule variable ou une expression. Toutefois:

- Certaines équations estimées considèrent en effet une variable unique !
- La présence de la stationnarité est essentielle à la cointégration, qui prend en compte plusieurs variables.

La question est la suivante:

Comme nous l'avons vu précédemment, la méthode des moindres carrés exige l'absence d'autocorrélation des résidus, ou plutôt une preuve raisonnable que cela est vrai (comme une probabilité de 95 %). Nous allons examiner la situation inverse: pouvons-nous rejeter le fait que le résidu est parfaitement corrélé, donc le que le coefficient de corrélation est égal à 1 ? Appliquons cela au cas d'une variable constante:

$$x_t = a + e_t$$

$$e_t = \rho \cdot e_{t-1} + u_t$$

La question posée sera également différente. Nous ne considérerons pas une explication économique réelle (du moins pour le moment), mais voudrions savoir si la variance du processus croît indéfiniment, ou si elle est bornée dans le long terme.

La logique utilisée est assez simple. Appelons σ la variance de la période (l'erreur est considérée comme homoscédastique). À partir de la période 1, la variance de l'erreur croîtra jusqu'à la période T comme :

$$V_T = \sigma^2 + \rho^2 \cdot \sigma^2 + \rho^4 \cdot \sigma^2 + \dots = \sum_{i=1}^T \rho^{2(T-i)} \cdot \sigma^2$$

Si $\rho = 1$, l'erreur qui apparaît à une période donnée est transmise intégralement à la suivante, et la variance totale à la période T sera la somme des variances individuelles, soit T fois la variance pour une seule période. Elle va grandir jusqu'à l'infini avec t .

$$V_T = \sum_{i=1}^T 1^{2(T-i)} \cdot v_i = \sum_{i=1}^T v_i = T \cdot \sigma^2$$

Au contraire, si $\rho < 1$, la variance venant d'une période donnée se réduira avec le temps (en proportion $\rho^2 < 1$) et la variance V globale atteinte à la période t sera la somme pondérée des erreurs individuelles :

$$V_T = \sum_{i=1}^T \rho^{2(T-i)} \cdot v_i = \sum_{i=1}^T \rho^{2(T-i)} \cdot \sigma^2 = \sigma^2 \cdot \sum_{i=1}^T \rho^{2(T-i)} = \sigma^2 (1 - \rho^{2T}) / (1 - \rho^2)$$

qui converge quand t croît jusqu'à l'infini.

En d'autres termes, quand nous passons de la période t à $t + 1$, la croissance de la variance correspond à l'ajout d'un élément ayant subi t réductions. Ce terme tend vers zéro quand t tend vers l'infini.

Si ρ est inférieur à 1, nous appelons la série stationnaire (sinon, non stationnaire, bien sûr). La distance entre les valeurs réelles de x et sa moyenne est bornée.

Remarque: nous ne considérons pas $\rho > 1$, car cela conduirait à un processus explosif qui ne correspond pas à la réalité économique.

Pour vérifier la stationnarité, l'outil le plus couramment utilisé est le test de Dickey-Fuller, qui tente plutôt d'établir (de façon équivalente) si $\rho - 1$ est inférieur à zéro. Pour cela, il régresse la variation temporelle de la variable sur sa valeur retardée:

$$\Delta x_t = e_t - e_{t-1} = (\rho - 1) \cdot e_{t-1} + u_t$$

et compare la statistique de Student (en principe négative) de la variable explicative aux valeurs d'une table. En raison de la présence de la valeur décalée des deux côtés de l'équation, le test est plus exigeant, et une probabilité de 5 % demande une valeur d'environ 3.

Le test peut être «augmenté» en introduisant la variation retardée une ou plusieurs fois comme variables explicatives supplémentaires.

Maintenant, quel est le lien entre la stationnarité et l'estimation économétrique ? Nous verrons plus loin l'application de la cointégration, mais nous allons commencer par un cas plus simple: l'estimation d'une seule variable, ou plutôt l'estimation de l'évolution de la distance à une cible donnée.

En utilisant le cadre ci-dessus, cela signifie que nous considérons «a» comme une cible, et x comme une variable fluctuant autour de la cible (ce qui signifie que la distance à la cible est bornée).

Une autre façon d'envisager le problème est de construire à partir de la variable (par exemple les importations) une expression qui devait être fixe en moyenne (comme leur part dans la demande totale) et de rechercher la stationnarité de ce concept.

$$\text{Log}(M_t / TD_t) = a + e_t$$

6.1.3.1.2 Un premier modèle à correction d'erreur

Jusqu'à présent, nous n'avons considéré que le processus appliqué à un seul élément, avec un objectif constant. La seule correction pourrait provenir de la valeur résiduelle. Si cet élément est le ratio des importations à la demande locale, nous avons:

$$e_t = \text{Log}(M_t / FD_t) - a = \rho \cdot e_{t-1} + u_t = \rho \cdot [\text{Log}(M_{t-1} / FD_{t-1}) - a] + u_t$$

On peut supposer que le processus de correction sera mené, non par le rapport lui-même, mais par un élément de celui-ci, les importations dans notre cas. Le processus sera formulé comme suit:

$$\begin{aligned} \text{Log}(M_t) &= \text{Log}(FD_t) + a + \rho \cdot (\text{Log}(M_{t-1} / FD_{t-1}) - a) + u_t \\ \Delta \text{Log}(M_t) &= \Delta \text{Log}(FD_t) - (1 - \rho) \cdot (\text{Log}(M_{t-1} / FD_{t-1}) - a) + u_t \end{aligned}$$

Cela signifie qu'une variation de FD sera suivie par la même variation relative de M , et que le seul changement dans le rapport provient:

- du résidu u_t
- de la correction de l'écart préexistant entre le ratio et a , dans la proportion $(1 - \rho)$.

Mais un changement de la demande finale n'est pas forcément suivi immédiatement par une variation proportionnelle des importations. Nous aurons dans ce cas:

$$\Delta \text{Log}(M_t) = \alpha \cdot \Delta \text{Log}(FD_t) - (1 - \rho) \cdot (\text{Log}(M_{t-1} / FD_{t-1}) - a) + u_t$$

Ce que nous avons établi ici est la forme la plus simple d'un modèle à correction d'erreur. Nous pouvons constater que:

- comme dans le cas précédent, le processus corrige l'écart à la cible précédente et introduit un nouvel écart par le nouveau résidu,
- mais l'adaptation limitée³¹ d'un élément à l'autre introduit une nouvelle source d'écart.

Si $\alpha = 1$ on parlera d'«homogénéité dynamique». Nous verrons plus loin les conséquences de cette hypothèse.

³¹ Avec $\alpha > 1$, la variation de M dépassera la variation de FD , produisant un écart de signe opposé. Cela ne contredit pas la capacité de correction d'erreur du processus.

Plus généralement, on peut aussi considérer

$$A(L) \cdot \Delta \text{Log}(M_t) = B(L) \cdot \Delta \text{Log}(FD_t) - (1 - \rho) \cdot (\text{Log}(M_{t-1} / FD_{t-1}) - a) + u_t$$

où A et B sont des polynômes de matrices de retards³².

La forme la plus simple est finalement la suivante:

$$\Delta \text{Log}(M_t) = \alpha \cdot \Delta \text{Log}(FD_t) - \beta \cdot (\text{Log}(M_{t-1} / FD_{t-1}) - a) + u_t$$

L'interprétation en est facile. L'agent a un objectif à long terme, et agit selon une formulation comprenant un ou plusieurs termes explicatifs.

- Si la cible change, il va s'adapter partiellement à la nouvelle valeur, en proportion α
- Si, à la période précédente, il y avait un écart entre le ratio observé et la cible, il le comblera partiellement, dans la proportion β

Les valeurs de α et β devraient être positives (pour la formulation en matrices, le terme dynamique devrait conduire dans la bonne direction). En théorie, le coefficient α peut prendre n'importe quelle valeur (positive), mais une valeur plus élevée que l'unité suppose une sur-réaction initiale. Quant à β , sa valeur devrait également être inférieure à un. Entre 1 et 2, elle va créer un processus de sur-adaptation alternée, convergent cependant. Supérieure à deux, elle va produire une divergence.

³² Rappelons-nous le sens de ces matrices.

si $A(L) = a + b \cdot L + c \cdot L^2$

alors

$$A(L) \cdot x_t = a \cdot x_t + b \cdot x_{t-1} + c \cdot x_{t-2}$$

Si ces conditions sont remplies, l'écart produit par un choc sur la cible à une période donnée se réduit régulièrement, en proportion α pour la première période, et $\beta = (1 - \rho)$ pour les suivantes.

Cela garantit également la convergence à long terme de x_t vers un sentier de croissance équilibré, si les éléments explicatifs le suivent aussi. Revenons à une formulation sous forme de logarithmes.

$$\Delta \text{Log}(x_t) = \alpha \cdot \Delta \text{Log}(x_t^*) - \beta \cdot \text{Log}(x_{t-1} / x_{t-1}^*) + \gamma + u_t$$

Ceci peut être écrit sous la forme:

$$\text{Log}(x_t) = (1 - \beta) \cdot \text{Log}(x_{t-1}) + \alpha \cdot \Delta \text{Log}(x_t^*) + \beta \cdot \text{Log}(x_{t-1}^*) + \gamma + u_t$$

Si x^* atteint un sentier de croissance équilibrée (avec un taux de croissance constant q), nous avons

$$\text{Log}(x_t) = (1 - \beta) \cdot \text{Log}(x_{t-1}) + \alpha \cdot q + \beta \cdot (q \cdot (t-1) + d) + \gamma + u_t$$

La dérivation de l'expression donne:

$$\Delta \text{Log}(x_t) = (1 - \beta) \cdot \Delta \text{Log}(x_{t-1}) + \beta \cdot q + \Delta u_t$$

Cela signifie que le taux de croissance de x converge vers q , si β est inférieur à un (voire à deux, si nous acceptons une convergence alternée).

L'écart entre x et x^* converge vers

$$q = \alpha \cdot q + \beta \cdot (\text{Log}(x_{t-1}^*) - \text{Log}(x_{t-1}))$$

ou

$$\text{Log}(x_t) - \text{Log}(x^*) = (1 - \alpha) \cdot q / \beta$$

Nous pouvons observer que l'écart ne dépendra pas des valeurs de départ.

Il ne sera nul que si

- $\alpha = 1$, un cas que nous avons déjà appelé «homogénéité dynamique»
- $q = 0$

Ces raisons sont faciles à expliquer:

- Si $\alpha = 1$, toute nouvelle évolution de la cible sera corrigée immédiatement. Comme le processus corrige progressivement l'erreur initiale, l'erreur globale convergera vers zéro.
- Si $q = 0$, la cible ne bouge pas. Et encore une fois, le processus de correction d'erreur réduit avec le temps l'écart initial.

Si aucune de ces conditions n'est remplie, un écart à long terme s'accumulera, convergeant vers une valeur d'autant plus élevée que

- et β sont faibles
- q est grand.

6.1.3.2 Les avantages de cette formulation

Dans un modèle, l'utilisation d'une structure à correction d'erreur peut poursuivre plusieurs objectifs:

- Contrôler la stabilité des formulations

Les valeurs numériques des coefficients indiquent immédiatement si les équations donnent des solutions stables.

- Améliorer la stabilité des propriétés numériques

En particulier, les mécanismes dynamiques de correction vont freiner avec le temps les effets des erreurs d'estimation (sur les périodes passées) et du terme aléatoire³³.

³³ Mais pas de l'erreur sur le coefficient, malheureusement.

- Identifier immédiatement la formulation à long terme

Le modèle à long terme peut être produit directement par l'extraction de la partie correspondante de la formulation complète³⁴³⁵.

- Permettre une meilleure interprétation des comportements

En ce qui concerne la sensibilité de la trajectoire du modèle à des chocs sur les hypothèses, il sera plus facile de séparer les contributions des changements permanents dans la cible, et de la dynamique qui y ramène.

Cependant

- Il peut être considéré comme abusif de parler de correction d'erreurs. Le nombre réel d'observations auquel un modèle macroéconomique a accès est souvent très limité (surtout pour une périodicité annuelle). La précision de l'estimation dans la formulation des objectifs à long terme peut être trop faible pour que l'écart puisse être appelé « erreur ».
- La taille de l'échantillon doit être assez grande pour que l'on soit autorisé à effectuer ce genre de tests, bien que l'on soit tenté de le faire pour se conformer aux habitudes actuelles. Mais leurs conclusions peuvent de toute façon être utiles, même s'ils ne sont pas suffisamment précis pour être appliqués avec rigueur.

6.1.3.3 La cointégration

Les modèles à correction d'erreur ont été présents dans la théorie économique, et la modélisation en particulier, depuis un temps assez long, en fait bien avant que le nom ait été inventé.

Par exemple, les équations d'investissement employant le taux d'utilisation des capacités de production comme variable explicative, ou l'adaptation de l'emploi à un objectif de productivité structurelle, adoptent de fait un cadre à correction d'erreur. Mais c'est seulement au cours des trente dernières années que l'usage de ce type d'équation s'est généralisé, et que des méthodes économétriques spécifiques ont

³⁴ Deleau M., Malgrange P., « Méthodes d'analyse des modèles empiriques », *Annales de l'INSEE* n° 20, septembre 1975.

³⁵ Brillet J.L., « Propriétés de longue Terme de la maquette Micro-DMS », dans *Biologie et économie, les apports de la modélisation*, Demongeot J. et Malgrange P. (éd.), 1988.

été conçues, liées à ce cadre. Ces méthodes relèvent du concept général de cointégration, que nous allons résumer maintenant, en essayant d'éviter autant que possible les éléments trop techniques.

Prenons une variable temporelle unique, sa valeur sur une période donnée étant la somme d'un terme constant et d'un terme aléatoire dépendant de sa valeur précédente. Nous avons vu précédemment que sa variance peut être bornée, la rendant stationnaire. Si ce n'est pas le cas, on peut appliquer des différenciations temporelles successives à la variable. Si la différence devient stationnaire la n ème fois, la variable sera dite intégrée d'ordre n , ou $I(n)$.

Mais il est également possible qu'entre un ensemble de variables (ou expressions) non stationnaires, on puisse mettre en évidence une combinaison linéaire, présentant les caractéristiques de la stationnarité. En d'autres termes, l'impact de toute différence de cette expression à une cible constante donnée est réduit à la période suivante, ce qui rend la nouvelle variable composite stationnaire (ou $I(0)$). La relation est considérée comme «de cointégration». Bien sûr, avec un ensemble de variables, plusieurs relations de cointégration peuvent apparaître, le nombre maximum étant celui des éléments de l'ensemble.

Des exemples typiques de ce genre de relation de long terme pourrait se rencontrer entre:

- la part des salaires dans la valeur ajoutée et le taux de chômage ;
- la part des exportations dans la demande mondiale et la compétitivité des prix des exportations, et / ou le taux d'utilisation des capacités productives ;
- le ratio du travail au capital et le ratio de leurs coûts relatifs.

Ces éléments peuvent être non stationnaires individuellement, mais liés dans le long terme: si la valeur à long terme d'un élément change, l'autre changera dans une proportion donnée.

Dans notre formulation ECM précédente, exprimée ici en niveaux:

$$\Delta x_t = \alpha \cdot \Delta x_t^* + \beta \cdot (x_{t-1}^* - x_{t-1}) + \gamma + u_t$$

La relation de cointégration est bien sûr représentée par:

$$x_t - x_t^* = 0$$

mais maintenant, c'est une formule qui devrait contenir les coefficients estimés.

L'estimation de la formulation à correction d'erreur ne peut se faire directement (en utilisant les moindres carrés non linéaires). Il faut d'abord contrôler l'existence de la relation de cointégration, à l'aide d'un test spécifique. Alors, si (et seulement si) une équation de cointégration a été mise en évidence, il convient de procéder à l'estimation de la formule de correction d'erreur dynamique, en utilisant le résidu de la relation de cointégration comme terme d'erreur.

Bien sûr, les coefficients de l'équation de cointégration doivent également être considérés comme acceptables, tant d'un point de vue statistique (en utilisant la statistique T comme dans une régression normale) qu'économique. Ainsi, l'élasticité de la part des importations dans la demande doit être influencée négativement (et significativement) par le rapport des déflateurs des importations et de la production locale.

Dans les exemples suivants, nous allons aborder ce processus deux fois.

Malheureusement, un inconvénient de cette structure vient de ses exigences en terme de taille de l'échantillon (50 observations à tout le moins). Cela signifie en particulier que la cointégration ne doit pas être prise en considération pour les modèles macroéconomiques annuels: soit l'échantillon est trop petit, ou sa durée (cinquante ans) rend tout à fait douteuse la production de formules reposant sur des données acceptables. En particulier, la séparation des variables mesurées à prix courants entre éléments à prix constants et déflateurs perd beaucoup de sa précision à des décennies de l'année de base.

C'est ce qui explique certainement le fait que la plupart des modèles opérationnels utilisent la cointégration avec parcimonie. Très souvent, un cadre à correction d'erreur est en effet introduit (ne serait-ce que pour profiter des avantages ci-dessus), mais malheureusement, l'estimation se fera en une seule étape.

6.2 APPLICATIONS: NOTRE MODELE

Nous allons maintenant appliquer les principes exposés ci-dessus à notre exemple de modèle.

Dans notre modèle, nous devons estimer cinq équations, pour lesquelles nous avons déjà des idées sur leur logique:

- La variation des stocks, l'emploi et l'investissement devraient dépendre du PIB
- Les exportations et les importations devraient dépendre de la demande associée (mondiale et locale) et de la disponibilité de l'offre potentielle.

Nous allons utiliser chacune de ces équations pour illustrer un aspect spécifique de l'estimation.

- La variation des stocks: les éléments généraux, l'homoscédasticité, la présence d'un terme constant.

- L'emploi: la stationnarité, les modèles à correction d'erreur.
- L'investissement: la nécessité d'établir une équation théorique cohérente avant l'estimation.
- Les exportations: les processus autorégressifs, la cointégration, la stabilité à long terme.
- Les importations: aller plus loin sur la cointégration et la stabilité à long terme.

Chacune de nos formulations sera basée sur des idées économiques très simples, et nous allons établir un cadre qui respecte à la fois les tests économétriques et la cohérence économique. Elles seront également choisies d'une manière qui devrait leur permettre de s'intégrer harmonieusement dans le modèle que nous construisons. Cependant, il devrait être clair que

- D'autres formulations simples pourraient probablement être construites à partir du même échantillon, avec une qualité équivalente ou peut-être meilleure.
- En utilisant un autre échantillon (un autre pays, par exemple) les mêmes idées économiques pourraient conduire à des formulations différentes (et pas seulement pour les valeurs des coefficients) ou à un échec.
- D'autres idées économiques pourraient être appliquées, avec la même complexité.

Pour produire un modèle réellement opérationnel, le cadre actuel devrait être développé de manière importante. Nous allons présenter ces évolutions plus tard.

Toutefois, le modèle que nous construisons représente à notre sens un résumé simplifié mais représentatif de la classe générale de modèles de ce type. À la lecture des documents décrivant n'importe quel modèle structurel opérationnel, on va rencontrer un grand nombre des idées que nous allons développer.

Une note sur les estimations: on peut observer que nous utilisons des données relativement anciennes, avec une périodicité bisannuelle, à mi-chemin entre les deux plus courantes: trimestrielle et annuelle. Nous avons accès à la même information sur une plus longue période sur une base trimestrielle, ce qui aurait représenté une amélioration évidente. Cependant, notre cours propose également une série de leçons, suivant les mêmes orientations (et utilisant des concepts très similaires) que celles que nous allons développer ici.

Évidemment, les leçons (qui demandent aux utilisateurs de résoudre un ensemble de problèmes) ne peuvent pas utiliser les mêmes données que les exemples. Nous avons décidé qu'il était plus important de garder la périodicité plus courte pour ces dernières, qui représenteront un élément plus important dans le processus d'enseignement.

Une autre raison est que les comptes nationaux les plus récents produisent des séries chaînées, autrement dit les éléments «à prix constants» deviennent «aux prix de l'année précédente». Leur utilisation en tant que telles rend les estimations économétriques et la modélisation difficiles sinon presque impossibles. Et

les transformer en séries annuelles «traditionnelles» à année de base fixée ne peut se faire d'une manière rigoureuse.

6.2.1 LES VARIATIONS DE STOCKS

Nous utiliserons cette estimation simple pour présenter les caractéristiques de base de l'estimation sous EViews, et souligner la nécessité de l'homoscédasticité.

Notre formulation suppose simplement que les entreprises désirent rendre leur niveau de stocks proportionnel à leur production (ou PIB). Pour un producteur donné, cela doit être vrai aussi bien pour les biens qu'il produit que pour ceux qu'il va utiliser pour la production. Par exemple, un constructeur automobile prévoira un délai donné entre la production et la vente (par exemple trois mois, ce qui va conduire à un niveau de stock de 1/4 de la production annuelle). Et pour être sûr de la disponibilité des biens intermédiaires (comme l'acier, les pneus, les composants électroniques et le carburant pour les machines) il va acheter la quantité nécessaire (proportionnelle à la production) à l'avance.

Nous supposerons que les entreprises ont atteint, à la période précédente, un niveau de stocks IL représentant un nombre de semestres de production:

$$IL_{t-1} = a \cdot Q_{t-1}$$

Et qu'elles veulent conserver ce niveau à la période actuelle:

$$IL_t^* = a \cdot Q_t$$

$$IL_t^* = IL_t$$

La variation des stocks représentera:

$$IC_t = (IL_t - IL_{t-1}) = a \cdot \Delta Q_t$$

Cela signifie que, contrairement au cas général, cette équation ne doit pas inclure de terme constant. Sa présence pourrait correspondre à une tendance (et une constante) dans l'équation en niveau, sans aucune

justification économique. Elle introduirait également un problème: l'ajout d'une constante à une explication en euros constants rendrait l'équation non homogène.

Même dans ce cas, l'équation est confrontée à un autre problème, concernant la valeur du résidu: entre 1963 et 2004, le PIB français a été multiplié par 4. Nous pouvons supposer que le niveau des stocks a aussi augmenté (peut-être un peu moins avec les économies d'échelle et l'amélioration des techniques de gestion).

Il est difficile de croire que la partie inexpliquée de la variation des stocks n'est pas affectée par cette évolution. Comme la variable augmente, l'erreur doit suivre. Mais pour appliquer la méthode des moindres carrés (MCO), il est nécessaire que la valeur résiduelle ait un écart-type constant.

L'idée la plus simple consiste à supposer que l'erreur croît au même rythme que le PIB, ce qui signifie que si l'on mesure la variation des stocks en proportion du PIB, nous devrions obtenir un concept pour lequel l'erreur reste stable. Bien sûr, nous devons appliquer la même modification au terme de droite, qui deviendra la variation relative du PIB.

Pour éviter les problèmes de causalité (pour un semestre donné, la demande de IC est en partie satisfaite par Q), on utilisera la valeur précédente de Q .

L'équation devient:

$$IC_t / Q_{t-1} = a \cdot \Delta Q_t / Q_{t-1}$$

6.2.1.1 Les éléments de base de l'estimation sous EViews

Comme il s'agit de notre premier exemple, nous allons l'utiliser pour présenter les fonctionnalités de base de l'estimation.

En fait, la technique appliquée variera avec l'état du processus d'estimation: selon que nous explorons plusieurs formulations différentes, à la recherche de la meilleure option en termes statistiques et économiques, ou que nous avons déjà élu la meilleure, et que nous voulons l'assembler avec notre modèle.

Nous allons commencer par traiter la première situation.

La façon la plus simple d'estimer une équation sous EViews par menus utilisera en séquence:

Quick > Estimate equation

Une fenêtre apparaît, dans laquelle on doit taper la formule.

Dans le cas des moindres carrés ordinaires³⁶, cela peut être une liste d'éléments séparés par des espaces, dans notre cas :

IC/Q (-1) D(Q)/Q (-1)

Nous pouvons également utiliser

$IC/Q (-1) = c(1)*D(Q)/Q(-1)$

Les deux méthodes donnent les mêmes résultats (dans le premier cas, les coefficients estimés vont également être transférés dans le vecteur «c»).

La méthode par défaut sera les moindres carrés, ordinaires dans notre cas. Si l'équation n'est pas linéaire par rapport aux coefficients, la deuxième présentation sera appliquée, ceci de façon automatique.

On notera que

- Pour introduire un terme constant, il doit être spécifié explicitement (comme un élément supplémentaire appelé «C»).
- EViews permet de spécifier une période d'estimation, qui sera appliquée uniquement à l'équation en question (la période courante n'est pas modifiée). Ceci est très utile si certaines périodes doivent être exclues de l'estimation. Ce sera le cas par exemple lorsqu'elles sont supposées ne pas suivre le comportement estimé (des données pré-transition pour les pays d'Europe centrale, la Chine ou le Vietnam), ou si des observations sont fournies également sur le futur. Dans ses «Perspectives économiques», l'OCDE complète les données historiques avec les résultats de ses prévisions au cours des trois prochaines années.
- Au contraire, il n'est pas nécessaire de se soucier de la présence dans la période courante d'observations pour lesquelles l'estimation n'est pas possible, en raison d'éléments manquants ou de l'impossibilité de calculer un terme (par exemple le logarithme d'une valeur négative).

³⁶ Le lecteur attentif aura observé qu'en l'absence de terme constant, nous n'appliquons pas vraiment les MCO, mais plutôt une estimation « linéaire par rapport aux coefficients ».

EViews éliminera de lui-même les périodes correspondantes (et vous indiquera l'échantillon réduit).

Dans notre cas, nous pouvons utiliser la période:

```
smpl 1960s1 2002s1
```

ce qui signifie que nous considérons les données du premier semestre de 1960 au dernier de 2002 (nos données sont biannuelles dans ce cas particulier).

Si l'équation est linéaire par rapport aux coefficients, EViews reconnaît cette propriété, et ne cherche pas à itérer sur les coefficients, car il sait que les valeurs trouvées sont les bonnes.

En utilisant le bouton «Ok» on obtient les résultats suivants :

Dependent Variable: IC/Q(-1)				
Method: Least Squares				
Date: 11/06/12 Time: 19:37				
Sample (adjusted): 1963S2 2002S1				
Included observations: 78 after adjustments				
IC/Q(-1)=C_IC(1)*@PCH(Q)				
	Coefficient	Std. Error	t-Statistic	Prob.
C_IC(1)	0.262021	0.041507	6.312769	0.0000
R-squared	0.071191	Mean dependent var		0.004759
Adjusted R-squared	0.071191	S.D. dependent var		0.007484
S.E. of regression	0.007213	Akaike info criterion		-7.013135
Sum squared resid	0.004006	Schwarz criterion		-6.982920
Log likelihood	274.5123	Hannan-Quinn criter.		-7.001039
Durbin-Watson stat	0.754596			

Nous pouvons voir qu'EViews donne la période utilisée (les périodes pertinentes de notre échantillon). L'estimation commence en 1963S2, car la série Q ne commence qu'en 1963 et est retardée une fois.

Nous obtenons également le nombre de périodes, et la date et l'heure.

Les autres éléments sont les statistiques habituelles, décrites plus haut. Les plus importantes sont:

- le R2, le test de Durbin et Watson et l'écart-type de la régression pour les éléments globaux ;

- les coefficients, la statistique de Student et la probabilité d'atteindre la valeur du coefficient si le coefficient est nul, en considérant l'écart-type estimé.

Dans notre cas:

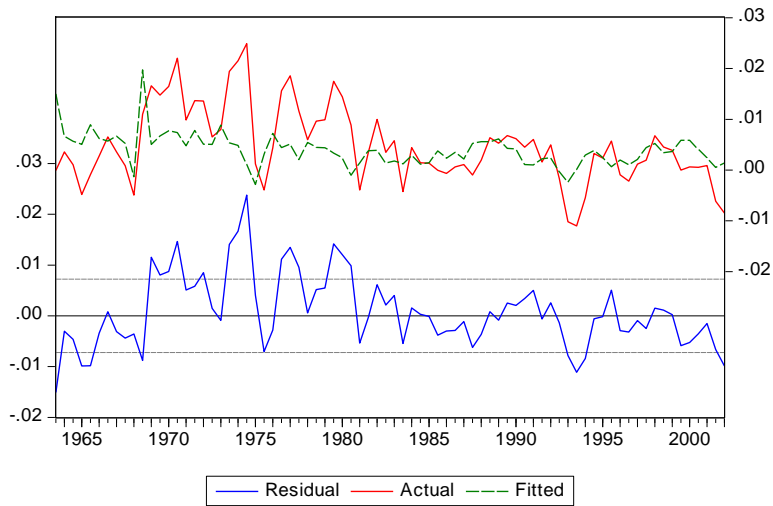
- Le R^2 est très faible, même si la variabilité extrême et l'absence de tendance de l'élément de gauche plaident en faveur de l'explication³⁷.

Cependant, comme pour presque toutes les estimations homogènes, une interprétation simple est disponible grâce à l'écart-type: la variable expliquée est mesurée en points de PIB, cette erreur représente en moyenne 0,72 point.

- Le coefficient est très significatif. La probabilité d'atteindre 0,26 pour une loi normale de moyenne 0 et d'écart-type 0,042 est indiquée comme étant nulle. Bien sûr, elle ne l'est pas vraiment, mais elle est inférieure à 0,00005, et probablement encore plus nettement.
- Mais le test de Durbin-Watson prend une valeur inacceptable, même si l'absence d'un terme constant (et donc des résidus de moyenne non nulle) rend son utilisation douteuse.

Le graphique des résidus est le deuxième élément important pour le diagnostic. Il montre l'évolution des séries observée et estimée (dans la partie haute du graphique, en utilisant l'échelle de droite) et de la valeur du résidu (en bas, en utilisant l'échelle de gauche, avec des lignes en pointillés représentant $+1$ et -1 écarts-types). Cela signifie que les résidus à l'intérieur de la bande sont inférieurs à la moyenne, et supérieurs dans le cas contraire. Bien sûr, ceci ne donne qu'un diagnostic relatif.

³⁷ Si on connaissait les valeurs de IL , son estimation donnerait un meilleur R^2 (en raison de la colinéarité de IL et Q). Mais nous serions amené à estimer un modèle à correction d'erreur sur IL , de toute façon. Nous avons vu l'avantage de cette formulation, mais pour étendre cette qualité à l'ensemble du modèle, toutes les équations doivent être de ce type.



Le graphique montre (à notre avis) que l'équation fournit une explication significative, mais que certaines périodes (1975-1980, en particulier) subissent une erreur importante et persistante, et que les résidus présentent après 1975 une tendance négative (et peut-être une tendance positive auparavant).

En plus de l'affichage des résultats de l'estimation et du graphique des résidus, EViews crée plusieurs objets:

- un vecteur de coefficients, contenu dans le vecteur «C». Les valeurs zéro ou les résultats de la régression précédente sont remplacés³⁸ ;
- une série de valeurs résiduelles, contenues dans la variable «RESID». Les «NA» des valeurs ou les résultats de la régression précédente sont remplacés³⁹ ;
- un texte d'équation, appelé «Untitled» pour le moment, et contenant la formule développée, avec «C» comme vecteur des coefficients, indiqué à partir de 1. Dans notre cas, la formule est évidemment

³⁸ Mais si la régression actuelle contient moins de coefficients que les précédentes, les éléments additionnels ne sont pas mis à zéro.

³⁹ Mais cette fois, toute valeur préexistante est remplacée soit par "NA" soit par un résultat calculé.

$$IC/Q(-1)=c(1)*D(Q)/Q(-1)$$

Toute estimation suivante va remplacer l'équation par la nouvelle version «Untitled».

EViews fournit également plusieurs options, accessibles depuis le menu, et qui peuvent se révéler utiles:

- «View > Representations » donne trois représentations de l'équation: la déclaration de départ, et deux formules, où les coefficients apparaissent par leurs noms (éléments du vecteur «C») ou par leurs valeurs numériques.
- «Print» permet d'imprimer la fenêtre courante: vers une imprimante, vers un fichier texte (en utilisant des caractères, ce qui économise l'espace, mais réduit la lisibilité, notamment pour les graphiques), ou vers un fichier graphique au format RTF. Cette dernière option peut faire appel à une présentation monochrome, qui est obtenue en décochant «Print in color» dans les options.
- «Name» permet la création de l'équation sous forme d'un objet nommé dans le fichier de travail, accompagné d'un commentaire éventuel. Il est important d'utiliser cette option immédiatement après l'estimation, car l'équation temporaire (nommée «Untitled») sera remplacée par l'estimation suivante.

EViews propose comme nom standard «EQ» suivi d'un nombre à deux chiffres, le plus bas utilisé jusqu'ici. Deux options sont disponibles:

- donner un nom représentatif de l'équation (comme «EQ_X3U» pour la troisième équation d'estimation de X, ici influencé par le taux d'utilisation) ;
- accepter la suggestion EViews et s'appuyer pour l'explication sur le commentaire ajouté.

Personnellement, nous préférons la seconde option:

- elle est plus simple et plus naturelle à utiliser ;
- elle permet de placer toutes les équations dans le fichier de travail (et la fenêtre) au même emplacement ;
- elle évite de définir une nomenclature complexe et peut-être difficile à appliquer rigoureusement ;
- la zone de commentaire est beaucoup plus large et peut suivre n'importe quel format, y compris les lettres accentuées et les caractères spéciaux.

Cependant, insérer un blanc souligné («_») devant le nom proposé permettra de placer les équations dans les premières positions de la fenêtre de travail.

En fait, l'élément enregistré est plus complexe que la formule elle-même. Double-cliquer dessus permet d'accéder à la formule, et éventuellement de la ré estimer ou de la modifier..

- Forecast produit une série pour la grandeur estimée (ou l'expression de gauche, généralement moins intéressante), ainsi qu'un graphique présentant un intervalle d'erreurs (et une table de statistiques).

6.2.1.2 Une technique alternative: utiliser la fenêtre des commandes

Au lieu d'utiliser Quick>Estimate, on peut travailler directement dans la fenêtre de commande. Il suffit d'ajouter «ls» devant la formule.

```
ls IC/Q (-1)=c(1)*D(Q)/Q(-1)
```

Cela présente plusieurs avantages:

- En copiant et éditant l'équation courante sur la ligne suivante de la fenêtre de commande, la saisie des changements est beaucoup plus facile.
- Après une séquence d'estimations, l'ensemble peut être copié dans un fichier de programme et réutilisé à volonté. La gestion d'un ensemble de versions alternatives est alors beaucoup plus facile.
- On peut contrôler la taille des caractères. Ceci est assez intéressant quand on travaille avec une équipe, ou que l'on réalise une présentation sur grand écran, car la méthode par menus utilise généralement une police assez petite.

Le seul inconvénient est la définition de la période: elle doit être entrée aussi par commande, et non comme un élément de l'écran «estimation». Elle deviendra alors la période courante.

6.2.1.3 D'autres spécifications possibles

Revenons à notre formule estimée. Si nous ne sommes pas satisfaits des résultats précédents, nous pouvons essayer d'autres options, sans changer l'arrière-plan économique:

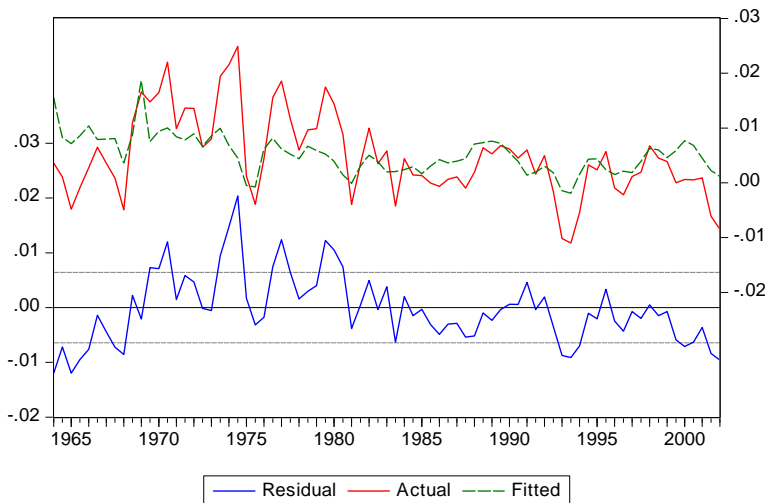
- Les entreprises pourraient tenir compte des variations du PIB pour les deux derniers semestres, avec des impacts différents.

Dependent Variable: IC/Q(-1)
 Method: Least Squares
 Date: 11/06/12 Time: 19:37
 Sample (adjusted): 1964S1 2002S1
 Included observations: 77 after adjustments
 IC/Q(-1)=C_IC(1)*@PCH(Q)+C_IC(2)*@PCH(Q(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
C IC(1)	0.130473	0.056396	2.313510	0.0234
C IC(2)	0.214330	0.053280	4.022712	0.0001

R-squared	0.280583	Mean dependent var	0.004821
Adjusted R-squared	0.270991	S.D. dependent var	0.007513
S.E. of regression	0.006414	Akaike info criterion	-7.234887
Sum squared resid	0.003086	Schwarz criterion	-7.174009
Log likelihood	280.5432	Hannan-Quinn criter.	-7.210536
Durbin-Watson stat	0.658613		

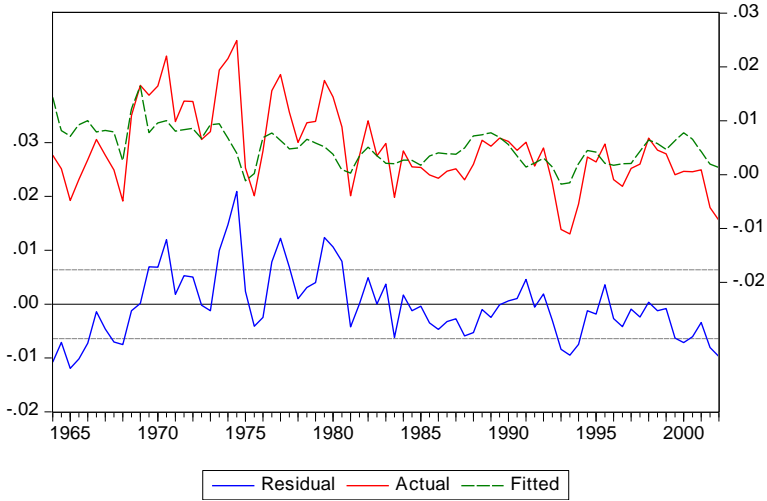
Les résultats sont plutôt meilleurs (mais pas le test de Durbin-Watson !).



- On peut aussi considérer le même coefficient:

Dependent Variable: IC/Q(-1)
 Method: Least Squares
 Date: 11/06/12 Time: 19:37
 Sample (adjusted): 1964S1 2002S1
 Included observations: 77 after adjustments
 IC/Q(-1)=C_IC(1)*PCH(Q)+C_IC(1)*PCH(Q(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
C_IC(1)	0.173786	0.020427	8.507613	0.0000
R-squared	0.274067	Mean dependent var		0.004821
Adjusted R-squared	0.274067	S.D. dependent var		0.007513
S.E. of regression	0.006401	Akaike info criterion		-7.251844
Sum squared resid	0.003114	Schwarz criterion		-7.221405
Log likelihood	280.1960	Hannan-Quinn criter.		-7.239669
Durbin-Watson stat	0.628277			



Cette restriction ne réduit pas la qualité (le R2 diminue mais l'écart-type aussi, un résultat étrange dû au plus grand nombre de degrés de liberté, avec un coefficient estimé en moins).

En fait, l'observation des résidus montre une tendance croissante avant 1975, et une tendance à la baisse à partir de cette date. On peut être tenté de prendre cela en compte, une perspective qui n'est pas tout à fait illégale: ces dernières années, la politique des entreprises a été de réduire le niveau des stocks (donc leur variation) et les possibilités techniques de sa mise en œuvre ont augmenté. Par contre, l'augmentation pré-1975 est plus difficile à expliquer.

Nous allons introduire un terme de correction, représenté par une constante, une tendance à partir de 1975, et une autre de 1975 à la fin.

6.2.1.3.1 L'introduction des tendances dans les équations

Pour ajouter une tendance à une équation (ou d'ailleurs n'importe quelle expression ne dépendant que du temps), deux solutions sont possibles:

- créer une variable (exogène),
- introduire directement la formule associée, en utilisant une variable de tendance.

Nous préférons clairement la deuxième option:

- elle réduit le nombre d'éléments dans le modèle ;
- elle ne nécessite aucune extrapolation pour les simulations sur le futur ;
- elle permet une interprétation directe par le modélisateur ou d'autres personnes, et ne réclame pas de documentation.

Nous ne voyons aucun avantage relatif pour la première option.

Pour définir la tendance, nous avons deux options:

- utiliser la fonction @trend ou @trendc de EViews, donnant une variable partant de 1 et augmentant de 1 à chaque période⁴⁰ ;
- créer une tendance temporelle spécifique, en prenant la valeur de l'année, et en augmentant la série de 1 à chaque année. Évidemment, en cas de série non-annuelle, il faudra distinguer les

⁴⁰ Pour toute périodicité.

sous-périodes. La solution évidente consiste à choisir pour la première période la valeur de l'année civile, et faire croître les valeurs ultérieures d'une fraction, qui donnera 1 pour l'année complète.

Pour le premier semestre de 1975, la valeur sera 1975, pour le second semestre 1975.50.

Malheureusement, cette technique rencontre des problèmes au-delà de la périodicité trimestrielle, cas cependant peu fréquent en modélisation.

Nous estimons indifféremment:

$$I_C/Q(-1) D(Q)/Q(-1) D(Q(-1))/Q(-2) (T-1975)^*(T \leq 1975) (T > 1975)^*(T-1975) C$$

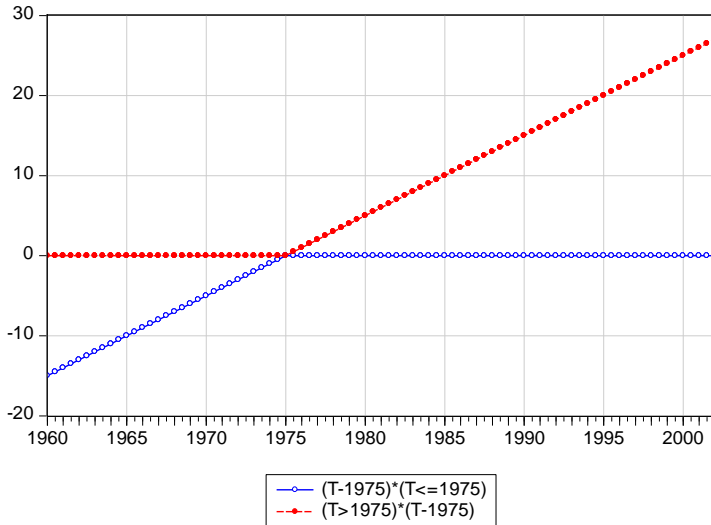
ou

$$I_C/Q(-1) = C(1)*D(Q)/Q(-1) + C(2)*D(Q(-1))/Q(-2) + C(3)*(T-1975)^*(T \leq 1975) + C(4)*(T > 1975)^*(T-1975) + C(5)$$

Décrivons en détail le calcul des tendances. La première: $(T-1975)^*(T \leq 1975)$ multiplie une tendance croissante prenant une valeur nulle au premier semestre de 1975, par une condition qui n'est vraie que jusqu'en 1975. Le résultat est une tendance croissante (prenant des valeurs de moins en moins négatives) jusqu'en 1975, quand elle prend définitivement la valeur 0.

La seconde: $(T-1975)^*(T > 1975)$ crée une tendance nulle jusqu'en 1975, qui augmente ensuite de 1 à chaque année.

En 1975, les deux tendances se rejoignent à 0.

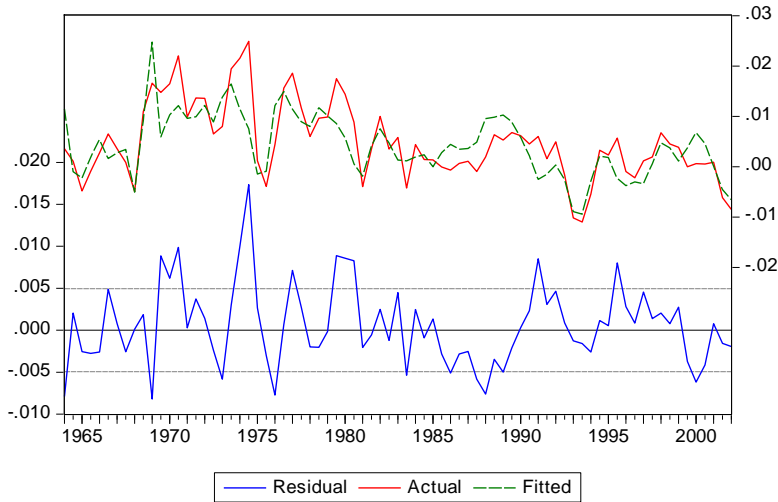


La formulation associée fonctionne très bien, avec des termes très significatifs. L'écart-type est considérablement réduit, le R2 est assez élevé pour une variable expliquée aussi erratique, et la qualité de l'explication est assez forte à la fin (pour les six dernières années, l'erreur est inférieure à la moyenne, et la dernière valeur est très faible). Cependant, la prévision de la tendance sera très problématique: si nous la maintenons dans le long terme, sa contribution deviendra infiniment négative, et le rapport explicatif aussi.

Et évidemment, il sera sans doute justifié de considérer notre formule comme établie ad hoc. De fait nous renoncerons à son utilisation.

Dependent Variable: IC/Q(-1)
 Method: Least Squares
 Date: 11/06/12 Time: 19:37
 Sample (adjusted): 1964S1 2002S1
 Included observations: 77 after adjustments
 IC/Q(-1)=C_IC(1)*@PCH(Q)+C_IC(2)*@PCH(Q(-1))+C_IC(3)*(T-1975)
 (T<=1975)+C_IC(4)(T-1975)*(T>=1975)

	Coefficient	Std. Error	t-Statistic	Prob.
C_IC(1)	0.286030	0.048544	5.892212	0.0000
C_IC(2)	0.383594	0.047563	8.064939	0.0000
C_IC(3)	0.001608	0.000253	6.362828	0.0000
C_IC(4)	-0.000327	5.37E-05	-6.083855	0.0000
R-squared	0.582654	Mean dependent var		0.004821
Adjusted R-squared	0.565502	S.D. dependent var		0.007513
S.E. of regression	0.004952	Akaike info criterion		-7.727463
Sum squared resid	0.001790	Schwarz criterion		-7.605707
Log likelihood	301.5073	Hannan-Quinn criter.		-7.678762
Durbin-Watson stat	1.214565			



6.2.1.4 La préparation des équations pour le modèle

Une fois qu'une équation a été choisie pour être introduite dans le modèle, une stratégie différente doit être utilisée.

Si nous utilisons directement la formule estimée, nous serons confrontés à plusieurs problèmes:

- il n'est pas simple de relier le nom de l'équation avec son objet, ce qui rend l'ensemble assez confus et interdit d'utiliser tout processus automatisé et systématique ;
- le vecteur C utilisé par toutes les équations est seulement compatible avec la dernière estimation ;
- les résidus ne peuvent pas être gérés simplement.

Au lieu de cela, nous proposons l'organisation suivante, en tirant tous les éléments du nom de la variable expliquée, par une transformation systématique:

- Nommer l'équation à partir de la variable estimée.

Par exemple nous pouvons appeler notre équation EQ_CI.

- Utiliser la spécification développée, avec des coefficients explicites.

- Mais nommer le vecteur de coefficients d'après la variable estimée.

Par exemple, nous pouvons l'appeler C_Ci. Bien sûr, cela nécessite sa création, en choisissant une dimension assez élevée:

```
coef(10) c_ci
```

(Nous avons choisi 10 comme un chiffre rond dont nous savons que nous ne l'atteindrons jamais.)

- Introduire un résidu additif explicite, nommé à partir de la variable estimée. La raison en est la suivante.

Il est essentiel pour un modèle d'estimer et de simuler la même équation. Bien entendu deux versions peuvent être gérées, l'une étant copiée dans l'autre après chaque nouvelle estimation. Ceci est cependant:

- fastidieux,
- difficile à gérer,
- source d'erreurs.

Il est préférable d'utiliser un seul élément. Cependant, cette stratégie est confrontée à un problème: on veut accéder à la valeur du résidu, en particulier pour les prévisions comme nous le verrons plus tard. Et l'estimation n'y fait pas appel.

La solution est simple: introduire un résidu formel, mais le mettre à zéro avant toute estimation.

6.2.1.4.1 Utiliser un programme.

Cela permet:

- un contrôle visuel de la spécification ;
- une répétition facile de l'estimation (par exemple, si les données ont changé) ;
- la mise en place facile de changements marginaux ;
- la documentation du contexte économique (par l'introduction de commentaires dans le programme).

Dans notre cas nous allons utiliser:

```
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
```

6.2.1.5 Les fonctionnalités introduites par EViews 8

EViews 8 améliore fortement la gestion des équations du modèle, en particulier des identités.

Jusqu'à EViews 7:

- La suppression d'une identité n'était pas possible.
- L'ajout d'une nouvelle version dupliquait de fait sa définition, ce qui donnait plus d'équations que d'endogènes.
- Supprimer une équation estimée ne pouvait se faire qu'à travers l'ordre EXCLUDE, difficile à gérer, et pour conserver l'exclusion, il fallait préciser à nouveau celle-ci dans toutes les déclarations d'exclusion suivantes.
- Pour remplacer une équation estimée vous deviez utiliser le même nom, sinon vous étiez confronté au même problème de duplication que ci-dessus.

Maintenant vous pouvez (si vous choisissez cette option):

- Supprimer n'importe quel type d'équation, en utilisant:

```
model_name.drop nom-de-variable (pour les identités)
model_name.droplink nom-d'équation (pour les équations estimées)
```

Par exemple, vous pouvez spécifier:

```
PIB_fra_1.drop PIB : PIB devient exogène
_fra_1.droplink eq_i : l'équation eq_i qui donne l'investissement est abandonnée, et I devient exogène
```

- Remplacer la formule à l'aide de:

```
nom-du-modèle.replace nom-de-variable (pour les identités)
```


nom-du-modèle.replacelink ancien-nom-équation nouveau-nom-équation (pour les équations estimées)

`_fra_1.replace PIB=`
`_fra_1.replacelink eq_i eq_i_new`

En outre, vous pouvez remplacer toutes les occurrences d'une variable du modèle en utilisant un nouveau nom, grâce à:

nom-du-modèle.replacevar ancien-nom-de-variable nouveau-nom-de-variable

Par exemple:

`_fra_1.replacevar PIB Q`

remplace PIB par Q dans l'ensemble du modèle (mais pas dans les relations estimées)

6.2.2 L'INVESTISSEMENT: LA NECESSITE D'ETABLIR UN CADRE THÉORIQUE COHERENT AVANT TOUTE ESTIMATION

Dans cette estimation, nous insisterons sur l'importance d'établir un cadre économique cohérent avant toute estimation.

L'idée de base économique est très simple: le but de l'investissement est

- de remplacer le capital déclassé ;
- de permettre d'augmenter la production, face à une augmentation de la demande.

Sans aller plus loin, en théorie, de nombreuses formulations peuvent être envisagées. Par exemple, l'investissement pourrait avoir une élasticité constante par rapport au PIB, peut-être avec un terme de correction d'erreur, comprenant le capital ...

À notre sens, rechercher la meilleure estimation sans tenir compte de la théorie économique sous-tendant la formule, et surtout de ses conséquences formelles pour les propriétés du modèle, est assez

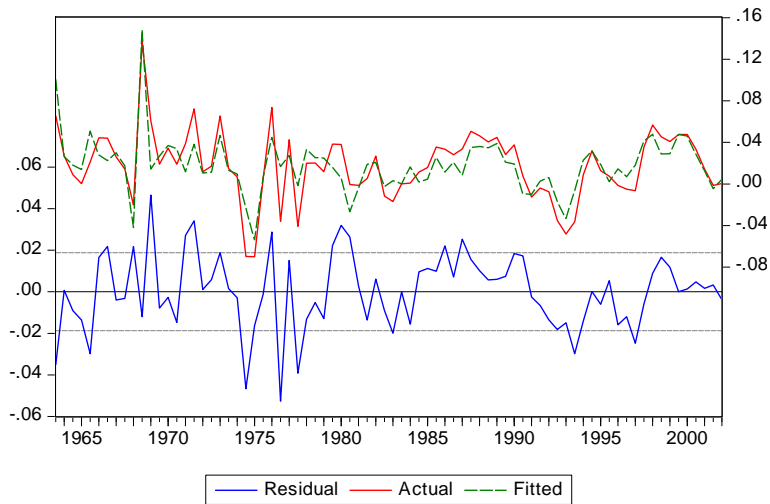
irresponsable. Par exemple, utiliser le logarithme de l'investissement est très dangereux. Représentant de fait une dérivée, sa valeur peut changer dans des proportions très élevées, et si l'on remonte au fondement microéconomique de ce comportement, sa valeur pourrait très bien être négative, car certaines entreprises sont amenées à désinvestir de temps en temps, en vendant plus de capital qu'ils n'en achètent.

Par exemple, l'équation suivante semble fonctionner assez bien:

$$\Delta \text{Log}(I_t) = a \cdot \Delta \text{Log}(Q_t) + b \cdot \text{Log}(I_{t-1} / Q_{t-1}) + c \cdot t + d$$

Les résultats sont:

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(Q)	2.425971	0.195149	12.43140	0.0000
LOG(I(-1)/Q(-1))	-0.076428	0.032747	-2.333903	0.0223
C	-2.458531	0.535731	-4.589116	0.0000
T	0.001145	0.000249	4.605258	0.0000
R-squared	0.684209	Mean dependent var		0.017985
Adjusted R-squared	0.671407	S.D. dependent var		0.032691
S.E. of regression	0.018740	Akaike info criterion		-5.066421
Sum squared resid	0.025987	Schwarz criterion		-4.945564
Log likelihood	201.5904	Hannan-Quinn criter.		-5.018040
F-statistic	53.44416	Durbin-Watson stat		1.658039
Prob(F-statistic)	0.000000			

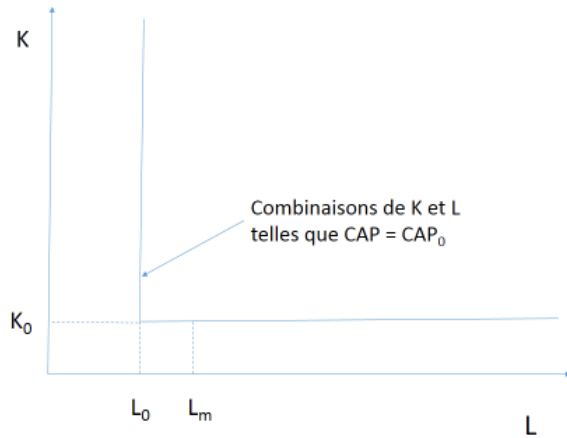


Tout semble aller bien: les statistiques sont assez bonnes (sauf peut-être pour le test de Durbin-Watson), les signes sont les bons, le graphique montre un ajustement très fort. Cependant, lorsque l'on assemble l'équation avec un modèle, ses propriétés de simulation seront affectées par la solution de base: par exemple, même une très forte augmentation du PIB aura un faible impact sur le niveau absolu de l'investissement si celui-ci était très faible à la période précédente.

On peut deviner que, bien que lier l'investissement (une variation du capital) à la variation de la production semble une idée naturelle, le passage à la formulation ci-dessus était un peu trop rapide. On doit être naturellement réticent à prendre le logarithme du taux de croissance.

Nous allons essayer de clarifier le processus économique à travers une formalisation logique complète.

Supposons que la production suive une fonction «à facteurs complémentaires», ce qui signifie que pour atteindre un niveau donné de capacité de production, des niveaux fixes de capital et de travail distincts sont requis, et qu'une réduction d'un facteur ne peut pas être compensée par une augmentation de l'autre. Cela signifie évidemment que le processus le moins coûteux (optimal) est celui qui respecte exactement ces conditions.



Avec « pk » productivité du capital, et « pl » productivité du travail, on obtient:

$$CAP_t = \min(pk_t \cdot K_{t-1}, pl_t \cdot L_t)$$

(le « $t-1$ » signifie que nous allons utiliser le niveau de capital atteint à la fin de la période précédente).*

Ainsi, sur notre graphique, K_0 et L_0 sont les facteurs les plus faibles permettant de produire CAP_0 .

En fait, pour un niveau donné de l'emploi, on peut toujours faire évoluer la production associée, du moins au niveau macroéconomique. Augmenter temporairement la productivité du travail de 2 % peut être facilement atteint par des heures supplémentaires, la réduction des vacances, des stages...

Sur notre graphique, L_m représente le niveau normal d'emploi associé au niveau minimal L_0 (et l'écart le gain de productivité possible à court terme).

Le capital sera donc le seul facteur limitatif à court terme.

L'équation donnant la capacité peut être simplifiée en:

$$CAP_t = pk_t \cdot K_{t-1}$$

Maintenant, définissons le taux d'utilisation des capacités:

$$UR_t = Q_t / CAP_t$$

$$CAP_t = pk_t \cdot K_{t-1} = Q_t / UR_t$$

Maintenant, supposons que les entreprises veulent réellement atteindre un objectif constant, le taux d'utilisation UR^* , et s'attendent à un niveau de production Q_{t+1}^a . Alors, par définition:

$$K_t^* = CAP_t^* / pk_t^* = Q_{t+1}^a / UR^* / pk_{t+1}$$

$$K_t = CAP_t / pk_t = Q_t / UR_t / pk_t$$

Et en définissant $tx(z)$ comme le taux de croissance de z :

$$tx^*(K_t) \approx tx^*(CAP_t) - tx(pk_t) \approx tx^a(Q_t) - tx^*(UR_t) - tx(pk_t)$$

Cela signifie que le taux de croissance du capital recherché peut être décomposé comme la somme de trois termes, l'un avec une influence positive:

- Le taux de croissance prévu de la production

et deux négatives:

- La cible de taux de croissance du taux d'utilisation: si les entreprises estiment que leurs capacités sont de 1 % trop élevées pour le niveau actuel de la production, elles peuvent atteindre la cible par une diminution du capital de 1 %, même si la production ne doit pas changer.
- Le taux de croissance de la productivité du capital: si elle augmente de 1 %, les besoins en termes de capital diminueront d'autant.

Mais l'élément que nous voulons déterminer est l'investissement. Pour l'obtenir, nous utiliserons la définition:

$$K_t = K_{t-1} \cdot (1 - dr_t) + I_t$$

ce qui peut s'écrire comme

$$tx(K_t) = -dr_t + I_t / K_{t-1}$$

ce qui donne finalement:

$$I_t^* / K_{t-1} = tx(K_t) = dr_t + tx^*(K_t) = dr_t + tx^a(Q_t) - tx^*(UR_t) - tx(pk_t)$$

En d'autres termes:

- Si les entreprises s'attendent à un taux de croissance de 4 %, les capacités devraient s'adapter à cette croissance.
- Mais si elles estiment que leurs capacités sont sous-utilisées de 1 %, leur capacité souhaitée n'augmentera que de 3 %.
- Si la productivité du capital va augmenter de 1 %, elles ont besoin de 1 % de capital en moins.
- Mais une fois la croissance du capital définie, elles ont aussi besoin de compenser la dépréciation.

Si nous supposons

- que le taux d'amortissement est constant, ainsi que le taux de croissance de la productivité du capital,
- que les prévisions de croissance de production sont basées sur la moyenne des taux antérieurs,

et si nous considérons que le taux d'utilisation est le ratio du PIB réel à une valeur obtenue avec utilisation normale des facteurs, ce qui conduit à une cible unitaire,

nous obtenons la formule simplifiée:

$$I_t^* / K_{t-1} = a + \sum_{i=0}^n \alpha_i \cdot tx^a(Q_{t-i}) - tx^*(UR_{t+1})$$

avec

$$\sum_{i=0}^n \alpha_i = 1$$

Enfin, on peut supposer, comme nous le ferons aussi pour l'emploi, que la cible de croissance du capital n'est que partiellement atteinte dans la pratique, soit parce que les entreprises réagissent avec prudence aux fluctuations de la demande, soit parce qu'elles sont contraintes par des programmes d'investissement couvrant plus d'une période.

Et nous allons laisser libres les coefficients:

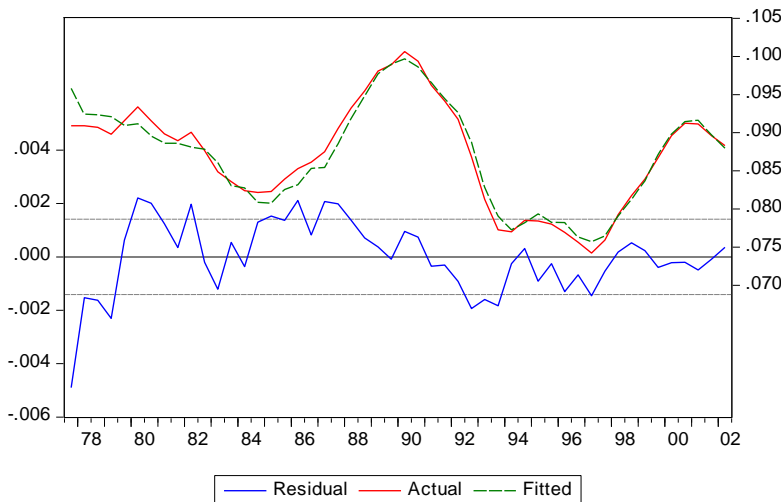
$$I_t^* / K_{t-1} = b \cdot I_{t-1} / K_{t-2} + (1-b) \cdot (a + c \cdot \sum_{i=0}^n \alpha_i \cdot tx^a(Q_{t-i}) - d \cdot tx^*(UR_t))$$

Les résultats sont plutôt satisfaisants, avec le signe correct et des statistiques acceptables pour tous les éléments explicatifs. Ce n'était pas évident, car leur forte corrélation (Q est utilisé deux fois au numérateur) **aurait pu rendre difficile pour le processus d'estimation de séparer leur rôle.**

Dependent Variable: VK(-1)
 Method: Least Squares
 Date: 11/07/12 Time: 12:03
 Sample (adjusted): 1977S2 2002S1
 Included observations: 50 after adjustments
 $VK(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.25*Q/Q(-4)+C_I(4)$

	Coefficient	Std. Error	t-Statistic	Prob.
C_I(1)	0.825643	0.033524	24.62809	0.0000
C_I(2)	0.027926	0.006984	3.998580	0.0002
C_I(3)	0.152870	0.058292	2.622495	0.0118
C_I(4)	-0.052515	0.010337	-5.080438	0.0000

R-squared	0.959393	Mean dependent var	0.087074
Adjusted R-squared	0.956745	S.D. dependent var	0.006785
S.E. of regression	0.001411	Akaike info criterion	-10.21218
Sum squared resid	9.16E-05	Schwarz criterion	-10.05922
Log likelihood	259.3046	Hannan-Quinn criter.	-10.15393
F-statistic	362.2686	Durbin-Watson stat	0.628797
Prob(F-statistic)	0.000000		



Le graphique des résidus montre que la qualité de l'explication croît avec le temps et qu'elle est particulièrement élevée pour les dernières périodes. Ceci est assez important pour les simulations sur le futur, et l'on peut se demander ce que nous aurions fait si l'échantillon avait été inversé, et les résidus initiaux s'étaient appliqués aux dernières périodes.

Nous allons traiter ce problème des erreurs croissantes sur les périodes récentes lorsque nous aborderons les prévisions.

L'équation que nous avons construite est non seulement satisfaisante par elle-même, mais on peut s'attendre à ce qu'elle donne au modèle des propriétés adéquates. En particulier, l'élasticité à long terme du capital à la production est maintenant établie comme unitaire par construction. À partir d'une simulation de base, un choc permanent de 1 % sur Q laissera la valeur à long terme de UR inchangée⁴¹. Cela implique des variations identiques pour la production, la capacité et (avec une productivité du capital constante) le capital lui-même.

Les coefficients « a » et « b » ne détermineront que la dynamique de la convergence vers cette cible.

En fait, nous avons estimé une sorte d'équation à correction d'erreur, dans laquelle l'erreur est l'écart entre la capacité observée et la cible (donc le taux d'utilisation, à une constante multiplicative près).

Nous espérons avoir montré que pour produire une formulation cohérente, en particulier dans un contexte de modélisation, il faut commencer par établir un cadre économique sain.

6.2.3 L'EMPLOI: LA STATIONNARITE, LES MODELES A CORRECTION D'ERREUR, LE TEST DE RUPTURE

6.2.3.1 Le cadre économique

Bien sûr, l'équation déterminant l'emploi devra suivre également un cadre à facteurs complémentaires.

Dans le paragraphe précédent, nous avons montré que dans ce cadre, l'élément déterminant est la capacité du seul capital, tandis que les entreprises pourraient demander à des travailleurs une augmentation temporaire de leur productivité, assez élevée pour assurer le niveau requis de production⁴².

⁴¹ Comme le côté gauche représente le taux de croissance à long terme (fixe) du capital.

⁴² Cela est vrai dans notre cadre macro-économique, dans lequel les variations de la production sont limitées, et une partie de la croissance est compensée par des augmentations de la productivité structurelle (en raison, par exemple, de processus plus intensifs en capital). Au niveau des entreprises, l'emploi peut produire des goulots d'étranglement. Ce sera le cas si une mode apparaît

L'adaptation de l'emploi au niveau correspondant à une cible de «productivité normale» se fera par étapes.

Cela signifie que l'estimation de l'emploi va nous permettre d'appliquer les éléments relatifs aux modèles de correction d'erreurs que nous avons présentés plus tôt, dans un cadre très simple.

Nous supposerons que les entreprises:

- connaissent le niveau de production qu'elles doivent réaliser ;
- savent également le niveau de la production qui devrait être atteint par chaque travailleur dans des circonstances normales (en d'autres termes sa productivité normale).

À partir de ces deux éléments, ils peuvent déterminer le nombre normal des travailleurs dont ils ont besoin.

Mais ils ne vont pas adapter immédiatement le niveau d'emploi réel à cet objectif, et ce pour deux types de raisons:

- Des raisons techniques: entre la conclusion que plus d'employés sont nécessaires et l'embauche effective⁴³, les entreprises ont à se prononcer sur le type d'emplois à proposer, mettre en place leurs offres, mener des entretiens, négocier les salaires, établir des contrats, obtenir des autorisations s'ils sont des citoyens étrangers, peut-être demander aux travailleurs présents de former les futurs employés... Bien sûr, ce délai dépend fortement du type de travail concerné. Des éléments du même type valent aussi pour les suppressions de postes.
- Des raisons de comportement: si face à une hausse de la production, les entreprises adaptent immédiatement leur niveau d'emploi à un objectif plus élevé, elles risquent d'être confrontées plus tard à un suremploi si la hausse n'est que temporaire. Les travailleurs qui ont reçu une formation, peut-être à un coût élevé, n'auront plus aucune utilité au moment où ils deviennent potentiellement efficaces. Et les licenciés appelleront généralement à des indemnités

soudain pour des biens particuliers nécessitant des artisans spécialisés, même si les outils et les machines sont disponibles sur le marché.

⁴³ Mais pas le début des travaux réels: ce que nous mesurons est le nombre de travailleurs employés, même s'ils sont encore en formation par exemple.

6.2.3.2 Les formules: stationnarité et correction d'erreurs

Il devrait être clair que nous sommes confrontés à un cadre de correction d'erreur, que l'on peut matérialiser de la façon suivante.

- La productivité «normale» du travail ne dépend pas des conditions économiques. Elle pourrait suivre une tendance constante au cours de la période, telles que:

$$\text{Log}(pl_t) = a + b \cdot t$$

- Les entreprises utilisent cet objectif pour définir l'emploi «normal»:

$$LE_t^* = Q_t / pl_t^*$$

- Elles adaptent l'emploi effectif à cet objectif avec une certaine inertie:

$$\Delta \text{Log}(L_t) = \alpha \cdot \Delta \text{Log}(L_t^*) + \beta \cdot \text{Log}(L_{t-1}^* / L_{t-1}) + \gamma + \varepsilon_t$$

Nous reconnaissons ici le cadre à correction d'erreur présenté plus tôt, ce qui nécessite la stationnarité de:

$$\text{Log}(L_t^* / L_t)$$

Mais α n'est pas nécessairement unitaire. Si l'on suit le raisonnement ci-dessus, sa valeur doit être comprise entre 0 et 1, et probablement éloignée de façon significative de chacune de ces limites.

Pour estimer ce système nous sommes confrontés à un problème évident: pl^* n'est pas une véritable série (LE^* non plus, mais si nous nous connaissons l'une nous connaissons l'autre).

Mais si nous appelons « pl^* » le niveau structurel de la productivité (Q / LE), nous pouvons observer que:

$$\text{Log}(L_t^* / L_t) = \text{Log}((Q_t / pl_t^*) / (Q_t / pl_t)) = -\text{Log}(pl_t^* / pl_t)$$

La stationnarité de $\text{Log}(L_t^* / L_t)$ est équivalente à celle de:

$$\text{Log}(pl_t^* / pl_t)$$

Il devrait maintenant être évident que, si pl et pl^* ont une tendance, elle doit être la même, en fait la tendance définissant complètement pl^* . Sinon, ils divergent sur le long terme, et nous ferons face à un sur ou sous-emploi infini. Ainsi, la tendance de la productivité cible peut être identifiée en utilisant la tendance de la valeur observée, si elle existe.

Cela signifie que nous pouvons tester la stationnarité du ratio par celle de la productivité observée autour d'une tendance, un test fourni directement par EViews.

On peut s'attendre à un cadre dans lequel la productivité réelle fluctue autour d'une cible évoluant régulièrement, avec des cycles que nous ne nous attendons pas à être trop longs, mais peuvent durer plusieurs périodes⁴⁴.

6.2.3.3 Les premières estimations

Tout d'abord, calculons la productivité du travail réel

$$\text{genr PROD} = Q / LE$$

et régressons la sur le temps:

$$\text{ls log (PROD) c t}$$

⁴⁴ Qui va créer une autocorrélation (acceptable) dans la différence à la tendance.

pour obtenir la tendance structurelle de la productivité.

Dependent Variable: LOG(Q/LE)

Method: Least Squares

Date: 11/07/12 Time: 12:07

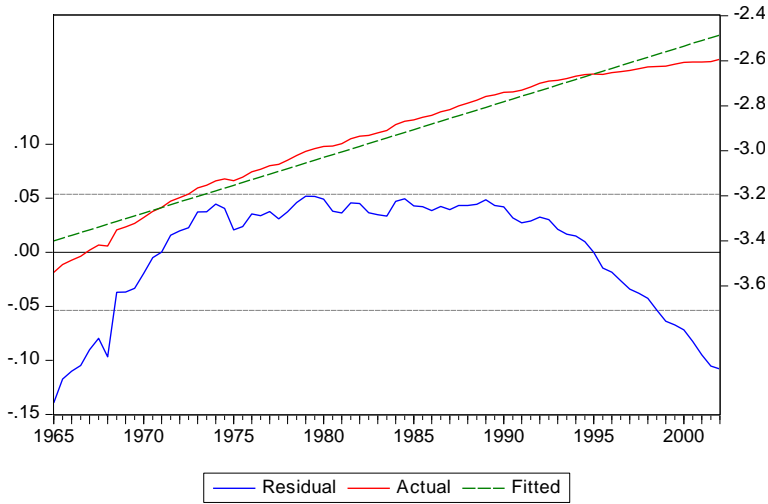
Sample: 1993S1 2002S1

Included observations: 19

LOG(Q/LE)=C(1)+C(2)*T

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-22.34121	0.603556	-37.01598	0.0000
C(2)	0.009865	0.000302	32.64715	0.0000
R-squared	0.984300	Mean dependent var		-2.636851
Adjusted R-squared	0.983377	S.D. dependent var		0.027976
S.E. of regression	0.003607	Akaike info criterion		-8.312619
Sum squared resid	0.000221	Schwarz criterion		-8.213204
Log likelihood	80.96988	Hannan-Quinn criter.		-8.295794
F-statistic	1065.836	Durbin-Watson stat		1.123297
Prob(F-statistic)	0.000000			

Les résultats sont très mauvais. Bien sûr, la productivité affiche une croissance significative, mais l'écart-type est très élevé (plus de 5 %). Plus important encore, le graphique des résidus et le test d'autocorrélation montrent que nous ne respectons pas la condition que nous avons posée: que la productivité observée fluctue autour d'une tendance, avec des cycles éventuels, mais pas excessivement longs.



Le problème réside apparemment dans le fait que le taux de croissance moyen est systématiquement plus élevé dans la première partie de la période, et inférieur plus tard. Si on les observe individuellement, chaque sous-période peut sembler satisfaisante à la condition.

D'après le graphique ci-dessus, nous avons clairement besoin de deux ruptures. On observera que la première arrive avec le premier choc pétrolier, et le début d'un ralentissement durable de l'économie mondiale. La raison de la seconde rupture est moins claire (certains pays comme les États-Unis et en Scandinavie montrent bien une rupture, mais en sens opposé).

Pour choisir les dates les plus appropriées, on peut utiliser deux méthodes:

- visuelle: 1973 et 1990 pourrait être choisies, à un an près ;
- statistique: le test le plus approprié est le test de rupture de Chow («Chow breakpoint test»), que nous avons expliqué plus tôt. Pour faire notre choix de façon automatique, nous allons considérer deux ruptures, et appliquer le test à toutes les combinaisons possibles de dates raisonnables associées. Comme on pouvait s'y attendre, tous les tests concluent à une rupture. Mais nous allons choisir le couple associé à la plus faible probabilité (de l'absence de rupture), autrement dit le plus haut ratio de vraisemblance. Bien sûr, ce critère ne fonctionne que parce que la période de test et le nombre de ruptures restent les mêmes.

Le meilleur résultat correspond en fait à 1973S1 et 1992S1, comme indiqué dans le tableau suivant, présentant les rapports de log-vraisemblance⁴⁵.

Ratio de Log-vraisemblance

dates	1991S2	1992S1	1992S2
1972S2	895	913	908
1973S1	904	928	925
1973S2	895	917	915

L'équation de productivité structurelle est:

$$\text{LOG}(\text{PRLE}) = \text{C_PRLE}(1) + \text{C_PRLE}(2) * (\text{T} - 2004) + \text{C_PRLE}(3) * (\text{T} - 1973) * (\text{T} < 1973) + \text{C_PRLE}(4) * (\text{T} - 1992) * (\text{T} < 1992)$$

On notera:

- que nous avons n'introduit aucun résidu, contrairement à notre habitude ;
- que nous avons introduit des tendances inverses, qui s'arrêtent après un certain temps au lieu de commencer à l'intérieur de la période.

L'objectif de productivité ne représente pas un comportement.

Le premier élément est tout à fait logique: ce que nous estimons pour le modèle n'est pas la productivité réelle (qui sera donnée dans le modèle par une identité, à partir du PIB observé et de l'emploi). Nous sommes à la recherche de la valeur exacte de la productivité cible, sujette à erreur seulement parce que

⁴⁵ Le F plus élevé donne la même conclusion.

nous n'avons pas assez d'informations pour en produire la valeur réelle. Si l'échantillon augmente, ou la périodicité croît, la précision s'améliorera sans cesse, même si le résidu ne diminue pas (il ne représente pas une erreur mais l'écart à la cible). Alors que, dans une équation de comportement normale, les résidus correspondent à une erreur sur la variable et ne peuvent pas être réduits indéfiniment, car l'identification du rôle des éléments explicatifs devient de moins en moins fiable avec leur nombre.

Les tendances partielles devraient s'appliquer à des périodes passées.

La raison en est ici purement technique. Notre modèle sera principalement utilisé sur l'avenir. Il est donc essentiel de rendre le processus de prévision aussi facile que possible.

Si les tendances partielles sont toujours actives dans l'avenir, nous aurons à les gérer simultanément. On peut s'attendre à ce que nous voulions contrôler la tendance globale de la productivité du travail, si ce n'est que pour le rendre conforme à nos évolutions à long terme du PIB (qui devra suivre la croissance mondiale) et de l'emploi (qui devrait suivre la tendance des populations). De toute évidence, il est plus simple de contrôler une seule tendance globale que la combinaison de trois tendances.

En outre, la dernière tendance est la plus importante pour l'interprétation des propriétés du modèle, et il est préférable de la rendre la plus facile à observer.

La technique que nous proposons n'a pas de défauts, une fois qu'elle a été assimilée.

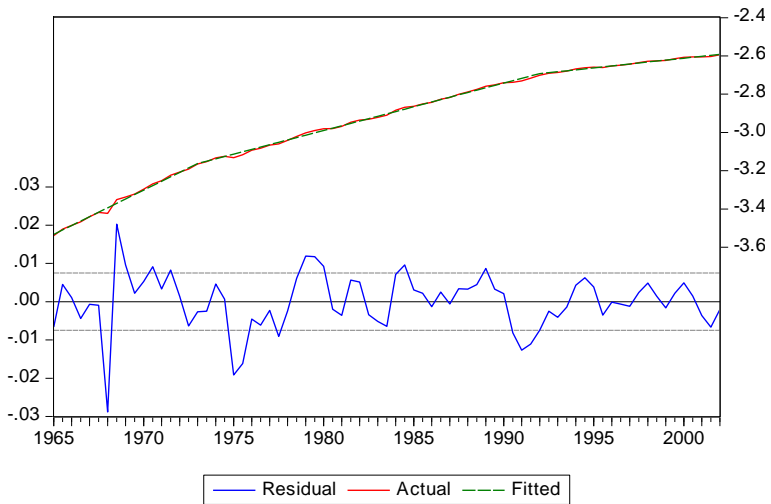
Enfin, faire commencer la tendance à partir de 2002 est également lié à la manipulation de ses valeurs futures. Si le coefficient global est modifié, ce sera la période d'une nouvelle rupture, et c'est la meilleure période pour l'introduire.

Les résultats semblent acceptables, de la validation des coefficients aux graphiques (nous vous présentons la version «programme», telle qu'elle sera introduite dans le modèle)⁴⁶.

⁴⁶ Ceci n'est pas absolument nécessaire, car une variable dépendant uniquement de temps peut être considérée comme exogène et calculée en dehors du modèle. Mais nous voulons être en mesure de changer l'hypothèse dans les prévisions, et ceci est le moyen le plus facile.

Dependent Variable: LOG(PRLE)
 Method: Least Squares
 Date: 11/07/12 Time: 11:59
 Sample (adjusted): 1965S1 2002S1
 Included observations: 75 after adjustments
 $LOG(PRLE)=C_PRLE(1)+C_PRLE(2)*(T-2002)+C_PRLE(3)*(T-T1)*(T<T1)$
 $+C_PRLE(4)*(T-T2)*(T<T2)$

	Coefficient	Std. Error	t-Statistic	Prob.
C_PRLE(1)	-2.591659	0.002928	-885.1806	0.0000
C_PRLE(2)	0.010124	0.000394	25.66585	0.0000
C_PRLE(3)	0.021567	0.000636	33.91560	0.0000
C_PRLE(4)	0.014596	0.000511	28.58584	0.0000
R-squared	0.999288	Mean dependent var	-2.943006	
Adjusted R-squared	0.999258	S.D. dependent var	0.274576	
S.E. of regression	0.007478	Akaike info criterion	-6.901879	
Sum squared resid	0.003970	Schwarz criterion	-6.778280	
Log likelihood	262.8205	Hannan-Quinn criter.	-6.852527	
F-statistic	33233.07	Durbin-Watson stat	1.410774	
Prob(F-statistic)	0.000000			



Nous devons maintenant tester la stationnarité du résidu. Nous allons utiliser le test de Dickey-Fuller (ou Phillips-Perron).

Nous avons d'abord besoin de générer une variable contenant le résidu courant (comme le test va calculer son propre RESID, il n'est donc pas possible de faire le test sur celui-ci).

Sous forme de programme, le test est effectué par:

```
genr res_prle = resid
uroot (1, p) res_prle
uroot (h, p) res_prle
```

En utilisant les menus, il faut

- afficher la variable,
- sélectionner View>Unit root test,
- choisir la méthode. Nous avons choisi ici celle de Dickey-Fuller augmentée, sans doute la plus commune, même s'il en existe de plus récentes (Phillips-Perron, KPSS...)

Null Hypothesis: RES_PRLE has a unit root
 Exogenous: Constant
 Lag Length: 1 (Automatic - based on SIC, maxlag=11)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.191811	0.0000
Test critical values:		
1% level	-3.522887	
5% level	-2.901779	
10% level	-2.588280	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(RES_PRLE)
 Method: Least Squares
 Date: 11/07/12 Time: 12:10
 Sample (adjusted): 1966S1 2002S1
 Included observations: 73 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RES_PRLE(-1)	-0.735169	0.141602	-5.191811	0.0000
D(RES_PRLE(-1))	0.048991	0.118711	0.412687	0.6811
C	-2.46E-06	0.000833	-0.002958	0.9976

R-squared	0.352694	Mean dependent var	-8.95E-05
Adjusted R-squared	0.334200	S.D. dependent var	0.008722
S.E. of regression	0.007117	Akaike info criterion	-7.012437
Sum squared resid	0.003546	Schwarz criterion	-6.918309
Log likelihood	258.9540	Hannan-Quinn criter.	-6.974926
F-statistic	19.07026	Durbin-Watson stat	1.993440
Prob(F-statistic)	0.000000		

Le test conclut très fortement à la stationnarité du résidu.

Les valeurs de la productivité structurelle et de l'emploi cible sont données par:

```

genr          log(prle_t)=c_prle(1)+c_prle(2)*t+c_prle(3)*(t-1973)*(t<1973)+c_prle(4)*(t-
1992)*(t<1992)
genr led = q / prle_t
    
```

LE lui-même sera estimé (en utilisant ici la forme développée) par:

```

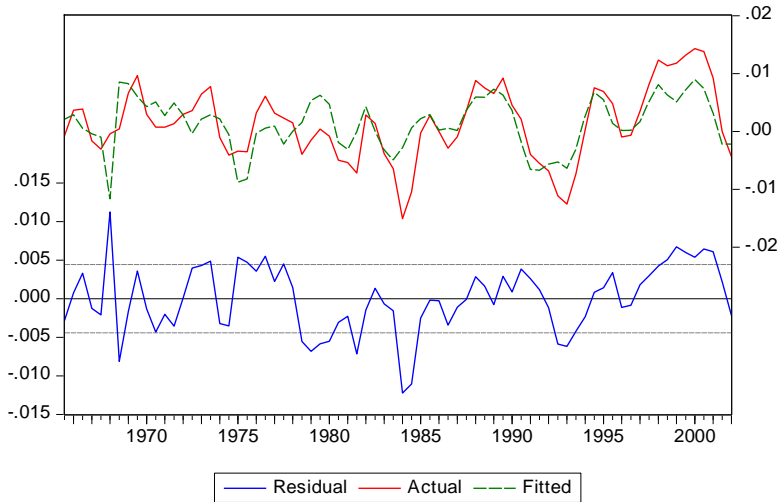
equation eq_le.lsdlog(LE)=c_le(1)*dlog(LED)+c_le(2)*log(LED(-1)/LE(-1))+c_le(3)
    
```

où LED / LE est égale à (Q / LE) / prle, le résidu de l'équation précédente.

Les résultats sont assez satisfaisants, à l'exception du dernier coefficient.

Dependent Variable: DLOG(LE)
 Method: Least Squares
 Date: 11/07/12 Time: 12:55
 Sample (adjusted): 1965S2 2002S1
 Included observations: 74 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)	0.422879	0.054347	7.781036	0.0000
C_LE(2)	0.455672	0.074240	6.137818	0.0000
C_LE(3)	0.000732	0.000520	1.407567	0.1636
R-squared	0.514224	Mean dependent var		0.001335
Adjusted R-squared	0.500540	S.D. dependent var		0.006259
S.E. of regression	0.004423	Akaike info criterion		-7.964257
Sum squared resid	0.001389	Schwarz criterion		-7.870849
Log likelihood	297.6775	Hannan-Quinn criter.		-7.926996
F-statistic	37.57891	Durbin-Watson stat		1.015944
Prob(F-statistic)	0.000000			



Si l'on suit le raisonnement précédent, $c_le(3)$ (ou plutôt $c_le(3)/c_le(2)$) doit représenter le logarithme de l'écart à long terme entre l'emploi cible et le niveau atteint. Cet écart sera significatif si à la fois:

- L'emploi affiche une tendance (la cible est en mouvement), ce qui signifie que la productivité et le PIB présentent des tendances différentes.
- Une différence entre les croissances de la productivité et du PIB cible n'est pas compensée immédiatement (la valeur de α est différente de un)

La seconde condition est clairement remplie, mais pas la première.

Dans le cas présent, l'emploi ne montre aucune tendance évidente (positive peut-être ?). Et le dernier coefficient n'est pas vraiment nécessaire. Mais il n'y a aucune raison de l'éliminer (comme il n'y a aucune raison de renoncer à une constante non significative dans une équation linéaire).

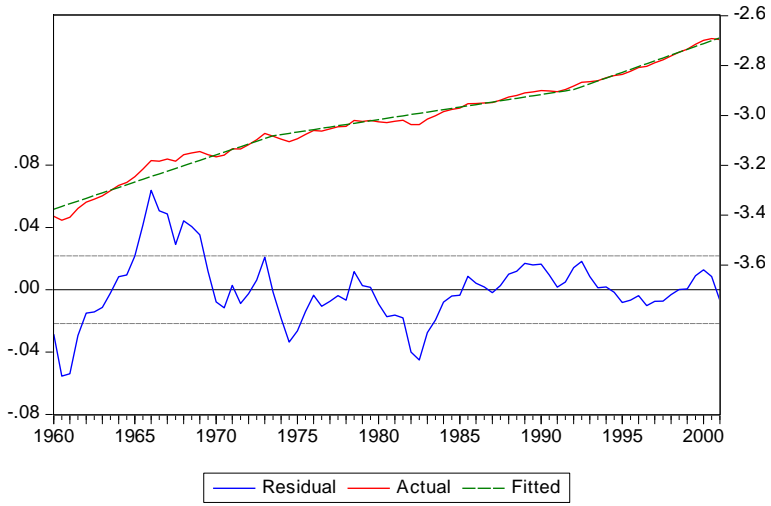
En ce qui concerne les premiers coefficients, ils semblent très différents. Le premier est plus significatif et aussi plus fort: cela pourrait signifier qu'il est plus facile (et plus souhaitable) de combler la première partie de l'écart entre l'emploi désiré et réel.

Ce ne sera pas vrai dans le cas américain.

Tout d'abord, la rupture «1992» existe aussi, mais elle est maintenant positive, comme indiqué ici:

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)

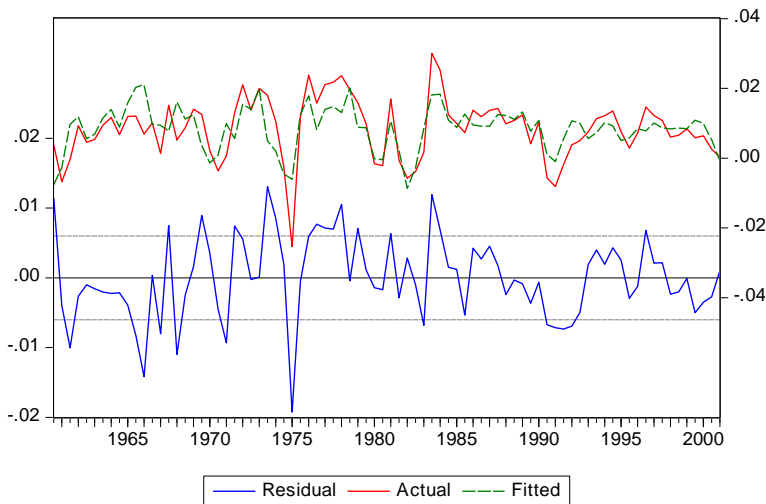
	Coefficient	Std. Error	t-Statistic	Prob.
C_PRLE(1)	-48.83508	2.620014	-18.63924	0.0000
C_PRLE(2)	0.023062	0.001313	17.55985	0.0000
C_PRLE(3)	0.011769	0.001133	10.38824	0.0000
C_PRLE(4)	-0.013038	0.001639	-7.954501	0.0000
R-squared	0.985989	Mean dependent var		-3.025035
Adjusted R-squared	0.985457	S.D. 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			



Deuxièmement, l'emploi a augmenté considérablement au cours de la période étudiée, ce qui signifie que l'introduction d'un terme constant sera nécessaire:

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			



6.2.3.3.1 Une variable muette vertueuse

Revenant sur le cas français, on observera une aberration locale: l'année 1968 présente un fort résidu négatif au premier semestre, et une valeur négative pour le dernier. Les Français (et des personnes familières avec la France de l'après-guerre) se souviennent certainement de la «révolution étudiante» de mai 1968, qui a duré à peu près de mars à juin. Pendant cette période, le fonctionnement de l'économie française a été fortement perturbé, en particulier le système de transports, et le PIB a diminué (de 0,5 % pour le semestre⁴⁷). Si l'équation avait fonctionné, l'emploi aurait diminué aussi, d'autant plus que la croissance de la productivité était très élevée à cette époque. Au contraire, il est resté à peu près stable.

L'explication est évidente: les entreprises ont considéré que l'effondrement était purement temporaire, et que l'activité allait repartir après un certain temps (en fait, elles avaient raison, et le PIB a augmenté de 7,5 % au cours du semestre suivant, en partie grâce à l'augmentation de la consommation permise par les négociations salariales dites «Grenelle⁴⁸», très favorables aux travailleurs). Elles étaient peu portées à licencier (à un coût élevé) les travailleurs dont elles auraient besoin plus tard, sans aucune garantie de retrouver les mêmes personnes, familières avec les techniques qu'elles employaient. Ainsi, le niveau de l'emploi a été très peu affecté.

Cela signifie que le comportement global ne s'applique pas ici, et la période doit être soit éliminée de l'échantillon, soit traitée par un élément spécifique, par exemple une variable prenant la valeur 1 au premier semestre et -1 au second (lorsqu'une augmentation de l'emploi inférieure à la croissance du PIB l'exigerait).

⁴⁷ Alors qu'on se trouvait encore dans une période de croissance forte.

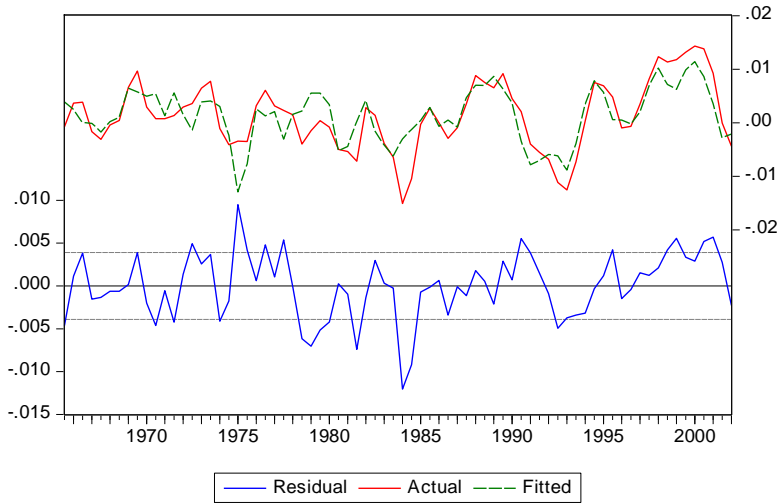
⁴⁸ Nom de l'emplacement du ministère de l'Emploi où les négociations ont été menées.

Dependent Variable: DLOG(LE)
Method: Least Squares
Date: 11/07/12 Time: 12:53
Sample (adjusted): 1965S2 2002S1
Included observations: 74 after adjustments
DLOG(LE)=C_LE(1)*DLOG(LED)+C_LE(2)*LOG(LED(-1)/LE(-1))+C_LE(3)
+C_LE(4)*((T=1968.5)-(T=1968))

	Coefficient	Std. Error	t-Statistic	Prob.
C_LE(1)	0.587984	0.059833	9.827155	0.0000
C_LE(2)	0.411309	0.066185	6.214541	0.0000
C_LE(3)	0.000502	0.000461	1.089292	0.2798
C_LE(4)	-0.016716	0.003625	-4.611097	0.0000
R-squared	0.627400	Mean dependent var		0.001335
Adjusted R-squared	0.611431	S.D. dependent var		0.006259
S.E. of regression	0.003901	Akaike info criterion		-8.202472
Sum squared resid	0.001065	Schwarz criterion		-8.077928
Log likelihood	307.4915	Hannan-Quinn criter.		-8.152790
F-statistic	39.28961	Durbin-Watson stat		1.076553
Prob(F-statistic)	0.000000			

Ce cas est assez intéressant: un économiste peut être tenté d'introduire une variable muette simplement lorsque l'équation ne fonctionne pas pour une période donnée, sans y identifier de problème particulier. Ceci va généralement améliorer les résultats (y compris les statistiques générales pour les variables explicatives). Mais cela pourrait sans doute s'appeler tricherie. Au contraire, ne pas en introduire dans la situation actuelle peut être considéré comme une erreur: nous savons que le comportement local n'a pas suivi la formulation que nous avons choisie, il doit donc être modifié en conséquence.

Les résultats globaux sont légèrement améliorés, et le premier coefficient augmente de manière significative (ce qui était prévisible). L'introduction de l'élément n'était donc pas une question négligeable.



Nous utiliserons:

```
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
```

et nous devons introduire deux nouvelles variables:

```
g_vendo.drop led prle_t
g_vendo.add led prle_t
```

Remarque: la raison de la première instruction «drop» est d'éviter la duplication des éléments à l'intérieur du groupe, dans le cas où la procédure est répétée. Si les éléments ne sont pas présents, rien ne se passe.

6.2.4 LES EXPORTATIONS: LES PROCESSUS AUTOREGRESSIFS, LA COINTEGRATION, LA STABILITE DE LONG TERME

L'estimation des exportations sera plus simple du côté théorique. Nous allons l'utiliser comme un exemple pour introduire nos premiers processus autorégressifs, puis la cointégration.

Commençons d'abord par une idée élémentaire: les exportations montrent une élasticité constante à la demande mondiale. En d'autres termes:

$$\Delta X / X / (\Delta WD / WD) = a$$

ou par intégration:

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

L'estimation devrait donner à a une valeur proche de l'unité, compte tenu du fait que la demande mondiale est mesurée comme la demande normalement adressée à la France par ses clients, et prend en compte:

- l'expansion du commerce international,
- les types de biens exportés par la France,
- la structure des pays vers lesquels la France exporte naturellement.

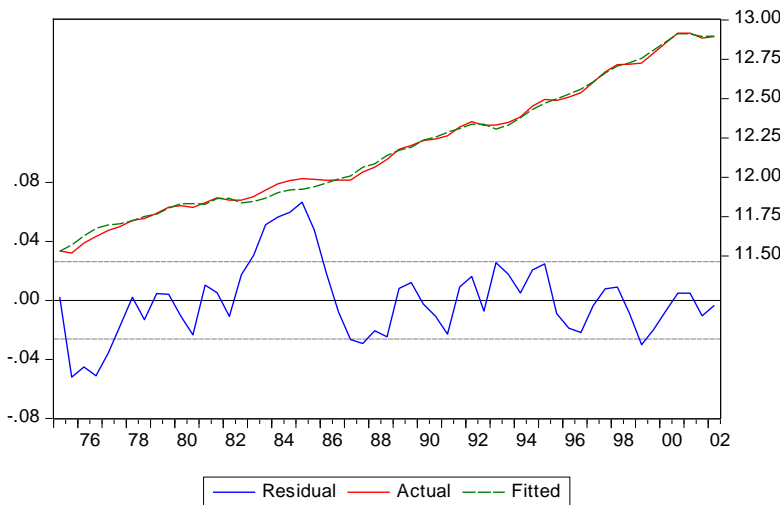
Par exemple, les produits de luxe et l'Allemagne ont tous deux une part plus élevée dans cet indicateur, par rapport au marché mondial. Comme une pondération des importations normales, il prend également en compte le développement du commerce international.

En effet, le coefficient que nous obtenons est proche de l'unité.

Dependent Variable: LOG(X)
 Method: Least Squares
 Date: 11/07/12 Time: 11:59
 Sample (adjusted): 1975S1 2002S1
 Included observations: 55 after adjustments
 LOG(X)=C_X(1)*LOG(WD)+C_X(2)

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	1.002987	0.008866	113.1316	0.0000
C_X(2)	-0.105243	0.108654	-0.968605	0.3371

R-squared	0.995876	Mean dependent var	12.18047
Adjusted R-squared	0.995798	S.D. dependent var	0.403814
S.E. of regression	0.026176	Akaike info criterion	-4.412289
Sum squared resid	0.036314	Schwarz criterion	-4.339295
Log likelihood	123.3380	Hannan-Quinn criter.	-4.384062
F-statistic	12798.76	Durbin-Watson stat	0.497272
Prob(F-statistic)	0.000000		



Cependant la faible valeur du test du Durbin-Watson indique une autocorrélation des résidus fortement positive, et invalide la formulation.

Cherchons à éliminer l'autocorrélation, en supposant que la valeur du résidu évolue comme:

$$e_t = \rho \cdot e_{t-1} + u_t$$

où ρ devrait être significatif (ici positif), et $u(t)$ doit être indépendant du temps.

L'idée la plus simple consiste à transformer l'équation

$$\text{Log}(X_t) = a \cdot \text{Log}(WD_t) + b + e_t$$

qui est également valable à la période précédente.

$$\text{Log}(X_{t-1}) = a \cdot \text{Log}(WD_{t-1}) + b + e_{t-1}$$

On peut multiplier la deuxième équation par ρ , et soustraire le résultat de la première:

$$\text{Log}(X_t) - \rho \cdot \text{Log}(X_{t-1}) = a \cdot (\text{Log}(WD_t) - \rho \cdot \text{Log}(WD_{t-1})) + b \cdot (1 - \rho) + e_t - \rho \cdot e_{t-1}$$

Le résidu de la nouvelle équation est u non corrélé:

$$\text{Log}(X_t) - \rho \cdot \text{Log}(X_{t-1}) = a \cdot (\text{Log}(WD_t) - \rho \cdot \text{Log}(WD_{t-1})) + b \cdot (1 - \rho) + u_t$$

6.2.4.1 Introduire un processus autorégressif

Pour estimer la formule ci-dessus, il n'est pas nécessaire d'établir l'équation complète (qui réclame explicitation, car elle n'est pas linéaire par rapport aux coefficients).

On peut très bien utiliser la même présentation que pour les moindres carrés ordinaires, en introduisant dans la fenêtre d'estimation un terme additionnel $AR(n)$, n représentant le décalage d'autocorrélation, 1 dans notre cas:

```
ls Log (X) log (WD) c ar(1)
```

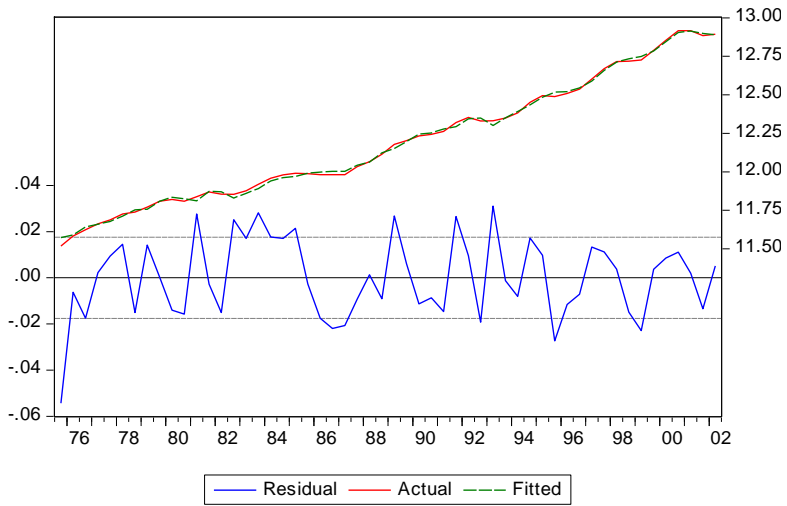
Mais l'application d'une formule complètement spécifiée est également très simple:

```
equation eq_lx.ls Log (X) = c_x (1) * log (WD) + c_x (2) + [ar (1) = c_x (3)]
```

Les résultats sont plutôt satisfaisants: le premier coefficient conserve la valeur théorique, le nouveau coefficient est significatif, la précision globale est nettement améliorée (voir aussi le graphique) et le test de Durbin - Watson s'est rapproché d'une valeur satisfaisante.

Dependent Variable: LOG(X)
 Method: Least Squares
 Date: 11/07/12 Time: 11:59
 Sample (adjusted): 1975S2 2002S1
 Included observations: 54 after adjustments
 Convergence achieved after 9 iterations
 LOG(X)=C_X(1)*LOG(WD)+C_X(2)+[AR(1)=C_X(3)]

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.999049	0.023932	41.74592	0.0000
C_X(2)	-0.057019	0.295426	-0.193006	0.8477
C_X(3)	0.752588	0.092241	8.158958	0.0000
R-squared	0.998113	Mean dependent var		12.19248
Adjusted R-squared	0.998039	S.D. dependent var		0.397575
S.E. of regression	0.017604	Akaike info criterion		-5.187439
Sum squared resid	0.015805	Schwarz criterion		-5.076940
Log likelihood	143.0609	Hannan-Quinn criter.		-5.144824
F-statistic	13491.02	Durbin-Watson stat		1.530307
Prob(F-statistic)	0.000000			
Inverted AR Roots	.75			

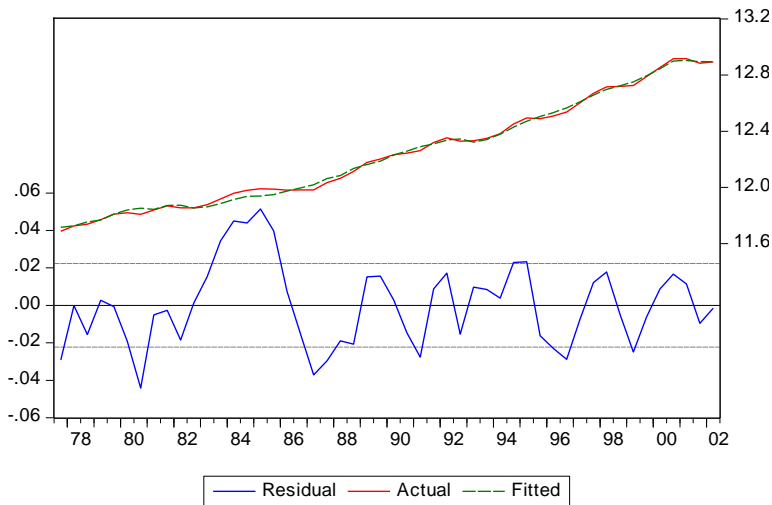


Cependant, notre formulation est un peu trop simpliste. Nous voulons que les exportations diminuent avec le taux d'utilisation des capacités, ce qui représente le fait que plus leurs capacités sont déjà utilisées, moins les entreprises locales vont être dynamiques dans leur recherche de marchés étrangers (voir plus loin).

Introduisons le taux d'utilisation UR dans la formule initiale (on ne sait jamais, il pourrait éliminer l'autocorrélation). Nous obtenons:

Dependent Variable: LOG(X)
 Method: Least Squares
 Date: 11/07/12 Time: 12:03
 Sample (adjusted): 1977S2 2002S1
 Included observations: 50 after adjustments
 LOG(X)=C_X(1)*LOG(WD)+C_X(2)+C_X(3)*UR

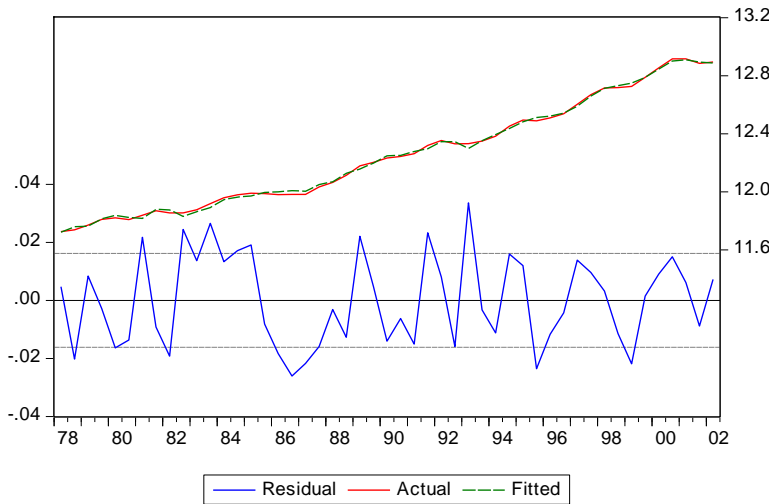
	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.989908	0.008545	115.8478	0.0000
C_X(2)	0.207375	0.120579	1.719825	0.0920
C_X(3)	-0.149706	0.076793	-1.949464	0.0572
R-squared	0.996566	Mean dependent var	12.24014	
Adjusted R-squared	0.996420	S.D. dependent var	0.373424	
S.E. of regression	0.022342	Akaike info criterion	-4.706530	
Sum squared resid	0.023462	Schwarz criterion	-4.591809	
Log likelihood	120.6632	Hannan-Quinn criter.	-4.662843	
F-statistic	6820.462	Durbin-Watson stat	0.752199	
Prob(F-statistic)	0.000000			



Sans surprise, la valeur du test de Durbin-Watson invalide les résultats à nouveau (le miracle ne s'est pas produit).

Utilisons la même méthode que précédemment:

Dependent Variable: LOG(X)				
Method: Least Squares				
Date: 11/07/12 Time: 12:03				
Sample (adjusted): 1978S1 2002S1				
Included observations: 49 after adjustments				
Convergence achieved after 10 iterations				
LOG(X)=C_X(1)*LOG(WD)+C_X(2)+C_X(3)*UR+[AR(1)=C_X(4)]				
	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.975444	0.023544	41.43150	0.0000
C_X(2)	0.161832	0.291397	0.555367	0.5814
C_X(3)	0.079844	0.073949	1.079726	0.2860
C_X(4)	0.742440	0.098471	7.539650	0.0000
R-squared	0.998194	Mean dependent var		12.25142
Adjusted R-squared	0.998073	S.D. dependent var		0.368588
S.E. of regression	0.016179	Akaike info criterion		-5.332101
Sum squared resid	0.011779	Schwarz criterion		-5.177666
Log likelihood	134.6365	Hannan-Quinn criter.		-5.273508
F-statistic	8289.252	Durbin-Watson stat		1.692421
Prob(F-statistic)	0.000000			
Inverted AR Roots	.74			



Nous avons éliminé l'autocorrélation, mais maintenant le coefficient de UR n'est plus significatif (heureusement car il a le mauvais signe).

Ne nous désespérons pas. Si cet outil ancien ne fonctionne pas, essayons-en un autre un peu plus moderne: la cointégration.

6.2.4.1.1 L'application de la cointégration sous EViews

Comme nous l'avons indiqué précédemment, pour établir la cointégration de deux éléments, on doit prouver que dans le long terme ces éléments se déplacent ensemble, maintenant une distance bornée (ou plutôt qu'une combinaison linéaire des deux éléments est bornée), tandis que la valeur de chacun ne l'est pas (condition nécessaire).

Pour qu'un groupe de plus de deux éléments soit cointégré, aucun sous-ensemble de ce groupe ne doit avoir cette propriété (stationnarité d'un seul élément ou cointégration d'un sous-ensemble).

Si nous voulons aller au-delà de l'intuition, la raison de cette dernière condition est que si une relation de cointégration est mise en évidence entre les éléments, dont certains sont déjà cointégrés, on peut toujours recomposer l'équation globale en la vraie équation de cointégration (considérée comme une nouvelle variable stationnaire) et d'autres variables.

Par exemple, si:

$$a \cdot x + b \cdot y + c \cdot z$$

est testé comme une équation de cointégration, mais:

$$a \cdot x + b' \cdot y$$

l'est aussi (on peut utiliser de même une forme d'une équation de cointégration est connue à un facteur donné), alors

$$a \cdot x + b \cdot y + c \cdot z$$

est égal à:

$$(a \cdot x + b' \cdot y) + (b - b') \cdot y + c \cdot z$$

Avec trois éléments nouveaux, dont l'un est stationnaire, qui interdit de tester sur la cointégration des trois.

Les deux propriétés doivent être vérifiées: bouger ensemble signifie à la fois «bouger» et «ensemble».

En utilisant des images plutôt liées à la stationnarité (elles s'appliquent alors à la variable mesurée par la différence de deux éléments, sans coefficients de pondération), nous pouvons illustrer le concept considérant les éléments comme:

- des corps en mouvement dans l'espace et liés ensemble par la gravité. Leur distance est limitée, mais leur position par rapport à l'autre est inconnue à l'intérieur de ces limites, et nous ne savons pas si l'un détermine l'autre ;
- des êtres humains: s'ils sont toujours proches les uns des autres, cela peut vouloir dire qu'ils sont liés (par l'amour, la haine, une relation professionnelle). Mais seulement s'ils se déplacent: s'ils sont en prison, la faible distance ne signifie rien.

Dans notre exemple, la première idée pourrait être de tester la cointégration entre X , WD et UR . Mais pour assurer la stabilité de nos simulations à long terme, nous avons besoin que les exportations présentent une élasticité unitaire relativement à WD . Si ce n'est pas le cas, si X atteint un taux de croissance constant, il sera différent de celui de WD : soit la France deviendra le seul exportateur dans le monde (en termes relatifs) soit le rôle de la France dans le marché mondial deviendra infiniment négligeable. Les deux perspectives sont inacceptables (la première plus que la seconde, nous devons l'admettre).

Cette contrainte peut être appliquée très facilement en ne considérant dans le long terme (équation de cointégration) que le rapport de X à WD , que nous allons lier au taux d'utilisation. Nous allons donc tester la cointégration entre ces deux éléments.

Testons d'abord leur stationnarité. Nous savons comment le faire (en nous souvenant de l'estimation de l'emploi).

Dans un programme, ceci peut être fait par:

```
UROOT (1, p) log (X/WD)
UROOT (1, p) Log (UR)
```

Remarque: si l'on utilise les menus, il faut d'abord afficher l'ensemble (soit en sélectionnant les éléments dans la fenêtre du fichier de travail ou en créant un groupe). Le menu propose alors l'option «Cointegration test» avec comme sous-option «Johansen System Cointegration Test».

Null Hypothesis: LOG(X/WD) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.709700	0.2370
Test critical values:		
1% level	-4.133838	
5% level	-3.493692	
10% level	-3.175693	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(LOG(X/WD))
 Method: Least Squares
 Date: 11/07/12 Time: 13:04
 Sample (adjusted): 1975S2 2002S2
 Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(X(-1)/WD(-1))	-0.249865	0.092211	-2.709700	0.0091
C	-0.019128	0.011568	-1.653591	0.1042
@TREND("1960S1")	3.10E-05	0.000149	0.207341	0.8366

R-squared	0.123932	Mean dependent var	-0.000178
Adjusted R-squared	0.090237	S.D. dependent var	0.018311
S.E. of regression	0.017465	Akaike info criterion	-5.204224
Sum squared resid	0.015862	Schwarz criterion	-5.094733
Log likelihood	146.1162	Hannan-Quinn criter.	-5.161883
F-statistic	3.678052	Durbin-Watson stat	1.532444
Prob(F-statistic)	0.032062		

Null Hypothesis: LOG(UR) has a unit root				
Exogenous: Constant, Linear Trend				
Lag Length: 0 (Automatic - based on SIC, maxlag=10)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-3.181020	0.0999
Test critical values:		1% level	-4.152511	
		5% level	-3.502373	
		10% level	-3.180699	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LOG(UR))				
Method: Least Squares				
Date: 11/07/12 Time: 13:05				
Sample (adjusted): 1978S1 2002S2				
Included observations: 50 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(UR(-1))	-0.355378	0.111718	-3.181020	0.0026
C	-0.006570	0.020409	-0.321922	0.7489
@TREND("1960S1")	3.81E-05	0.000326	0.116806	0.9075
R-squared	0.177436	Mean dependent var		1.24E-05
Adjusted R-squared	0.142433	S.D. dependent var		0.035782
S.E. of regression	0.033136	Akaike info criterion		-3.918294
Sum squared resid	0.051604	Schwarz criterion		-3.803573
Log likelihood	100.9574	Hannan-Quinn criter.		-3.874608
F-statistic	5.069199	Durbin-Watson stat		1.579826
Prob(F-statistic)	0.010151			

Ces premiers tests montrent qu'à la fois *UR* et le ratio des exportations à la demande mondiale ne peuvent pas être considérés comme stationnaires, même autour d'une tendance: la statistique *T* est trop faible, et la probabilité estimée pour que le coefficient vaille zéro est trop élevée^{49, 50}.

Voyons maintenant si les deux éléments sont cointégrés, en utilisant le test de Johansen.

Pour EViews cela s'exprime par:

```
coint(option,p) liste-des-variables ou nom-de-groupe
```

«option» représente le type de cointégration testé, selon la présence de constantes et de tendances:

- a aucune tendance déterministe dans le VAR, et aucune constante ou tendance dans l'équation de cointégration.
- b aucune tendance déterministe dans le VAR, et une constante, mais aucune tendance dans l'équation de cointégration.
- c une tendance linéaire dans le VAR, et constante, mais pas de tendance dans l'équation de cointégration.
- d une tendance linéaire dans le VAR, et à la fois une constante et une tendance dans l'équation de cointégration.
- e une tendance quadratique dans le VAR, et à la fois une constante et une tendance dans l'équation de cointégration.
- s un résumé des résultats pour les 5 options.

Dans notre cas, nous allons utiliser l'option d (tendance dans l'équation de cointégration, aucune tendance dans le VAR)⁵¹

⁴⁹ Pas très élevée, cependant.

⁵⁰ Nous pouvons observer que dans une estimation traditionnelle par les moindres carrés, la même valeur de *T* donnerait le diagnostic inverse.

⁵¹ Ceci est bien sûr cohérent avec nos tests de stationnarité.

coint(d,p) log (x/wd) log(ur)

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)				
Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.366312	20.98510	19.38704	0.0291
At most 1	0.186003	9.466749	12.51798	0.1535
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/TD)	LOG(UR)	@TREND(60S2)		
-7.930144	24.22342	-0.030087		
22.38533	-17.70598	-0.342308		
Unrestricted Adjustment Coefficients (alpha):				
D(LOG(M/TD))	-0.013891	-0.005565		
D(LOG(UR))	-0.020030	0.008593		
1 Cointegrating Equation(s): Log likelihood 203.3606				
Normalized cointegrating coefficients (standard error in parentheses)				
LOG(M/TD)	LOG(UR)	@TREND(60S2)		
1.000000	-3.054601	0.003794		
	(0.48502)	(0.00368)		
Adjustment coefficients (standard error in parentheses)				
D(LOG(M/TD))	0.110155			
	(0.02737)			
D(LOG(UR))	0.158844			
	(0.04042)			

De toute évidence, la cointégration est acceptée (un message le dit). Mais nous pouvons aussi:

- Interpréter le processus logique:

EViews teste d'abord s'il n'y a pas de cointégration. Si cela est accepté (si sa probabilité est assez élevée), le processus s'arrête. Mais ici la cointégration n'est pas refusée, car la probabilité (qu'il n'y a pas de cointégration) est trop faible.

Dans ce cas, il existe au moins une équation de cointégration, et EViews va tester s'il y en a au plus une. Cela ne sera pas refusé ici, car cette fois la probabilité est trop élevée.

Nous avons au moins une relation, et au plus une: nous avons une.

Si la deuxième hypothèse (au plus 1 relation) avait été rejetée, il y en aurait au moins deux et nous devrions continuer (il ne peut y avoir plus de relations que de variables, toutefois).

La mise en évidence de plus d'une relation est problématique, peut-être pire pour le constructeur de modèles que de n'en trouver aucune (dans ce cas, nous pouvons toujours continuer à ajouter des éléments). Même si une équation de cointégration n'a aucune incidence sur la causalité entre les éléments, nous avons généralement l'intention de l'inclure dans une formule unique dynamique (VAR), qui expliquera un des éléments utilisés. Avec deux équations, nous devons faire face à une relation parasite, qui sera difficile, voire impossible à gérer dans le contexte du modèle (si nous nous arrêtons à la phase économétrique, le problème est moins grave).

- Observer les probabilités:

Nous pouvons observer si l'existence (ou le rejet) d'au moins une relation est à peine ou fortement acceptée (ainsi que d'une seule d'ailleurs).

- Étudier les coefficients de l'équation de cointégration:

L'équation introduit un arbitrage entre plusieurs concepts (ici, la part des exportations françaises dans la demande mondiale et le taux d'utilisation). Nous avons une idée sur le signe de la relation, ainsi que sur un intervalle de validité économique. Il n'y a aucune garantie que la valeur suivra ces contraintes. Il peut même arriver que le bon signe obtenu par les moindres carrés ordinaires devienne mauvais quand la cointégration est testée sur la même relation.

Ici, il n'est pas trop difficile de juger du bien-fondé de l'explication.

Tout d'abord, le signe est juste: les exportations se réduisent lorsque le taux d'utilisation augmente (le signe est positif, mais les deux éléments sont du même côté de la formule).

La valeur du coefficient est plus difficile à juger. La dérivation de l'équation par rapport à Q donne:

$$\frac{\Delta(X)}{X} = -0,62 \frac{\Delta(UR)}{UR} = -0,62 \left(\frac{\Delta(Q)}{Q} - \frac{\Delta(CAP)}{CAP} \right)$$

À court terme, CAP ne change pas:

$$\begin{aligned} \frac{\Delta(X)}{X} &= -0,62 \frac{\Delta(UR)}{UR} = -0,62 \frac{\Delta(Q)}{Q} \\ \Delta(X) &= -0,62 (X / Q) \Delta(Q) \end{aligned}$$

Supposons une augmentation de Q de 1 milliard d'euros provenant uniquement de la demande locale FD . En 2004, la part des exportations dans le PIB français était de 32 %. La cible d'exportations diminuera de

$$X = 0,62 * 0,32 = 205 \text{ millions d'euros.}$$

L'effet de substitution semble tout à fait raisonnable.

Bien sûr:

- À long terme, les capacités vont s'accumuler, UR va revenir à sa valeur de base (comme nous le dit l'équation d'investissement) et la perte disparaîtra.
- Les changements de Q peuvent aussi venir de X , introduisant une boucle. Cela signifie que UR pourrait ne pas être l'élément le plus représentatif. Peut-être devrions nous limiter UR à la satisfaction de la demande locale (mais cela semble difficile à formuler).

On peut aussi observer la significativité des coefficients, pour mesurer la fiabilité de la relation. Elle est satisfaisante.

Lorsque la cointégration a été mise en évidence

Deux choses doivent être faites:

- stocker l'équation de cointégration et ses paramètres ;
- estimer le VAR (l'équation dynamique) et créer l'élément associé.

La première tâche devrait être facile, comme EViews affiche l'équation demandée, avec ses valeurs. Toutefois:

- cette équation n'est pas disponible comme un objet ;
- les coefficients ne sont pas disponibles en tant que vecteur (ou scalaires).

Il existe une méthode pour résoudre ce problème. On peut estimer un VAR, en d'autres termes un système qui comprend à la fois une équation dynamique et une équation de cointégration. Ceci ne paraît pas directement utile pour nous. En effet:

- Cela nécessite d'utiliser les mêmes variables dans les deux formes, en d'autres termes d'étendre l'hypothèse d'élasticité unitaire à l'équation dynamique, ce qui n'est ni nécessaire sur le plan statistique ni réaliste d'un point de vue économique (comme nous l'avons vu lors de l'estimation de l'emploi).
- La sortie ne fournit aucune information sur la qualité de la cointégration, un élément essentiel dans notre processus.

Mais le point positif est que dans l'estimation d'un VAR utilisant les mêmes éléments que la cointégration testée, nous obtenons les mêmes coefficients de cointégration, et cette fois ils sont stockés dans un vecteur ! Nous pouvons ensuite définir l'équation de cointégration utilisant ces éléments.

Ceci a l'avantage extrêmement élevé de permettre d'établir un programme qui s'adaptera automatiquement à tout changement dans les données, un élément essentiel dans tous les projets de modélisation opérationnels.

Pour cet exemple particulier, après:

```
COINT(d,p) log(x/wd) log(ur)
```

nous utiliserons

```
_var_x.ec(d,p) 1 1 log(x/wd) log(ur)52
```

⁵² Les valeurs 1 et 1 indiquent l'étendue des variations retardées de la variable de gauche du VAR, qui sera ajoutée aux termes de droite. Ici, il sera de 1 à 1 (ou 1 avec un décalage de 1). Les options utilisées sont conformes à celles que nous avons utilisées dans COINT (qui ont été déterminées automatiquement par EViews).

pour créer et d'estimer le VAR, et récupérer les coefficients dans un vecteur en accédant à la première ligne de la matrice générée `_var_x.b` (disponible, bien qu'elle ne s'affiche pas dans la fenêtre de fichier de travail):

```
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)
```

L'équation de cointégration sera la suivante:

```
0= p_x(1)*log(x/wd)+p_x(2)*log(ur)+p_x(3)*@trend(60S1)
```

En fait, le premier paramètre n'est pas vraiment nécessaire, car il est égal à 1 par construction. Nous pensons que son utilisation rend l'équation plus claire.

On peut alors calculer la valeur du résidu de l'équation de cointégration:

```
genr res_x= p_x(1)*log(x/wd)+p_x(2)*log(ur)+p_x(3)*@trend(60S1)
```

Il faut ensuite estimer le VAR, en relâchant la contrainte sur l'élasticité de X à WD . En principe, le coefficient devrait être positif pour WD , négatif pour UR , et nous pouvons introduire une structure de retards pour les deux éléments.

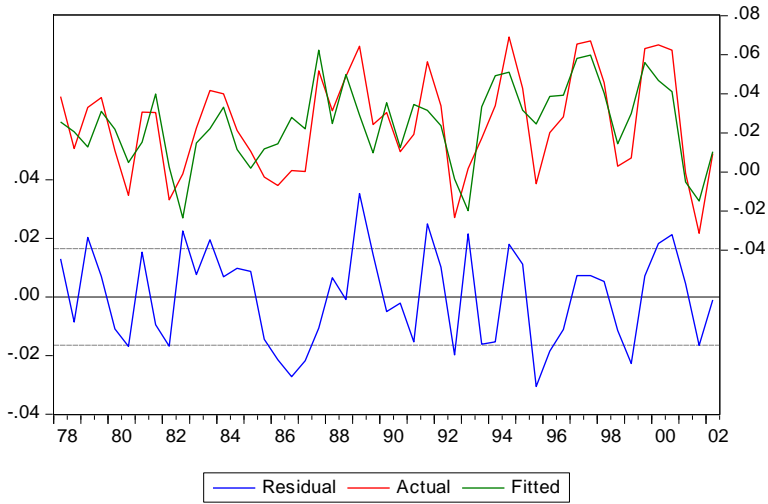
Le principal problème ici est la validité économique: nous devons évaluer, par exemple, si une croissance donnée de la demande mondiale adressée à la France (disons de 1 %) en l'absence de problèmes de capacité (ce sera le rôle de UR) va augmenter les exportations de plus ou moins de 1 %.

Dans notre cas, malheureusement, la variation du taux d'utilisation n'est pas significative⁵³, ou toute autre transformation de cet élément (il faut vraiment le mauvais signe). Nous l'avons éliminé.

⁵³ Ceci pourrait être expliqué par le fait que l'impact du taux d'utilisation est sur les contrats d'exportation, les exportations réelles suivant avec un certain retard. Dans ce cas, l'équation de cointégration aurait pu considérer le taux d'utilisation décalé.

Dependent Variable: DLOG(X)
 Method: Least Squares
 Date: 02/21/12 Time: 19:11
 Sample (adjusted): 1978S1 2002S1
 Included observations: 49 after adjustments
 DLOG(X)=C_X(4)*DLOG(WD)+C_X(6)+C_X(7)*RES_X(-1)+EC_X

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(4)	0.940211	0.115121	8.167185	0.0000
C_X(6)	-0.012945	0.006877	-1.882371	0.0661
C_X(7)	-0.195067	0.075051	-2.599137	0.0125
R-squared	0.607911	Mean dependent var		0.024624
Adjusted R-squared	0.590864	S.D. dependent var		0.025734
S.E. of regression	0.016460	Akaike info criterion		-5.316476
Sum squared resid	0.012463	Schwarz criterion		-5.200650
Log likelihood	133.2537	Hannan-Quinn criter.		-5.272532
F-statistic	35.66021	Durbin-Watson stat		1.662233
Prob(F-statistic)	0.000000			



6.2.5 LES IMPORTATIONS: ALLER PLUS LOIN DANS LA COINTÉGRATION ET LA STABILITE DE LONG TERME

Dans un modèle à un seul pays, le reste du monde est exogène, et les importations et les exportations doivent être estimées séparément, avec bien sûr les mêmes lignes directrices :

Les importations dépendront de la demande et de l'utilisation des capacités, avec des élasticités constantes.

Toutefois, la définition de la demande n'est pas si simple. Pour les exportations, nous avons eu simplement à tenir compte des importations globales de chaque pays partenaire dans chaque produit, et calculer une moyenne à l'aide d'une double pondération: le rôle de ces partenaires dans les exportations françaises, et la structure des produits exportés par la France.

Cela n'a été possible que parce que nous avons considéré le reste du monde comme exogène, et n'avons pas essayé de retracer l'origine de ses importations.

Dans notre cas, les importations peuvent provenir de quatre éléments endogènes:

- La demande locale finale, telle que les automobiles étrangères.
- Les biens intermédiaires nécessaires aux entreprises locales pour satisfaire cette demande locale. Pour les automobiles, ce seront les composants électroniques, l'acier, l'énergie utilisée pour faire fonctionner les machines ...
- De manière identique, les biens intermédiaires nécessaires à la fabrication des biens exportés.
- Mais pas des produits finis utilisés pour satisfaire la demande étrangère: la France ne réexporte pas les marchandises sans transformation significative (contrairement à Hong Kong par exemple).

Fondamentalement, deux méthodes peuvent être envisagées:

- À ce stade, notre modèle ne considère que la demande finale et les exportations. De toute évidence, elles n'ont pas le même impact sur les importations (en raison de l'absence de réexportations). Nous pouvons générer une demande totale synthétique en appliquant un facteur de correction aux exportations. Sous des hypothèses plutôt simples (même part des importations dans tous les usages, ratio unitaire de la consommation intermédiaire à la valeur ajoutée), ce facteur peut être fixé à 0,5.
- On peut aussi définir les consommations intermédiaires, et les ajouter à la demande finale pour obtenir la demande totale du pays, dont une partie sera importée. Cette méthode est évidemment plus acceptable du point de vue économique. Malheureusement, elle repose sur le calcul des consommations intermédiaires, une variable mesurée avec moins de précision, et sensible aux nomenclatures et au degré d'intégration verticale du processus de production⁵⁴.

Nous avons cependant choisi cette dernière méthode, favorisant les propriétés économiques par rapport à la fiabilité statistique.

Bien sûr, comme vous pouvez le deviner, nous allons tester la cointégration, comme nous l'avons fait pour les exportations. Voyons cependant ce qui arrive avec un processus autorégressif, juste pour le plaisir de constater sa dangerosité.

⁵⁴ Par exemple, si le bien A (disons le coton) est utilisé pour produire le bien B (les tissus non imprimés) qui donne un bien C (le tissu imprimé), A et B seront considérés comme consommations intermédiaires. Si le tissu est imprimé en même temps qu'il est produit, un seul sera pris en compte. Mais si l'on considère la valeur ajoutée, le montant total ne change pas, simplement le nombre d'éléments.

Notre formule fera dépendre les importations de la demande totale et du taux d'utilisation:

$$\Delta M / M = a \cdot \Delta TD / TD + b \cdot \Delta UR / UR$$

où

$$TD = FD + IC$$

$$IC = tc \cdot Q$$

et tc est le nombre d'unités de consommation intermédiaires nécessaires pour produire une unité de PIB.

Par intégration, nous obtenons

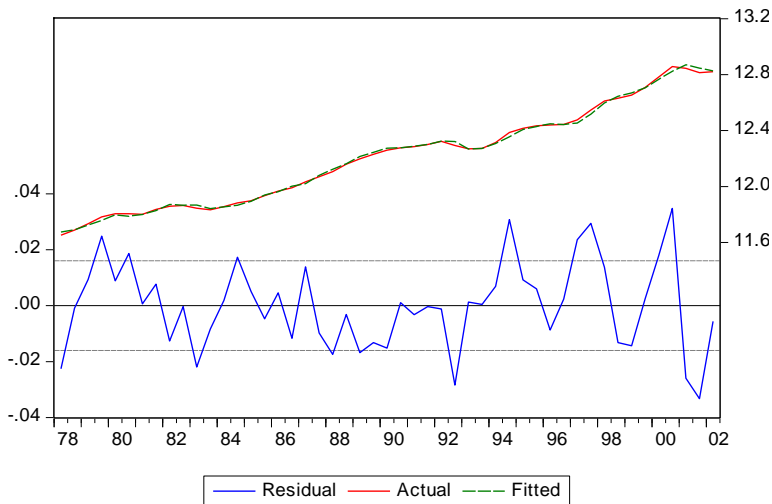
$$\text{Log}(M) = a \cdot \text{Log}(TD) + b \cdot \text{Log}(UR) + c$$

Comme on pouvait s'y attendre, la formule de base est confrontée à une autocorrélation élevée (ce qui rend non pertinente la question de la non significativité de UR):

Dependent Variable: LOG(M)
 Method: Least Squares
 Date: 11/07/12 Time: 13:20
 Sample (adjusted): 1977S2 2002S1
 Included observations: 50 after adjustments
 LOG(M)=C_M(1)*LOG(TD)+C_M(2)*LOG(UR)+C_M(3)

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(1)	2.395145	0.042125	56.85867	0.0000
C_M(2)	0.162941	0.145531	1.119634	0.2686
C_M(3)	-22.70890	0.614266	-36.96918	0.0000

R-squared	0.986068	Mean dependent var	12.19882
Adjusted R-squared	0.985475	S.D. dependent var	0.355330
S.E. of regression	0.042824	Akaike info criterion	-3.405322
Sum squared resid	0.086192	Schwarz criterion	-3.290601
Log likelihood	88.13305	Hannan-Quinn criter.	-3.361635
F-statistic	1663.282	Durbin-Watson stat	0.165492
Prob(F-statistic)	0.000000		

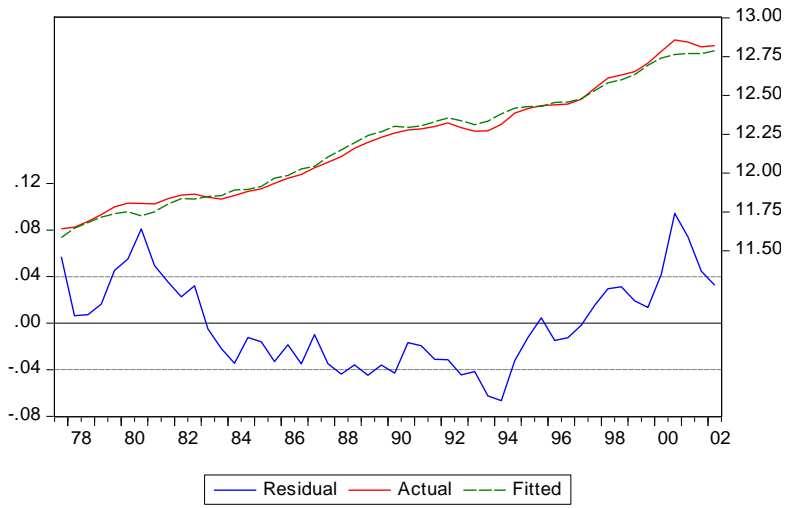


En plus des mauvais résultats statistiques, l'équation ci-dessus fait face à un problème économique: le coefficient de la demande est toujours très élevé. Ceci était prévisible, car dans la période d'estimation, les importations ont augmenté plus rapidement que la demande, et le coefficient représente la comparaison entre les croissances.

Maintenant la question suivante se pose: la croissance de la demande est-elle la seule explication de la croissance des importations ? En d'autres termes, n'existe-t-il pas une force autonome augmentant régulièrement le poids du commerce extérieur, indépendamment de la croissance elle-même ? Ou encore: si la demande n'a pas augmenté pendant une période donnée, les importations vont-elles rester stables, ou vont-elles conserver une partie de leur dynamique ?

La meilleure façon de répondre à cette question est sans doute d'appliquer l'économétrie, en introduisant une tendance temporelle additionnelle. Cette option obtient un succès partiel sur le plan économique, mais l'autocorrélation évidente interdit son utilisation:

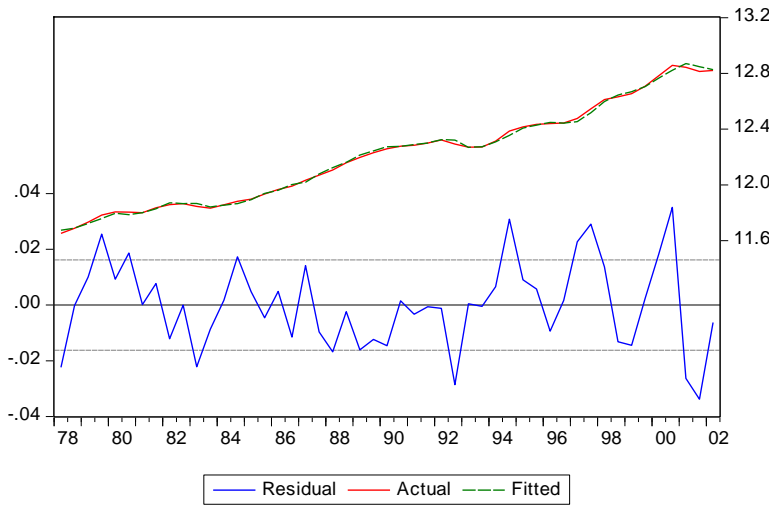
Dependent Variable: LOG(M)				
Method: Least Squares				
Date: 11/07/12 Time: 13:23				
Sample (adjusted): 1977S2 2002S1				
Included observations: 50 after adjustments				
LOG(M)=C_M(1)*LOG(TD)+C_M(2)*LOG(UR)+C_M(3)+C_M(4)*T				
	Coefficient	Std. Error	t-Statistic	Prob.
C_M(1)	1.472617	0.321303	4.583269	0.0000
C_M(2)	0.380206	0.154757	2.456791	0.0179
C_M(3)	-46.22968	8.150882	-5.671739	0.0000
C_M(4)	0.018580	0.006423	2.892783	0.0058
R-squared	0.988212	Mean dependent var		12.19882
Adjusted R-squared	0.987444	S.D. dependent var		0.355330
S.E. of regression	0.039816	Akaike info criterion		-3.532460
Sum squared resid	0.072926	Schwarz criterion		-3.379498
Log likelihood	92.31150	Hannan-Quinn criter.		-3.474211
F-statistic	1285.479	Durbin-Watson stat		0.273643
Prob(F-statistic)	0.000000			



Si l'on essaie d'éliminer l'autocorrélation, tout s'écroule:

Dependent Variable: LOG(M)
 Method: Least Squares
 Date: 11/07/12 Time: 13:25
 Sample (adjusted): 1978S1 2002S1
 Included observations: 49 after adjustments
 Convergence achieved after 19 iterations
 LOG(M)=C_M(1)*LOG(TD)+C_M(2)*LOG(UR)+C_M(3)+C_M(4)*T
 +[AR(1)=C_M(5)]

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(1)	2.557629	0.350384	7.299512	0.0000
C_M(2)	0.007373	0.076966	0.095791	0.9241
C_M(3)	-28.34794	12.80653	-2.213554	0.0321
C_M(4)	0.001626	0.008525	0.190675	0.8497
C_M(5)	0.903205	0.054688	16.51554	0.0000
R-squared	0.998033	Mean dependent var		12.21019
Adjusted R-squared	0.997854	S.D. dependent var		0.349705
S.E. of regression	0.016200	Akaike info criterion		-5.311125
Sum squared resid	0.011548	Schwarz criterion		-5.118083
Log likelihood	135.1226	Hannan-Quinn criter.		-5.237885
F-statistic	5580.647	Durbin-Watson stat		1.352968
Prob(F-statistic)	0.000000			
Inverted AR Roots	.90			



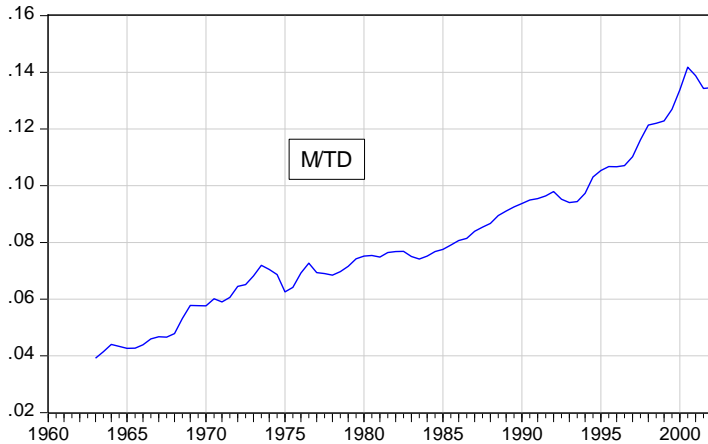
Mais si nous décidons néanmoins d'utiliser la formule, le modèle devra faire face à un problème important:

Prenons la dérivée de la formule sans tendance, corrigée (en partie) de l'autocorrélation:

$$\Delta(M)/M = 2.62 \Delta(TD)/TD$$

$$\Delta(M) = 2.62 (M/TD) \Delta(TD)$$

Étudions maintenant l'évolution de la part des importations dans la demande totale: en 2004, elle atteint 15 %:



En 2004 cela donne:

$$\begin{aligned}\Delta(M) &= 2.62 \times 0.15 \Delta(TD) = 0.39 \Delta(TD) \\ \Delta(M) &= 2.62 \times 0.15 \Delta(TD) = 0.39 ((FD) + \Delta(Q)) \\ \Delta(Q) &= \Delta(FD) - \Delta(M) = 0.61 \Delta(TD) - 0.39 \Delta(Q) \\ \Delta(Q) &= 0.61/1.39 \Delta(TD)\end{aligned}$$

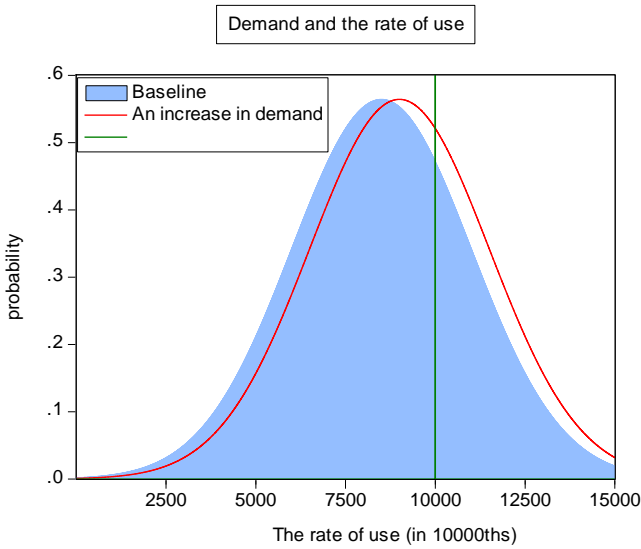
ou plus généralement en utilisant « a »:

$$\Delta(Q) = (1 - 2.62 a) / (1 + 2.62 a) \Delta(TD)$$

Cela signifie que si « a » ne cesse de croître (disons jusqu'à 0,4) le coefficient sera négatif, et il deviendra possible qu'une augmentation de la demande finale réduise le PIB, avec un multiplicateur négatif.

Le problème de notre formulation est effectivement très clair: en termes de propriétés du modèle, il est raisonnable de supposer que, dans le court terme, une augmentation de la demande finale augmente les importations au-delà de leur contribution normale, en générant des goulots d'étranglement locaux sur l'offre intérieure. Mais l'élément explicatif devrait être le taux d'utilisation des capacités. Comme il est fixé sur le long terme, la part reviendra à la normale avec le temps.

Un taux d'utilisation de 85 % (une valeur normale sur toute l'économie) ne signifie pas que toutes les entreprises travaillent à 85 % de leurs capacités, auquel cas elles peuvent facilement passer à 86 % en cas de besoin. Cela signifie que les taux d'utilisation suivent une distribution donnée, certains au moins de 85 %, certains plus, et un nombre fini à 100 % (voir graphique).



La courbe représente (en 10 000ièmes) la probabilité associée au rapport de la demande adressée aux entreprises de l'échantillon, rapportée à leur capacité. Ce rapport représente le taux d'utilisation des capacités, sauf évidemment si la demande est supérieure à l'offre, auquel cas il vaut 1.

La zone bleue représente la situation de base, la courbe rouge indique son déplacement suite à une augmentation de la demande. Davantage d'entreprises devront alors faire face à la limite: une augmentation de 1% des taux sera atteinte à mi-chemin par les entreprises partant d'un taux de 99,5 %. La demande supplémentaire, si les clients n'acceptent pas des substituts locaux, devra être satisfaite par des importations.

Toutefois, les entreprises locales vont réagir à cette situation, et essayer de regagner les parts de marché perdues en augmentant leurs capacités grâce à l'investissement: c'est le mécanisme que nous avons décrit

plus haut. Dans notre petit modèle, le taux à long terme de l'utilisation est fixé: le partage de la demande supplémentaire va revenir aux valeurs de base. Ces valeurs peuvent cependant augmenter avec le temps en raison de l'expansion du commerce mondial.

Pour représenter ce processus, nous avons besoin d'une formule qui:

- introduit une élasticité unitaire des importations à la variable de demande totale, avec un effet positif supplémentaire du taux d'utilisation ;
- laisse les élasticités libres dans le court terme.

Nous reconnaissons dans ce qui précède le même cadre que dans l'équation des exportations, traité par la cointégration.

Remarque: comme dans l'équation d'exportations, tenir compte de l'ensemble du modèle rendra l'élasticité à long terme des importations à la demande réellement unitaire, comme le taux d'utilisation est fixé par l'équation d'investissement. Tout choc sur un élément exogène (y compris la productivité du capital) va ramener UR à sa valeur de base, interdisant tout arbitrage. Ceci va changer lorsque l'on considèrera la rentabilité du capital, ce qui permettra à la cible de UR de varier.

Essayons d'estimer une telle équation.

Nous allons commencer par tester la cointégration entre la part des importations dans la demande: M/TD et le taux d'utilisation.

Avant, nous testerons la stationnarité de M/TD , ou plutôt son logarithme (UR a déjà été testé):

Elle est fortement contredite par le test de Dickey-Fuller:

Null Hypothesis: LOG(M/TD) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 2 (Automatic - based on SIC, maxlag=11)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.023358	0.5793
Test critical values:		
1% level	-4.081666	
5% level	-3.469235	
10% level	-3.161518	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(LOG(M/TD))
 Method: Least Squares
 Date: 11/07/12 Time: 13:29
 Sample (adjusted): 1964S2 2002S2
 Included observations: 77 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(M(-1)/TD(-1))	-0.097183	0.048031	-2.023358	0.0467
D(LOG(M(-1)/TD(-1)))	0.439724	0.104688	4.200318	0.0001
D(LOG(M(-2)/TD(-2)))	-0.337369	0.107119	-3.149483	0.0024
C	-0.294686	0.154078	-1.912578	0.0598
@TREND("1960S1")	0.001282	0.000669	1.915312	0.0594

R-squared	0.277500	Mean dependent var	0.014636
Adjusted R-squared	0.237361	S.D. dependent var	0.029908
S.E. of regression	0.026119	Akaike info criterion	-4.389613
Sum squared resid	0.049117	Schwarz criterion	-4.237418
Log likelihood	174.0001	Hannan-Quinn criter.	-4.328737
F-statistic	6.913494	Durbin-Watson stat	1.974691
Prob(F-statistic)	0.000091		

Maintenant, testons la cointégration de LOG(UR) et LOG(M / TD)

Elle échoue.

Date: 11/07/12 Time: 13:31
 Sample (adjusted): 1979S1 2002S2
 Included observations: 48 after adjustments
 Trend assumption: Linear deterministic trend (restricted)
 Series: LOG(MTD) LOG(UR)
 Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None	0.154667	12.10335	25.87211	0.8048
At most 1	0.080687	4.038169	12.51798	0.7366

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	0.154667	8.065182	19.38704	0.8156
At most 1	0.080687	4.038169	12.51798	0.7366

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(MTD)	LOG(UR)	@TREND(60S2)
-5.221110	36.73565	0.057533
-26.68206	16.01556	0.406982

Unrestricted Adjustment Coefficients (alpha):		
D(LOG(M/TD))	0.000665	0.004179
D(LOG(UR))	-0.010282	0.003897
1 Cointegrating Equation(s): Log likelihood 242.4388		
Normalized cointegrating coefficients (standard error in parentheses)		
LOG(M/TD)	LOG(UR)	@TREND(60S2)
1.000000	-7.035985 (2.33234)	-0.011019 (0.00511)
Adjustment coefficients (standard error in parentheses)		
D(LOG(M/TD))	-0.003472 (0.01192)	
D(LOG(UR))	0.053682 (0.02230)	

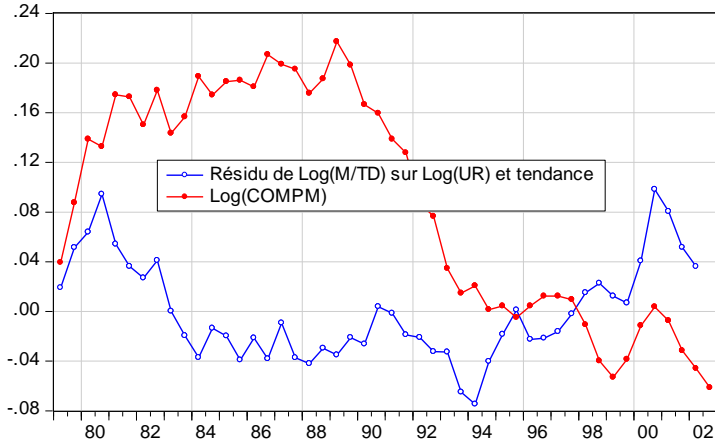
Ne nous désespérons pas. L'absence de cointégration n'est pas une si mauvaise nouvelle⁵⁵, car elle nous permet d'aller plus loin. Si un ensemble de deux variables ne fonctionne pas, pourquoi pas un ensemble de trois ?

Maintenant quel élément supplémentaire pouvons-nous considérer ? Le candidat naturel provient à la fois de la théorie et des données.

Si la demande est présente, mais les producteurs locaux n'ont pas de problèmes de capacité, comment les exportateurs étrangers peuvent-ils pénétrer un marché ? Bien sûr, grâce à la compétitivité de leurs prix, en d'autres termes en diminuant le prix des importations par rapport à celui des producteurs locaux.

Régressons le ratio des importations à la demande sur le taux d'utilisation, considérons le résidu (la partie inexplicquée) et comparons-le au rapport des prix à l'importation et locaux (que nous appellerons COMPM).

⁵⁵ L'identification des deux équations serait encore pire, surtout dans un cadre de modélisation.



Nous observons une corrélation évidente, surtout si l'on considère les variations autour d'une tendance négative.

Après avoir testé:

- La non-stationnarité de Log (COMPM)
- La non-cointégration de Log (COMPM) à la fois avec Log (UR) et log (M / TD) individuellement,

Nous pouvons tester la cointégration des trois éléments:

Elle fonctionne !

Date: 11/07/12 Time: 13:37
 Sample (adjusted): 1979S1 2002S2
 Included observations: 48 after adjustments
 Trend assumption: Linear deterministic trend (restricted)
 Series: LOG(M/TD) LOG(COMPM) LOG(UR)
 Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.430972	46.69131	42.91525	0.0200
At most 1	0.282157	19.62767	25.87211	0.2454
At most 2	0.074486	3.715476	12.51798	0.7828

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 *	0.430972	27.06363	25.82321	0.0342
At most 1	0.282157	15.91220	19.38704	0.1490
At most 2	0.074486	3.715476	12.51798	0.7828

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/TD)	LOG(COMPM)	LOG(UR)	@TREND(60S2)
-26.25607	-17.29497	41.40372	0.295194
18.37876	-2.562664	-13.07275	-0.337567
15.08409	9.794651	15.94489	-0.167276

Unrestricted Adjustment Coefficients (alpha):			
D(LOG(M/TD))	0.004500	-0.003607	-0.002787
D(LOG(COM...)	0.005905	0.008244	-0.002935
D(LOG(UR))	-0.010552	-0.001976	-0.006576
<hr/>			
1 Cointegrating Equation(s):	Log likelihood	374.4392	
<hr/>			
Normalized cointegrating coefficients (standard error in parentheses)			
LOG(M/TD)	LOG(COMPM)	LOG(UR)	@TREND(60S2)
1.000000	0.658704	-1.576920	-0.011243
	(0.12206)	(0.23954)	(0.00078)
Adjustment coefficients (standard error in parentheses)			
D(LOG(M/TD))	-0.118147		
	(0.05526)		
D(LOG(COM...)	-0.155046		
	(0.08330)		
D(LOG(UR))	0.277042		
	(0.11304)		
<hr/>			
2 Cointegrating Equation(s):	Log likelihood	382.3953	
<hr/>			
Normalized cointegrating coefficients (standard error in parentheses)			
LOG(M/TD)	LOG(COMPM)	LOG(UR)	@TREND(60S2)
1.000000	0.000000	-0.862522	-0.017123
		(0.39271)	(0.00092)
0.000000	1.000000	-1.084551	0.008926
		(0.66393)	(0.00156)
Adjustment coefficients (standard error in parentheses)			
D(LOG(M/TD))	-0.184434		-0.068581
	(0.06492)		(0.03542)
D(LOG(COM...)	-0.003529		-0.123256
	(0.09270)		(0.05057)

D(LOG(UR))	0.240731	0.187552
	(0.13762)	(0.07508)

Bien sûr, nous devons considérer les coefficients de l'équation. Ils décrivent:

- Une sensibilité apparemment élevée des importations au taux d'utilisation (mais rappelez-vous que l'équation d'investissement le stabilisera à la fin).

Ceci dit, le véritable effet n'est pas si élevé. Si l'équation a été appliquée à court terme, avec le même coefficient de 0,15 que précédemment, nous obtenons:

$$\Delta(M) = 1 \times 0.15 \Delta(TD) + 1.57 \times 0.15 \Delta(Q)$$

$$\Delta(M) = 0.15 \Delta(TD) + 0.23 \Delta(Q)$$

et si nous ne considérons pas l'évolution des exportations:

$$\Delta(FD - Q) = 0.15 \Delta(FD + ct. Q) + 0.23 \Delta(Q)$$

$$\Delta(Q) = 0.85 / 1.38 \Delta(FD) = 0.62 \Delta(FD)$$

- Un multiplicateur tout à fait acceptable pour la France.
- Une sensibilité beaucoup plus élevée que dans la formule autorégressive.

Nous pouvons maintenant tester le VAR:

Dependent Variable: DLOG(M)

Method: Least Squares

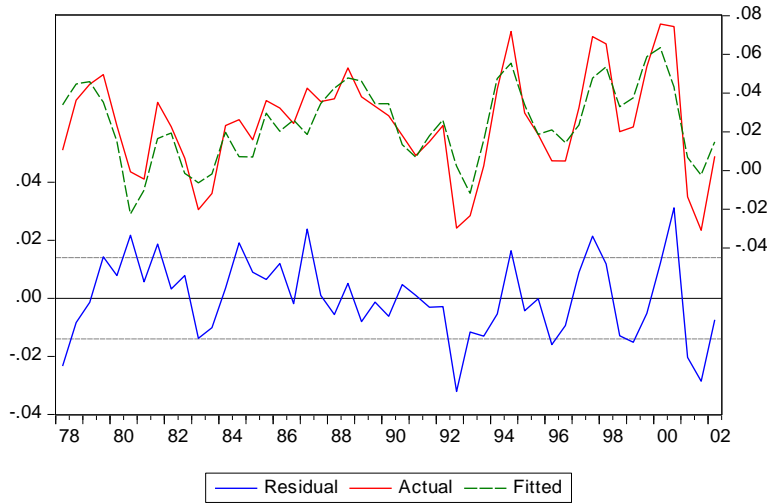
Date: 11/07/12 Time: 14:03

Sample (adjusted): 1978S1 2002S2

Included observations: 50 after adjustments

$$\text{DLOG(M)} = \text{C_M(1)} * \text{DLOG(TD)} + \text{C_M(2)} * \text{DLOG(UR)} + \text{C_M(3)} * \text{RES_M(-1)} + \text{C_M(4)}$$

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(1)	1.635016	0.411194	3.976267	0.0002
C_M(2)	0.170456	0.083792	2.034268	0.0477
C_M(3)	-0.152136	0.044768	-3.398328	0.0014
C_M(4)	-0.452681	0.132508	-3.416262	0.0013
R-squared	0.684182	Mean dependent var		0.023935
Adjusted R-squared	0.663585	S.D. dependent var		0.025806
S.E. of regression	0.014968	Akaike info criterion		-5.489229
Sum squared resid	0.010305	Schwarz criterion		-5.336267
Log likelihood	141.2307	Hannan-Quinn criter.		-5.430980
F-statistic	33.21779	Durbin-Watson stat		1.408067
Prob(F-statistic)	0.000000			



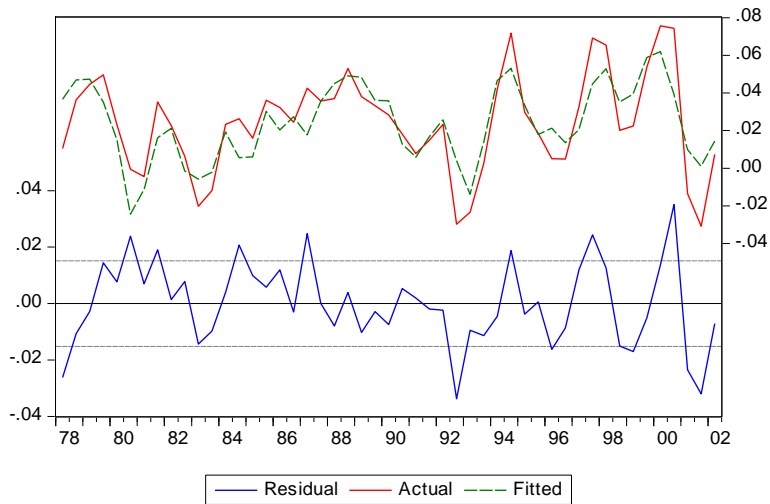
Les résultats sont assez acceptables, même si le coefficient de UR est à peine significatif⁵⁶. Mais le graphique est plutôt favorable. Par ailleurs, il montre que le rejet d'un R^2 de 0,69 n'est pas justifié lorsque l'élément expliqué montre une grande variabilité, et aucune tendance.

Forcer le coefficient de la demande à une valeur plus faible (par exemple 1,2) augmente le coefficient du taux d'utilisation, sans réduire trop la qualité. La fiabilité de UR augmente, en raison principalement de la hausse du coefficient (bien sûr UR et TD sont corrélés car ils dépendent tous deux de Q) et un peu de celle de l'écart-type.

⁵⁶ Nous voyons que la statistique de Student est supérieure à 2, mais la probabilité supérieure à 5 % !

Dependent Variable: DLOG(M)
 Method: Least Squares
 Date: 11/07/12 Time: 14:04
 Sample (adjusted): 1978S1 2002S2
 Included observations: 50 after adjustments
 DLOG(M)=1.2*DLOG(TD)+C_M(2)*DLOG(UR)+C_M(3)*RES_M(-1)
 +C_M(4)

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(2)	0.230294	0.061900	3.720416	0.0005
C_M(3)	-0.183253	0.033793	-5.422878	0.0000
C_M(4)	-0.542178	0.102118	-5.309318	0.0000
R-squared	0.676498	Mean dependent var		0.023935
Adjusted R-squared	0.662731	S.D. dependent var		0.025806
S.E. of regression	0.014987	Akaike info criterion		-5.505189
Sum squared resid	0.010556	Schwarz criterion		-5.390468
Log likelihood	140.6297	Hannan-Quinn criter.		-5.461503
F-statistic	49.14241	Durbin-Watson stat		1.380208
Prob(F-statistic)	0.000000			



6.2.6 RETOUR A LA VERIFICATION RESIDUELLE

La méthode que nous utilisons pour sauvegarder les équations présente un avantage supplémentaire. Maintenant que les résidus ont été introduits avec leurs valeurs estimées, toutes les équations doivent tomber juste. Le processus de vérification peut être étendu à toutes les variables endogènes.

En théorie, les équations estimées (et fusionnées avec le modèle) devraient être compatibles avec les données, Toutefois:

- si le logiciel ne fournit pas ce stockage direct, ou si l'équation a dû être normalisée par le modélisateur, l'édition de la formulation a pu introduire des erreurs ;
- le stockage des coefficients a peut-être été mal fait ;
- le texte, les séries ou les coefficients peuvent avoir été modifiés par l'utilisateur après l'estimation ;
- l'utilisateur a pu accéder à d'autres séries ou coefficients que ceux utilisés par l'estimation (par exemple les chercher dans une autre banque comprenant une série de noms similaires).

Les raisons pour lesquelles une équation de comportement présente une valeur résiduelle non nulle sont moins nombreuses que pour les identités. Elles ne peuvent venir que du fait que les éléments de l'équation ont changé depuis la dernière estimation.

De toute évidence, le principal suspect est les données. De nouvelles séries de données peuvent être incompatibles avec la précédente estimation, si elles ont été mises à jour (en établissant une version plus précise) ou corrigées (une erreur a été repérée).

En fait, dans EViews, l'application d'une nouvelle version d'une équation d'un modèle nécessite, en plus de son estimation, de la fusionner à nouveau au modèle. Cela va créer une nouvelle version compilée, sans qu'il soit besoin de mettre à jour explicitement celui-ci.

Quoi qu'il en soit, à notre avis, l'application d'une nouvelle estimation devrait conduire à une nouvelle création du modèle complet. C'est la seule façon de garantir une mise à jour claire et sûre.

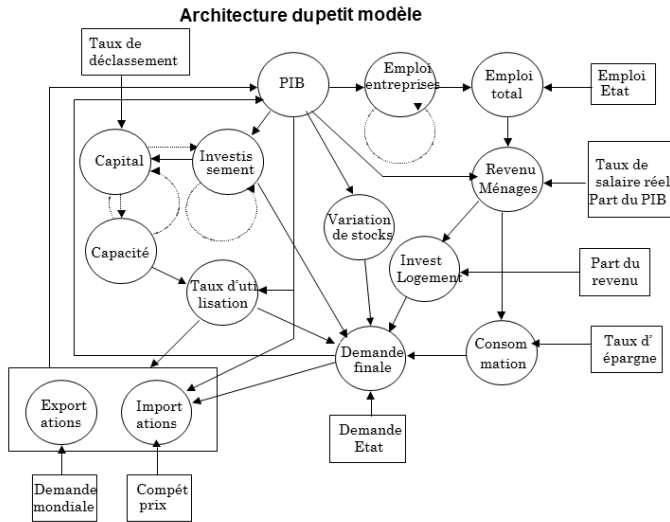
Dans notre modèle, les instructions pour la vérification résiduelle seront les suivantes:

```
' Nous vérifions les résidus
smpl 1980S2 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
```

La série solution prendra le suffixe «_c», et les résidus le préfixe «dc_» pour les erreurs dans les niveaux, et «pc_» pour les erreurs relatives.

Nous pouvons maintenant présenter le cadre du modèle.

6.2.7 LE MODÈLE ACTUEL



En notations EViews, ses spécifications sont:

- [1] $CAP = pk * K(-1)$
- [2] $Q + M = FD + X$
- [3] $UR = Q / CAP$
- [4] $CI = tc * Q$
- [5] $IC/Q(-1) = 0.130 * PCH(Q) + 0.214 * PCH(Q(-1)) + EC_IC$
- [6] $IP/K(-1) = 0.825 * I(-1)/K(-2) + 0.0279 * UR + 0.152 * .25 * Q/Q(-4) - 0.0525 + EC_I$
- [7] $\log(PRLE_T) = c_prle(1) + c_prle(2) * (t - 2002) + c_prle(3) * (t - t1) * (t < t1) + c_prle(4) * (t - t2) * (t < t2)$
- [8] $LED = Q / PRLE_T$

[9] $DLOG(LE) = 0.587 * DLOG(LED) + 0.411 * LOG(LED(-1) / LE(-1)) + 0.000502 - 0.0167 * ((T = 1968.5) - (T = 1968)) + EC_LE$

[10] $LT = LE + Iq$

[11] $RHI = wr * LT + r_rhiq * Q$

[12] $IH = r_ih * RHI$

[13] $CO = RHI * (1 - sr)$

[14] $FD = CO + IP + CI + IH + gd$

[15] $TD = FD + tc * Q$

[16] $RES_M = \log(M / TD) - 1.322108 * \log(UR) + 0.419518 * \log(compm) - 0.012582 * (@trend(60:1) * (t <= 2002) + @elem(@trend(60:1), "2002S2") * (t > 2002))$

[17] $DLOG(M) = 1.2 * DLOG(TD) + 0.282 * DLOG(UR) - 0.212 * RES_M(-1) - 0.629 + EC_M$

[18] $RES_X = \log(X / wd) + 0.686 * \log(UR) - 4.87E-05 * (@trend(60:1) * (t <= 2002) + @elem(@trend(60:1), "2002S2") * (t > 2002))$

[19] $DLOG(X) = 0.940 * DLOG(wd) - 0.0129 - 0.195 * RES_X(-1) + EC_X$

[20] $K = K(-1) * (1 - dr) + IP$

6.2.8 ASPECTS LOGICIELS

Nous allons considérer le problème en général, tout en observant si EViews fournit les caractéristiques et les qualités que nous attendons (et il le fait généralement).

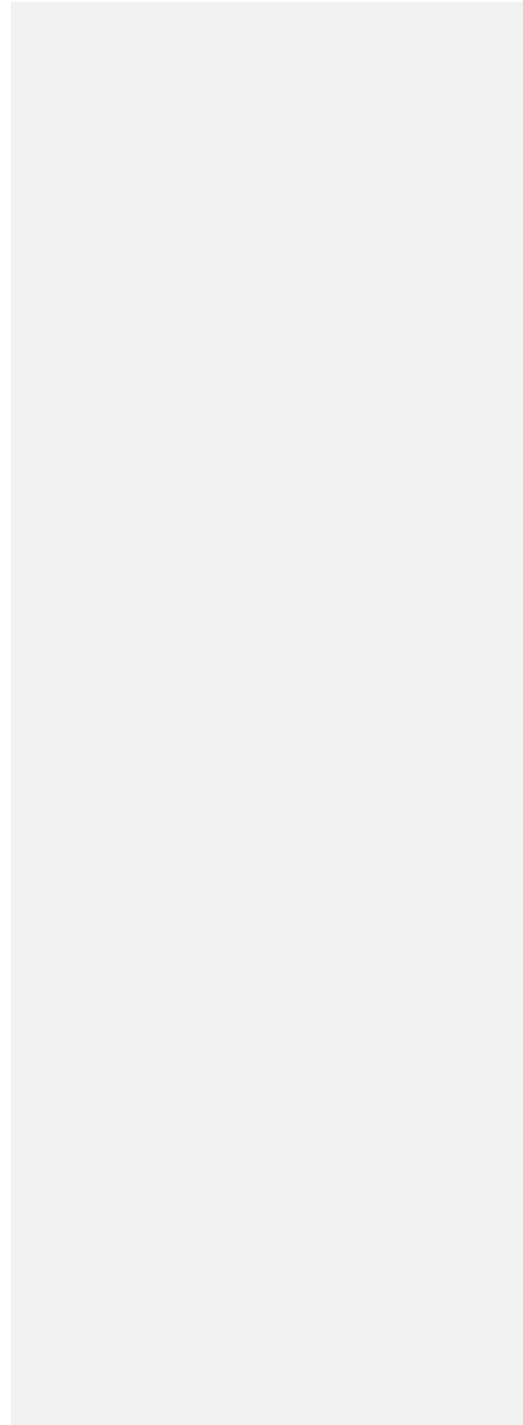
Le choix du logiciel d'estimation, ou d'un logiciel de modélisation comprenant des fonctions d'estimation, doit tenir compte des critères suivants:

- le champ d'application des méthodes d'estimation proposées ;
- l'interactivité entre la spécification des équations et le processus d'estimation permettant:
 - de définir l'équation de comportement (y compris les coefficients définis en tant que paramètres scalaires), sous un nom d'équation, dans un format spécifique. Cela peut être fait à la suite du processus d'estimation,
 - de stocker les coefficients estimés,
 - de stocker les résidus et les valeurs estimées.

Ceci assurera (ou du moins, permettra de vérifier facilement) la compatibilité permanente entre les formules estimées et le modèle que nous allons simuler à l'étape suivante ;

- un accès implicite (les séries seront chargées automatiquement dès qu'elles seront réclamées par une formule) à une ou plusieurs banques, dans un ordre choisi ;
- la fourniture de statistiques adéquates, de graphiques clairs et faciles à produire ;
- un langage complet de spécification des équations, comprenant tous les opérateurs logiques et mathématiques nécessaires ;
- une fonction d'édition conviviale des programmes (plein écran) ;
- la normalisation automatique des équations (cf. ci-dessus) ;
- des algorithmes d'estimation rapides et, qualité nécessaire pour les méthodes les plus complexes, avec une bonne probabilité de convergence ;
- la fourniture de diagnostics explicites et complets en cas de problèmes ;
- des méthodes permettant de résoudre ces problèmes.

En utilisant un logiciel ayant les qualités requises, on sera en mesure de spécifier un modèle complet, syntaxiquement correct, économiquement cohérent avec les équations estimées. Pour contrôler complètement cette cohérence, nous devons appliquer une vérification résiduelle.



CHAPITRE 7: LE TEST DU MODÈLE PAR DES SIMULATIONS SUR LE PASSE

Nous avons maintenant achevé la production d'un modèle, pour lequel:

- toutes les identités sont compatibles avec les données ;
- toutes les estimations sont à notre avis acceptables d'un point de vue statistique ;
- toutes les valeurs des coefficients et les spécifications des équations semblent compatibles avec la théorie économique⁵⁷ ;
- tous les liens nécessaires entre éléments du modèle sont présents ;
- les équations estimées ont été construites de telle manière qu'elles devraient fournir une solution à long terme pour le modèle complet.

Cela ne veut pas dire que notre modèle est acceptable:

- la cohérence des données peut cacher des erreurs se compensant mutuellement ;
- certaines équations complexes peuvent avoir caché de mauvaises propriétés individuelles ;
- des liens entre concepts auraient pu être oubliés ou mal spécifiés ;
- les taux de croissance prévus naturellement par des équations pourraient être erronés ;
- certains processus de correction d'erreur pourraient être divergents ;
- mettre sur pied des formulations acceptables individuellement peut créer un système aux propriétés inacceptables.

Donnons un exemple: une augmentation de 1 000 de la demande du gouvernement crée à court terme 800 de consommation et 600 d'investissement, tandis que les exportations ne changent pas et les importations augmentent de 300. Sur le plan individuel les équations semblent acceptables, mais le modèle divergera immédiatement suite à un effet multiplicateur explosif ($800 + 600 - 300 = 1\ 100$).

Notre prochain objectif sera de contrôler:

- que le modèle peut en effet être résolu,
- qu'il peut être utilisé pour les prévisions,
- qu'il peut être utilisé pour l'analyse des politiques économiques.

⁵⁷ Dans certains cas, nous aurions été obligés de calibrer les valeurs.

En fait, nos premiers tests vont plutôt répondre à la question inverse:

- Y a-t-il un indice que le modèle ne convienne pas pour les prévisions et l'analyse des politiques économiques ?

À notre avis, ce n'est que par des simulations sur l'avenir (son domaine de fonctionnement normal, en fait) que nous pouvons vraiment valider l'utilisation d'un modèle. Mais comme d'habitude, les problèmes doivent être diagnostiqués le plus tôt possible. Et la disponibilité des données observées favorise fortement certains tests.

Les erreurs détectées à ce stade pourront aider à construire une meilleure structure de prévisions.

Intéressons-nous d'abord au processus de résolution du modèle.

Pour résoudre le modèle, nous avons besoin d'appliquer une méthode. Présentons maintenant les différents algorithmes.

7.1 LA RESOLUTION

7.1.1 GAUSS-SEIDEL

C'est l'algorithme le plus naturel: on utilise souvent Gauss-Seidel, sans le savoir, comme Monsieur Jourdain (le Bourgeois Gentilhomme) fait de la prose.

Le processus part de valeurs initiales. Elles peuvent être les valeurs historiques sur le passé, sur l'avenir les valeurs calculées pour la période précédente ou par une simulation de base. L'ensemble des équations est appliqué dans un ordre donné, en utilisant les informations les plus récentes (les calculs remplaçant les valeurs initiales). Ceci donne un nouvel ensemble de valeurs de départ. Le processus est répété, en utilisant toujours la dernière information disponible, jusqu'à ce que la distance entre les deux dernières solutions soit suffisamment faible pour être considérée comme négligeable. On considère alors que la solution a été trouvée.

Formalisons ce processus.

Considérons le modèle

$$y_t = f_i(y_t, y_{t-1}, x_t, \hat{\alpha})$$

avec y_t vecteur de $y_{i,t}$ et $i \in [1, n]$

Comme seules les valeurs actuelles vont changer au cours du calcul, nous ne considérerons pas les autres éléments, et supprimerons l'indice de temps.

$$y = f(y)$$

Nous allons utiliser cet indice pour définir une endogène particulière, tandis que l'exposant définira le numéro de l'itération.

a - Nous partons de y^0 , valeur à l'itération 0.

b - Nous ajoutons 1 au nombre d'itérations (que nous noterons k), ce qui donne à la première itération le nombre 1.

c - Nous calculons les équations de $i = 1$ à n , en prenant en compte les valeurs de 1 à $i-1$ que nous venons de produire. Cela signifie que nous calculons:

$$y_i^k = f(y_1^k, \dots, y_{i-1}^k, y_i^{k-1}, \dots, y_n^{k-1})$$

(à la première itération, les éléments explicatifs prendront la valeur initiale y^0 si leur indice est plus élevé que celui de la variable calculée)⁵⁸.

⁵⁸ Ceci signifie que les valeurs des variables qui sont utilisées avant qu'elles ne soient calculées sont nécessaires pour l'initialisation. Nous y reviendrons plus tard.

d - À la fin du processus, nous comparons y^k et y^{k-1} . Si la distance est assez faible pour chaque élément (nous verrons plus loin quels sont les critères utilisés) nous arrêterons le processus, et prendrons comme solution la dernière valeur. Sinon, nous vérifierons si nous avons atteint le nombre maximum d'itérations, auquel cas nous acceptons l'échec de l'algorithme, et nous nous arrêterons. Sinon, nous reprendrons le processus à l'étape b.

De toute évidence, cet algorithme exige un modèle identifié.

7.1.2 RITZ-JORDAN

La méthode de Ritz-Jordan est similaire à la précédente: elle s'abstient tout simplement d'utiliser les valeurs calculées à l'itération courante:

$$y^k = f(y^{k-1})$$

Refusant de prendre en compte les dernières informations, il semble moins efficace que Gauss-Seidel. À notre avis, son seul intérêt apparaît lorsque le modèle est confronté à des problèmes de convergence: elle rend leur interprétation plus facile en réduisant les interférences entre les variables.

Cette méthode n'est pas proposée par EViews.

7.1.3 NEWTON ET SES VARIANTES

Contrairement aux deux précédentes, la méthode de Newton s'applique naturellement aux formulations non identifiées. Elle représente en fait la généralisation à un problème de dimension n de la méthode bien connue, qui utilise une séquence de linéarisations pour résoudre une équation unique.

Considérons le modèle:

$$f_t(y_t, y_{t-1}, x_t, \hat{\alpha}) = 0$$

que nous allons simplifier comme ci-dessus en:

$$f(y) = 0$$

La linéarisation de f autour d'une solution de départ donnée, en appelant « fl » la valeur de f linéarisée:

$$(\partial f / \partial y)_{y=y^0} \cdot (y - y^0) = fl(y) - f(y^0)$$

Résoudre le système pour $fl(y) = 0$ conduit à:

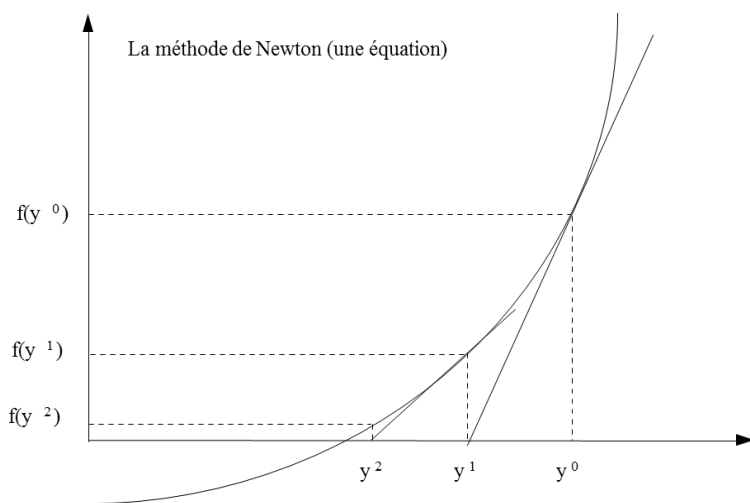
$$y = y^0 - (\partial f / \partial y)_{y=y^0}^{-1} \cdot f(y^0)$$

Et avec un système identifié:

$$y - f(y) = 0$$

nous aurions naturellement:

$$y = y^0 - (I - \partial f / \partial y)_{y=y^0}^{-1} \cdot (y^0 - f(y^0))$$



En linéarisant le modèle à nouveau, autour de la nouvelle solution y^1 , et en résolvant le nouveau modèle linéarisé, nous définissons un processus itératif qui, comme le précédent, s'arrête lorsque la distance entre les deux dernières valeurs est assez faible. L'application de cette méthode est plus complexe: en plus d'inverser une matrice, chaque itération implique le calcul d'un Jacobien. Ceci peut être fait en pratique de deux manières:

- analytiquement, en déterminant dès le début les expressions formelles de dérivées. À chaque itération, nous les recalculerons à partir des valeurs courantes des variables. Cette méthode réclame soit d'entreprendre le calcul «à la main» avec une forte probabilité d'erreurs, ou d'avoir accès à un dérivateur formel automatique, un programme analysant le texte des équations pour

produire l'expression formelle de leurs dérivées. À un coût initial élevé, cette méthode oppose une simplification des calculs au cours du processus itératif⁵⁹.

À partir de la version 6, EViews permet d'effectuer la dérivation analytique⁶⁰.

- par différences finies, en déterminant séparément chaque colonne de la matrice jacobienne par le calcul numérique d'un développement limité du premier ordre, appliqué successivement à chaque variable. On calcule le vecteur y en utilisant les valeurs de base, puis à partir de valeurs qui n'en diffèrent que par une seule variable, et on compare les deux résultats pour obtenir une colonne de la matrice jacobienne. On parlera alors d'une méthode de sécantes, ou pseudo-Newton.

La formulation dérivée devient alors:

$$\sum_{j=1}^n \left\{ \left[f_i(y^k + e_j \Delta y_j) - f_i(y^k) \right] / \Delta y_j \right\} (y_j - y_j^k) = f_i(y) - f_i(y^k)$$

où e_j est un vecteur de dimension n (nombre d'endogènes), avec 1 en position i et 0 sinon.

En d'autres termes, les éléments de la matrice jacobienne:

$$\partial f_i / \partial y_j$$

seront approximés par

⁵⁹ Toutefois, la modification de certaines caractéristiques du modèle appelle à une nouvelle dérivation globale (ou une dangereuse mise à jour manuelle).

⁶⁰ Malheureusement, le code associé n'est apparemment pas disponible pour l'utilisateur. Il permettrait des calculs intéressants.

$$\left[f_i(y^k + e_j \Delta y_j) - f(y^k) \right] / \Delta y_j$$

On aura seulement à calculer le vecteur y_{n+1} fois: une fois sans aucune modification et une fois pour chacune des variables endogènes .

La partie coûteuse de cet algorithme étant clairement le calcul du Jacobien et son inversion, une variante consistera à ne réaliser le calcul que toutes les m itérations. La convergence sera plus lente en termes de nombre d'itérations, mais le coût global pourrait diminuer.

7.1.3.1 La méthode de Broyden

Depuis la version 6, EViews offre une autre alternative: la méthode de Broyden, une méthode de sécante qui ne nécessite de calculer le Jacobien qu'à la première étape. Il est ensuite simplement réactualisé en tenant compte de la direction menant de la nouvelle solution à la précédente.

$$J_n = J_{n-1} + \frac{\Delta \vec{F}_n - J_{n-1} \Delta \vec{x}_n}{\|\Delta \vec{x}_n\|^2} \Delta \vec{x}_n^T$$

Pour une explication complète, voir: http://en.wikipedia.org/wiki/Broyden's_method.

Comme nous le verrons plus tard, cette méthode se révèle souvent très efficace.

7.1.3.2 Le cas d'une forme identifiée

Si le modèle est «identifié» sous la forme:

$$y = f(y)$$

l'algorithme de Newton est appliqué à

$$y - f(y) = 0$$

et la formule de Newton devient

$$y = y^0 - (I - \partial f / \partial y)_{y=y^0}^{-1} \cdot (y^0 - f(y^0))$$

ou

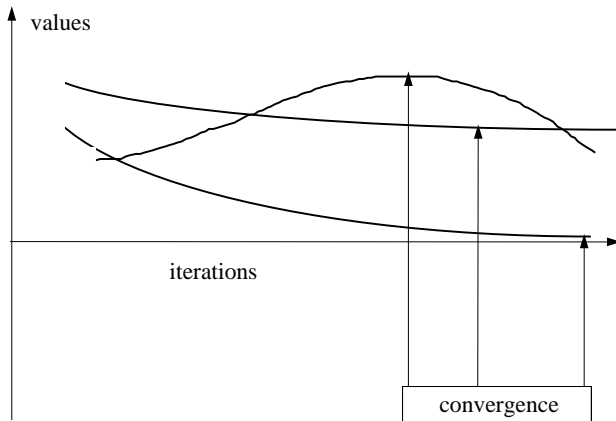
$$y = (I - \partial f / \partial y)_{y=y^0}^{-1} \cdot (f(y^0) - (\partial f / \partial y)_{y=y^0} \cdot y^0)$$

Cela ne change pas le processus technique.

7.1.4 LES ITÉRATIONS ET LE TEST DE CONVERGENCE

Les méthodes décrites ci-dessus ont une caractéristique commune: à partir de valeurs initiales, elles appliquent des formulations pour obtenir un nouvel ensemble. Le processus est répété jusqu'à ce que les deux derniers ensembles soient suffisamment proches pour être considérés comme la solution du système.

On ne peut pas considérer la différence entre deux itérations comme la précision effectivement atteinte (ou la différence à la solution). Ce n'est valable que pour les processus alternés. Pour les monotones, cela peut être l'inverse: plus la convergence est lente, plus la modification du critère d'une itération à l'autre sera faible, et plus la probabilité sera élevée qu'elle sera respectée assez loin de la solution. Quant aux processus cycliques, ils peuvent atteindre la convergence tout en haut ou en bas d'un cycle.



On pourrait donc critiquer ce type de méthode, en soulignant que la stabilité relative des valeurs ne signifie pas que la solution a été trouvée. Cependant, on peut observer que si les valeurs ne changent pas, cela signifie que le calcul qui a donné une variable donnerait le même résultat avec les nouvelles valeurs de ses variables explicatives: cela signifie aussi que l'équation tombe presque juste.

Cependant, il est clair que nous n'avons pas la solution exacte. Cette critique ne doit pas être trop mise en exergue: la précision des modèles est de toute façon limitée, et même les algorithmes soi-disant exacts sont limités par la précision des ordinateurs.

7.1.4.1 Les options générales

Pour savoir à quel moment arrêter les calculs de l'algorithme, nous devons mettre en place un test.

En fait, le seul critère utilisé dans la pratique étudiera la variation de l'ensemble des variables de la solution, d'une itération à l'autre.

Elle peut être mesurée, pour chaque variable:

- en valeurs relatives:

$$d_i = \left| \frac{y_i^k - y_i^{k-1}}{y_i^{k-1}} \right|$$

- ou en niveau:

$$d_i = |y_i^k - y_i^{k-1}|$$

Quant à la condition de l'acceptation de convergence, elle peut être définie:

par variable: $d_i < c_i, \forall i$

pour l'ensemble: $d_i < c_i, \forall i$

ou par une mesure globale: $f(d) < c$

En général, on choisira un critère en valeur relative, chaque erreur étant comparée à un critère global. Cette valeur devra être petite par rapport à la précision attendue du modèle (de sorte que l'erreur de simulation ne puisse pas vraiment contribuer à l'erreur globale), et au nombre de chiffres utilisés pour l'interprétation des résultats.

L'exception la plus fréquente correspond à des variables qui, comme la balance commerciale, fluctuent fortement et peuvent même changer de signe: le choix d'un critère en niveau semble une solution naturelle, ce qui permettra d'éviter un diagnostic de non-convergence dû à des fluctuations négligeables d'une variable si sa solution est (par pur hasard) très faible.

Par exemple, si le seuil de convergence est de 0.0001 en valeur relative, la convergence sera refusée si des solutions pour la balance commerciale américaine alternent par hasard entre - 1 milliard de dollars

US courants et – 1,0002⁶¹, tandis qu'un écart de 200 000 dollars, à l'échelle du commerce extérieur américain, est évidemment très faible. Et cette différence, ce qui représente moins d'un millionième des exportations et des importations des États-Unis, pourra ne jamais être éliminée si la précision informatique ne garantit que 8 chiffres significatifs⁶².

Dans la pratique, nous verrons que le test pourrait être limité à un sous-ensemble de variables du modèle, dont la convergence mathématiques peut s'étendre à la convergence globale.

La valeur donnée au critère peut dépendre:

- De l'algorithme utilisé:
 - pour Gauss-Seidel, chaque chiffre additionnel présente à peu près le même coût ;
 - pour Newton, le nombre de chiffres gagnés augmente avec les itérations: au-delà du niveau minimum (disons 0,01 %) un gain donné est de moins en moins cher (ceci sera développé plus loin). On peut donc être plus exigeant à moindre coût.

Sur le type de simulation:

- pour une prévision, on ne sera pas trop rigoureux, car comme nous le savons tous, la précision est assez faible de toute façon. Une prévision de croissance de 2,05 % et 2,07 % d'ici trois ans délivre le même message, en particulier si la croissance réelle risque d'être de 1 % (nous donnerons plus d'éléments sur la précision des prévisions plus tard) ;
- pour une analyse des chocs, surtout si le choc est faible, l'évaluation de la différence entre les deux simulations est évidemment plus affectée par l'erreur: selon que des décisions données augmentent le PIB de 0,07 % et de 0,09 %, il ne leur sera pas attribué la même efficacité.
- Et peut-être du caractère stochastique:

Dans une simulation stochastique, il est essentiel que la conséquence sur la solution de l'introduction de petits résidus aléatoires soit précisément associée au choc, et non au processus de simulation.

En ce qui concerne le nombre d'itérations, il sera utilisé comme une limite, au-delà de laquelle nous supposons que le modèle n'a aucune chance de converger. Dans la pratique on n'envisage jamais d'arrêter un processus apparemment convergent, simplement parce qu'il a pris trop de temps. Donc le seul cas

⁶¹ Il n'y a aucun risque à la période actuelle.

⁶² Exportations et importations sera précises jusqu' à la 8^e décimale, mais la différence, un million de fois plus petite, à la 2^e seulement.

pratique correspond à un processus qui ne progresse pas, parce qu'il oscille entre deux ou plusieurs solutions, et il faut débloquer la situation. L'utilisation de facteurs de relaxation (décrits plus loin) devrait résoudre le problème dans le cas de Gauss-Seidel.

7.1.4.2 Les options EViews

Le test de convergence proposé par EViews n'est pas très souple: la seule option autorisée est le niveau du critère de convergence (relative), et il s'appliquera à toutes les variables.

On peut aussi choisir le nombre maximum d'itérations. Pour la plupart des modèles, après 1 000 itérations, la convergence devient plutôt improbable. Mais juste pour être sûr, on peut fixer d'abord un nombre plus élevé. Constaté le nombre réel dont on a besoin permet ensuite de préciser le chiffre.

7.1.5 ETUDE DE LA CONVERGENCE

Nous allons maintenant montrer comment le choix de l'algorithme affecte le processus de convergence.

Commençons par poser le problème, en introduisant quelques définitions.

7.1.5.1 La matrice d'incidence

La matrice d'incidence d'un modèle à n équations

$$f(y, \dots) = 0 \text{ (n variables endogènes, n équations)}$$

sera définie comme la matrice booléenne A , de dimension $(n \times n)$ telle que

$A_{i,j} = 1$ si la variable y_j apparaît formellement, à travers sa valeur instantanée, dans l'équation de rang i

$$A_{i,j} = 0 \text{ sinon}$$

Nous supposons que le modèle est normalisé, donc peut se mettre sous la forme:

$$y - f(y) = 0$$

où la variable y_i apparaîtra naturellement à gauche de l'équation de rang i : la diagonale principale de la matrice sera composée de 1.

La définition de la matrice d'incidence, comme on peut le voir, ne dépend pas seulement du modèle, mais aussi de l'ordre des équations, en fait de celui dans lequel elles vont être calculées.

La présence formelle d'une variable dans une équation ne signifie pas nécessairement une influence numérique: elle pourrait être affectée par un coefficient potentiellement nul, ou n'intervenir que dans une branche d'une alternative. Souvent, nous ne serons pas en mesure d'associer à un modèle ni une matrice d'incidence unique, ni une matrice constante dans le temps, sauf si l'on considère l'ensemble complet des influences potentielles (la matrice sera alors la somme booléenne des différentes matrices booléennes).

On remarquera également que la définition de la matrice d'incidence n'exige pas les formulations complètes, ou la connaissance des valeurs des variables. Nous avons simplement besoin de connaître la liste des variables qui apparaissent dans chaque explication, ainsi que leur caractère instantané ou retardé⁶³.

7.1.5.2 Application à notre modèle

Pour appliquer cette technique à notre modèle, nous pouvons utiliser la structure par blocs fournie par EViews, grâce à:

L'accès au modèle par:

```
model (double-click)>View>Block structure
```

qui donne dans notre cas:

⁶³ Selon notre méthodologie, la matrice d'incidence peut être produite avant toute estimation.


```

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)                x(19)                prle_t(7)

Block 2: 14 Simultaneous Equations
ur(2)                q(3)                ic(4)
  ci(5)                i(6)
  led(8)                le(9)
  lt(10)                rhi(11)            ih(12)
  co(13)                fd(14)
  td(15)                m(17)

Block 3: 3 Recursive Equations
res_m(16)            res_x(18)            k(20)

```

Tous ces éléments sont compatibles avec le graphique précédent.

Nous pouvons utiliser ces éléments pour améliorer l'ordonnement.

Tout d'abord, nous pouvons utiliser la séparation ci-dessus pour déplacer les trois variables prédéterminées au début, et placer à la fin les variables du troisième bloc, ce qui donne la matrice suivante:

res_m		1		1		1
k			1			1

Nous pouvons voir que le modèle a été découpé en trois parties:

- Un bloc de trois équations, avec des éléments qui ne dépendent pas du complément, ou des variables suivantes dans le même bloc. Les variables de ce bloc peuvent alors être calculées une fois pour toutes, en une seule itération, au début du processus de résolution. En fait, elles ne dépendent pas des autres variables du même bloc, mais ce n'est pas obligatoire.

Cette propriété est appelée récursivité, et le bloc est habituellement appelé le prologue.

Nous pouvons voir que ces variables peuvent appartenir à ce bloc pour diverses raisons:

- prle_t dépend uniquement du temps. La seule raison pour introduire son équation est de faciliter sa modification lors des prévisions,
- CAP dépend seulement d'une exogène et d'une variable prédéterminée,
- X devrait dépendre le reste de l'équilibre (par UR), mais le lien instantané n'a pas été mis en évidence statistiquement, ne laissant que l'influence instantanée de l'exogène WD.

En pratique, cependant, respecter le seuil de convergence nécessitera deux itérations, la valeur de départ étant différente de la solution trouvée, à moins que la récursivité ne soit connue dès le début, et la première solution acceptée sans contrôle (ce qui est le cas pour EViews).

- Un bloc de trois équations, dans lequel se trouvent des éléments du modèle qui n'affectent pas leur complément, et ne dépendent pas des variables suivantes du même bloc. Les éléments de ce bloc peuvent être calculés après tous les autres, une fois pour toutes en un seul passage.

Dans ce bloc, on trouve:

- les résidus des équations de cointégration, qui ne seront corrigés qu'à la période suivante ;
- le capital en fin de période, qui ne peut évidemment pas affecter l'équilibre de la période.

Nous verrons plus loin une autre catégorie importante: les variables ayant un rôle purement descriptif, telles que le déficit des administrations publiques en points de PIB.

- Le reste du modèle est interdépendant, et parfois appelé le cœur. On peut vérifier sur le graphique que pour tout couple de variables de l'ensemble, il existe au moins une séquence de relations de causalité menant de la première à la seconde, et vice versa. Cela signifie également

que si l'on exogénéise tout élément (à une valeur différente de la solution du modèle bien sûr) tous les autres éléments seront affectés.

Nous pouvons maintenant chercher à mieux interpréter la simultanéité dans le cœur. La première étape est l'observation de la présence de variables «de bouclage».

La matrice d'incidence permet de définir les variables de bouclage, comme des variables qui entrent par leur valeur instantanée dans le calcul d'une équation de rang inférieur à celle qui les définit, ou vont être utilisées avant d'être calculées. En notations matricielles, nous aurons:

Pour la variable j , $\exists A_{i,j} = 1$ tel que $i < j$

Les variables apparaissant comme un facteur explicatif de leur propre équation de définition devront également être ajoutées à cet ensemble. Mais dans la pratique, ce cas est plutôt rare.

Penchons-nous sur notre matrice d'incidence. Deux variables de bouclage sont présentes: FD et M . La raison en est qu'elles sont utilisées pour calculer Q , dans une équation qui apparaît à la fin (du cœur).

En fait, X devrait également être présent, mais comme UR n'intervient que par sa valeur retardée, et WD est exogène, sa valeur exacte peut être calculée immédiatement, ce qui place la variable dans le prologue. En un sens, cet élément est désormais techniquement exogène (on ne considère que les relations de la même période).

Bien sûr, un modèle peut contenir une séquence de blocs non récursifs. Ce sera le cas par exemple pour une séquence de deux blocs non récursifs si des éléments du second dépendent d'éléments du premier, mais pas l'inverse. Entre les deux blocs, un bloc récursif peut être introduit.

Nous verrons des exemples de cette situation lorsque nous traiterons avec des modèles plus complexes.

La définition de l'ensemble des variables de bouclage présente l'intérêt suivant: si cet ensemble est vide, le modèle est récursif, ce qui signifie que le calcul séquentiel de chacune des équations (dans l'ordre adéquat) donne la solution du système. Les valeurs obtenues à la fin de la première itération satisfont l'ensemble des équations, aucune de ces valeurs n'étant remise en cause par une modification ultérieure d'un élément explicatif. Et une seconde itération ne modifiera pas le résultat.

Ce cas favorable (mais pauvre économiquement) est assez rare. Cependant, on peut souvent identifier des sous-ensembles simultanés (ou «blocs») ayant une structure récursive rapport à l'autre, tels que les p premières équations du modèle ne sont pas influencées par les dernières $n - p$. Le processus de simulation peut être alors être amélioré, car il suffit de résoudre en séquence deux systèmes de taille

réduite, ce qui permet de gagner du temps car le coût de résolution augmente plus que proportionnellement avec le nombre d'équations. Cette propriété est évidente pour Newton, où le coût de calcul vient à la fois du Jacobien et de l'inversion, moins pour Gauss-Seidel et Broyden, où la seule preuve vient de la pratique.

Il est évident que la découverte des propriétés ci-dessus, et le réordonnement associé sont intéressants pour le constructeur du modèle, car ils permettent d'améliorer l'organisation du processus de solution, et donc de réduire les temps de calcul. Ce processus permettra également de détecter les erreurs logiques, par exemple en établissant la détermination récursive d'un élément connu comme appartenant à une boucle (comme les investissements dans la boucle keynésienne). La plupart des logiciels, y compris EViews, traitent cette recherche et la réorganisation associée, mais une amélioration supplémentaire peut être poursuivie dans le processus de résolution en modifiant à la main l'ordre des calculs d'équations.

À la lumière des observations précédentes, on peut rechercher:

- la meilleure séquence de blocs ;
- à l'intérieur de chaque bloc, l'ordre qui permet la convergence le plus rapide ou qui l'assure avec la plus forte probabilité.

Le premier objectif est incontestable, et en fait le plus facile à réaliser du point de vue algorithmique. La séparation trouvée est unique, mais certains ordres de blocs peuvent être équivalents (par exemple les équations utilisant uniquement des éléments exogènes ou retardés peuvent être placés dans n'importe quel ordre).

Le second est beaucoup moins évident et en tout cas plus complexe. On cherchera généralement à réduire au minimum le nombre de variables de bouclage. Le coût de cette technique dépend de l'ambition: la recherche d'un ensemble de variables de bouclage, dans lequel on ne peut pas éliminer de variable (Nepomiaschy et Ravelli), est moins coûteuse que la recherche de tous les ordres contenant le plus petit nombre possible d'éléments (Gilli et Rossier). Le premier type d'ensemble sera appelé minimal, le second minimum. En fait, la minimisation du nombre de variables de bouclage n'est pas forcément une bonne option pour l'utilisation de l'algorithme de Gauss-Seidel, comme nous le verrons plus tard.

EViews détermine automatiquement la structure par bloc du modèle (qui est en fait optimale, même si d'autres organisations existent). Quant au réordonnement des blocs simultanés, s'il ne va pas appliquer un algorithme d'optimisation, il détermine les variables de bouclage associées à l'un d'entre eux (associé en fait à l'ordre initial) et place les équations associées à la fin du bloc (en signalant leur nombre).

L'efficacité de cette dernière action est douteuse, car cela signifie que dans une itération donnée tous les calculs utiliseront la valeur précédente de variables de bouclage, ce qui retarde quelque peu l'impact de la «nouvelle» information.

Par exemple, dans notre modèle, nous pouvons réduire le nombre de variables de bouclage par le transfert de l'équation déterminant Q à la fin du cœur:

res_		1		1		1
k			1			1

Maintenant, Q est la seule variable de bouclage (mais au rôle accru car elle apparaît 6 fois en cette qualité).

Le nouveau classement est le suivant:

Prologue:

$$[1] \text{CAP} = \text{pk} * \text{K}(-1)$$

$$[2] \log(\text{PRLE_T}) = \text{c_prle}(1) + \text{c_prle}(2) * (\text{t} - 2002) + \text{c_prle}(3) * (\text{t} - \text{t1}) * (\text{t} < \text{t1}) + \text{c_prle}(4) * (\text{t} - \text{t2}) * (\text{t} < \text{t2})$$

$$[3] \text{DLOG}(\text{X}) = 0.940 * \text{DLOG}(\text{wd}) - 0.0129 - 0.195 * \text{RES_X}(-1) + \text{EC_X}$$

cœur:

$$[4] \text{UR} = \text{Q} / \text{CAP}$$

$$[5] \text{CI} = \text{tc} * \text{Q}$$

$$[6] \text{IC}/\text{Q}(-1) = 0.130 * \text{PCH}(\text{Q}) + 0.214 * \text{PCH}(\text{Q}(-1)) + \text{EC_IC}$$

$$[7] \text{I}/\text{K}(-1) = 0.825 * \text{I}(-1)/\text{K}(-2) + 0.0279 * \text{UR} + 0.152 * .25 * \text{Q}/\text{Q}(-4) - 0.0525 + \text{EC_I}$$

$$[8] \text{LED} = \text{Q} / \text{PRLE_T}$$

$$[9] \text{DLOG}(\text{LE}) = 0.587 * \text{DLOG}(\text{LED}) + 0.411 * \text{LOG}(\text{LED}(-1)/\text{LE}(-1)) + 0.000502 - 0.0167 * ((\text{T} = 1968.5) - (\text{T} = 1968)) + \text{EC_LE}$$

$$[10] \text{LT} = \text{LE} + \text{lg}$$

$$[11] \text{RHI} = \text{wr} * \text{LT} + \text{r_rhiq} * \text{Q}$$

$$[12] \text{IH} = \text{r_ih} * \text{RHI}$$

$$[13] \text{CO} = \text{RHI} * (1 - \text{sr})$$

$$[14] \text{FD} = \text{CO} + \text{IP} + \text{CI} + \text{IH} + \text{gd}$$

$$[15] \text{TD} = \text{FD} + \text{CI}$$

$$[16] \text{DLOG}(\text{M}) = 1.2 * \text{DLOG}(\text{FD} + \text{CI}) + 0.282 * \text{DLOG}(\text{UR}) - 0.212 * \text{RES_M}(-1) - 0.629 + \text{EC_M}$$

$$[17] \text{Q} + \text{M} = \text{FD} + \text{X}$$

épilogue:

[18] $RES_M = \log(M / (FD+CI)) - 1.322108 * \log(UR) + 0.419518 * \log(compm) - 0.012582 * (@trend(60:1) * (t \leq 2002) + @elem(@trend(60:1), "2002S2") * (t > 2002))$

[19] $RES_X = \log(X / wd) + 0.686 * \log(UR) - 4.87E-05 * (@trend(60:1) * (t \leq 2002) + @elem(@trend(60:1), "2002S2") * (t > 2002))$

[20] $K = K(-1) * (1 - dr) + IP$

et la résolution du modèle se fera de la manière suivante (nous verrons plus tard les déclarations correspondantes):

Model: _FRA_1

Date: 11/15/12 Time: 16:52

Sample: 2003S1 2500S2

Solve Options:

Dynamic-Deterministic Simulation

Solver: Gauss-Seidel

Max iterations = 5000, Convergence = 1e-006

Scenario: Scenario 1

Solve begin 16:52:42

2003S1	Block 1 - 3 eqns	Solved (recursive block)
2003S1	Block 2 - 13 eqns	Convergence after 10 iterations
2003S1	Block 3 - 3 eqns	Solved (recursive block)
2003S2	Block 1 - 3 eqns	Solved (recursive block)
2003S2	Block 2 - 13 eqns	Convergence after 10 iterations
2003S2	Block 3 - 3 eqns	Solved (recursive block)
2004S1	Block 1 - 3 eqns	Solved (recursive block)
2004S1	Block 2 - 13 eqns	Convergence after 10 iterations
2004S1	Block 3 - 3 eqns	Solved (recursive block)
2004S2	Block 1 - 3 eqns	Solved (recursive block)
2004S2	Block 2 - 13 eqns	Convergence after 11 iterations

2004S2 Block 3 - 3 eqns Solved (recursive block)

Il est maintenant possible de comparer les propriétés de convergence des différentes méthodes, en commençant par l'algorithme de Gauss-Seidel.

On peut voir que, pour cet algorithme, les seules variables de bouclage influencent (par leurs valeurs de départ) le résultat de l'itération courante. Notons y_b le vecteur des variables de bouclage (et son complément y_c), associés à la version en cours. On pourrait envisager de remplacer tour à tour chaque occurrence d'une variable déjà calculée par l'expression courante (éventuellement déjà transformée). Cela donnera un modèle dans lequel seules les variables de bouclage apparaissent à droite:

$$y = g(y_b)$$

Et pour le processus de convergence, nous n'avons qu'à prendre en compte les éléments y_b du résultat:

$$y_b = g(y_b)$$

Il suffit de résoudre le nouveau (petit) modèle, et y_c sera obtenu en un seul passage à la fin du processus.

Notre modèle donne:

```
[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))
```

$$[12] \text{ ih} = f_{12}(f_{11}(f(f_{10}(f_9(f_8(q))), q)))$$

$$[13] \text{ co} = f_{13}(f_{12}(f_{11}(f(f_{10}(f_9(f_8(q))), q))))$$

$$[14] \text{ fd} = f_{14}(f_{13}(f_{12}(f_{11}(f(f_{10}(f_9(f_8(q))), q))), f_7(q, f_4(q)), f_6(q), f_{12}(f_{11}(f(f_{10}(f_9(f_8(q))), q))))$$

$$[15] \text{ td} = f(f_{14}(f_{13}(f_{12}(f_{11}(f(f_{10}(f_9(f_8(q))), q))), f_7(q, f_4(q)), f_6(q), f_{12}(f_{11}(f(f_{10}(f_9(f_8(q))), q))), f_5(q)))$$

$$[16] \text{ m} = f_{15}(f_{14}(f_{13}(f_{12}(f_{11}(f(f_{10}(f_9(f_8(q))), q))), f_7(q, f_4(q)), f_6(q), f_{12}(f_{11}(f(f_{10}(f_9(f_8(q))), q))), f_5(q), f_4(q)))$$

Même pour ce simple exemple, la production de l'ensemble complet des relations explicites aurait été difficile, voire impossible. Mais si g est formellement inconnu, ses valeurs peuvent être calculées (par une itération typique de Gauss-Seidel).

Ainsi, dans notre exemple, l'application successive (dans l'ordre) des formules f associées à une valeur initiale de Q donne une valeur numérique de la fonction g associée (mais pas sa solution, bien sûr). Aucune autre valeur initiale ne doit être prise en considération.

Encore une fois, cela montre que, pour la convergence, il suffit de considérer le modèle réduit:

$$y_b = g(y_b) \text{ ou } y_b^{k+1} = g(y_b^k)$$

et le contrôle de la convergence peut être limité à ces variables.

Servons-nous de cette observation pour interpréter la convergence de l'algorithme. Nous allons linéariser le processus autour de la solution:

$$y^k - y^* = g(y^{k-1}) - g(y^*) \approx \left(\frac{\partial g}{\partial y}\right)_{y=y^{k-1}} (y^{k-1} - y^*)$$

Cette approximation, si elle peut être utilisée dans la pratique (ce qui suppose soit que l'on est proche de la solution, soit que le modèle est suffisamment linéaire), montre que le vecteur des erreurs est multiplié à chaque itération par le Jacobien. Si celui-ci est assez stable au cours du processus, ceci rend la convergence presque linéaire, et le coût d'obtention d'un chiffre supplémentaire dans la solution sera indépendant de sa position dans la représentation de la variable. Et pour que l'algorithme converge correctement, la condition nécessaire et généralement toujours suffisante est que lorsque l'on élève la

matrice jacobienne à une puissance tendant vers l'infini, sa valeur tend vers zéro. Une condition équivalente est que chaque valeur propre de la matrice soit strictement inférieure à 1 en module (on dira que le rayon spectral est inférieur à l'unité). Ainsi, dans notre exemple, la convergence sera assurée si:

$$\Delta(g(Q))/Q < 1$$

On voit que la réduction du nombre de variables de bouclage, si elle réduit la taille du Jacobien, n'a pas de raison spécifique de réduire le rayon spectral⁶⁴.

En fait, il arrive fréquemment que la réorganisation ne modifie en aucune façon le processus de convergence, car elle ne concerne que le point d'entrée dans une boucle par ailleurs inchangée. C'est le cas dans notre exemple: en transférant à la fin l'équation donnant Q , nous commençons par la valeur de Q venant du fichier de travail, et non celle qui équilibre l'équilibre offre-demande en utilisant les valeurs initiales de X , M et FD . Mais à partir de là le processus est exactement le même. Plus encore: si les valeurs des données satisfont l'équilibre (elles le feront la plupart du temps), le calcul de Q ne change pas sa valeur, et le processus sera identique y compris dans ses chiffres.

On peut même avoir (sans plus de preuves) l'intuition inverse: en concentrant le processus sur un nombre limité de variables, on peut augmenter leur rôle individuel.

De plus, l'appel à un algorithme de réorganisation automatique transfère le contrôle de la logique de l'ordonnancement établie par l'économiste (qui a sans doute choisi un ordre pertinent en termes de causalités) vers une méthode mathématique aveugle.

Évidemment, si le nombre de variables de bouclage est élevé, la probabilité de convergence serait très faible si les coefficients avaient été choisis au hasard. Dans la pratique, heureusement, on peut associer au processus de convergence mathématique une autre plus économique, le progrès itératif vers un équilibre combinant les comportements des agents (des exemples peuvent être trouvés dans la boucle keynésienne, ou dans la boucle prix-salaires). La probabilité de succès est beaucoup améliorée, et la plupart des modèles convergent s'ils reposent sur des hypothèses et un point de départ raisonnables⁶⁵.

⁶⁴ En termes plus simples: le module le plus élevé de l'ensemble des valeurs propres.

⁶⁵ Ne pas converger pour des valeurs invraisemblables des hypothèses doit plutôt être considéré comme positif.

Nous allons illustrer ce point en simplifiant à l'extrême notre modèle habituel, en prenant seulement en compte le PIB (Q), la demande privée (C) et de la demande du gouvernement (g). Notre modèle ne tient pas compte du commerce extérieur.

$$\begin{aligned} C &= Q \\ Q &= C + g \end{aligned}$$

et

$$0 < a < 1$$

Avec l'algorithme de Gauss-Seidel, ce modèle sera toujours convergent vers la solution Q^* , comme pour la variable de bouclage unique (Q):

$$Q^k - Q^* = a \cdot (Q^{k-1} - Q^*)$$

si $|a| < 1$ ⁶⁶

Inversons maintenant les causalités:

$$\begin{aligned} C &= Q - g \\ Q &= C / a \end{aligned}$$

⁶⁶ Avec $a > 1$, le modèle serait divergent, ce à quoi on pouvait s'attendre d'un point de vue économique.

Cette fois le modèle diverge:

$$Q^k - Q^* = 1/a \cdot (Q^{k-1} - Q^*)$$

et $|1/a| > 1$

mais cela est normal, car à un cadre de convergence économique (le multiplicateur keynésien), nous avons substitué un autre où, pour simplifier:

- les ménages consomment ce qui reste de l'offre une fois la demande exogène satisfaite,
- les entreprises cherchent à maintenir l'offre dans un rapport constant (supérieur à 1) à la consommation.

Si un tel cadre ne converge pas mathématiquement, c'est logique: bien qu'il existe évidemment une solution (la même que plus haut), une erreur initiale sur Q s'appliquera à l'identique à C , puis va s'amplifier sur Q Et la même instabilité sera rencontrée avec le choc de demande (calcul du multiplicateur de dépenses publiques).

On peut observer aussi que, avec $a > 1$, le premier modèle serait divergent (car les effets secondaires sur la production seraient supérieurs à leur source). Et le second convergent (car un changement dans la consommation va maintenant se réduire sur la production): les deux aberrations se compensent.

Avec le premier modèle, un changement initial de la demande du gouvernement conduira à l'itération k , à la variation cumulée:

$$\Delta Q^k = (1 + a + a^2 + \dots + a^k) \cdot \Delta g = (1 - a^{k+1}) / (1 - a) \cdot \Delta g$$

qui converge vers:

$$\Delta Q^k = 1 / (1 - a) \cdot \Delta g$$

si $a < 1$

et avec la seconde:

$$\Delta Q^k = (1 + 1/a + (1/a)^2 + \dots + (1/a)^k) \cdot \Delta g = (1 - (1/a)^{k+1}) / (1 - (1/a)) \Delta g$$

qui converge que si $a > 1$

La convergence d'un modèle est équivalente à la divergence de l'autre.

7.1.5.3 Accélérer Gauss-Seidel: les facteurs de relaxation

Une technique efficace pour accélérer la convergence, ou pour forcer les modèles naturellement divergents à parvenir à une solution, consiste à introduire un ou plusieurs «facteurs de relaxation». Cette technique est basée sur l'observation élémentaire suivante, inspirée par ce qui précède.

Si y est solution de

$$y = f(y)$$

alors ce sera aussi la solution de

$$y = D \cdot f(y) + (I - D) \cdot y$$

où y représente le vecteur des endogènes et D une matrice diagonale.

Il va de même pour g :

si y_b est une solution de

$$y_b = g_b(y_b)$$

alors il sera également de:

$$y_b = D \cdot g_b(y_b) + (I - D) \cdot y_b$$

Mais la nouvelle formulation modifie le processus de résolution Gauss-Seidel, en introduisant une inertie sur les valeurs de départ y_b .

En pratique, cela se traduira par l'initialisation de l'itération k , et non pas par:

$$y_b^{k-1}$$

Mais par:

$$D_b \cdot y_b^{k-1} + (I - D_b) \cdot y_b^{k-2}$$

Le changement de valeurs de départ ne joue un rôle que si les variables sont utilisées avant qu'elles ne soient calculées lors de l'itération. Par conséquent les facteurs de relaxation ne seront appliqués qu'aux variables de bouclage.

On montre facilement que par l'application de facteurs de relaxation, nous nous orientons vers une méthode de type Newton, au moins dans les cas simples. En effet, dans un modèle avec une variable de bouclage unique, l'introduction de facteurs de relaxation consiste à utiliser:

$$y_b^k = D g_b(y_b^{k-1}) + (I - D) y_b^{k-1}$$

ou

$$y_b^k - y_b^{k-1} = D (g_b(y_b^{k-1}) - y_b^{k-1})$$

Ceci est équivalent à la méthode de Newton, si l'on choisit

$$D = (I - \frac{\partial g_b}{\partial y_b})^{-1}_{y_b=y_b^{k-1}}$$

ou l'inverse de la matrice jacobienne de:

$$y_b - g_b(y_b) = 0$$

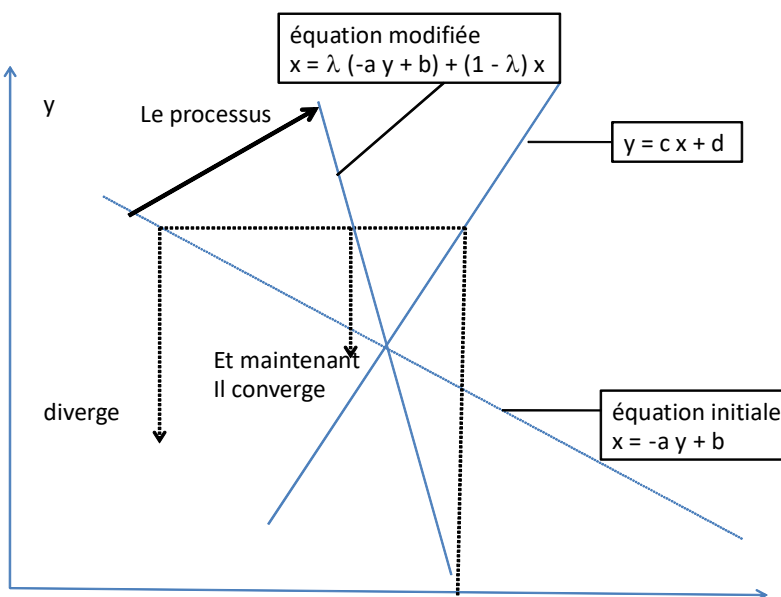
Cette méthode, plus ou moins efficace selon la stabilité du Jacobien, donnera la solution en une seule itération dans le cas d'un modèle linéaire.

Terminons notre exposé par un exemple concret.

Nous allons utiliser un modèle formel très simple

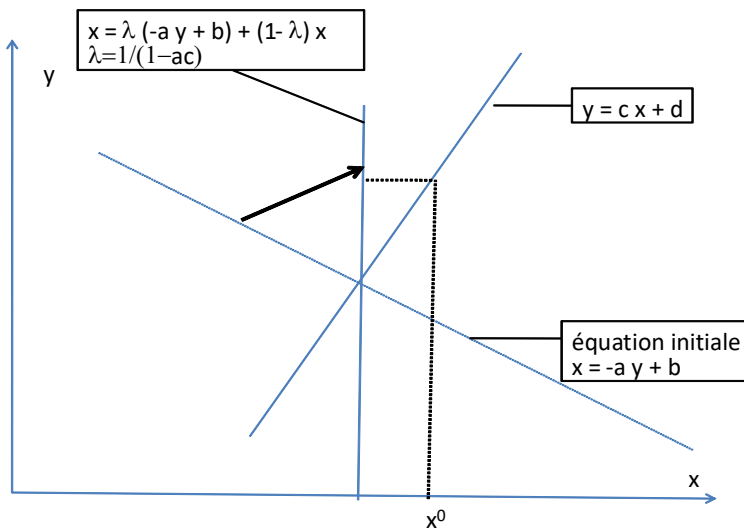
$$y = c x + d$$

$$x = -a y + b$$



Le processus de base diverge nettement, mais si l'on applique un facteur de relaxation adéquat (nous avons choisi une valeur proche de 0,5), nous pouvons le faire converger. Ce que nous avons fait, c'est tout simplement remplacer une des équations ($y = -ax + b$) par une autre qui associe à une valeur donnée de y la combinaison, avec des poids constants, les valeurs de x associées au même y par les deux équations. La solution resta la même, mais le processus de convergence est modifié.

On peut même atteindre la valeur exacte, en faisant de la nouvelle «courbe» une verticale



L'application la méthode des facteurs de relaxation n'est cependant pas si simple.

Elle ne peut se faire que par essais successifs, sauf si l'on résout le problème décrit ci-dessus (en inversant une matrice), ce qui est en fait plus complexe que la résolution du modèle lui-même par la méthode de Newton⁶⁷.

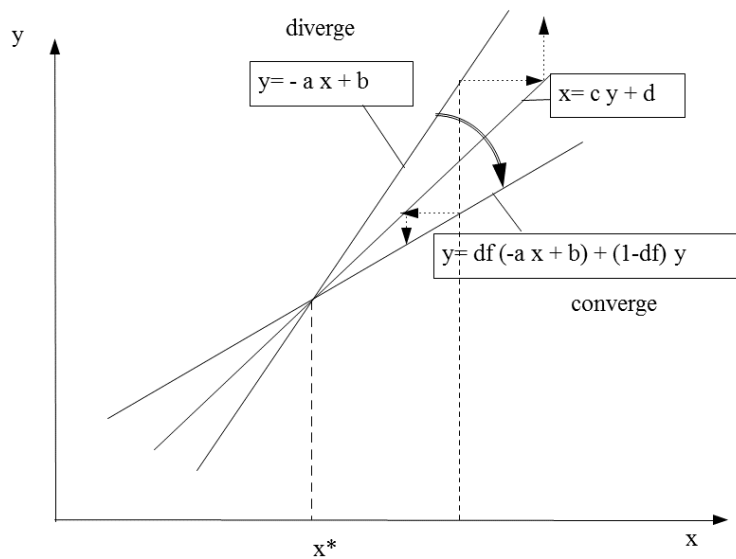
Elle n'est vraiment efficace pour des valeurs comprises entre 0 et 1 (convergence alternée). Si les pentes des deux courbes ont la même orientation, ce qui rend la convergence monotone, la méthode devient plus difficile à mettre en œuvre, car le facteur de relaxation efficace doit maintenant être choisi en dehors de l'intervalle (même si cela peut fonctionner, comme le montre le schéma ci-dessous).

⁶⁷ On peut aussi se concentrer sur une variable de bouclage, considérée comme la plus importante, et de mesurer son évolution sur trois itérations consécutives. Une approximation très grossière du facteur de convergence peut être obtenue par :

$$a = (x^k - x^{k-1}) / (x^{k-1} - x^{k-2})$$

et on peut utiliser comme un facteur de relaxation la valeur :

$$\lambda = 1 / (1 - a)$$



Et lorsque la taille du modèle augmente, le processus de recherche devient rapidement complexe.

7.1.5.3.1 Des modèles plus complexes

Si l'équivalence de Newton ne peut pas être étendue au cas de plusieurs variables de bouclage, c'est seulement parce que la méthode retient une matrice D diagonale. On pourrait toutefois considérer:

- la recherche de la matrice diagonale «le plus proche» du Jacobien, selon une mesure donnée ;
- la décomposition du vecteur y sur la base des vecteurs propres de la matrice jacobienne, ce qui permettrait de diagonaliser le processus.

$$y^k - y^{k-1} \approx (\partial g / \partial y)_{y=y^{k-1}} (y^{k-1} - y^{k-2}) = V D V^{-1} (y^{k-1} - y^{k-2})$$

$$V^{-1} (y^k - y^{k-1}) \approx D V^{-1} (y^{k-1} - y^{k-2})$$

Avec D matrice diagonale avec les valeurs propres de la matrice jacobienne.

Une relative stabilité du Jacobien (donc une quasi-linéarité, ou une valeur d'initialisation proche de la solution) est alors nécessaire⁶⁸.

Voyons maintenant comment les considérations ci-dessus influent sur la convergence de l'algorithme de Newton.

L'utilité des variables de bouclage, et la réduction de leur nombre, est désormais incontestable, si l'on décide de calculer le Jacobien par différences finies. En effet il suffit de connaître les valeurs de départ de y^b pour calculer $g(y^b)$ et le Jacobien $\partial g_b / \partial y_b$ (ce dernier par itérations de type Gauss-Seidel),

seule la partie $\partial g / \partial y_b$ n'est pas identiquement nulle,

et seule la partie $\partial g_b / \partial y_b$ va affecter le calcul de la nouvelle valeur de y_b .

Nous n'avons qu'à appliquer la formule de Newton aux seules variables de bouclage. Encore une fois, seules les valeurs des variables de bouclage de l'itération précédente (et leur influence sur le Jacobien de g) joueront un rôle dans le processus itératif.

Ceci a évidemment l'avantage de limiter les calculs, pour le Jacobien lui-même (seulement pour le calcul des colonnes non nulles) ainsi que pour son inversion.

L'utilité de la prise en compte des variables de bouclage semble claire. La réalité l'est un peu moins:

⁶⁸ L'expérience montre que le Jacobien de la différenciation linéaire est beaucoup moins stable que la valeur relative:

$$D = (I - \partial g_b / \partial y_b)_{y_b=y_b^{k-1}}^{-1}$$

- la matrice jacobienne associée à la formulation initiale f est très creuse (le nombre moyen de variables explicatives instantanées dans l'équation ne croît pas trop rapidement avec la taille du modèle, et reste généralement de l'ordre de 3 ou 4). Cela signifie que la méthode de Newton appliquée à f peut utiliser un algorithme d'inversion spécifiquement adapté aux matrices creuses, nettement plus efficace. La formulation en g concentre le processus de détermination sur certaines variables, mais la complexité de la formulation implicite produit une matrice jacobienne très dense, qui ne peut être inversée que par un algorithme général.
- Quant au calcul de la matrice elle-même, il sera évidemment plus coûteux pour f que pour g , si l'on utilise les différences finies. Mais la matrice jacobienne associée au vecteur f peut également être déterminée par analyse formelle: ensuite il suffira de calculer de manière indépendante chaque élément non nul du Jacobien, selon une formule généralement plus simple que les équations initiales⁶⁹. Si le nombre moyen de variables instantanées dans l'équation est inférieur au nombre de variables de bouclage, le coût global diminue. Ce raisonnement n'est cependant valable que si l'on fait abstraction du coût (important) de la dérivation formelle initiale⁷⁰, qui suppose également une certaine stabilité de la formulation du modèle.

Quant à la comparaison avec Gauss-Seidel, on peut s'attendre à une diminution du nombre d'itérations, mais chacune d'elles sera évidemment beaucoup plus coûteuse.

Voyons comment le processus converge avec Newton.

Pour cela développons au second ordre la formule de dérivation:

$$y^{k+1} - g(y^{k+1}) - (y^k - g(y^k)) = (I - \partial g / \partial y)_{y=y^k} (y^{k+1} - y^k) - {}^T(y^{k+1} - y^k) (\partial^2 g / \partial y^2)_{y=y^k} (y^{k+1} - y^k)$$

plus un terme du troisième ordre.

L'utilisation de la dérivation de y^{k+1} élimine le terme central, laissant:

⁶⁹ Par exemple, lorsque la dérivation est appliquée à une somme.

⁷⁰ Depuis la version 6, EViews est en mesure de calculer les dérivées analytiques. Toutefois, il ne fournit pas les formules.

$$y^{k+1} - g(y^{k+1}) = - {}^T(y^{k+1} - y^k)(\partial^2 g / \partial y^2)_{y=y^k}(y^{k+1} - y^k) + e^3(y^{k+1} - y^k)$$

Et comme $y^{k+1} - y^k$ est une fonction linéaire de $y^k - g(y^k)$ (à nouveau par construction),

l'erreur $y^{k+1} - g(y^{k+1})$ est une fonction quadratique de $y^k - g(y^k)$.

De même, la distance de y à la solution y^* est une fonction quadratique de la distance à l'itération précédente, car la dérivation de la première commande donne:

$$(y^{k+1} - g(y^{k+1})) - (y^* - g(y^*)) [= 0] = (I - \partial g / \partial y)_{y=y^k}(y^{k+1} - y^*)$$

La convergence de l'algorithme n'est donc pas uniforme: en fait, elle va s'accélérer avec le nombre d'itérations, comme la forme linéarisée se rapproche de la dérivée calculée à la solution. On dira qu'elle est quadratique dans le cas d'une dérivation formelle, mais seulement supra linéaire si l'on utilise des différences finies, parce qu'on ne calculera pas la dérivée exacte. L'utilisation de l'algorithme de Newton devrait être d'autant plus judicieuse que la précision demandée de la solution est plus élevée.

7.1.5.4 Comparaison des efficacités

Examinons d'abord la capacité à trouver la solution. La convergence de l'algorithme de Gauss-Seidel semble assez douteuse, si elle repose sur la combinaison de plusieurs processus itératifs simultanés (le nombre de variables de bouclage) qui doivent tous être convergents. En d'autres termes, une matrice doit avoir toutes ses valeurs propres inférieures à l'unité en module. Ceci est toutefois plus fréquent que l'on pouvait s'y attendre. Comme nous l'avons vu précédemment, le processus numérique est généralement lié à un ensemble de causalités économiques, déterminant un processus normalement convergent. En fait, si ce processus ne converge pas, la cohérence économique du modèle lui-même doit probablement être remise en cause.

L'algorithme de Newton au contraire ne met aucune condition sur l'intensité des relations numériques, seulement sur leur variabilité (dans la pratique, la stabilité du Jacobien et dans une moindre mesure, du Hessian). Cette condition est généralement vérifiée dans le cas des modèles économétriques. En particulier, la convergence sera immédiate pour un modèle linéaire présentant une solution unique. L'algorithme de Newton sera alors conseillé pour les modèles présentant des difficultés de convergence, ou pour les modélisateurs qui veulent se libérer de cette préoccupation, qui reste toujours présente et parfois cruciale (par exemple, s'ils sont souvent confrontés à des problèmes de délai).

D'autre part, concernant la vitesse de convergence, le choix est beaucoup moins évident. Si la méthode de Newton converge en moins d'itérations, généralement entre trois et cinq (rappelez-vous que la convergence s'accélère avec les itérations), chaque itération est beaucoup plus coûteuse. En fait, ce supplément peut être mesuré, dans le cas d'une dérivation à l'aide de différences finies: chaque itération comporte autant d'itérations de type Gauss-Seidel que le modèle contient des variables de bouclage (plus une), ainsi qu'une inversion de matrice (de coût croissant avec la dimension comme un polynôme de degré trois). On peut donc minorer le coût total, quelle que soit la méthode, par un certain nombre d'itérations de Gauss-Seidel.

Dans la pratique, on observe que le nombre d'itérations nécessaires à la convergence, quelle que soit la méthode employée, croît assez lentement avec la taille du modèle. Par contre le nombre de variables de bouclage est plus ou moins proportionnel à celle-ci.

Nous allons maintenant présenter le temps pris par chacune des méthodes (en cas de succès) pour traiter un cas pratique:

- Gauss-Seidel
- Newton avec dérivations analytiques
- Newton avec approximation numérique
- Broyden avec dérivations analytiques
- Broyden avec approximation numérique.

Le test a été effectué en 2011 sur un Toshiba Qosmio F60-10J, en utilisant une puce Intel I5 et 4 Go de mémoire.

Pour 10 modèles de taille croissante, tous construits par l'auteur seul ou en équipe, dans le cadre de projets de coopération et de séminaires:

- France: un très petit modèle français, semblable à celui que nous utilisons ici
- France cf: un petit modèle français avec fonction de production à facteurs complémentaires
- France cd: un petit modèle français avec fonction de production Cobb-Douglas. Ces deux modèles seront présentés plus tard
- Vietnam: un petit modèle vietnamien à un produit
- Algérie 1: un petit modèle algérien à un produit (hydrocarbures séparés dans certains cas)
- Algérie 5: un modèle algérien à 5 produits
- Algérie QAM: un modèle algérien quasi comptable à 19 produits
- Vietnam: un modèle vietnamien à 3 produits
- Monde: un modèle mondial à 11 pays (plus un reste du monde)
- Chine: un modèle à 3 produits et 4 régions

Les trois premiers modèles sont disponibles dans les fichiers accompagnant le livre.

Avec les caractéristiques suivantes:

- prod: nombre de produits
- pays: le nombre de produits ou de régions
- Pér: nombre de périodes
- Equ: nombre d'équations
- Bloc: taille du plus grand bloc non récursif
- Feedback: nombre de variables de bouclage de ce bloc

Pays	prod	pays	pér	equ	Bloc	Fback	Gauss-Sei	Newton A	Newton N	Broyden A	Broyden N
France	1	1	q	18	13	2	0,002	échec	échec	0,003	0,004
France fc	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	échec	0,017	0,017
Vietnam	1	1	a	94	55	4	0,030	0,026	0,045	0,019	0,020
Algérie	1	1	a	149	85	7	0,028	0,026	0,036	0,023	0,029
Algérie	5	1	a	630	408	23	échec	0,169	0,237	0,145	0,146
Algérie MQC	19	1	a	1361	922	19	échec	0,312	0,433	0,261	0,234
Monde	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	échec	échec
Chine	3	4	a	2648	1900	45	échec	échec	échec	0,349	0,321

On peut observer les éléments suivants:

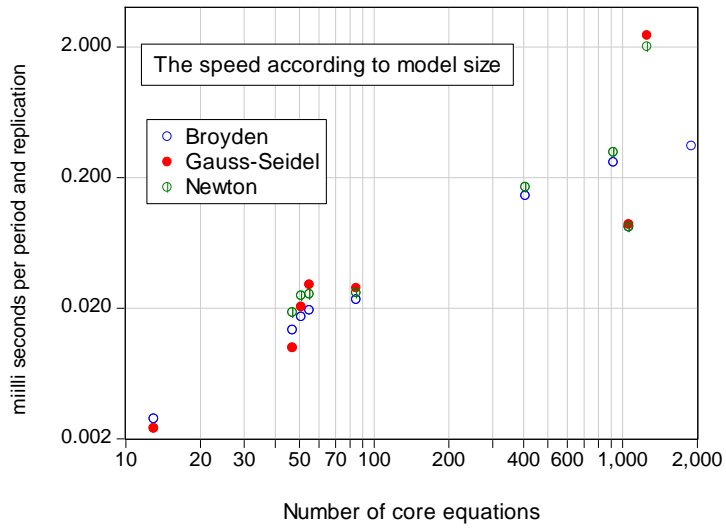
- La mauvaise nouvelle est que chaque méthode échoue au moins une fois, la bonne qu'au moins une fonctionne pour n'importe quel modèle. Si cela était vrai dans tous les cas, la convergence serait garantie.
- Gauss-Seidel fonctionne mieux pour les petits modèles. Ceci est tout à fait contraire à nos attentes, puisque le coût de l'inversion d'une matrice de taille croissante devrait devenir de plus en plus coûteux au fur et à mesure que la taille augmente. Il faut sans doute y voir la présence d'un coût d'entrée pour Newton.
- Broyden est globalement meilleur pour les grands modèles, sans raison particulière. Cependant, il échoue une fois. Mais il est le seul à réussir pour le modèle chinois.
- Pour le grand modèle vietnamien, Broyden échoue, et Newton est meilleur que Gauss-Seidel.

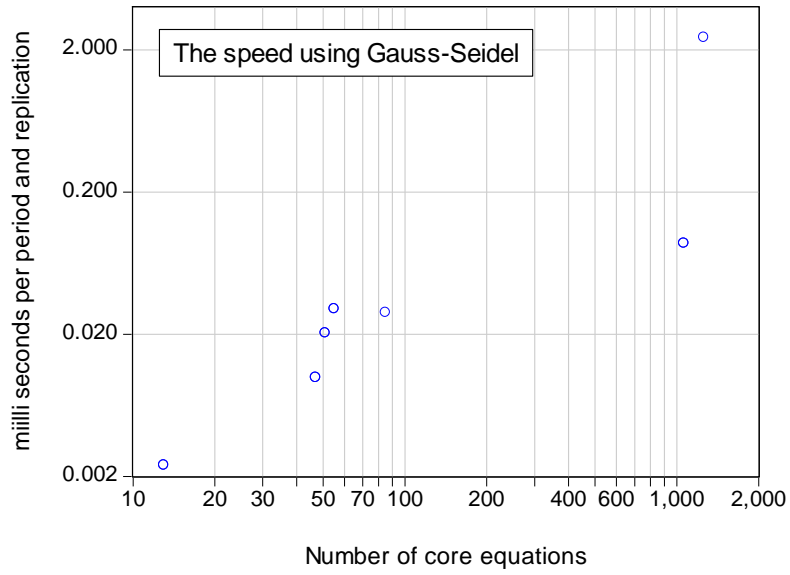
Considérons maintenant le lien entre la taille du modèle et le coût de calcul.

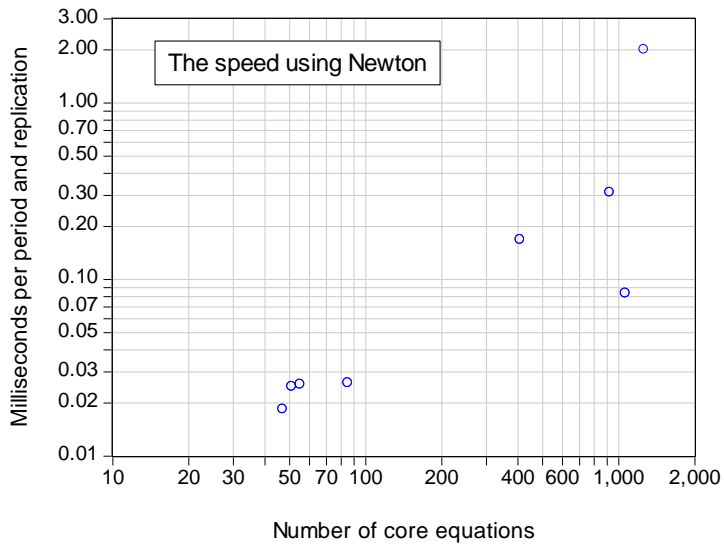
Nous présenterons quatre graphiques, chacun considérant le logarithme des deux éléments:

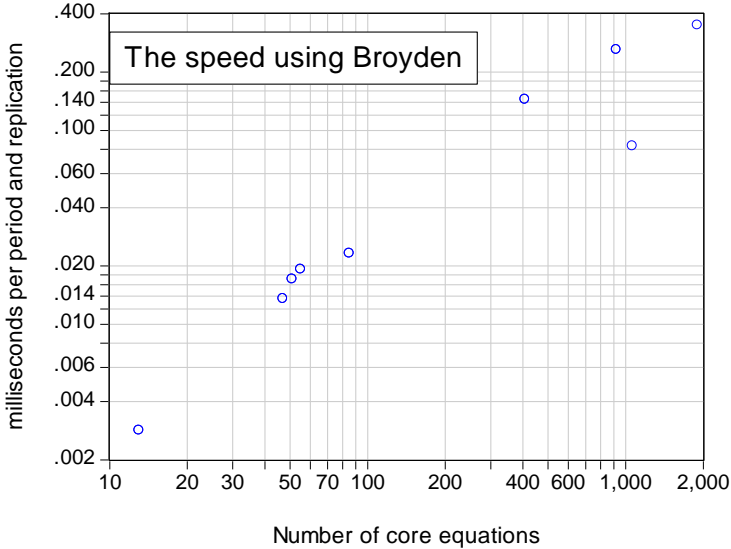
- le premier comparant toutes les méthodes,
- les trois autres consacrés à chaque méthode.

Ils montrent que le lien est globalement log-linéaire, avec une légère inflexion à la baisse pour les méthodes de type «inversion». Quant à l'élasticité elle-même, une estimation séparée donne une valeur proche de 1, et une estimation globale une valeur de .99 (avec une statistique T de 14,7).









Pays	Newton analytique	Newton numérique	Newton ratio	Broyden analytique	Broyden numérique	Broyden ratio
France	échec	échec	échec	0,00	0,00	1,31
France cf	0,02	0,03	1,62	0,01	0,01	1,05
France cd	0,02	échec	échec	0,02	0,02	0,96
Vietnam	0,03	0,05	1,78	0,02	0,02	1,06
Algérie	0,03	0,04	1,40	0,02	0,03	1,26
Algérie	0,17	0,24	1,40	0,14	0,15	1,01
Algerie MQC	0,31	0,43	1,39	0,26	0,23	0,89
Monde	0,08	0,08	0,97	0,08	0,08	0,99
Vietnam	2,01	2,52	1,26	échec	échec	échec
Chine	échec	échec	échec	0,35	0,32	0,92

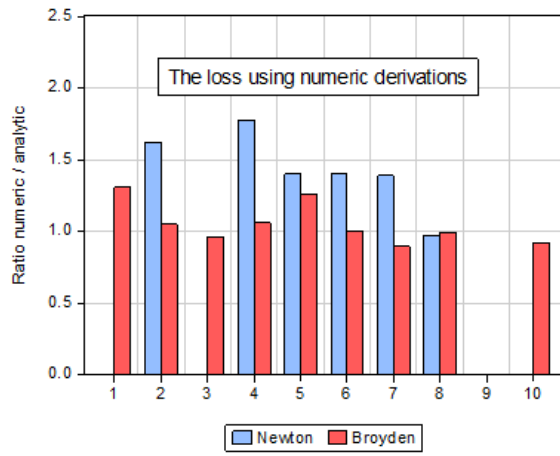
Les deux derniers graphiques comparent les méthodes analytiques et numériques, pour les deux algorithmes associés.

Le diagnostic est complété par un histogramme présentant le rapport entre les coûts des simulations numériques et analytiques (lorsque les deux options ont conduit à la convergence).

- En bleu: Newton,
- En rouge: Broyden.

Ils montrent clairement:

- Que l'établissement des dérivées analytiques est tout à fait efficace pour la méthode de Newton, avec un gain moyen de 40 %.
- Que pour Broyden (où les dérivées sont calculées une fois seulement), il n'y a pas de gain, mais pas de perte non plus.



7.1.6 LA RÉOLUTION DU MODÈLE: LES TECHNIQUES EIEWS ELEMENTAIRES

Le programme précédent a rendu les données du modèle et les équations cohérentes. Le suivant va le résoudre.

7.1.6.1 La syntaxe

Encore une fois, nous allons résoudre le modèle, mais cette fois, les options sont plus nombreuses.

Résoudre le modèle sera généralement réalisé dans un programme. Cependant, on peut utiliser les menus, d'abord en accédant à l'élément «modèle», puis en utilisant l'ordre SOLVE, qui donne accès à toutes les fonctionnalités que nous allons présenter.

Dans un programme, ces options peuvent être établies soit avant tout calcul, en utilisant l'instruction SOLVEOPT, soit au moment de la simulation (SOLVE).

La syntaxe est la suivante:

```
nom-du-modèle.solveopt(options)
```

où les options sont (tirées du fichier d'aide EViews):

m = arg (valeur par défaut = 5 000): Le nombre maximum d'itérations pour trouver une solution (maximum 100 000).

c = arg (valeur par défaut = 1e-8): Critère de convergence. Ce sera la variation maximale relative de chacune des variables endogènes du modèle, d'une itération à l'autre. Vous pouvez définir un nombre entre 10^{-15} et 0,01.

s = arg: Type de solution: «d» (déterministe, valeur par défaut), «m» (stochastique – création des seules moyennes), «s» (stochastique – création des moyennes et écarts-types), «b» (stochastique – création des moyennes, et des intervalles de confiance), «a» (stochastique – création de tous les éléments: moyennes, écarts-types et intervalles de confiance).

d = arg: Dynamique de la résolution: «d» solution dynamique, par défaut, «s» (solution statique), «f» (valeurs ajustées – Solution par équations séparées).

La prévision fait appel à une solution dynamique, qui est l'option par défaut. Il y a peu de raisons d'en choisir une autre (à l'exception de d = f présentée plus haut).

n = arg: Comportement en présence d'observations indisponibles («NA»): «n» (arrêt), «t» (pas d'arrêt). S'applique uniquement à la solution déterministe ; EViews s'arrête toujours s'il rencontre des valeurs «NA» en solution stochastique.

a = arg: Traitement des scénarios alternatifs: «t» (résolution à la fois de l'actif et l'alternatif et stockage des écarts pour une simulation stochastique), «f» (par défaut, résolution du seul scénario actif).

o = arg: Méthode de résolution: «g» (Gauss-Seidel), «n» (Newton), «b» (Broyden).

i = arg: Pour les valeurs d'initialisation: «a» (observées), «p» (solution de la période précédente). La valeur par défaut est «a» si les valeurs sont disponibles, «p» sinon. Dans les prévisions, «p» doit être appliqué, mais on peut initialiser les endogènes, ce qui donne les deux options («a» devient alors la valeur par défaut).

Pour les simulations stochastiques:

- r = nombre entier (par défaut = 1000) Nombre de répétitions stochastiques (utilisé avec l'option «s=»).
- b = nombre (défaut = .95) Taille des intervalles de confiance stochastiques (utilisé avec l'option «s =»).

7.1.6.1.1 Le suffixe

En plus des options ci-dessus, EViews demande à l'utilisateur de définir le nom des variables calculées.

Bien sûr, il ne faut pas envisager d'utiliser les noms réels en tant que tels. Cela remplacerait les valeurs d'origine en détruisant donc l'information correspondante. Cependant, le lien doit être aussi facile que possible. La solution évidente est d'utiliser les noms originaux, mais de les modifier par un ajout: un préfixe ou un suffixe donné.

Il faudra pour cela une déclaration spécifique:

```
nom-du-modèle.assign @all suffixe
```

Cette instruction ajoutera au texte du modèle *nom-du-modèle* l'élément:

```
assign @all suffixe
```

où *suffixe* est une chaîne de caractères.

Par exemple:

```
_fra_1.assign @all _b
```

donnera aux variables obtenues par la prochaine résolution du modèle un nom composé de la variable d'origine et du suffixe «_b». Dans ce cas, la solution pour Q sera appelée Q_B.

7.1.7 LES TECHNIQUES DE RESOLUTION DES PROBLEMES

Comme tous les logiciels de ce type, EViews ne parvient pas toujours à une solution, même si celle-ci existe. Nous allons donner ici une typologie des techniques que l'on peut utiliser pour résoudre les problèmes de convergence, en nous concentrant sur celles qui sont disponibles sous EViews, et sur la façon de les appliquer.

Dans tous les cas, il faut s'assurer d'abord que la vérification résiduelle ne présente aucune erreur. Cela est essentiel, et peu coûteux (au moins dans l'observation des erreurs, et l'identification de leur emplacement).

7.1.7.1 Les outils proposés par EViews

En plus de la présentation des spécifications du modèle, EViews fournit un certain nombre d'outils donnant des informations sur le processus. Certains d'entre eux ne peuvent être spécifiés dans les menus, mais ils restent actifs si l'on exécute dans la même session un programme qui les contient.

- Affichage du nombre d'itérations nécessaires à la convergence (utile uniquement bien sûr si la convergence du modèle a été obtenue au moins sur une partie de l'échantillon). Mais si le modèle commence à diverger après quelques périodes, il est intéressant de savoir si le processus est progressif (le nombre d'itérations est de plus en plus élevé jusqu'à l'échec) ou soudain (le nombre est stable jusqu'à l'échec).

Ceci est obtenu en

- double-cliquant sur l'élément de modèle
- utilisant:

Solve > Diagnostic

- spécifiant: «display detailed messages including iteration count by blocks».

Une fenêtre présentera le nombre d'itérations nécessaires pour résoudre chaque bloc (plusieurs messages par période si le modèle contient plus d'un bloc non récursif).

Cette fenêtre restera ouverte après la résolution, et on peut la faire défiler en arrière, jusqu'au premier message si nécessaire.

```

Model: FRA_1
Date: 11/07/12 Time: 14:41
Sample: 2001S1 2001S2
Solve Options:
Dynamic-Deterministic Simulation
Solver: Gauss-Seidel
Max iterations = 5000, Convergence = 1e-06
Scenario: Scenario 1
Solve begin 14:41:07
2001S1      Block 1 - 3 eqns      Solved (recursive block)
2001S1      Block 2 - 14 eqns     Convergence after 14 iterations
2001S1      Block 3 - 3 eqns      Solved (recursive block)
2001S2      Block 1 - 3 eqns      Solved (recursive block)
2001S2      Block 2 - 14 eqns     Convergence after 14 iterations
2001S2      Block 3 - 3 eqns      Solved (recursive block)
Solve complete 14:41:07

```

- Afficher les valeurs obtenues à chaque itération pour un ensemble de variables.

Ceci est obtenu en

- double-cliquant sur l'élément de modèle
- utilisant:

Solve>diagnostics

- spécifiant des noms des variables dans la fenêtre «Trace variables»

après avoir résolu, en utilisant:

View>trace output

Une fenêtre s'ouvrira avec les valeurs prises par les variables à chaque itération, pour chacune des périodes (une colonne de valeurs par variable).

Si la variable appartient à un bloc récuratif, une seule valeur sera affichée. Dans notre cas, les valeurs uniques pour CAP et X seront précédées de la description du processus itératif, et suivies de la valeur K unique.

La première valeur (itération 0) correspond à la valeur initiale.

Les colonnes suivent l'ordre utilisé dans le processus de résolution, et pas celui dans lequel les équations ont été introduites. Ceci est très utile pour suivre les calculs du processus de Gauss-Seidel.

	CAP	X	UR	Q	I	FD	M	K
0	1477674.	411669.1						
0			0.967825	1430727.	181579.8	1388969.	381013.8	
1			0.968230	1419625.	179655.5	1344976.	370204.8	
2			0.960716	1386440.	177326.4	1339545.	368264.6	
3			0.940504	1383299.	176118.0	1338384.	366440.3	
4			0.936155	1383611.	175753.1	1337710.	365701.2	
5			0.936344	1383678.	175766.0	1337723.	365717.2	
6			0.936389	1383675.	175768.5	1337726.	365721.7	
7			0.936388	1383673.	175768.2	1337725.	365721.5	

8		0.936386	1383673.	175768.1	1337725.	365721.3
9		0.936386	1383673.	175768.1	1337725.	365721.3
0						2085298.
0	1495279.	421079.3				
0		0.968590	1453147.	183168.3	1409358.	386757.6
1		0.971823	1443679.	178010.0	1350618.	372634.4
2		0.965492	1399063.	175065.9	1343571.	370409.1
3		0.938637	1394724.	173433.1	1342002.	367972.8
4		0.932781	1395107.	172935.2	1341078.	366972.9
5		0.933008	1395185.	172950.4	1341093.	366991.3
6		0.933060	1395181.	172953.2	1341095.	366996.5
7		0.933057	1395178.	172952.9	1341095.	366996.1
8		0.933055	1395178.	172952.8	1341095.	366995.9
9		0.933055	1395178.	172952.8	1341095.	366995.9
0						2105233.

Mais l'interprétation des évolutions comme une colonne de chiffres est souvent difficile, en particulier si l'on tente d'évaluer la présence et la nature de cycles. Il est préférable de transformer ces chiffres en graphiques. Cela peut se faire de deux manières:

- Sélectionner avec la souris la zone demandée (par exemple les valeurs d'une période donnée) et la copier dans un logiciel graphique comme Excel.
- La même chose peut être faite en EViews, mais c'est un peu plus complexe. Il faut:
 - sélectionner avec la souris la zone demandée,
 - créer des variables vides associées aux éléments affichés (avec des noms différents, par exemple en insérant un préfixe commun),
 - afficher les variables, comme une feuille remplie de valeurs «NA»,
 - copier la sélection dans la feuille,
 - appliquer View>Graph.
- Éliminer des équations du processus de résolution

Ceci est très facile. En utilisant des menus, il faut

- accéder au modèle,
- utiliser:

Basic options>Edit scenario options>Exclude

- dans la fenêtre, spécifier les séries à exogénéiser.

En utilisant un programme, il faut déclarer

```
nom-du-modèle.EXCLUDE liste-des-séries
```

L'équation correspondante ne sera pas appliquée, et la variable conservera sa valeur initiale.

Remarque: la liste restera active pour toutes les solutions produites ultérieurement dans le même scénario, jusqu'à ce qu'elle soit modifiée à l'aide d'une liste différente. Revenir au modèle complet réclamera une nouvelle déclaration avec une liste vide.

On peut envisager d'éliminer une variable unique, comme les variations de stocks (si elles semblent introduire un processus divergent), ou un ensemble complet, comme tous les déflateurs du modèle (si le système des prix semble exploser, ou perturber le calcul du secteur réel).

Cependant, il faut être conscient que si une variable déterminée par une identité est exclue, l'identité n'est plus prise en compte. En particulier, les variables d'équilibrage ne doivent jamais être exclues. Par exemple, si l'on exclut Q dans notre modèle, l'offre n'est pas égale à la demande (et le modèle devient récursif...).

7.1.7.2 Autres outils

Certains outils peuvent être considérés par l'utilisateur:

7.1.7.2.1 Utiliser des facteurs de relaxation

Ceci a été décrit précédemment. Techniquement, vous devrez réécrire l'équation. Par exemple

$$Q + M + X = FD$$

deviendra:

$$Q = \text{rel_q} * (FD + X - M) + (1 - \text{rel_q}) * Q$$

où la relaxation est appliquée en rendant `rel_q` différent de 1 (plus faible en général).

Il pourrait être utile de définir le facteur sous forme d'une série, même si elle est constante au cours du temps. Pour appliquer une modification à un scalaire, le modèle doit en effet être recompilé (par `UPDATE`).

En outre, cette méthode est un peu différente de celle décrite plus haut. L'application des facteurs ne se fait pas à la fin de l'itération, mais au moment où la variable est calculée. Si la variable est également utilisée après avoir été calculée, cela affectera le processus de convergence.

7.1.7.2.2 Réordonner le modèle

Si l'on utilise Gauss-Seidel, la réorganisation va changer la résolution. Pas pour Newton car le Jacobien est basé sur la dérivation des équations individuelles.

7.1.7.2.3 modifier les valeurs des paramètres

Modifier des valeurs de paramètres va évidemment affecter le processus de résolution. En particulier, un coefficient qui semble trop élevé peut être réduit, voire mis à zéro. Cela pourra être le cas par exemple pour l'intensité de l'effet accélérateur à court terme dans l'équation d'investissement.

Cette technique peut être utilisée pour éliminer une influence spécifique d'une variable, en laissant inchangés ses autres rôles.

Ceci réclame évidemment une bonne connaissance des propriétés du modèle.

7.1.7.2.4 modifier la spécification des équations

Un mécanisme qui semble créer des problèmes peut être remplacé par un autre. Par exemple, l'hypothèse de parité de pouvoir d'achat pour le taux de change peut être lissée ou remplacée par un taux fixe.

7.1.7.3 L'application de ces outils

Nous allons maintenant proposer une méthodologie.

1 - Le modèle converge pour l'ensemble de la période: OK

2 - Le modèle ne converge pas à la première période.

2.1 Le nombre maximum d'itérations a été atteint (le nombre m = après la déclaration SOLVE ou SOLVEOPT, défaut = 5 000).

Ce n'est pas trop mal: faire converger le modèle ne devrait pas être trop difficile, la plupart du temps.

Examiner les variables qui n'ont pas convergé, en sélectionnant certaines d'entre elles (les variables de bouclage dans le cas de Gauss-Seidel) et en affichant leurs valeurs.

Techniquement, cela se fait en:

- accédant au modèle (double-cliquant sur l'élément du modèle dans le fichier de travail),
- sélectionnant «Diagnostic»,
- indiquant les noms des variables dans la fenêtre «Trace variables»,
- résolvant le modèle,
- sélectionnant View>Trace output.

2.1.1 Aucune variable n'a convergé (ou quelques-unes seulement)

Dans le cas de quelques-unes, elles doivent appartenir au prologue.

- Si Gauss-Seidel:

vérifier que nous ne rencontrons pas des oscillations entre deux valeurs très proches (problème purement numérique en raison de la précision des calculs informatiques), auquel cas le seuil choisi est trop petit.

- Dans le cas contraire:
 - rechercher un autre ordonnancement, en vérifiant d'abord la logique économique de l'actuel ;
 - mettre en place un ou plusieurs facteurs de relaxation (calculés par essais successifs) sur les boucles principales ;
 - si Newton (rare): vérifier que toutes les valeurs de départ sont raisonnables ; rechercher d'autres valeurs de départ.

2.1.2 Une seule variable (ou quelques-unes) ne satisfait pas au critère.

Normalement, cela provient de:

- La pure malchance: une variable qui peut changer de signe, ce qui signifie qu'elle est calculée par différence (comme la variation des stocks ou de la balance commerciale) et prend des valeurs beaucoup plus faibles que ses composantes, perdant des chiffres de précision dans le processus. La même chose s'observe pour les taux de croissance, lorsque l'élément associé est stable dans le temps.
- Un petit bloc non récursif ne converge pas: nous verrons que dans les modèles opérationnels le budget de l'État appartient à l'épilogue sous forme généralement récursive, mais il inclut souvent la boucle suivante:

déficit -> intérêts payés -> déficit

qui pourrait ne pas converger.

Remarque: dans ce dernier cas, si le bloc est assez petit, on peut le transformer en un bloc récursif en éditant soi-même les équations.

Dans le premier cas, la situation est frustrante, car nous estimons que nous avons un modèle convergent naturellement, mais avec EViews nous ne pouvons pas:

- adapter le critère à la variable, en niveau et en mode de calcul. Sur les soldes ou les taux de croissance, il sera préférable de contrôler les différences en niveaux, correspondant à la présentation normale de la variable: une variation de l'inflation sera présentée en points, une hausse du déficit public en milliards, tandis que pour un gain sur le PIB on pensera à des pourcentages.
- adapter le critère à la période: préciser la valeur pour la période de non-convergence et revenir à la normale par la suite.

Donc on se retrouve avec les méthodes de 2.1.1

2.2 Le modèle ne peut être résolu en raison de données manquantes

De toute évidence, les données devraient être disponibles. Dans l'exemple suivant, l'estimation n'a pas été effectuée sur l'échantillon complet, et le résidu RES_M n'a pas été prolongé. La détermination de M ne peut se faire.

Model: _FRA_1

Date: 11/07/12 Time: 16:15

```
Sample: 2003S1 2050S2
Solve Options:
Dynamic-Deterministic Simulation
Solver: Gauss-Seidel
Max iterations = 5000, Convergence = 1e-006
Scenario: Scenario 1
Solve begin 16:15:25
2003S1 NA generated for M
2003S1 NA generated for Q
2003S1 NA generated for I
2003S1 NA generated for LED
2003S1 NA generated for LE
2003S1 NA generated for LT
2003S1 NA generated for RHI
2003S1 NA generated for IH
2003S1 NA generated for CO
2003S1 NA generated for IC
2003S1 NA generated for FD
2003S1 NA generated for UR
Solve terminated - Unable to compute due to missing data in "UR = Q / CAP"
```

2.3 L'interruption des calculs, car certaines variables atteignent une valeur inacceptable (par exemple l'argument d'un logarithme prend une valeur négative).

EViews fournit à l'utilisateur un message indiquant le lieu où l'erreur survient, par exemple:

```
Log of non positive number in DLOG(LE) = C_LE(1) * DLOG(LED) + C_LE(2) * LOG(LED( -
1) / LE( - 1)) + C_LE(3) + ec_le
" in "SOLVE(N=T M=10001 C=1E-6 O=G D=D I=A) _FRA_1"
```

Tout d'abord, cela ne veut pas dire que l'équation elle-même est en cause: si l'équation de l'emploi échoue suite à un argument négatif dans un logarithme, cela signifie que l'emploi cible LED est négatif, ce qui peut venir que du fait que Q (le PIB) est lui-même négatif. Comme il est calculé par:

$$[4] Q = M + X + FD$$

cela peut signifier que FD ou X est trop faible ou négatif, ou M trop élevé.

Remonter la chaîne des causalités n'est pas toujours simple, cependant, et parfois nous ne recevons aucune information du tout, comme avec Newton:

```
near singular matrix in "SOLVE(N=T M=10001 C=1E-6 O=N D=D I=P) _FRA_1"
```

Dans la pratique, il faut d'abord regarder le nombre d'itérations effectuées avant l'échec.

2.3.1 S'il est petit (2, 3 ou même un seul) avec Gauss-Seidel, ou 1 pour Newton, il s'agit probablement d'un problème technique, car un modèle apparemment acceptable ne devrait pas se comporter si mal.

Les sources de problèmes techniques peuvent être:

- Votre modèle est en effet très faux. Pour Gauss-Seidel, les valeurs propres sont très supérieures à 1⁷¹. Pour Newton, le Jacobien conduit dans une direction très mauvaise (peut-être aussi parce que le point de départ n'est pas réaliste).
- Votre modèle ne peut pas répondre à des valeurs anormales de l'endogène. Par exemple, vous prenez le logarithme d'une variable que vous vous attendez à voir positive, mais à cause de mauvaises propriétés du modèle elle est calculée comme négative. L'investissement est un bon candidat ici, si le PIB baisse trop.

$$I/K(-1)=0.825*I(-1)/K(-2)+0.0279*UR+0.152*.25*Q/Q(-4)-0.0525+EC_I$$

⁷¹ Par exemple, si un mécanisme de correction d'erreur est fort, et son signe est erroné.

- Votre modèle n'est pas compatible avec les données. Bien sûr, vous avez effectué un contrôle résiduel, mais peut-être que quelque chose a changé depuis. Il faut effectuer le contrôle à nouveau.

Les autres cas s'appliquent surtout à des simulations en dehors de la période de l'échantillon:

- Les hypothèses exogènes ont reçu des valeurs anormales.

Par exemple, avec $WD = 0$, on obtient:

```
Solve terminated - Log of non positive number in "DLOG(X) = @COEF(1) * DLOG(WD) +
@COEF(2) + @COEF(3) * RES_X(-1) + EC_X"
```

Les valeurs initiales sont anormales, et les calculs échouent avant que les valeurs n'atteignent une zone acceptable.

Nous pouvons afficher les valeurs associées des variables. Au-delà de la variable erronée elle-même, on s'intéressera à chacune de ses déterminants, comme la source logique du problème sera souvent située loin en amont de l'équation diagnostiquée. Un nouvel exemple: un taux de chômage négatif (utilisé dans un logarithme) pourrait provenir d'une surestimation de la demande donc de la production et du niveau de l'emploi lui-même (qui devient supérieur à la force de travail).

Pour résoudre le problème, il faudra dans ce cas, retracer l'enchaînement causal.

Par exemple, si M a été initialisé à 10^{15} , on obtient:

```
Solve begin 18:07:52
2003S1      Block 1 - 3 eqns      Solved (recursive block)
2003S1      Block 2 - 13 eqns     No convergence after 3 iterations
Solve terminated - Log of non positive number in "DLOG(LE) = @COEF(1) * DLOG(LED) +
@COEF(2) * LOG(LED(-1) / LE(-1)) + @COEF(3) ..."
```

Et nous pouvons suivre l'évolution des variables dans la fenêtre de «Trace output»:

Date	Block	Iteration	X	UR	Q	LED	LE	FD	M
2003S1	0	0	411669.1						
2003S1	1	0		0.964223	-2.50E+14	-8.11E+14	18608119	1.23E+12	1.00E+15

2003S1	1	1	-1.69E+08	-9.99E+14	-1.30E+16	18608119	4.42E+12	1.00E+15
2003S1	1	2	-6.76E+08	-9.96E+14	-1.29E+16	18608119		

Pour Gauss-Seidel, une méthode efficace (non fournie par EViews) consiste à passer à l'algorithme de Ritz-Jordan, qui ne prend en compte que les valeurs de l'itération précédente, ce qui élimine les interférences et permet d'améliorer l'individualisation des processus divergents. Au début d'un processus, cette méthode identifie immédiatement l'équation défectueuse (moins efficacement qu'une vérification résiduelle, cependant).

2.3.2 Si le nombre d'itérations avant l'échec est plus élevé, deux situations sont à considérer:

A: Le modèle n'a pas de solution.

B: Le modèle a une solution que le logiciel n'est pas en mesure d'atteindre.

Avec deux options pour ce dernier, dans le cas de Gauss-Seidel:

B1: Vous avez ordonné les équations de la même manière que le processus d'équilibrage qu'il décrit.

B2: L'ordre que vous avez choisi n'est pas cohérent avec ce processus.

Malheureusement, il n'existe aucun moyen de séparer les cas. Le premier conseil est simple et économique à appliquer: essayez l'autre algorithme !

Il y a une bonne chance que:

- Newton ne converge pas en raison de colinéarités dans la matrice jacobienne, ou de valeurs initiales trop éloignées de la solution, ce qui rend inefficace la linéarisation. Cela n'affecte pas trop Gauss-Seidel.
- La matrice jacobienne associée à Gauss-Seidel contient des valeurs propres de module supérieur à 1, ce qui n'est pas un problème pour Newton.

Cela pourrait effectivement résoudre votre problème. Toutefois, dans le cas B1, vous devriez vous préoccuper de la stabilité de l'équilibre que votre modèle décrit.

Si cela ne fonctionne pas, la solution est partiellement différente pour les deux algorithmes:

- Pour les deux algorithmes, vous pouvez débrancher une ou plusieurs équations, et essayer de résoudre à nouveau le modèle. Le choix dépendra des variables qui semblent conduire à la divergence: soit celles qui prennent réellement les pires valeurs, ou les éléments qui les influencent. En forçant une variable à prendre une valeur donnée, on peut souvent mesurer sa contribution au problème (surtout si la variable passe par des valeurs inacceptables) en

observant si le problème disparaît. Cette méthode a cependant la même limite que les autres: l'exogénéité de la variable détruit toutes les influences transmises par celle-ci, et faire disparaître le message d'erreur ne signifie pas que nous avons exogénéisé l'équation défectueuse.

- Vous pouvez également modifier les valeurs des paramètres, ou éliminer une influence individuelle en mettant son coefficient à zéro. Ceci est différent de l'élimination complète de la variable: dans notre modèle, vous pouvez supprimer l'influence du taux d'utilisation sur les exportations, mais la maintenir pour les importations. En modifiant certains coefficients stratégiques, choisis soit parce qu'ils jouent un rôle important dans le modèle (dans le cas d'une absence générale de convergence), ou parce qu'ils semblent liés à l'erreur diagnostiquée (dans le cas d'une erreur locale), on pourrait établir un lien entre ces coefficients et l'intensité du problème. Cependant cette méthode, une fois de plus, ne localise pas nécessairement l'origine du problème, qui peut venir de loin en amont. Et il reste à déterminer comment utiliser l'information, si l'on exclut la fixation arbitraire du coefficient à la valeur assurant la convergence.
- Vous pouvez essayer de substituer une autre formulation (probablement plus simple⁷²). Par exemple, vous pouvez remplacer une fonction complexe de type CES utilisant une détermination sophistiquée du coût relatif des facteurs, par une fonction très simple à facteurs complémentaires. Bien sûr, vous pouvez commencer par exogénéiser le coût relatif.
- Si l'argument d'un logarithme prend une valeur négative, vous pouvez introduire un test pour éviter cette situation (par exemple en remplaçant $\text{Log}(x)$ par $\text{Log}(\max(x, a))$, avec $a > 0$. Le système pourrait en fait converger vers une valeur acceptable ($x > 0$)⁷³, mais trouver une solution avec $x < 0$ vous donnera des informations de toute façon⁷⁴.
- Pour Newton, vous pouvez tester d'autres points de départ. Vous devriez essayer de deviner aussi bien que possible le domaine dans lequel réside la solution, et initialiser les variables à l'intérieur de celui-ci. Cela devrait être assez facile sur le passé.
- Vous pouvez vérifier les dérivées de vos équations. Vous pouvez le faire par un examen visuel ou par des calculs (un peu plus difficile, mais EViews fournit les outils élémentaires). Si vous commencez le processus près de la solution, mais que la pente est mauvaise, vous pouvez partir dans la mauvaise direction, peut-être par une large marge.

⁷² Si le modèle que vous souhaitez produire est assez complexe, Cela pourrait être une bonne idée de commencer avec une version plus simple, et d'introduire progressivement des niveaux de complexité.

⁷³ Une situation rare dans laquelle l'étude de l'erreur résout le problème.

⁷⁴ Lorsque les choses tournent mal, les moyens de résoudre le problème sont bien sûr tout à fait différents si vous avez une solution (mauvaise) ou rien du tout. Cela signifie qu'il est généralement fondamental de réussir à parvenir à une solution, même dans un cadre non acceptable.

- Pour Gauss-Seidel, introduisez des facteurs de relaxation. Techniquement, vous devrez réécrire l'équation.
- Si vous avez essayé de le faire, mais les choses ont empiré, il y a une bonne chance que votre modèle ne converge pas naturellement pour des raisons économiques, mais il peut aussi n'avoir aucune solution.
- Mais en tout cas, le changement apporté ne doit pas réduire la validité théorique du modèle: on ne cherche pas à imposer une solution grâce à une formulation *ad hoc*, mais plutôt à corriger une incohérence.
- Enfin, vous pouvez modifier la période de simulation (dans tous les cas). Il est intéressant de savoir si le problème est général, ou n'apparaît que pour certaines périodes. L'examen des aspects particuliers des périodes défectueuses (croissance lente, forte inflation, chômage élevé ...) peut donner des informations utiles sur les origines de l'erreur. Son utilisation peut cependant être difficile. Un cas particulier est le rôle du commerce extérieur: dans les prévisions, une expansion trop rapide par rapport à la production et à la demande locales peut conduire à des valeurs anormales du multiplicateur, qui disparaîtront si nous simulons le modèle sur les premières périodes disponibles.

3 Le modèle converge pendant un certain temps, puis se bloque.

Cette situation a des bons et des mauvais côtés:

- Le modèle n'a pas de défauts techniques fondamentaux, sauf si vous avez introduit une erreur dans les hypothèses futures, ou utilisé une formulation qui se dégrade au fil du temps (une variable muette avec un mauvais coefficient, une condition mal formulée qui ne peut se produire que dans le futur, par exemple quand le commerce extérieur atteint un trop grand rôle ...).
- Mais ces défauts non techniques sont les plus difficiles à résoudre. Votre modèle peut décrire un processus d'équilibrage qui converge au début, mais que la dynamique des évolutions rend divergent après un certain temps. Par exemple, l'effet de l'expansion économique sur le multiplicateur keynésien peut être plus élevé que l'unité, ce qui rend le processus explosif (la même chose peut se produire avec la boucle prix-salaires). Ou le processus de correction d'erreur augmente en fait l'erreur, même si le coefficient a le bon signe.

Prenons l'exemple d'une voiture: si elle ne démarre pas, vous pouvez avoir oublié de mettre de l'essence, ou la batterie pourrait être à plat. Ceci peut être résolu facilement. Si elle démarre mais ne va pas assez vite, ou vire à gauche systématiquement, le problème pourrait être bien pire (et plus difficile à localiser).

- Il y a aussi moins de chance que le modèle soit vraiment mauvais. Et vous pouvez toujours recommencer sous des hypothèses nouvelles. Le moteur de la voiture peut exploser, à un coût de remplacement élevé. Pour un modèle, vous pouvez recréer une nouvelle version gratuitement, et répéter le processus indéfiniment.

Mais la meilleure nouvelle est la suivante :

- Nous obtenons plus d'informations

Nous pouvons :

- Observer la façon dont le modèle diverge, et quelles variables (ou ensembles) semblent en être à l'origine. Par exemple, les variables à prix constants pourraient suivre une évolution acceptable, et les prix exploser. Ou le chômage peut exploser alors que l'emploi est acceptable, entraînant les salaires et la répartition de la valeur ajoutée entre les travailleurs et les entreprises
- Nous pouvons voir si la divergence est monotone ou cyclique, régulière ou irrégulière, soudaine ou progressive.
- Nous pouvons choquer des hypothèses et en examiner les conséquences. Elles peuvent être acceptables au début, ce qui signifie que le problème réside dans la dynamique. Elles peuvent se dérégler d'une manière régulière, ou exploser soudain.

Par exemple, une demande croissante de main-d'œuvre dans le processus de production peut rendre le chômage négatif avec le temps, alors que son logarithme est utilisé dans l'équation de salaire⁷⁵. Ou, comme nous l'avons vu précédemment, un niveau accidentellement très faible de l'investissement peut soudain faire exploser l'équation associée. Ou une augmentation des exportations beaucoup plus lente que celle des importations peut rendre à long terme le PIB négatif.

Tout cela peut être combiné avec l'exclusion d'équations, afin de mettre en évidence la propagation d'une erreur introduite par un mécanisme spécifique.

On devine déjà que l'interprétation de cette information sera plus facile avec une simulation longue, à partir d'une base suivant une trajectoire régulière. En d'autres termes, une prévision, ou plutôt une simulation sur le futur.

- Enfin, le calcul des valeurs propres de $((I - \hat{\alpha}_y / \hat{\beta}_y)^{-1})$ lorsque l'on utilise l'algorithme de Gauss-Seidel peut aider à comprendre le processus de convergence, et mettre en évidence les caractéristiques mathématiques d'une divergence possible (intensité et le nombre de processus divergents, présence de cycles). Le traitement de cette information peut se faire par une élimination progressive des influences de bouclage, analogue à celle qui est recommandée par Deleau et Malgrange (1978) pour étudier les propriétés dynamiques des modèles. Dans le meilleur des cas, une association sera obtenue entre chaque processus divergent et une variable

⁷⁵ Par exemple, remplacer le logarithme par un niveau ne change pas la situation car un chômage négatif est évidemment inacceptable.

de bouclage unique (ou deux pour un couple de valeurs propres conjuguées, à l'origine d'un processus cyclique).

L'interprétation et la solution d'un problème de convergence ne sont donc pas simples, essentiellement à cause des interférences dues à la nature interactive du système. Il sera souvent nécessaire d'utiliser plusieurs méthodes, et de répéter chacune d'elles à plusieurs reprises, ce qui ne dispensera pas d'une réflexion approfondie sur la structure logique du modèle. Il est donc essentiel pour le constructeur de modèles de connaître son outil, ce qui sera d'autant plus difficile que sa taille est importante.

7.2 UNE PREMIERE VALIDATION

Bien que la convergence d'un modèle ne soit pas sans lien avec sa validité théorique, elle ne suffit pas pour l'assurer. Pour le faire, un certain nombre de méthodes plus ou moins pertinentes sera appliqué.

7.2.1 SIMULATIONS EX POST

Nous voulons d'abord savoir si le modèle est capable d'effectuer des prévisions réalistes. Pour cela, deux techniques sont disponibles: les simulations ex post (ou prévisions ex post) sur le passé, et les simulations stochastiques, généralement sur l'avenir. Pour l'instant, nous nous limiterons à la première.

Pour répondre à la question, nous avons un problème: nous ne connaissons pas l'avenir (si c'était le cas, nous n'aurions pas besoin du modèle de toute façon), il faut donc s'appuyer sur le passé, qui fournit des informations limitées. Nous ne sommes pas dans la situation d'un scientifique qui peut produire des relations aussi exactes qu'il désire en répétant les expériences.

Donc, nous allons effectuer des simulations sur le passé, mais placer le modèle dans une situation aussi proche que possible de celles qu'il rencontrera dans l'avenir. Et il y a un élément qui ne sera certainement pas à sa disposition: les résidus des équations estimées, en d'autres termes la partie inexpliquée des comportements.

Ce que l'on connaît, c'est leurs valeurs les plus probables, la moyenne de leur distribution: zéro.

Donc, nous allons simuler le modèle sur le passé, avec des résidus nuls. Ce test prend le nom de simulations «ex post».

Elles sont effectuées sur l'ensemble complet des périodes utilisées pour l'estimation des coefficients: il peut sembler que plus les résultats de la simulation seront proches des valeurs observées, meilleur devrait être le modèle.

On compare ensuite les résultats du modèle à la réalité, à travers des tableaux de résidus, des graphiques montrant les deux courbes, ou des résumés statistiques.

Dans la pratique, on peut effectuer :

- des simulations statiques, qui utilisent des valeurs historiques des variables retardées. Nous aurons en fait une séquence de simulations indépendantes à horizon d'une période ;
- des simulations dynamiques, qui s'appuient sur les résultats de la simulation des périodes précédentes (sauf bien sûr au début).

En général, ce dernier critère sera favorisé, car il est plus représentatif de l'utilisation future du modèle. On peut aussi supposer qu'il est plus difficile d'y répondre, mais rien ne le prouve réellement.

En plus de l'observation visuelle des erreurs (facilitée par une présentation graphique) les informations peuvent être synthétisées en utilisant les critères suivants :

- erreur absolue moyenne,
- pourcentage absolu moyen d'erreur,
- écart-type (racine carrée de l'erreur quadratique moyenne),
- écart-type divisé par la moyenne des observations,
- biais.

On commence généralement par l'un des quatre premiers, que nous allons maintenant présenter. Soient

\hat{x}_t la valeur simulée, et x_t la valeur observée.

L'erreur quadratique moyenne :

$$\text{RMSE} = \sqrt{1/T \sum_{t=1}^T (x_t - \hat{x}_t)^2}$$

L'erreur absolue moyenne :

$$\text{MAE} = 1/T \sum_{t=1}^T |x_t - \hat{x}_t|$$

Le pourcentage d'erreur absolue moyen:

$$\text{MAPE} = 100 / T \sum_{t=1}^T |(x_t - \hat{x}_t) / x_t|$$

L'erreur quadratique moyenne, normalisée:

$$\text{RMSE} = \sqrt{1/T \sum_{t=1}^T (x_t - \hat{x}_t)^2} / \bar{x}$$

Le choix entre un critère en niveau ou en valeur relative viendra comme d'habitude de la nature de la variable. En général, le critère relatif sera plus pertinent (et plus parlant): une erreur de 1 % sur la consommation des ménages (CO) est plus facile à interpréter qu'une erreur de 3 milliards de dollars base 2000. Le critère absolu sera utilisé pour les éléments présentant une forte variabilité, surtout si leur signe n'est pas déterminé: l'erreur sur la balance commerciale sera mesurée naturellement en milliards d'unités courantes (Euros ou Dollars US), mais il sera préférable de toute façon de la normer (par le PIB observé par exemple). Le niveau peut également être utilisé si la notion de relativité est déjà présente dans la variable elle-même (taux de croissance, ratios): l'erreur absolue sur le «taux de croissance des salaires» donne la précision relative sur le salaire lui-même. L'erreur sur le taux de profit sera également mieux interprétée en points⁷⁶.

En ce qui concerne le choix entre les deux types de moments: premier et deuxième ordre, il sera souvent forcé, les logiciels fournissant en général une seule des deux statistiques. Il faut noter que:

- le premier moment est toujours inférieur au second, lorsque la formule est du même type (inégalité de Cauchy-Schwartz);
- cette différence est d'autant plus importante que les erreurs individuelles sont de taille inégale: une simulation contenant une ou deux observations particulièrement mal simulées sera fortement affectée par le second critère.

Enfin, une autre question apparaît, dans le cas des simulations dynamiques des quantités ou en particulier des indices: doit-on considérer l'erreur entre les niveaux simulés et observés de la variable, ou l'erreur additionnelle introduite par la simulation de la période? En ce qui concerne les indices, doit-on considérer

⁷⁶ Nous revenons à la même notion que d'habitude: pour choisir le meilleur critère, il suffit de considérer la manière dont la variable (et ses modifications) sont présentées normalement.

l'erreur sur le niveau ou le taux de croissance ? Dans le cas d'un indice de prix, les deux peuvent avoir un sens: dans les prévisions économiques, ou lorsqu'il s'agit de juger de l'efficacité d'une politique économique, on utilise généralement le niveau de l'inflation, mais l'évolution cumulée sur les niveaux de prix est le facteur déterminant de la compétitivité des prix et du pouvoir d'achat.

Et le résultat peut être complètement différent, comme l'exemple suivant va le montrer.

La simulation de l'indice des prix p , à partir de $t = 1$ à T (supposons T pair), utilise une formulation économétrique du taux de croissance, où la formule:

$$p_t = p_{t-1} \cdot (1 + f_t) + e_t$$

peut être affecté par deux ensembles d'erreurs

$$(1) \quad e_1 = a, e_T = -a, e_t = 0 \quad \text{pour les autres valeurs de } t.$$

$$(2) \quad e_t = a \quad \text{quand } t \text{ est pair,} \quad e_t = -a \quad \text{quand } t \text{ est impair}$$

la somme des erreurs est égale à zéro dans les deux cas.

Nous ne discuterons pas la façon dont ces résultats ont été obtenus: la première estimation semble biaisée, au moins pour les petits échantillons, tandis que la seconde doit être corrigée de l'autocorrélation. Nous devons admettre que les problèmes que nous allons rencontrer viennent au moins partiellement d'estimations erronées.

Le critère «erreur absolue relative moyenne» donnera les résultats:

	Erreur sur le niveau	Erreur sur le taux de croissance
cas (1)	$a(T-1)/T$	$2a/T$
cas (2)	$a/2$	a

L'ordre s'inverse fortement selon le critère utilisé.

En termes plus généraux, la question est la suivante: est-ce qu'une simulation qui présente un écart durable à la réalité observée, mais une tendance relativement stable parallèlement à celle-ci, est meilleure qu'une simulation qui oscille autour des vraies valeurs, avec le même intervalle moyen que la première⁷⁷ ?

Plus généralement, la différence entre les deux critères vient du caractère dynamique du modèle. Lors d'une période donnée, l'erreur provient de l'accumulation des éléments présents et passés, et l'impact de ces derniers se réduit en principe avec le temps.

La croissance de la variance de l'erreur globale est réduite par le fait que la somme des résidus d'origine est nulle, si ceux-ci sont positifs et négatifs en alternance. Cependant, cette condition n'est remplie que sur l'ensemble de l'échantillon. Il est tout à fait possible, même en l'absence d'autocorrélation apparente, que la somme courante ait une valeur significativement positive ou négative.

Il n'est donc pas surprenant pour des équations hautement dynamiques que la plus petite erreur de simulation soit assez souvent trouvée à la dernière période, et la plus grande erreur au milieu.

Mais dans certains cas, une accumulation d'erreurs peut entraîner la variable loin de sa valeur observée. Prenons l'exemple suivant, tiré d'une simplification du modèle DMS.

Le taux de salaire dépend du prix à la consommation (en taux de croissance) et du taux de chômage (chômage divisé par la population active totale).

$$tx(W_t) = a \cdot tx(PC_t) - b \cdot UN_t / (L_t + UN_t) + c$$

Le taux de chômage diminue avec la création d'emplois, avec une intensité plus faible:

$$\Delta UN_t = -d \cdot \Delta L_t$$

⁷⁷ Nous préférons le premier cas.

Les deux équations sont estimées et donc sujettes à erreurs. Supposons que la deuxième erreur n'est pas auto corrélée. Sur UN l'erreur sera la somme courante des erreurs non corrélées: elle peut très bien garder le même signe sur la période, par exemple, si les premières erreurs sont majoritairement positives. En fin de compte l'erreur s'annulera, mais elle aura introduit sur W lui-même une erreur cumulée positive qui peut faire diverger sensiblement son niveau de sa valeur historique⁷⁸.

7.2.1.1 Réévaluer ce critère

En fait, le critère n'est pas vraiment pertinent, et son utilisation est très discutable, pour une raison évidente: la réalité numérique que l'on a cherché à approcher était connue à l'estimation. Elle a été utilisée pour déterminer les valeurs des coefficients, mais aussi pour choisir parmi les formulations possibles. Même un constructeur de modèles scrupuleux sera conduit parfois à favoriser les équations qui donnent des résultats proches de la réalité, au détriment de la qualité théorique.

Un cas extrême est l'introduction de variables muettes, prenant la valeur 1 pour une ou plusieurs périodes, et zéro ailleurs: elles permettent à l'équation de s'adapter parfaitement à un point assez mal expliqué. On ne voit pas pourquoi cette introduction, qui améliore le critère, pourrait améliorer le modèle et ses propriétés.

Il est clair dans ce cas que l'explication supplémentaire n'ajoute rien au modèle, mais plutôt masque les problèmes (d'autant plus que ces variables sont attachées aux périodes les plus mal expliquées). Elles ne peuvent être acceptées que si la période présente une caractéristique spécifique, impossible à traduire en un élément d'explication ou de comportement (comme pour le cas de l'année 1968, que nous avons vu plus haut). Mais, même alors, on court le risque, en éliminant les fluctuations principales de la série expliquée, de rendre l'interprétation de son comportement plus difficile. Et le problème de la prévision de ces variables demeure.

Plus généralement, il n'y a pas de lien strict entre la proximité des résultats de la formule à la réalité observée et sa qualité intrinsèque (on peut observer que des estimations très précises peuvent fournir des coefficients aux valeurs tout à fait inacceptables par une théorie quelconque, par exemple présentant le mauvais signe, tout en étant significatifs). Un modèle produit en sélectionnant systématiquement les formulations donnant la plus petite distance par rapport à des valeurs observées, au détriment possible

⁷⁸ Bien sûr, cela n'arriverait pas avec des formulations à correction d'erreur.

de la signification économique, pourrait donner de bonnes simulations ex post⁷⁹. En poussant le raisonnement à l'extrême, on peut se retrouver avec un modèle décrivant exactement la réalité, à condition que suffisamment de variables aient été ajoutées⁸⁰.

Mais un modèle déterminé en fonction de ces principes échouera probablement à l'épreuve la plus difficile des chocs analytiques, que nous allons présenter plus tard.

Il est néanmoins clair qu'un modèle ne parvenant pas à décrire les évolutions passées devra être rejeté ou moins corrigé (mais par des contributions plus théoriques que techniques), car la qualité de ses prévisions devrait être encore plus faible. Passer le test présenté est donc nécessaire, mais non suffisant, pour garantir la qualité du modèle.

Intéressons-nous à la formation des erreurs de simulation, venant des résidus (nous examinerons plus tard les conséquences des erreurs sur les coefficients et sur la forme réelle des équations). Ils ont trois sources:

- Les erreurs d'estimation originales: il est clair que plus les erreurs d'estimation sont faibles, meilleure sera l'adaptation attendue des résultats de simulation à la réalité observée.
- Les effets multiplicateurs des liens entre les variables instantanées. En tenant compte pour les variables explicatives d'éléments affectés par des erreurs d'estimation (directement ou indirectement) le nombre de sources d'erreurs va augmenter et aussi dans la plupart des cas l'erreur globale (en l'absence de corrélation, la variance d'une somme est la somme des variances).

Ce n'est pas garanti cependant car certaines erreurs peuvent se compenser par un pur hasard, et certains mécanismes du modèle peuvent jouer un rôle modérateur. Par exemple, une erreur positive sur la demande verra son effet sur la production amorti par l'augmentation des importations qu'elle va générer. Ce sera le cas dans notre modèle:

⁷⁹ Bien que il soit fort possible que dans un modèle incohérent les liens instantanés et dynamiques entre les variables exercent une forte influence sur l'expansion des erreurs d'estimation.

⁸⁰ Mais la petite explication donnée par les dernières variables introduites devrait leur donner des statistiques d'acceptabilité très faibles.

$$[14] \text{FD} = \text{CO} + \text{IP} + \text{IC} + \text{IH} + \text{gd}$$

$$[15] \text{TD} = \text{FD} + \text{CI}$$

$$[16] \text{DLOG}(\text{M}) = 1.2 * \text{DLOG}(\text{TD}) + 0.282 * \text{DLOG}(\text{UR}) - 0.212 * \text{RES_M}(-1) - 0.629 + \text{EC_M}$$

$$[17] \text{Q} + \text{M} = \text{FD} + \text{X}$$

Une erreur positive sur IP et FD augmentera M, ce qui limitera l'impact sur Q.

- L'erreur dynamique (dans les simulations dynamiques), venant de variables retardées qui sont désormais le résultat d'une simulation. L'erreur devrait croître à chaque période simulée, avec le nombre de sources d'erreurs. Toutefois, si nous imposons à toutes nos équations estimées une structure à correction d'erreurs, le problème est partiellement résolu: si une variable s'éloigne d'une valeur cible, le comportement des agents du modèle va la ramener à cet objectif (il n'est pas surprenant qu'un modèle «à correction d'erreurs» corrige les erreurs mieux que d'autres). Avec ces modèles, les simulations à moyen et long terme peuvent effectivement donner de meilleurs résultats qu'à court terme et moyen terme, et les simulations ex post peuvent être plus précises à la fin qu'au milieu⁸¹.

Nous allons améliorer notre compréhension des implications des correction d'erreur lorsque nous aborderons les simulations stochastiques.

7.2.1.1.1 Essais partiels

Pour mieux interpréter les erreurs de simulation, certaines de leurs sources peuvent être éliminées

- En exogénéisant des variables

Les variables prennent leurs valeurs exactes, ce qui devrait réduire l'erreur globale. La mesure de la diminution donne une idée de la contribution de la variable à l'erreur globale.

⁸¹ Sur le passé, cela est favorisé par le fait que la somme des erreurs historiques est égale à zéro.

Cependant, cette technique va mélanger deux effets, dans le cas d'une variable de comportement: les suppressions de l'erreur d'estimation et des erreurs de modélisation qui transitent par cette variable. Elle sera ainsi plus difficile à interpréter. Dans le cas d'une identité, l'évolution de l'erreur fera apparaître plus clairement la contribution de la variable.

On peut considérer:

- l'élimination d'une seule variable (ou d'un ensemble réduit) et l'observation de la diminution de la valeur des résidus par rapport au modèle complet;
- l'élimination d'un ensemble important de variables (par exemple tous les prix) et l'observation des conséquences sur les mécanismes du modèle (prix exogènes, absence de l'impact des prix et des bouclages transitant par les prix) .

Avec EViews ces opérations se réalisent simplement par l'instruction EXCLUDE.

- En réintroduisant les résidus (à partir d'un ensemble totalement nul)

Cela se fera en réintégrant les résidus d'estimation avec leurs valeurs observées, dans deux directions:

- maintien à zéro de chaque résidu, sauf pour une équation. La réduction (probable) de l'erreur montrera les gains à attendre de l'amélioration de l'équation ;
- retour de chacun à sa valeur historique, sauf pour une seule équation. L'erreur restante montrera comment la dynamique du modèle élargit l'erreur d'estimation originale.

Bien sûr, des combinaisons de variables peuvent également être utilisées: par exemple on observera l'erreur globale provenant du commerce extérieur, ou l'erreur restante si tous les prix n'introduisent aucune erreur résiduelle ...

Avec EViews, nous pouvons simplement mettre les résidus demandés à zéro (pour éviter de perdre leur valeur estimée, ils peuvent être stockés dans une variable intermédiaire). Cependant, il existe une bien meilleure technique, qui utilise l'option SCENARIO. Nous allons la présenter sous peu.

- À l'aide de simulations statiques, et en les comparant aux dynamiques

Ceci permet de séparer les effets de l'erreur de première période de ses effets dynamiques.

7.2.2 LES PREVISIONS EX POST

Pour faire perdre au critère ci-dessus son caractère artificiel, il semble plus approprié de simuler le modèle sur une période qui n'a pas été utilisée pour l'estimation. Cela peut se faire en excluant de l'échantillon d'estimation les dernières périodes connues, et en les utilisant pour contrôler la qualité des «prévisions».

Mais cela introduit plusieurs problèmes:

- Cela réduit la taille de l'échantillon, et donc limite la pertinence des coefficients estimés. Pour obtenir un test représentatif, il est nécessaire de retirer un grand nombre de périodes (5 au moins dans le cas d'un modèle annuel à moyen terme).
- L'évaluation du résultat est confrontée aux mêmes problèmes de subjectivité que le test précédent.
- Mais le principal problème est que si nous choisissons le modèle sur la base de ces tests, on utilise en fait la même technique que sur l'échantillon complet, et la même critique de choix «ad hoc» s'applique.

Tout dépend de la façon dont le test est effectué.

Toutes les équations ont d'abord été validées sur l'échantillon complet. Lorsque nous limitons la période, nous obtenons d'autres équations qui pourraient ou pourraient ne pas être satisfaisantes sur le plan économique et statistique. Si les équations sont restées stables, les résidus sont probablement de même taille, et il est tout à fait improbable que les effets multiplicateurs des propriétés du modèle vont beaucoup changer, donc l'erreur *a posteriori* ne devrait pas être beaucoup plus grande.

Par contre, si les résultats se modifient, on remettra en cause les spécifications elles-mêmes.

Cela signifie que le seul message de ce nouveau test concerne la stabilité de l'ensemble des équations, ce qui aurait été pu être mis en évidence par un test de rupture simple, comme le test de Chow, un critère plus objectif. Ce que nous obtenons est encore une approche plus globale (et peut-être utilitariste) mesure.

Bien sûr, si le processus de modélisation a été appliqué depuis le début à l'échantillon réduit, le processus sera beaucoup plus valable, à condition que l'on utilise le résultat comme une mesure absolue de la qualité du modèle, et non comme un moyen de sélectionner le meilleur modèle dans un ensemble. Dans ce cas, la logique du processus n'est pas très différente de l'estimation sur l'échantillon complet. Le choix repose sur un test (obtention de bons résultats sur des simulations de l'échantillon) qui demande essentiellement la stabilité de l'équation, ce qui signifie que les équations obtenues ne sont pas très éloignées de celles de l'échantillon complet.

En d'autres termes, le succès de la prévision *ex post* requiert pour les estimations de travailler sur l'échantillon complet, et de rester stables face à l'élimination des dernières périodes. Bien sûr, la stabilité est une condition pour la fiabilité du modèle. Mais ce n'est pas une preuve: une équation inversant d'un lien véritable et fort entre deux variables (faire dépendre la demande des importations) devrait rester stable avec n'importe quel échantillon.

7.2.3 LA RESOLUTION DES MODELES: LES SCENARIOS

Pour appliquer les techniques ci-dessus, une fonction proposée par EViews va se révéler très utile: les scénarios.

Quand on veut changer la façon dont le modèle est résolu, cela peut signifier:

- Résoudre le même modèle sous
 - des hypothèses différentes pour les variables exogènes,
 - des horizons différents,
 - des options différentes pour les formulations autorisant des versions alternatives (par exemple l'équation de taux de change peut proposer un taux nominal constant, un taux réel constant ou la parité non couverte des taux d'intérêt).
 - résoudre différentes versions coexistantes du même modèle.

L'un peut être officiel, ayant passé avec succès tous les tests requis, et familier à ses utilisateurs (en principe des prévisionnistes opérationnels), et l'autre peut être plus expérimental, en cours de test.

L'option «scenario» fournit un moyen efficace de faire face à ce problème.

On commence par la création d'un scénario.

Dans un programme, la déclaration est la suivante:

```
nom-du-modèle.scenario(n) nom-du-scénario
```

Le «(n)» crée un nouveau scénario, ou remplace un scénario existant de même nom.

Ensuite, il faut:

- Attribuer un suffixe au scénario / solution du modèle:

```
nom-du-modèle.assign @all suffixe
```

```
' nous résolvons sur le passé
' avec les résidus mis à zéro
```

```
' en utilisant le suffixe «_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
```

Pour remplacer les variables, deux opérations sont nécessaires:

- déclarer que les variables doivent être remplacées ;
- créer des hypothèses alternatives en utilisant le suffixe du scénario.

Si une des étapes est absente, les conséquences sont différentes.

- Si une variable créée avec le suffixe est omise de la liste, l'opération de remplacement ne sera pas appliquée, sans aucune conséquence négative (mais sans message). Cela permet de définir un ensemble de variables modifiables, entre lesquelles le choix sera fait plus tard.
- Si une variable de la liste n'a pas été créée, avec le suffixe adéquat, EViews refuse de résoudre le modèle, comme si la variable d'origine n'existait pas.

Si le programme passe à un autre scénario, les options associées au scénario précédent seront mises de côté, mais lui restent associées, jusqu'à ce qu'elles soient modifiées dans ce scénario particulier. Une déclaration de substitutions nouvelles annule toutes les précédentes.

L'option «scenario» peut également être utilisée pour exogénéiser des variables, en utilisant l'instruction EXCLUDE. Ceci peut servir à:

- obtenir des informations sur les raisons de l'échec de la convergence. Par exemple, si l'on soupçonne que le problème vient du système de prix, on peut exclure les prix et voir si le modèle peut maintenant être résolu ;
- résoudre des sous-modèles. Par exemple, si l'on veut éliminer l'effet de compétitivité-prix, une solution (pas la meilleure) est d'exclure les variables associées dans les calculs.

EViews 8 offre une autre possibilité: vous pouvez éliminer individuellement les remplacements par la commande REVERT appliquée à une liste de séries devant revenir à la version de base:

```
model_name.revert liste-des-séries
```

Mais l'ajout le plus important est la possibilité de spécifier directement les changements dans les hypothèses, en utilisant ADJUST.

La syntaxe est:

```
nom-du-modèle.adjust(init=série_initiale) nom-de-série ajustement
```

Par exemple, les instructions suivantes:

```
_mod_1.scenario(a=_2) "scenario 2"  
_mod_1.adjust(init="scenario 1") pib ==+10000
```

vont créer une série appelée `pib_2` avec une valeur supérieure de 10 000 à la valeur de «scénario 1» (peut-être appelé `gdp_1`). La série sera ajoutée par EViews à la liste «OVERRIDE», et à la liste «EXCLUDE» également si la variable est endogène.

Prenez soin d'introduire un espace avant le signe «=», sinon PIB prendra la valeur 10 000 !

L'introduction d'un changement comportant des éléments, tels que les paramètres ou des séries, est possible, mais plus difficile. On pourra préférer utiliser la génération directe de variables, comme nous l'avons fait dans nos exemples.

7.2.4 APPLICATION A NOTRE EXEMPLE

Nous allons maintenant simuler le modèle avec des résidus fixés à zéro, sur la période 1978S2 - 2002S1, la plus longue pour laquelle toutes les estimations peuvent être menées, et donc les simulations réalisées.

```
' simulation ex-post  
' Nous résolvons cours de la dernière  
' avec des résidus mis à zéro
```

```

smpl 1978S2 2002S1
for !i=1 to g_vbeha.@count
%2=g_vbeha.@seriesname(!i)
genr ec_{%2}_s=0
next
_fra_1.scenario "scenario 1"
_fra_1.append assign @all _s
_fra_1.override ec_i ec_ci ec_le ec_m ec_x
solve(d=d) _fra_1

' Nous calculons les résidus et leurs moyennes absolues
' - Les erreurs relatives (psa_*)
' - Les erreurs absolues sur les taux de croissance (psd_*)
' Nous les avons mis dans une matrice (19 x 2)

matrix(20,2) v_psm_a
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_a(!i,1)=(@mean(@abs(psa_{%st1})))
v_psm_a(!i,2)=(@mean(@abs(psd_{%st1})))
next

```

Ces déclarations peuvent sembler un peu complexes. Elles calculent la moyenne de la valeur absolue de ces deux critères.

Nous obtenons les résultats suivants:

	Erreurs de simulation	
	Pourcentage d'erreur sur le niveau	Erreur absolue sur le taux de croissance
CAP	1.69	0.54
CO	1.43	0.38
FD	1.36	0.49
I	2.85	0.47
IH	1.30	0.34
K	2.70	0.43
LE	1.68	0.45
LED	1.36	0.40
LT	1.29	0.35
M	2.81	1.40
PRLE_T	-	-
Q	1.36	0.40
RHI	1.30	0.34
TD	1.36	0.45
UR	1.00	0.46
X	1.77	1.36

La qualité semble assez acceptable, compte tenu de la taille du modèle⁸². La seule source d'incertitude provient des résidus d'estimation, qui ont une somme nulle sur la période. Et nos estimations, à l'exception de la variation des stocks, suivent un cadre à correction d'erreurs. Cela signifie que tout résidu apparaissant à une période donnée est corrigé avec le temps, parfois très rapidement. La variance de

⁸² Mais on peut aussi penser qu'une taille de modèle limitée, et un domaine d'hypothèses étendu, réduira les sources d'incertitude.

l'erreur globale va augmenter, mais converger sur le long terme vers une valeur finie, comme nous allons le constater plus tard.

7.2.5 CHOCS ANALYTIQUES

Si les simulations *ex post* fournissent un critère (plus ou moins valable) sur la qualité du modèle en prévision, les chocs analytiques permettent de juger de ses propriétés (donc de la fiabilité des études de politique économique) en étudiant en termes de logique économique quantifiée les résultats du modèle en simulation. Ainsi, le modèle indiquera si ses propriétés quantitatives sont conformes à la théorie économique qui a présidé à sa construction, et pas trop atypiques par rapport aux autres modèles de la même catégorie (la deuxième condition n'est bien sûr pas nécessaire, mais l'ignorer réclame une étude approfondie et des arguments solides).

Pour ce faire, un ensemble de chocs simples sera mené de manière indépendante, sur les principales hypothèses, en essayant de comprendre l'évolution de chacun des principaux domaines décrits par le modèle. Chaque type de domaine doit être traité: les comportements des agents, la fonction de production, l'offre et la demande. Cela signifie que nous allons étudier la sensibilité de l'économie à une variation de la demande exogène (par un investissement de l'Etat par exemple), le comportement des ménages (par une variation de l'impôt sur le revenu), des entreprises (par une variation de leurs cotisations sociales), le commerce extérieur (par une dévaluation ou un changement dans les taux de droits de douane), le rôle de la productivité du capital, d'une hausse exogène des prix. L'analyse quantitative et logique des évolutions des principaux éléments économiques donnera des indications sur la qualité du modèle, par comparaison avec la théorie économique et les propriétés des autres modèles.

Si les informations obtenues à partir de ce critère sont sans doute plus instructives, elles perdent aussi leur caractère objectif: un modèle qui affirmera que les ménages diminuent leur consommation si l'emploi s'améliore sera manifestement erroné, mais l'observation d'un comportement original par rapport à l'ensemble des autres modèles (par exemple, si les exportations montrent une sensibilité particulièrement forte à la compétitivité des prix, ou dépendent de l'âge du capital) ne sera pas concluante, en l'absence d'un consensus général sur sa valeur théorique. On peut même, dans certains cas, s'interroger sur le signe du lien: face à une montée de l'inflation, les ménages peuvent augmenter leur épargne (pour maintenir son pouvoir d'achat) ou la diminuer (acheter immédiatement ce qui deviendra bientôt plus cher).

Dans la pratique, pour valider un comportement atypique, celui-ci devra identifier la source de l'originalité. Parfois, elle aura été attendue avant toute simulation, car elle provient d'une caractéristique originale introduite volontairement dans le modèle.

Par exemple, la plupart des modèles supposent que l'augmentation de la demande locale sera partiellement satisfaite en limitant les exportations. Mais notre modèle peut dire aussi que les efforts d'investissement qu'elle réclame, en plus d'augmenter la capacité de production par le capital, permet

aux producteurs locaux de fournir des produits mieux adaptés à la demande potentielle, ce qui les rend plus attrayants pour les importateurs étrangers (c'est ce qu'on appelle la compétitivité hors prix). Un choc sur la demande pourrait donc après un certain temps se traduire par une augmentation des exportations, compensant l'effet de substitution initial⁸³.

En général, pour rendre l'interprétation plus facile, la variation des hypothèses ne fluctuera pas avec le temps. On distinguera néanmoins trois types de chocs, selon la nature de la variation:

- Des chocs instantanés, lorsque la modification disparaît après une seule période. Par exemple, l'État va augmenter sa consommation pendant une seule année budgétaire. Dans ce cas, on sera souvent intéressé, non seulement à l'évolution des effets (qui devraient disparaître avec le temps), mais aussi au cumul de ces effets dans le temps (par exemple, le gain total sur la production).
- Des chocs durables, où la modification est maintenue pendant un certain nombre de périodes, qui peuvent représenter la période entière utilisée dans la simulation. Par exemple, une réduction du taux d'impôt sur le revenu sera maintenue sur la période de simulation.
- Des chocs cumulatifs, amplifiés à chaque période d'un montant identique. Leur interprétation sera souvent difficile.

Il est clair que si l'on veut interpréter correctement ces chocs, ils doivent rester logiquement acceptables. La conséquence pour le type de choc n'est pas trop claire.

Pour déterminer le choc, il faut donc tenir compte de deux critères:

- L'acceptabilité du choc.
- La lisibilité des résultats.

Cela éliminera en général les chocs ponctuels, en particulier pour les variables de politique économique. Ils échouent la plupart du temps sur les deux critères. Sauf pour un investissement supplémentaire ou une subvention temporaire, les décisions politiques sont généralement prises pour durer pendant un certain temps, surtout si elles s'appliquent à des taux d'imposition ou aux prestations de sécurité sociale. Les changements structurels (gains sur la productivité des facteurs) sont également permanents. En ce qui concerne les chocs externes (comme une augmentation de la demande mondiale), ils pourraient très bien être n'apparaître qu'une seule fois.

⁸³ Nous observons ici une caractéristique importante: l'originalité peut se trouver non pas dans le signe à court terme, mais dans son renversement à long terme. Plus généralement, des effets originaux de long terme (favorisés par le format à correction d'erreur) peuvent n'apparaître qu'après un certain temps.

Mais l'argument le plus décisif est plus terre à terre: la plupart des interprétations se feront à travers des graphiques, et un choc instantané ramène toutes les variables au scénario de référence. Après quelques périodes, les courbes vont se rapprocher de l'axe du temps, et il sera impossible de les séparer les unes des autres, ce qui crée la confusion, alors que dans le cas d'un choc soutenu la stabilisation des conséquences créera un ensemble de courbes plus ou moins parallèles (ou à divergence lente). Quant aux chocs cumulés, ils vont produire des graphiques divergents et rendre l'interprétation des changements initiaux (associés à la plus petite échelle) plus difficile.

Donc on se retrouve généralement avec des chocs durables, qui sont en effet les plus faciles à traiter, à condition qu'ils soient raisonnables. Ceci s'applique d'abord à l'intensité: moins celle-ci est acceptable, plus difficile sera l'interprétation. Cela signifie que le choc doit être assez petit (par exemple abordable en termes de coût budgétaire) et assez grand (pourquoi le gouvernement devrait-il se préoccuper de ces décisions marginales ?).

Mais cela vaut aussi pour le coût actualisé, dans le cas de décisions politiques. Si la période est suffisamment longue, et le choc suffisamment important, on peut considérer que son coût pour l'État va devenir insupportable après un certain nombre de périodes. Et les agents économiques le savent. Ainsi, non seulement nous devrions arrêter le choc après un certain temps, mais le modèle devrait adapter le comportement des agents à cet événement futur. Ceci n'est pas possible avec notre modèle actuel, dans lequel nos agents ne basent leur comportement que sur le présent et le passé.

C'est le principe le plus fondamental de la critique de Lucas, que nous allons aborder plus tard, lorsque nous traiterons des anticipations rationnelles.

Bien sûr, ce problème de long terme sera plus présent si l'on considère les prévisions: on pourrait supposer que la période échantillon est suffisamment petite pour permettre des changements de politiques durables.

7.2.5.1 Le choix des hypothèses modifiées

En ce qui concerne les variables modifiées, on va rechercher un ensemble qui met à contribution tous les mécanismes du modèle: comme le but est de valider le modèle complet, aucun comportement anormal ne doit subsister. Un investissement initial dans l'examen global de ces éléments évitera tout problème ultérieur, dont l'identification se révélera beaucoup plus coûteuse (à condition même qu'elle puisse être faite).

En particulier, on cherchera à mettre en jeu l'ensemble des agents et l'ensemble des domaines: production, marché du travail, prix et salaires, commerce extérieur, mécanismes financiers. Les influences orientées offre et demande doivent toutes être étudiées. Toutefois, s'il y avait un seul test à réaliser, le plus représentatif d'un modèle habituel consisterait à augmenter la part exogène de la demande, ce qui détermine le multiplicateur keynésien. Nous reviendrons sur ce point dans le paragraphe suivant.

7.2.5.2 Le choix des résultats examinés

Tous les chocs possibles sur les éléments endogènes peuvent être ou ne pas être examinés, selon la taille du modèle. Dans de nombreux cas, les premières analyses porteront sur des résultats agrégés, en n'ayant recours à des valeurs détaillées que si un problème apparaît.

Dans le cas d'un modèle macroéconomique, les principaux éléments sont les suivants:

- l'équilibre de l'offre et de la demande: le produit intérieur brut, la demande et sa décomposition en produits, les exportations et les importations (le rapport de la variation du PIB à l'augmentation de la demande exogène définit le multiplicateur keynésien) ;
- les prix et les salaires, les taux d'intérêt ;
- l'emploi et le chômage ;
- les soldes commerciaux et budgétaires ;
- certains ratios: taux d'épargne, de profits, d'utilisation des capacités de production, les soldes et dettes financières en pourcentage du PIB.

7.2.6 NOTRE EXEMPLE

Dans notre exemple, le nombre d'hypothèses disponibles est très faible. Les choix les plus raisonnables sont les suivants:

- la demande du gouvernement GD
- la demande mondiale WD.

Et peut-être aussi:

- la productivité du capital PK
- le taux d'épargne SR
- le taux d'amortissement DR⁸⁴
- la part des revenus des ménages non salariaux dans le PIB R_RHIQ

⁸⁴ Un choc sur DR introduit un problème intéressant, car il est *de facto* interdit. En effet, le développement de notre équation d'investissement est basé sur le fait que DR est constant, ce qui permet de l'intégrer au terme constant, donc le changer viole les spécifications du modèle. C'est le seul choc qui va effectivement modifier le taux d'utilisation à long terme, et ce, pour de mauvaises raisons. Pour lui permettre de changer, on peut simplement estimer la variation du capital au lieu de l'investissement (l'hypothèse de la stabilité de DR disparaît) et réarranger l'équation actuelle de capital afin qu'il détermine les investissements.

- la part de l'investissement logement des ménages dans le revenu R_HI
- les résidus de chacune des cinq équations estimées.

Les principaux éléments sont clairement axés sur la demande. Nous choisirons le plus largement accepté: la demande du gouvernement GD.

Nous allons utiliser les déclarations suivantes, prises à partir d'un programme (nous avons laissé les commentaires).

```
' Nous produisons un choc
' Tout d'abord, une simulation non choquée
' Le suffixe est _b
' Nous ne remplaçons aucun élément (nous nous assurons que la liste est vide).
Smpl 1980S2 2002S1
_fra_1.scenario "scenario 1"
_fra_1.append assign @all _b
_fra_1.override
_fra_1.solveopt(n=t m=1000,c=1e-6,o=n,d=d)
solve _fra_1
' Maintenant, la simulation choquée
' Le suffixe est _v
' Nous calculons un choc sur gd (demande du gouvernement) de 1 point de PIB
' Nous l'appliquons à l'hypothèse du choc
' Nous remplaçons gd
smpl 1980S2 2002S1
genr dv_gd=.01*q_b*(t>=1981)
genr gd_v=gd+dv_gd
_fra_1.scenario "scenario 1"
_fra_1.append assign @all _v
_fra_1.override gd
solve _fra_1
' Nous calculons les différences
' En termes absolus et relatifs
' Ils seront appelés dv * et * pv_
for !i=1 to g_vendo.@count
%2=g_vendo.@seriesname(!i)
series dv_{%2}={%2}_v-{%2}_b
```

```
series pv_{%2}=100*dv_{%2}/({%2}_b+({%2}_b=0))
next
```

Quelques commentaires supplémentaires.

Nous avons commencé le choc un peu après le début de la simulation. Cela permet de vérifier que, en l'absence de choc, les deux simulations donnent le même résultat, donc que la différence peut être interprétée comme provenant uniquement du choc. D'autres différences peuvent provenir:

- d'une autre date de départ,
- de différences dans le modèle,
- de différences dans les autres hypothèses.

Ceci n'est pas une preuve absolue, cependant:

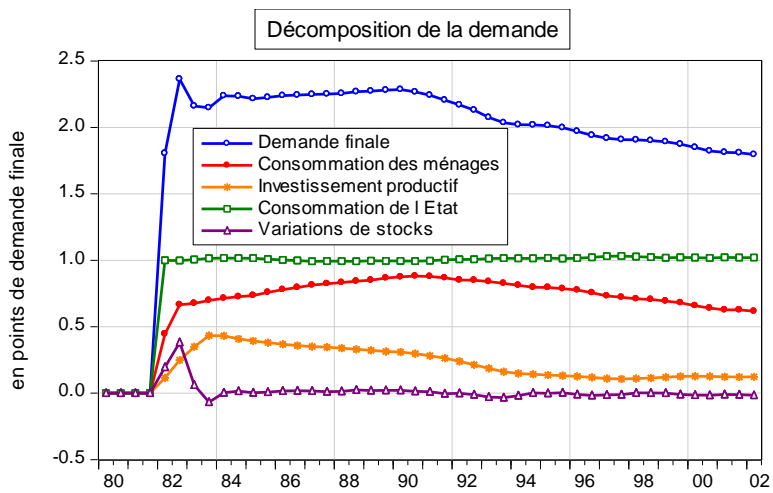
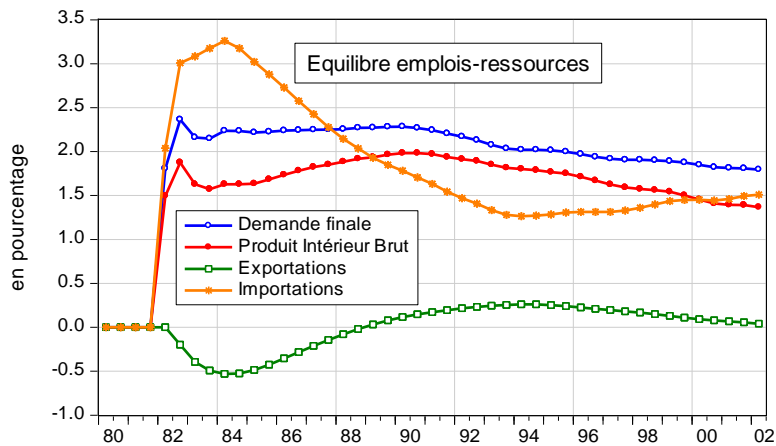
- Un changement supplémentaire dans les séries pourrait intervenir après 1981.
- La spécification du modèle peut contenir des modifications qui s'appliquent uniquement après 1981 (par exemple, un nouvel impôt) ou une condition sur les formulations qui ne sera vérifiée que plus tard (comme un niveau élevé de l'inflation).

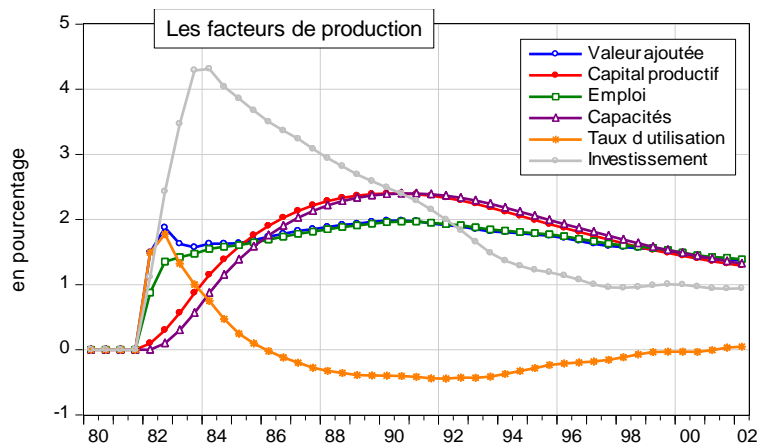
Afin de garantir l'absence de telles erreurs, il est conseillé de simuler à nouveau le scénario de base immédiatement avant le choc, même s'il a été produit auparavant. L'efficacité des algorithmes actuels et la vitesse des ordinateurs actuels rendent le coût négligeable dans la plupart des cas.

Nous avons produit l'hypothèse alternative en deux étapes, d'abord en calculant le niveau du choc lui-même, avec le même format («DV_» suivi du nom de la variable) que les écarts absolus endogènes. Cela nous permettra de comparer les changements ex ante et ex post du PIB, et de créer un graphique présentant toutes les modifications.

7.2.6.1 Les résultats

Nous allons baser notre commentaire sur trois graphiques montrant: l'évolution de l'équilibre offre-demande en termes relatifs, la décomposition de la demande finale en termes absolus (y compris le choc initial), et l'évolution des éléments liés à la production.





Les propriétés affichées par ces graphiques sont tout à fait habituelles pour ce type de modèle (compte tenu de la taille réduite de notre version).

La variation de la demande ex ante est amplifiée par l'augmentation des facteurs de production (investissement et l'emploi), ce dernier affectant le revenu des ménages et leur consommation. Ces deux éléments ne cessent de croître pendant un certain temps, avec une certaine inertie, traditionnelle pour un processus de correction d'erreur.

Plus tard, les deux changements se stabilisent à un niveau proportionnel au PIB. Mais, alors que les ménages obtiennent sous forme de revenus une part importante de l'augmentation, l'incitation à investir diminue avec l'adaptation du capital au PIB. À long terme, la seule motivation restante est le remplacement d'un niveau plus élevé de capital déclassé.

Le capital et la capacité s'adaptent lentement, mais cette inertie fait dépasser la cible. Un processus cyclique apparaît sur le taux d'utilisation, avec une intensité en diminution assez rapide.

Le commerce extérieur joue un rôle simple.

L'évolution des exportations suit le taux d'utilisation, avec une évolution négative limitée aux premières périodes.

Les importations représentant au début une part importante de la demande initiale, tant que les capacités ne sont pas adaptées. Elles contrebalancent l'augmentation ex post de la demande, en réduisant le multiplicateur à une valeur proche de l'unité. Au cours des périodes suivantes, les capacités vont

s'adapter, et le multiplicateur augmentera régulièrement. La sur-adaptation du capital à moyen terme favorise les entreprises locales pendant quelques années.

Dans l'ensemble, bien que ce modèle reste simpliste, il semble plutôt acceptable si l'on considère sa taille.

7.2.6.2 Les défauts des chocs sur le passé

Bien que la production de chocs sur le passé d'un modèle qui sera utilisé à l'avenir ne soit pas confrontée aux mêmes critiques que les tests *a posteriori*, elle doit faire face à plusieurs problèmes.

- Les tests sont menés sur une période limitée. Si l'on veut observer les propriétés de long terme du modèle, le temps disponible sera généralement trop court. Le problème va croître si la production de données a commencé récemment, ou si le pays a subi une transformation récente qui rend l'utilisation de données supplémentaires inacceptable. Pour les pays d'Europe centrale et orientale, les séries utilisables commencent en 1995, au mieux.
- Ils sont menés sur les mauvaises périodes. Si l'on veut interpréter les propriétés dynamiques d'un modèle sur une période suffisamment longue, il est nécessaire de commencer les chocs assez loin dans le passé. À cette époque, l'économie du pays présentait certainement des caractéristiques différentes. Et même si le comportement économique suivait une logique constante, certains paramètres structureaux étaient très différents, comme l'importance du commerce extérieur, ou les rôles relatifs du travail et du capital dans le processus de production.
- Ils sont fondés sur une base plus ou moins instable. Aucun modèle n'est complètement linéaire (il inclura toujours à la fois des sommes et des logarithmes). Cela signifie que la sensibilité aux chocs dépend de la trajectoire de base, qui peut présenter des irrégularités sur le passé. Par exemple, une augmentation donnée de la consommation publique va générer plus de croissance du PIB aux périodes où l'activité est historiquement faible, car le pays peut utiliser ses capacités inutilisées pour répondre à la demande.

Cela signifie qu'il sera difficile de décider si les variations irrégulières dans les résultats des tests proviennent des propriétés du modèle ou d'irrégularités dans la base.

Évidemment, tous ces problèmes disparaissent quand le choc s'applique aux prévisions:

- La durée est illimitée, ce qui permet de mettre en évidence les processus de convergence et en particulier les cycles.
- Comme les tests sont réalisés sur la même période que les futures applications opérationnelles, l'évaluation sera plus pertinente.
- Le contexte actuel et le futur proche sont plus familiers au modélisateur.
- Comme la trajectoire de base est régulière (en particulier dans le moyen et long terme) toute irrégularité pourra être attribuée aux propriétés du modèle.

C'est pourquoi nous suggérons de limiter à un minimum les études sur le passé, et de passer rapidement aux simulations sur l'avenir. Cela est vrai même pour le contrôle de la convergence: c'est sur l'avenir et non sur le passé que le modèle devra converger. Mais une série limitée d'essais immédiats peut mettre en évidence tout de suite des problèmes, ce qui fera gagner du temps dans le processus global en le maintenant sur la bonne voie.

7.2.7 ASPECTS LOGICIELS

Les éléments ci-dessus montrent que c'est dans le domaine de la simulation que l'utilisation d'un logiciel adapté sera le plus profitable. Il s'agit de la phase la plus coûteuse en temps de calcul, et la plus répétitive. C'est également là que l'expertise informatique devrait être la moins nécessaire. Si la gestion de la base de données suppose logiquement une certaine connaissance des méthodes informatiques, et l'économétrie la connaissance du langage associé, produire des simulations devrait être assez simple, une fois que la qualité du modèle est suffisamment élevée pour assurer la convergence vers un modèle correctement spécifié.

Ceci va conditionner les qualités demandées au logiciel de création de modèle:

- La présence d'algorithmes performants, à la fois en temps de résolution (nous avons vu comment cet aspect peut être amélioré) et en probabilité de convergence. Nous avons vu qu'EVIEWS est plutôt efficace à cet égard.

Mais en cas de problèmes, corriger le processus de convergence est une tâche difficile et fastidieuse qui demande des connaissances particulières et peut prendre beaucoup de temps⁸⁵. Elle est favorisée par la présence d'un large éventail d'options conviviales décrivant les causes des problèmes de convergence, et permettant de les corriger. Ces options ont été décrites ci-dessus. EVIEWS fournit une grande quantité d'informations, même si leur utilisation est parfois un peu difficile.

Le seul problème pourrait résider dans l'absence d'une stratégie automatisée, car EVIEWS s'arrête généralement à la première erreur, sans chercher à la gérer. En outre, comme la plupart des langages de programmation, EVIEWS pourrait autoriser l'introduction de points d'arrêt fixes ou conditionnels, auxquels on pourrait observer la situation numérique, et modifier à la main certains de ses éléments. Par exemple, quand un logarithme s'applique à une valeur négative, EVIEWS pourrait interrompre le processus et transférer le contrôle à l'utilisateur, lui permettant de la remplacer par une valeur positive, ou d'exclure

⁸⁵ EVIEWS autorise également à procéder sans corriger l'erreur, une stratégie dangereuse, mais parfois efficace.

la variable des calculs suivants, ou tout simplement de permettre l'accès à plus d'informations sur l'état du processus.

- Une gestion complète des critères de convergence. L'option de base: une seule valeur, mesurée en termes relatifs et appliquée à l'ensemble des variables, peut conduire à des problèmes. Comme nous ne pouvons pas envisager de définir un critère pour chaque variable, la meilleure solution est sans aucun doute: un critère général unique, modifié dans des cas spécifiques, avec exclusion totale pour certaines variables⁸⁶. Comme nous l'avons vu précédemment, le contrôle peut être limité aux éléments de bouclage. Malheureusement EViews n'autorise qu'un seul critère global relatif.
- Un stockage des résultats avec accès facile. Avec EViews, on peut identifier les résultats dans le fichier de travail par le biais d'un suffixe. Mais on peut aussi créer une base de données spécifique, ou une nouvelle feuille dans le même fichier de travail, ou exporter les données vers Excel.
- Un langage permettant la construction de macro-commandes, établies par le constructeur de modèles / informaticien. Ces commandes permettront la réalisation de prévisions ou de chocs sans connaître le langage du logiciel, en répondant simplement à des questions explicites, éventuellement sous forme de menus, avec une aide en ligne et la détection automatique des erreurs de syntaxe (l'environnement Windows fournit les outils nécessaires, dont l'usage pourra être compris facilement). Pendant longtemps, EViews⁸⁷ a seulement permis la production de programmes en mode texte, en s'appuyant sur des commentaires insérés (et un guide de l'utilisateur séparé) pour expliquer leur utilisation. Cependant, il était déjà facile d'organiser et de documenter ces programmes de telle manière que l'utilisateur ait besoin d'un minimum d'informations.

La version 7 a partiellement résolu ce problème en permettant d'entrer les informations dans des menus définis par l'utilisateur, même si les fonctionnalités autorisées sont limitées. Par exemple, le format de l'écran est en grande partie prédéfini, son organisation étant décidée par EViews⁸⁸, et il n'est pas possible de fournir de l'aide en ligne.

- L'accès à des fonctions écrites dans un autre langage (comme C + +) ou construit à l'aide du logiciel. Ceci est permis par EViews, mais seulement pour MatLab et R (dans les deux sens). Les

⁸⁶ Si exportations et importations ont convergé, on ne devrait pas avoir à contrôler la balance commerciale.

⁸⁷ Au moins jusqu'à la version 5.

⁸⁸ Par exemple, on ne peut pas définir interactivement la taille des fenêtres.

nouvelles versions présentent des progrès dans ces domaines, avec l'introduction dans les menus de fonctions établies par l'utilisateur («add ins»).

- Un processus permettant le ciblage automatique des résultats endogènes aux valeurs connues. EViews le permet pour une seule variable, **mais cette propriété peut être étendue** à plusieurs variables grâce à des programmes EViews simples (comme nous le verrons).
- La disponibilité d'ensembles spécifiques d'un choc, contenant toutes les modifications associées à la simulation alternative (les variables, les coefficients, les chocs sur les comportements), facilement établis et nommés. Cela se fait très facilement grâce aux fonctions SCENARIO et OVERRIDE, mais uniquement pour les variables.
- La visualisation facile des résultats. EViews est très bon dans ce domaine, à la fois sous forme de tableaux et de graphiques, comme nous le verrons plus tard.

7.2.7.1 Deux solutions alternatives

Considérons maintenant deux autres façons de résoudre des modèles économétriques. Bien qu'elles n'utilisent pas d'outils spécialement conçus pour cela, elles peuvent être efficaces dans certaines circonstances.

Nous avons déjà abordé ce sujet, mais nous allons maintenant le développer.

7.2.7.1.1 Les tableurs

On peut imaginer de simuler des modèles dans des feuilles de calcul, et cette option est effectivement utilisée par certains modélisateurs opérationnels. Mais même si aucune impossibilité technique n'existe (à l'exception de la taille des problèmes à gérer), les inconvénients sont trop élevés dans la plupart des cas.

Commençons par les points positifs. On peut citer

- la visualisation en temps réel des résultats ;
- la présence de la formulation dans la cellule elle-même ;
- la production immédiate de tableaux et graphiques associés, de très bonne qualité ;
- l'apprentissage d'un logiciel unique au lieu de deux, et d'aucun si l'on est déjà familier avec un tableur donné ;
- la possibilité d'appliquer sa propre technologie, développée au fil du temps, et ses propres macros instructions ;
- la baisse du coût financier (en particulier dans le cas ci-dessus) ;
- la convivialité de l'ensemble, et surtout l'intégration dans l'environnement Windows ;

- le développement plus rapide et plus fréquent des logiciels, grâce au marché beaucoup plus important pour ce type de produit ;
- la plus grande facilité de communication avec les non-constructeurs de modèles.
- On peut également citer la possibilité de faire appel directement à des méthodes d'optimisation-maximisation (voir « contrôle optimal »).

À l'opposé, en faveur d'un logiciel de modélisation, nous pouvons citer:

- une meilleure intégration de la dimension temporelle ;
- l'identification claire des objets (en particulier par type: les variables, les équations, les paramètres, les modèles), et la création de listes et index associés ;
- un affichage plus clair des spécifications des modèles ;
- la gestion sélective des objets utilisés, ce qui réduit la mémoire nécessaire en réduisant la taille des problèmes à gérer. Par exemple, l'estimation des équations appelle uniquement les séries associées, les équations et les paramètres. La simulation d'un modèle n'a besoin à un stade donné du processus que d'un sous-ensemble de périodes (la période concernée et les valeurs retardées des variables) ;
- la gestion intégrée des données, tant pour les données historiques que les résultats ;
- l'accès à des méthodes économétriques sophistiquées, y compris récentes⁸⁹, permettant l'identité des spécifications estimées et simulées, et simplifiant la gestion des paramètres ;
- l'accès implicite aux données utilisées pour les estimations ;
- une plus grande facilité d'interprétation des spécifications des équations ;
- le stockage des équations sous forme d'éléments complets ;
- des méthodes de résolution adaptées, qui rendent la convergence plus rapide et plus probable, en fournissant des diagnostics précis et des outils pour résoudre les problèmes ;
- la facilité de gestion des scénarios, comme des ensembles spécifiques ;
- une plus grande facilité de communication avec les autres constructeurs de modèles (surtout s'ils utilisent le même logiciel).

À notre avis, ces arguments font clairement pencher la balance en faveur d'un logiciel spécialisé, à l'exception des applications brèves à un petit modèle, ou si un constructeur de modèles à temps partiel fait par ailleurs un usage intensif des feuilles de calcul de type Excel.

⁸⁹ Y compris des méthodes développées par des utilisateurs extérieurs, mises à disposition sur le site EViews et intégrables sous forme d'« add-ins ».

7.2.7.1.2 La programmation directe

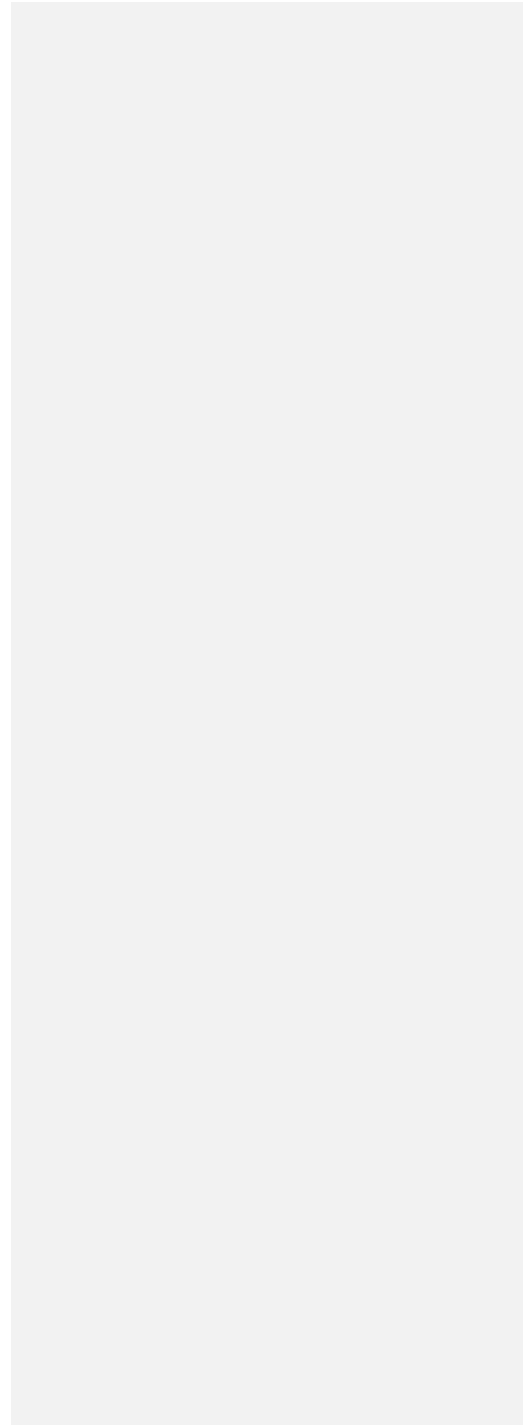
Une autre technique consiste à programmer directement les simulations du modèle à travers un langage comme C ++ ou Fortran. Cette option a été utilisée dans les années 1960, mais a depuis été quasiment abandonnée pour plusieurs raisons:

- Le travail était long et fastidieux.
- Prouver l'exactitude des calculs était presque impossible.
- Des logiciels spécialisés de plus en plus efficaces sont apparus sur le marché.
- Et probablement pour cette raison, les économistes pratiquent de moins en moins les langages de programmation informatique.

Ceci dit, cette option retrouve étrangement une certaine rationalité. Une raison est que les langages de programmation (en particulier sous Windows) offrent des options de transfert, assez faciles à mettre en œuvre, et ont acquis une certaine convivialité. Bien sûr, la gestion des données et les phases d'estimation sont difficiles à programmer. Mais une fois que le modèle est stabilisé, on peut envisager d'écrire un programme de simulation, qui permet la résolution du modèle grâce à un algorithme dédié, et fournit (à un coût de programmation élevé) la gestion des hypothèses et des résultats dans les menus dédiés, avec une facilité d'utilisation allant au-delà des options normalement proposées par les logiciels de modélisation. On peut aussi, dans ce cas, tirer parti des spécificités de son modèle, bien que cela interdise le transfert de l'environnement à d'autres cas.

Mais l'intérêt principal de cette technique est qu'elle permet d'introduire dans le programme des calculs complexes, d'un type interdit par des logiciels de modélisation. Le cas le plus évident est la maximisation d'une fonction sous contraintes, par exemple la recherche du plus bas taux de chômage ne produisant pas de déficits budgétaires et commerciaux trop élevés.

La contrainte principale est sans doute la nécessité d'une connaissance approfondie de la programmation et de ses caractéristiques mathématiques.



CHAPITRE 8: LES TESTS DU MODELE SUR LE FUTUR

Les tests effectués sur la période d'estimation ne sont pas vraiment satisfaisants, pour plusieurs raisons:

- Pour les simulations, même si c'était la seule façon de vérifier les résultats par comparaison aux données réelles, le fait que ces données ont été effectivement utilisées pour produire les estimations ne peut pas éviter aux tests d'être biaisés, quelles que soient les précautions prises.
- Pour l'analyse des chocs, la période est généralement trop courte pour mettre en évidence les propriétés à long terme et mesurer les cycles, et les irrégularités (pour le moins) dans la trajectoire de base sont transmises aux résultats (en raison de la non-linéarité du modèle), ce qui rend incertain le diagnostic sur la stabilité du modèle.
- Par ailleurs, les résultats sont appliqués à la période historique, ce qui n'est pas le véritable champ des futures utilisations opérationnelles du modèle.
- Pour obtenir des informations suffisantes, les chocs doivent être réalisés sur au moins dix ans, ce qui rend la période de départ assez éloignée du présent et encore plus de l'horizon des prévisions.

Cela conduit à l'idée naturelle: tester le modèle sur l'avenir.

- Nous aurons les premières informations sur la fiabilité des prévisions spontanées, et sur leur distance par rapport à ce que nous attendons de l'avenir proche.
- Les résultats seront plus représentatifs d'une utilisation opérationnelle.
- Les résultats s'appliquent à une période connue du modélisateur, facilitant leur interprétation.
- Les premières périodes du choc seront représentatives de l'efficacité actuelle des politiques portant sur la période actuelle (et donc envisageables en pratique).
- Les tests peuvent être réalisés sur une période de n'importe quelle longueur, ce qui permet d'observer la convergence et les cycles.
- Avec des hypothèses en croissance régulière, nous pouvons tester si la simulation est régulière, et si elle converge à la fois vers une croissance équilibrée et une solution de long terme stable.
- Appliquant la solution à ces chocs réguliers constants, nous pouvons vérifier que nous obtenons des évolutions lissées, et nous pouvons les interpréter plus facilement.
- Nous avons suffisamment d'observations pour traiter la critique de Lucas.

On ne peut voir que deux inconvénients:

- Nous ne pouvons pas vérifier les résultats de la simulation par rapport aux valeurs réelles⁹⁰. Nous avons déjà vu que ce test n'est pas si pertinent, et nous verrons peut être remplacé par d'autres tests.
- Nous devons produire une simulation sur l'avenir, un domaine inconnu où la convergence pourrait être plus difficile à atteindre.

Mais cette convergence devra être obtenue de toute façon. Donc, notre premier test sera :

- Le modèle peut-il converger vers n'importe quel horizon ?

Dans notre sens, il s'agit d'une condition *sine qua non* pour les modèles économétriques structurels. C'est là que ce type de modèle possède un avantage sur les CGE et les VAR: ils produisent des simulations dynamiques qui montrent la façon dont l'économie converge vers un équilibre de long terme, et permettent une interprétation complète du processus. Si cela manque (si le modèle ne converge que sur le moyen terme) cet avantage est perdu.

Donc, notre premier objectif sera :

- obtenir la convergence à long terme du modèle

Soulignons tout d'abord que ce n'est pas la même chose que la production de véritables prévisions à moyen-long terme. Ce que nous envisageons est le très long terme (la première version convergente du modèle peut prendre plusieurs siècles à se stabiliser). Nous n'allons évidemment pas supposer que le modèle reste valable pendant une telle période. Ce que nous voulons réaliser, c'est :

La reproduction technique sur une longue période des mécanismes associés aux équations que nous avons mises en évidence, comme si elles restaient valables indéfiniment⁹¹. La période est beaucoup plus longue que la plus longue sur laquelle le modèle sera utilisé.

Par exemple, de nombreux modèles présentent un comportement cyclique, liant les investissements, le capital, l'écart entre production et capacités, et la rentabilité. La durée de ce cycle peut dépasser 30 ans. Pour vérifier que sa période est constante, et qu'il converge régulièrement, nous avons besoin d'au moins trois cycles, ce qui signifie que nos simulations doivent durer jusqu'en 2100.

⁹⁰ Sauf éventuellement pour les toutes premières périodes, ou l'on dispose d'informations partielles sur les grands agrégats.

⁹¹ Mais nous savons que ce n'est pas le cas.

Ainsi, au début des essais, on doit considérer une très longue période.

Ceci pourrait être comparé à l'essai d'une voiture (en fait un prototype) sur une piste artificielle, aussi longtemps que il peut fonctionner, ce qui dépasse de loin sa durée de vie normale (par exemple plusieurs millions de kilomètres), à une vitesse constante. Et aussi le changement de l'un des éléments de la voiture, ou ses conditions de conduite, pour en observer les conséquences.

Ce que nous voulons obtenir est une information qui va nous aider à construire un meilleur modèle, qui sera utilisé plus tard dans des circonstances tout à fait différentes (environnement plus court mais moins régulier).

Nous devons garder cela à l'esprit lorsque nous entreprenons notre simulation.

8.1.1 LES HYPOTHESES

Pour produire une simulation à long terme, nous avons besoin d'hypothèses sur les éléments exogènes.

Le choix de ces hypothèses peut avoir plusieurs objectifs:

- La production d'évolutions de court terme qui coïncident avec ce que nous savons déjà de l'avenir proche (ou du passé récent que nous avons dû simuler en l'absence d'informations complètes).
- L'établissement de solutions raisonnables à moyen terme, conformément aux objectifs du gouvernement, ou à ce que nous attendons de l'économie locale (un déficit budgétaire élevé, mais en baisse par exemple).
- Permettre au modèle de converger régulièrement vers un équilibre de long terme, que nous allons examiner pour comprendre les propriétés du modèle.

Pour le moment, nous allons nous concentrer sur le dernier but. Bien sûr, nous allons contrôler que les deux premiers objectifs sont au moins partiellement remplis, car une simulation de base déraisonnable ne va pas réagir aux chocs de façon fiable, en raison des non-linéarités des modèles, ce qui rendra les résultats plus difficiles à interpréter sinon inutilisables.

Cela signifie que nous ne différencierons pas les hypothèses qui partagent la même dimension, mais leur appliquerons des taux de croissance communs. Dans la pratique, trois taux doivent être définis (en plus d'un taux nul):

- un taux pour les populations (y compris l'emploi),
- un taux pour les variables à prix constants,
- un taux pour les déflateurs.

Bien sûr, le deuxième taux pourrait être remplacé par la productivité du travail. En le combinant avec celui de l'emploi que nous aurions celui des volumes.

En fait, nous verrons que la plupart des modèles comprennent essentiellement des hypothèses sans dimension. Cela doit être recherché: cela supprime la nécessité d'obliger des hypothèses normalement indépendantes à croître à la même vitesse (comme, dans notre modèle, la demande du gouvernement *GD* et la demande mondiale *WD*).

8.1.1.1 Un cas particulier: les estimations

En général (ce qui est le cas pour toutes les équations de notre modèle) l'expression estimée a été définie sans dimension, ne serait-ce que pour éliminer le risque d'hétéroscédasticité. Et de toute façon, la valeur la plus probable d'un résidu est supposée être égale à zéro. Mais on peut supposer que la dernière valeur obtenue est un meilleur indicateur de l'avenir, en particulier si les valeurs précédentes ont présenté le même signe. Pour le moment, le choix n'est pas si important, car nous ne sommes pas trop préoccupés par les solutions à court terme. Mais il va devenir crucial plus tard, lorsque nous aborderons les véritables prévisions.

8.1.2 ADAPTER LES FORMUMATIONS

Pour faire converger le modèle à long terme, l'application d'une formule à des variables qui croissent à la vitesse théorique devrait produire un résultat qui croît à son taux théorique propre (ce qui est facile à contrôler comme nous le verrons plus tard).

Les identités fourniront automatiquement cette propriété, si elles sont correctement spécifiées et si les hypothèses sont bien définies.

Pour les équations estimées, le format à correction d'erreur permet d'introduire les contraintes nécessaires dans l'équation de long terme. On n'a pas à se préoccuper de la partie VAR, qui comprend seulement des taux de croissance, stables sur le long terme par définition. Mais il faut gérer (en général supprimer) toute tendance supplémentaire.

Expliquons-le sur notre équation d'exportations. Nous avons:

$$\begin{aligned} \text{dlog}(x) &= 0,940 * \text{dlog}(wd) - 0,0129 * \text{res}_x(-1) + \text{ec}_x \\ \text{res}_x &= \text{Log}(x / wd) + 0.686 * \log(ur) - 4,87 \text{ e-}05 * @ \text{ trend}(60:1) \end{aligned}$$

Supposons que le reste du modèle (en particulier l'équation d'investissement) a permis de stabiliser UR dans le long terme. Par la suite, toute croissance de res_x viendra:

- de valeurs de X croissant à un rythme différent de WD,
- de la tendance.

Nous allons maintenant bloquer la tendance, par exemple en remplaçant son expression par:

```
@trend(60:1)*(t<=2002)+@elem(@trend(60:1),"2002S1")*(t>2002)
```

À partir du second semestre de 2002 (représenté par t>2002), la variable ne croîtra plus.

Nous avons (avec $\Delta \log(WD)$ et ec_x constants)

$$\Delta \log(X_t) = a - b \cdot res_x_{t-1}$$

$$res_x_t = \log(X_t / WD_t) + c$$

ce qui donne

$$\log(X_t) = \log(X_{t-1}) + a + c - b \cdot (\log(X_{t-1}) - \log(WD_{t-1}))$$

et

$$\log(X_t) = (1-b) \cdot \log(X_{t-1}) + a + c - b \cdot \log(WD_{t-1})$$

Le taux de croissance de X converge vers celui de WD.

Dans des cas très particuliers, la tendance ne doit pas être supprimée, mais remplacée par une valeur donnée, car elle s'applique à un ratio qui évolue (entre deux termes de dimension différente).

Par exemple, dans l'équation donnant l'emploi, la productivité structurelle du travail suit une tendance, estimée à partir de:

$$\text{Log}(Q_t / LE_t) = a + b \cdot t$$

À long terme, les taux de croissance des quantités et de l'emploi sont fixes, ce qui définit le taux de croissance de la productivité du travail. Le coefficient estimé doit être remplacé par cette valeur.

8.1.2.1 comment améliorer les chances (et la vitesse) de convergence

Nous avons présenté plus haut quelques techniques. Elles peuvent être appliquées indifféremment aux simulations sur le passé ou le futur. Peu d'éléments nouveaux seront spécifiques aux prévisions.

En fait la seule différence importante concerne les valeurs d'initialisation des endogènes. EViews permet de faire commencer le processus de résolution, soit par les valeurs «historiques» ou celles obtenues à la période précédente. Cela se fait grâce à l'option: $i = a$ (observées ou «actual») ou $i = p$ (précédentes) dans l'ordre SOLVE ou SOLVEOPT. Dans les prévisions, seule la dernière devrait être disponible.

Mais en réalité, on peut définir des valeurs sur l'avenir:

- En calculant les valeurs théoriques, en appliquant aux variables endogènes leur taux de croissance théorique.
- En donnant aux séries la valeur d'une solution précédente, qui utilise des hypothèses comparables au cas actuel.

Ces valeurs seront considérées comme «observées» par EViews, même si elles correspondent à des périodes futures⁹². Elles peuvent être utilisées comme points de départ, et il est possible (en particulier dans le second cas) que cela permette d'améliorer la vitesse et la probabilité de convergence. La question est de savoir si le taux de croissance des variables est plus élevé que la différence pour la même période

⁹² Eviews laisse à l'utilisateur la gestion de la date courante.

entre les deux solutions. Les deux cas sont tout à fait possibles, en fonction de la différence entre les hypothèses, dans le second cas.

Dans tous les cas, le principe général s'applique: si une option ne fonctionne pas et un autre est disponible, essayez-la !

8.1.3 RESOUDRE DES PARTIES DU MODELE

Comme présenté précédemment, une façon efficace de diagnostiquer l'origine des problèmes de convergence est d'exclure une ou plusieurs variables endogènes de la résolution. Ceci nécessite de générer des valeurs pour ces variables. Construire au départ des valeurs pour toutes les variables permet d'utiliser cette fonction sans coût ultérieur.

8.1.4 VERIFIER L'EXISTENCE D'UNE SOLUTION DE LONG TERME

Nous allons maintenant présenter une façon très efficace de vérifier que les spécifications du modèle permettent une solution à long terme.

Pour que cette solution existe, les conditions suivantes doivent être remplies:

- Toutes les hypothèses exogènes qui partagent la même dimension doivent croître sur le long terme à un taux commun.

Bien sûr, leur évolution peut différer de ce taux pour les premières périodes.

Comme indiqué précédemment, dans la pratique, trois taux de base apparaissent:

- les variables à prix constants (nous appellerons leur taux txq),
- les déflateurs (appelons le txp),
- les populations (appelons le txn).

Les autres taux seront des combinaisons des trois (en plus de la stabilité). Les principales combinaisons concerneront la productivité du travail et du pouvoir d'achat des salaires (txq - txn).

- Les tendances doivent être arrêtées, ou parfois fixées aux valeurs théoriques (comme pour l'évolution de la productivité du travail).
- Dans les formules estimées, les équations à long terme (en principe les relations de cointégration) doivent utiliser des éléments sans dimension, et la partie dynamique aussi (ou alors des logarithmes). Cette dernière condition n'est pas absolument requise, mais cela rend les choses beaucoup plus simples, et c'est de toute façon nécessaire si nous voulons que les équations soient homoscédastiques.

Dans les conditions ci-dessus, toute formule doit donner une solution croissant à un rythme associé à l'élément qu'elle définit. Par exemple, l'équation d'importations à prix constants doit donner un résultat croissant à taux txq .

Il suffit qu'une seule équation ne remplisse pas ces conditions pour empêcher l'ensemble du modèle d'avoir une solution à long terme, sans parler de l'atteindre.

- En outre, les formules estimées doivent conduire à la solution et pas s'en éloigner. En particulier, le signe des coefficients associés à la correction d'erreurs doit être le bon, et la valeur ne doit pas être suffisamment élevée pour créer un processus divergent.

Si ces conditions ne sont pas remplies, le modèle va probablement exploser dans le long terme, et il est fort possible que ce processus affecte déjà les solutions à moyen terme. Pour être sûr que ce n'est pas le cas, la seule option est de vérifier la présence de cette solution à long terme.

Bien sûr, on peut commencer par la simulation de l'ensemble du modèle sur une période assez longue. Malheureusement, même si les précautions citées plus haut ont été prises, il restera généralement au moins une erreur, qui se propagera à l'ensemble du système, sans généralement donner des indices clairs quant à l'origine du problème.

Cependant, une technique très simple peut être appliquée: la production d'une vérification résiduelle sur l'avenir!

Il suffit de générer (comme proposé plus tôt) un ensemble complet d'éléments exogènes et endogènes respectant (en principe) les contraintes ci-dessus. Ensuite le modèle est résolu en utilisant l'option option «fit» autrement dit «d=f» (calcul séparé de chaque équation) et le taux de croissance du résultat est comparé à celui de la variable que l'équation définit. Les deux doivent être identiques. Sinon, les éléments de l'équation ne sont pas compatibles avec sa formulation.

Les causes des problèmes sont un peu différentes d'une vérification résiduelle normale. Les plus fréquentes sont:

- Une tendance a été oubliée ou mal calculée dans l'équation, par exemple l'ouverture du commerce mondial dans l'équation d'importations⁹³.

⁹³ Cette tendance peut se poursuivre au cours des périodes de prévision initiales, mais il faut la diminuer et la faire disparaître après un certain temps. Dans ce cas, il faut vérifier que les résidus suivants remplissent cette condition.

- La dimension d'un ou plusieurs éléments présents à droite n'est pas conforme à son rôle. Cela peut s'appliquer à des identités ou des formules estimées.

Par exemple, la variable exogène associée aux prestations sociales est calculée en pouvoir d'achat par tête, tandis que l'équation correspondante multiplie simplement cet élément par le déflateur. Ou une estimation linéaire ajoute des éléments de dimensions différentes.

- La dimension de la variable endogène est incompatible avec la formule.

Par exemple, une fonction de production CES utilise de mauvais coefficients, utilisés comme exposants.

- Nous avons décidé d'appliquer des formules différentes pour le passé ou l'avenir, par exemple en bloquant la tendance après la période d'estimation, le deuxième type pouvant contenir des erreurs.

Par exemple, un terme de tendance tel que:

$$C(1)*((t-2007) (t<2013)+0.02*t*(t>=2013))$$

au lieu de:

$$C(1)*((t-2007) (t<2013)+0.02*(t-2013)*(t>=2013))$$

passera immédiatement de 0 à $0,02*2013$ lorsque les prévisions débiteront en 2007.

L'avantage évident de cette technique est d'identifier en une seule étape la plus grande partie des équations erronées. Mais, comme dans l'application traditionnelle de la méthode, un résidu nul ne prouve pas l'absence d'erreur, et une valeur non nulle peut provenir de plusieurs erreurs. Cela signifie qu'en général atteindre une version (apparemment) propre prendra plusieurs itérations.

Le danger le plus immédiat est une erreur dans les hypothèses: le programme ci-dessus, qui fixe les taux de croissance de toutes les variables exogènes, devrait éviter ce problème. Nous avons vérifié que les taux de croissance sont corrects et conformes aux formules dans lesquelles ils sont utilisés. La seule erreur pourrait provenir de valeurs précisées de manière explicite (par exemple, parce que nous avons des informations préalables sur certains éléments, ou parce que nous essayons de spécifier une cible endogène en manipulant des instruments exogènes).

Les programmes proposés à la fin de l'ouvrage utiliseront systématiquement cette technique.

8.2 LES PROBLÈMES DE CONVERGENCE DANS LE COURT TERME

Le succès complet du test ci-dessus ne garantit cependant pas de trouver une solution cohérente: le premier problème réside dans les périodes initiales, et en particulier dans la toute première⁹⁴, où la méthode précédente ne peut être complètement appliquée. En effet les termes retardés de droite croissent encore au taux historique.

Enfin, il est possible que la gestion des résidus dans les premières périodes de prévision apporte de fortes variations, et en particulier les cycles importants, qui

- feront passer l'un des éléments par des valeurs inacceptables (comme des niveaux négatifs d'une variable représentant un bien, dont le modèle utilisera le logarithme) ;
- produiront un modèle aux propriétés divergentes. Cela peut être vrai pour Gauss-Seidel, si le modèle linéarisé est très différent de l'original, et peut alors contenir des valeurs propres avec un module supérieur à un. Et la méthode de Newton peut faire partir la résolution dans la mauvaise direction.

Quant aux remèdes, ils ne sont pas vraiment différents de ceux développés dans le chapitre précédent.

8.3 LES PROBLÈMES DE CONVERGENCE DANS LE MOYEN ET LONG TERME

Si le modèle a commencé par converger, et contient un processus à correction d'erreur, il doit donner des solutions sur toute la période.

Bien sûr, ce n'est pas vrai si la seconde condition n'est pas remplie. Dans ce cas, nous ne savons rien sur les propriétés de long terme, et la divergence peut survenir à tout moment (en fait le plus grand danger intervient lorsque la convergence est obtenue jusqu'à la fin de la période de prévision, mais les propriétés numériques sont déjà mauvaises, pas encore assez mauvaises cependant pour être remarquées).

Dans le cas contraire, le seul danger qui demeure est l'apparition de cycles trop forts, ce qui ne peut se produire que si le modèle est erroné. Mais maintenant nous avons des informations, comme si nous avions testé un moteur qui tourne pendant un certain temps avant de casser: il est beaucoup plus facile de comprendre la raison que s'il ne démarre pas du tout.

Les principales solutions ont été développées précédemment:

⁹⁴ De même que le décollage est la partie la plus dangereuse d'un vol en avion.

- L'application de chocs sur les principales hypothèses et l'observation de leurs conséquences. Le processus de divergence pourra débuter avant l'explosion, et ses raisons seront rendues plus explicites.
- Exogénéiser certaines variables, ou changer (augmenter ou diminuer) les explications localisées. Nous pouvons alors observer comment ces changements affectent le nombre de périodes avant l'accident, et les propriétés des évolutions: moins cycliques ou plus conformes à la théorie économique. Cela est beaucoup plus instructif que d'essayer de forcer l'obtention d'une solution.

8.4 LE TESTS DES RESULTATS

Supposons maintenant que notre modèle converge sur le long terme.

8.4.1 LES RESULTATS DES SIMULATIONS

Ce que nous devons vérifier à ce stade est la manière dont le modèle se comporte à court et moyen termes. Bien sûr, nous avons introduit des hypothèses trop simplistes pour fournir une prévision réaliste. En particulier, la fixation immédiate de la croissance des biens à la valeur q va stopper l'expansion des exportations (les importations devraient également suivre). Une option plus réaliste serait de la réduire progressivement. Toutefois:

- L'impact de cette diminution progressive va changer avec les évolutions du modèle, son traitement devrait donc être reconsidéré périodiquement.
- Cela va à l'encontre de l'objectif principal des simulations du modèle de l'avenir, qui est d'obtenir des trajectoires stables, donc plus faciles à interpréter, dans un avenir aussi proche que possible.

Dans la pratique, la qualité de nos simulations devra être assez raisonnable pour ne pas remettre en question les enseignements de sensibilité aux chocs. Cela représentera le diagnostic le plus important, du moins pour l'instant.

8.4.2 LES RESULTATS DES CHOCS

Pour les chocs, nous allons utiliser la même fonctionnalité «scenario» que dans le chapitre précédent. Nous en avons déjà décrit les avantages à attendre ce stade pour les tests. La situation n'a pas changé.

Comme nous l'avons dit plus haut, c'est sur l'avenir que la réalisation des chocs est la plus pertinente:

- Ils peuvent durer plus longtemps.
- Ils partent d'une base régulière.
- La période est la même que pour les utilisations futures.

En particulier, on peut vérifier les valeurs exactes de long terme et des ratios. La convergence est elle-même garantie par la production préalable d'un scénario de long terme, à condition que le choc ne soit pas trop grand.

En ce qui concerne l'horizon de la période elle-même, il est préférable de commencer par une très grande valeur (comme deux ou trois siècles). Ce choix pourrait sembler anormalement élevé, mais il permet:

- de vérifier qu'en effet le long terme est atteint numériquement ;
- d'interpréter la convergence (ou la divergence) des cycles potentiels, qui doivent être observés en nombre suffisamment élevé. Des cycles avec une période de 50 ans ne sont pas rares (surtout si le modèle exagère légèrement celui-ci) et l'observation de 4 cycles nécessite alors 200 ans ...

Bien sûr, une fois que le modèle est stabilisé, les cycles vont probablement converger plus rapidement, et la période de test peut être réduite, ce qui permet de se concentrer sur les propriétés à court et moyen terme (qui seront beaucoup plus visibles dans des graphiques raccourcis).

8.5 LES ASPECTS EViews

Nous allons maintenant voir comment nous pouvons appliquer les principes ci-dessus sous EViews.

8.5.1 LA PRODUCTION DES SIMULATIONS

Nous allons maintenant décrire successivement les phases du processus de simulation.

8.5.1.1 La préparation du fichier de travail («Workfile»)

Voyons comment nous pouvons gérer ceci sur notre modèle simple. Nous nous concentrerons sur les caractéristiques spécifiques (le programme complet sera présenté plus tard).

La première tâche est la production d'un fichier de «prévisions». La période doit être assez grande

Nous avons défini `_fra_1` comme fichier de travail

```
close _fra_1.wf1
close _fra_1.wf1
wlopen _fra_1.wf1
wfsave _fra_1.wf1
```

Nous étendons le fichier de travail jusqu'à 2200 ans (période réduite plus tard)

```
pagestruct(end=2200) *  
scalar txq=0.015  
scalar txn=0.01
```

Nous extrapolons la tendance

```
smpi 2002S2 2200S2  
genr T=t(-1)+0.5
```

Maintenant, nous allons produire les hypothèses.

Heureusement, la plupart des hypothèses n'ont pas de dimension, et sont donc indépendantes des taux de croissance.

Les seules exceptions sont les suivantes:

- la demande du gouvernement et la demande mondiale, gd et wd , quantités qui croissent à txq ,
- l'emploi du gouvernement LG qui croît à txn ,
- le taux de salaire réel qui augmente avec la productivité du travail, à $txq-txn$, ou plutôt $(1+txq)/(1+txn)-1$.

Les exogènes

```

for %2 gd wd
genr {%2}={%2}(-1)*(1+txq)
next
for %2 rfdx r_rhiq r_ih dr sr rvat pk compm tc ec_i ec_le ec_m ec_x ec_ci
genr {%2}={%2}(-1)
next
for %2 lg
genr {%2}={%2}(-1)*(1+txn)
next
for %2 wr
genr {%2}={%2}(-1)*(1+txq)/(1+txn)
next

```

Nous pouvons également initialiser les endogènes (comme cela a été justifié précédemment).

```

for %2 ic i m x co fd k ih rhi q cap ci
genr {%2}={%2}(-1)*(1+txq)
next
for %2 ur res_m res_x
genr {%2}={%2}(-1)
next
for %2 le lt led
genr {%2}={%2}(-1)*(1+txn)
next
for %2 prle_t
genr {%2}={%2}(-1)*(1+txq)/(1+txn)
next

```

Nous devons aussi parfois modifier les coefficients

Pour que notre simulation fournisse une solution croissant régulièrement, nous avons seulement besoin de changer un coefficient: l'évolution de la productivité du travail qui doit suivre $(1+txq)/(1+txn) - 1$ à

chaque période. Comme il s'applique à t , qui croît de 0,5, il faut multiplier le coefficient par 2^{95} . Et comme nous l'avons vu, notre manipulation des tendances fait que le changement de coefficient s'applique seulement à partir de 2002, sans remettre en cause les valeurs précédentes de la productivité tendancielle.

$$c_prle(2)=2*\log((1+txq)/(1+txn))$$

8.5.2 LA PRODUCTION D'UNE SIMULATION DE BASE

Pour produire une solution de base, nous utilisons la même technique que d'habitude. Nous calculerons les taux de croissance des endogènes pour vérifier qu'elles convergent vers les valeurs de long terme (zéro, txq , txn ou une combinaison des deux).

Deux remarques:

- Nous allons également calculer les taux de croissance théoriques, et la différence entre les taux réel et théorique, qui devrait se réduire à zéro dans le long terme.
- L'ordre «exclude» assure que toute exclusion précédente associée au scénario actuel est abandonnée.

```
wfsave _fra_1
smpl 2003S1 2200S2
_fra_1.scenario "Scenario 1"
_fra_1.append assign @all _p
_fra_1.override
_fra_1.solveopt(n=t m=1000,c=1e-6,o=g,d=d)
solve _fra_1
```

Nous calculons le taux de croissance de toutes les variables et le comparons à la valeur théorique

⁹⁵ 4 pour un modèle trimestriel.

```

for !i=1 to g_vendo.@count
%st2=g_vendo.@seriesname(!i)
genr tc_{%st2}=100* ({%st2}_p-{%st2}_p(-1))/((st2)_p(-1)+({%st2}_p(-1)=0))
genr tc0_{%st2}=100* ({%st2}-{%st2}(-1))/((st2)(-1)+({%st2}(-1)=0))
genr dtc_{%st2}=tc_{%st2}-tc0_{%st2}
next

```

8.5.3 LA PRODUCTION DES CHOCS

Pour produire des chocs, nous utiliserons la même technique que précédemment. Nous appliquons à gd une augmentation de 1 point de PIB (en utilisant Q_P , le niveau du PIB calculé par la simulation de référence). Nous résoudrons le modèle en utilisant le suffixe «_v». Ensuite, nous calculerons les différences absolues et relatives (détruisant d'abord toute variable de ce type, afin d'éviter les confusions).

Enfin, nous allons exporter les résultats au format Excel (cela devient de moins en moins nécessaire).

Une astuce: quand une simulation se décompose en un programme qui contient plusieurs déclarations SOLVE pour un modèle donné, dont l'une échoue, le message d'EViews ne vous permettra pas d'identifier laquelle d'entre eux a produit l'erreur. Une façon simple de réaliser cette identification est de modifier le nombre maximal d'itérations dans la déclaration (de 1000 en 1001, 1002, ...), car cette déclaration sera affichée dans le message.

```

genr gd_v=gd+.01*q_p*(t>=2004)
_fra_1.solveopt(n=t m=1000,c=1e-6,o=g,d=d)
_fra_1.scenario "Scenario 1"
_fra_1.append assign @all _v
_fra_1.override gd
smpl 2003S1 2200S2
solve(m=1001) _fra_1
delete dv_* pv_*
for !i=1 to g_vendo.@count
%st1=g_vendo.@seriesname(!i)
series dv_{%st1}={%st1}_v-{%st1}_p
series pv_{%st1}=100*dv_{%st1}/{%st1}_p
next
group shocks pv_* dv_*

```



```
write(t=xls) varia_pfra.xls chocs
```

8.5.4 LA MODIFICATION DES SPÉCIFICATIONS DU MODELE

Pour tester les propriétés du modèle à long terme, il peut être intéressant

- de tester des formules alternatives estimées, dont on comparera les propriétés ;
- de permettre à certaines de ses équations de suivre des formulations différentes de celles estimées, par un changement de valeurs de coefficients, voire un changement de la formule.

Cela peut être fait bien sûr en produisant un nouveau programme de génération de modèle. Mais cela présente plusieurs inconvénients:

- cela prendra plus de temps et impliquera d'autres déclarations ;
- il n'y aura plus une seule version « officielle » ;
- on devra traiter plusieurs programmes de gestion de modèles, ce qui rend la gestion des fichiers moins claire ;

le programme de prévision devra être exécuté de toute façon.

Dans notre sens, il vaut mieux:

- conserver une version officielle du modèle, qui représente la meilleure version testée selon les informations actuelles ;
- introduire les changements dans le programme de prévision.

Cela peut se faire de deux manières.

- Si une nouvelle version a été estimée pour une équation, de façon définitive, le jeu complet des éléments qui la génèrent devrait être spécifié (mise de la série des résidus à zéro, définition de l'équation et de l'estimation, stockage des résidus estimés, fusion de l'équation avec le modèle).
- Si l'on ne fait que tester une version alternative avec au moins un coefficient fixe, avec ou sans ré estimation du reste, nous suggérons la procédure suivante, que nous allons présenter en prenant l'investissement comme exemple:
 - ★ Copier dans le programme « prévision » l'ensemble des déclarations, comme ci-dessus.
 - ★ Spécifier un vecteur de paramètres (qui ne sera donc pas estimé).

```
vector(10) p_i
```

- ★ Estimer l'équation, comme d'habitude, avec un nom différent (ici: le nom de la normale suivie par «1»).

```
coef(10) c_i
smpl 1977S1 2002S1
genr ec_i=0
equation eq_i1.ls i/k(-1)=c_i(1)*i(-1)/k(-2)+c_i(2)*ur+c_i(3)*.25*q/q(-4)+c_i(4)+ec_i
genr ec_i=resid
```

Cette étape peut être ignorée si vous êtes sûr que l'équation actuelle est correcte, et que vous connaissez les résultats de l'estimation. Cependant, nous vous la conseillons car elle affiche la situation et vous aide à contrôler le processus.

- ★ Établir un transfert entre les coefficients et les éléments du vecteur, prenant en compte tout changement souhaité dans les valeurs. Ici, nous voulons fixer le coefficient de tension à 0,5.

```
p_i(1)=c_i(1)
p_i(2)=0.5
p_i(3)=c_i(3)
```

- ★ Remplacer tous les coefficients «C_» par des paramètres «P_», à l'exception du terme constant (ce qui permettra à EViews d'estimer l'équation, un processus qui nécessite au moins un coefficient⁹⁶).
- ★ Estimer l'équation en utilisant le nom normal.
- ★ Stocker la nouvelle valeur résiduelle.

```
equation eq_i.ls i/k(-1)=p_i(1)*i(-1)/k(-2)+p_i(2)*ur+p_i(3)*.25*q/q(-4)+c_i(4)+ec_i
```

⁹⁶ Nous n'avons pas de solution simple si l'équation n'a pas de terme constant (comme pour les variations de stocks dans notre exemple).

```
_fra_1.merge eq_i
genr pq_ec=resid
```

Nous pourrions modifier plus d'un coefficient, bien sûr, ou aucun, si nous voulons simplement être prêts pour une modification future.

Pour appliquer cette technique, on peut considérer deux stratégies générales.

- Commencer sans aucun changement, et copier successivement chaque équation modifiée, au fur et à mesure des besoins.
- Copier dès le début toutes les équations de comportement dans le programme de prévision, sans changement de valeur (tous les éléments «C_» sont transférés en leurs équivalents «P_»). Les spécifications ont changé, mais pas les propriétés numériques.

Les paragraphes demandés seront modifiés tour à tour plus tard, au fur et à mesure des besoins. Cela demande un fort investissement initial, mais rend les modifications ultérieures plus faciles et moins sujettes à l'erreur.

Comme vous pouvez le deviner, nous sommes favorables à la deuxième option. Mais il faut tenir compte de la part d'équations qu'on estime avoir l'intention de changer⁹⁷.

Remarque: lorsque vous testez de nouvelles valeurs, nous vous conseillons de garder l'ancienne comme un commentaire, ce qui vous permettra de revenir facilement en arrière si la situation se détériore, et vous rappellera l'ampleur de votre changement.

8.6 LES ANTICIPATIONS RATIONNELLES

Elles sont appelées par EViews «model consistent expectations» et leur raisonnement est un peu différent de l'habitude.

8.6.1 LE CADRE

Dans un cadre d'anticipations rationnelles, nous supposons que certains des agents (au moins) sont conscients des hypothèses futures, et sont en mesure d'évaluer leur influence sur certaines (au moins)

⁹⁷ Qui est difficile à prévoir.

des variables du modèle. Ils n'ont pas besoin de la connaissance formelle des équations du modèle réel, simplement d'être capables de réaliser l'application mathématique menant des hypothèses aux endogènes.

Comme ils vont utiliser ces connaissances dans leurs décisions actuelles, certaines des variables actuelles dépendront de valeurs futures, que ce soit des hypothèses ou des éléments endogènes déterminés par ces hypothèses.

Pour prendre en compte les anticipations rationnelles on n'a pas besoin d'y croire. L'interprétation des différences dans le comportement économique (et de leurs conséquences pour l'équilibre) entre les agents ayant cette possibilité est très intéressante au moins d'un point de vue théorique. L'exemple suivant devrait faire la lumière sur ce point.

8.6.2 LES CONSEQUENCES POUR LES SIMULATIONS DU MODELE

Dans ce contexte, on ne peut plus calculer la solution pour chaque période séparément, en progressant du passé au futur. La solution pour une période future donnée dépendra de valeurs pour lesquelles la solution n'a pas encore été obtenue.

Cela introduit deux problèmes:

- trouver une façon de prendre en compte les valeurs futures, appartenant à la période de prévision ;
- trouver un moyen pour traiter les dernières simulations, qui tiennent compte de valeurs en dehors de la période de prévision. Cela implique la définition des «conditions terminales».

8.6.2.1 à l'intérieur de la période de prévision

Le problème est purement technique. Deux stratégies sont considérées à l'heure actuelle.

- L'algorithme de Fair-Taylor (1983): il résout le modèle pour chaque période successive (en utilisant l'algorithme de Gauss-Seidel, Newton ou Broyden à chaque fois), puis revient à la première période et continue à itérer (cette fois selon un processus de type Gauss-Seidel) jusqu'à la convergence de l'ensemble du système. C'est la technique utilisée par EViews.
- L'algorithme de Laffargue (1990): en gros, il ajoute une dimension temporelle aux équations du modèle (en les dupliquant autant de fois qu'il y a de périodes) et résout le modèle associé dans son ensemble, en utilisant l'algorithme de Newton avec des améliorations possibles (en utilisant le fait que la matrice est bande diagonale, le nombre de matrices non nulles sur une ligne donnée étant fonction du nombre de retards positifs et négatifs).

De toute évidence, l'application du dernier algorithme n'a été rendue possible que par l'augmentation de la puissance de calcul des ordinateurs, et surtout de la capacité de la mémoire, car un très grand nombre de valeurs doit être géré simultanément.

8.6.2.2 au-delà de la période de prévision

EViews offre les possibilités habituelles de conditions terminales, qui sont décidées par

```
t = arg
```

dans la déclaration SOLVEOPT

arg peut prendre les valeurs:

```
«u» (fourni par l'utilisateur – par défaut: les chiffres réels)
«l» (niveau inchangé)
«d» (variation absolue inchangée)
«g» (taux de croissance inchangé)
```

8.6.3 ÉLÉMENTS TECHNIQUES

EViews détecte lui-même la présence dans les équations de valeurs futures des variables endogènes, et basculera automatiquement vers l'algorithme associé (en fait, Fair-Taylor). On peut utiliser les mêmes instructions que d'habitude, et modifier les conditions terminales par défaut, si cela est jugé utile.

8.6.4 APPLICATION A NOTRE EXEMPLE

Dans notre exemple, nous limiterons nos observations à l'équation d'investissement.

Dans le modèle actuel, nous utilisons:

```
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				
Date: 11/07/12 Time: 12:03				
Sample (adjusted): 1977S2 2002S1				
Included observations: 50 after adjustments				
I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.25*Q/Q(-4)+C_I(4)+EC_I				
	Coefficient	Std. Error	t-Statistic	Prob.
C_I(1)	0.825643	0.033524	24.62809	0.0000
C_I(2)	0.027926	0.006984	3.998580	0.0002
C_I(3)	0.152870	0.058292	2.622495	0.0118
C_I(4)	-0.052515	0.010337	-5.080438	0.0000
R-squared	0.959393	Mean dependent var		0.087074
Adjusted R-squared	0.956745	S.D. dependent var		0.006785
S.E. of regression	0.001411	Akaike info criterion		-10.21218
Sum squared resid	9.16E-05	Schwarz criterion		-10.05922
Log likelihood	259.3046	Hannan-Quinn criter.		-10.15393
F-statistic	362.2686	Durbin-Watson stat		0.628797
Prob(F-statistic)	0.000000			

Maintenant, nous allons examiner quatre options:

- Le cas de base (utilisation des seules variables du présent et du passé)

$$I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.25*(Q-Q(-4))/Q(-4)+C_I(4)$$

- Un cas mixte:

$$I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.25*(Q(2)-Q(-2))/Q(-2)+C_I(4)$$

- Une prise en compte du seul futur (et présent):

$$I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.25*(Q(4)-Q)/Q+C_I(4)$$

- Un autre cas de prise en compte du futur qui étend l'horizon des connaissances:

$$I/K(-1)=C_I(1)*I(-1)/K(-2)+C_I(2)*UR+C_I(3)*.05*(Q(20)-Q)/Q+C_I(4)$$

L'influence globale de Q est la même dans chaque cas, seuls les décalages changent.

8.6.5 LE TEST

Nous considérons comme d'habitude que l'État augmente sa demande de 1 % du PIB, mais seulement de 2010 à 2050. La durée de cette période peut sembler trop élevée⁹⁸, mais sera nécessaire pour l'analyse.

Bien sûr, on pourrait aussi appliquer la technique à d'autres éléments: l'emploi et la variation des stocks en particulier. Cela ne serait pas difficile.

8.6.6 LES RESULTATS

Nous nous concentrerons sur quelques graphiques. Tous les changements sont mesurés en pourcentage par rapport à la simulation de base.

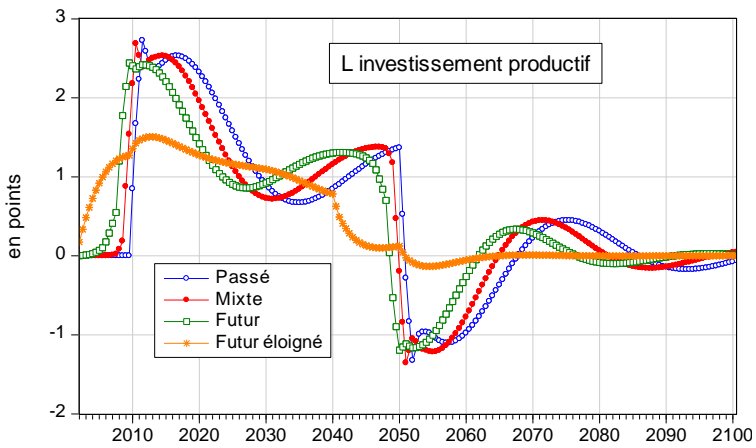
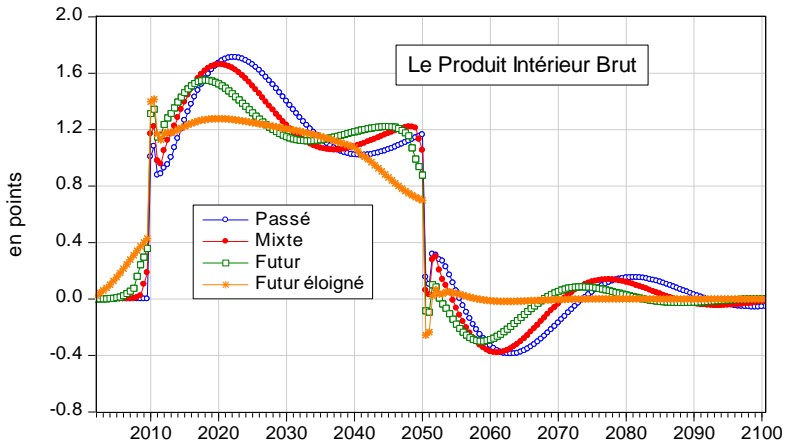
Commençons par le produit intérieur brut, avec les quatre options:

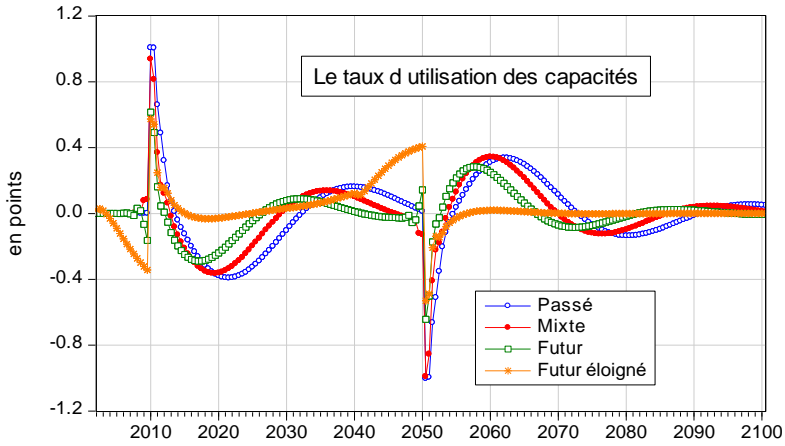
- passé: 4 périodes passées donc deux années.
- mixte: 2 périodes passées, 2 périodes futures.
- futur: 4 périodes futures.
- futur éloigné : 20 périodes futures.

Les résultats sont présentés en points.

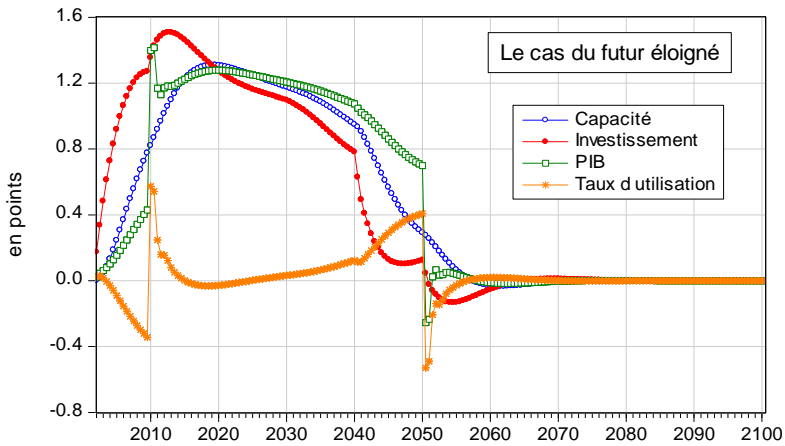
⁹⁸ En effet cette mesure conduira à l'augmentation de l'endettement primaire de l'État de 40 points de PIB (sans tenir compte des intérêts supplémentaires). Le niveau de dette atteint est certainement insupportable, réduisant l'intérêt de l'étude, en particulier en termes de réaction des autres agents.

Pour les mêmes simulations, nous présenterons maintenant l'écart sur l'investissement productif et le taux d'utilisation des capacités.





Le dernier graphique présentera les évolutions de la capacité de production, de l'investissement, du PIB et du taux d'utilisation pour le cas du futur éloigné.



Nous pouvons observer très clairement des éléments généralement logiques:

- Plus la décision d'investissement est fondée sur l'avenir, plus l'augmentation des investissements s'adapte tôt.
- Ceci est également vrai si l'horizon augmente. Dans ce cas, le poids des périodes associées à un changement de l'augmentation de décisions croît.
- Il en résulte une adaptation plus lisse à un changement dans les décisions du gouvernement, dans les deux sens. L'intensité des cycles diminue, et ils convergent plus rapidement.
- Avec un horizon de 10 ans (cas 4), l'écart de production ne change pas trop. Ses valeurs les plus élevées se présentent avant les changements, lorsque les décisions d'investissement (et les modifications de capital qui en résultent) sont dues aux changements futurs, pas aux conditions présentes. La sur réaction à la baisse des capacités observées dans les autres cas n'est pas présente, et la convergence des taux d'utilisation est monotone.
- Les modifications apparaissent plus tôt que l'horizon des anticipations lui-même. Les premières réactions des entreprises anticipatrices sont prises en compte encore plus tôt par d'autres entreprises, introduisant un effet en cascade (plutôt limité, cependant).

8.7 LES SIMULATIONS STOCHASTIQUES

Utilisée à des fins scientifiques, mais aussi pour l'analyse d'un modèle donné, l'incertitude du modèle en prévision peut être mesurée directement. Cette erreur peut provenir de plusieurs sources.

8.7.1 DES ERREURS PUREMENT STATISTIQUES

- la présence d'un terme aléatoire résiduel dans les équations de comportement ;
- l'incertitude sur les coefficients estimés, même avec une spécification exacte.

Quand un modèle est produit, certaines équations sont estimées en utilisant une technique économétrique, allant des simples moindres carrés ordinaires à des méthodes plus sophistiquées comme la cointégration. Ce processus permet:

- une estimation de la loi déterminant l'écart-type de l'estimation,
- l'estimation de l'écart-type des coefficients estimés.

La précision de ces deux éléments dépend de la taille de l'échantillon, d'une manière différente.

L'augmentation de la taille de l'échantillon doit améliorer l'estimation des coefficients.

Mais pour l'estimation des résidus, une fois que l'échantillon est suffisamment grand pour appliquer la méthode, son augmentation apporte essentiellement une meilleure connaissance de sa distribution (donc en pratique de son écart-type). Cela est dû au fait que, même avec un grand échantillon, l'introduction de

nouvelles explications, au-delà de cinq ou six variables, conduit très rapidement à des problèmes de colinéarité.

8.7.2 DES ERREURS SUR LES FORMULATIONS ET LES HYPOTHÈSES

En fait, ces deux erreurs ne sont que la moitié des types affectant les prévisions des modèles. Il faut également considérer:

- Le fait que la formule estimée **elle-même** peut être fausse. Même si une formule est acceptée par tous les tests, on peut toujours trouver une autre formule (parfois avec des propriétés très différentes) qui montre une qualité statistique identique ou même meilleure. Et même si la formule s'applique bien au passé, les agents peuvent très bien modifier leur comportement par la suite.
- L'erreur sur les hypothèses retenues pour les prévisions, dont le modèle ne doit pas être tenu pour responsable.

8.7.3 LE TRAITEMENT DES ERREURS

Dans la pratique, seuls les deux premiers types peuvent être mesurés sans trop de difficultés, en utilisant des statistiques (et en supposant que les deux derniers types sont absents): l'estimation des coefficients donne sous certaines hypothèses une estimation de la loi des résidus d'estimation, ainsi que de la loi des coefficients eux-mêmes. Les moments des variables projetées peuvent ensuite être calculés, soit:

- en utilisant une loi de probabilité (en principe la loi normale) pour tirer au hasard des échantillons de résidus, en réalisant les simulations associées, et en observant les caractéristiques statistiques des résultats (technique «Monte Carlo»);
- en déterminant, analytiquement ou numériquement, la transformation linéaire menant des résidus vers la solution du modèle, et en appliquant cette transformation directement à la loi des résidus, pour obtenir la loi des variables prévues;
- plus simplement et sous nettement moins d'hypothèses restrictives, par tirage au hasard d'un échantillon dans la séquence de résidus observés (avec ou sans remise des éléments sélectionnés dans l'ensemble). Cette technique est appelée «bootstrap».

Le même type de méthode s'applique à l'incertitude due aux coefficients:

- en répétant suffisamment de fois la séquence suivante d'opérations: élaboration d'un vecteur de résidus sur la période d'estimation (en utilisant la loi estimée ou la technique «bootstrap»), introduction dans le modèle, simulation de celui-ci pour obtenir un nouvel échantillon de variables cohérentes (en termes de modèle), ré-estimation des coefficients sur la base de cet

échantillon, et prévision du modèle ré-estimé. Il faudra ensuite mesurer les moments de l'échantillon des projections ;

- en utilisant la loi des coefficients pour générer un échantillon de modèles, qui, par leur prévision fournira un échantillon de valeurs des variables ;
- en déduisant analytiquement de la loi de coefficients la loi de la variable projetée, et en utilisant la matrice de transformation (jacobienne) pour relier les variables aux coefficients.

Trois types d'erreurs principaux devraient être envisagés:

- Le biais: dans le cas d'un modèle aux propriétés non-linéaires (par rapport aux endogènes, à la fois instantanées et retardées), la moyenne de la solution ne sera pas identique à la solution déterministe (obtenue avec une valeur résiduelle nulle): cette nouvelle prise en compte introduit un biais.

La différence vient bien sûr de la non-linéarité des équations, et tout modèle économique présente des non-linéarités. Les cas les plus évidents sont la présence de variables à prix courants, le produit d'une variable à prix constants par un déflateur, ou des variables calculées en appliquant un taux de croissance de leur valeur passée.

Si l'on est seulement intéressé par cet élément, une seule expérience (avec un grand nombre de répétitions) mesurera l'amélioration fournie par les simulations stochastiques. Si elle est suffisamment faible, on peut s'en tenir par la suite à des simulations déterministes beaucoup moins coûteuses.

Il faut être conscient que l'incertitude sur les coefficients peut introduire des non-linéarités, même si le modèle est linéaire par rapport à ceux-ci (en fait il n'est pas linéaire à l'ensemble: variables + coefficients réunis).

Prenons le modèle le plus simple possible (avec Q = PIB, C = la demande privée, g = demande du public):

$$\begin{aligned} Q_t &= C_t + g, \\ C_t &= a \cdot Q_t \end{aligned}$$

qui peut être écrit comme:

$$Q_t = g_t / (1 - a)$$

qui n'est pas linéaire par rapport à a .

- L'écart-type: cette fois l'intérêt réside dans la dispersion des solutions autour de leur moyenne. Ce critère permet d'évaluer la fiabilité des résultats (en partant peut-être de la moyenne des simulations stochastiques) en évaluant un intervalle de confiance, ou une plage de valeurs possibles.
- La distribution: ce que nous voulons est un graphique de la distribution de probabilité des solutions aléatoires. Évidemment, la production d'un graphique cohérent appelle à davantage de simulations que dans les cas précédents. Un message intéressant est son caractère symétrique. En fait, EViews prend en compte la dissymétrie, même sans un graphe, par la production d'un intervalle de confiance qui élimine les 2,5 % des valeurs plus grandes et plus petites.

L'analyse de l'incertitude n'est pas nécessairement limitée aux valeurs des variables. Elle peut également être mesurée sur les multiplicateurs (ou l'ensemble des dérivées de variables relatives aux exogènes), donc sur l'efficacité des politiques économiques, ou sur les valeurs propres de la forme structurelle, et donc sur la probabilité de convergence du modèle (voir le paragraphe suivant).

Enfin, on pourrait envisager de traiter la covariance entre les équations. La cointégration devrait éliminer le problème, car elle individualise un groupe de variables ou concepts qui évoluent indépendamment du reste du modèle (à condition que chaque ensemble contienne une seule équation). Cependant, ceci n'est pas tout à fait vrai, comme une variable peut appartenir à plusieurs équations de cointégration, établissant un lien entre elles.

Si l'on considère également l'erreur sur les coefficients, le processus est un peu plus complexe. Encore une fois, les ensembles d'équations de cointégration sont indépendants, mais ils vont toujours utiliser plus d'un coefficient (on doit considérer l'équation dynamique et la relation à long terme). Dans ce cas, il est nécessaire de prendre en compte la corrélation entre les coefficients et faire appel à une loi normale multivariée, avec une matrice de covariance non diagonale.

Les deux dernières formes d'erreur, sur la forme fonctionnelle et les hypothèses sont plus difficiles à évaluer. Toutefois:

- Si l'on dispose déjà d'informations récentes sur certaines variables, les grands agrégats (équilibre offre demande, inflation) et surtout sur les hypothèses du modèle (environnement extérieur, politique publique), on peut simuler le modèle sur ces périodes et observer si leur évolution est respectée.
- Une évaluation approximative de l'erreur sur les formules peut être obtenue en réduisant l'échantillon d'estimation, et en simulant le modèle sur les périodes restantes (comme dans une prévision ex post). Ensuite la taille des résidus peut être comparée à l'écart-type théorique obtenu à partir de la méthode précédente (qui supposait l'utilisation des vraies équations). Si l'erreur est beaucoup plus élevée que cet écart-type (par exemple systématiquement plus de deux fois sa valeur) on peut supposer que les fonctions estimées sont incapables de prévoir

l'avenir, et la forme fonctionnelle doit être faussée (ou l'agent a changé son comportement, qui est en fait la même chose). Cette technique est similaire au test de Chow en prévision («Chow forecast test»).

La critique évoquée plus haut disparaît, car nous n'envisageons pas de modifier notre choix de modèle en fonction du résultat.

Quant à l'erreur sur les hypothèses, elle est en dehors de la responsabilité du modèle (uniquement des producteurs de la prévision). Mais une estimation de celle-ci peut être obtenue par l'application d'un processus ARMA aux hypothèses externes (pas aux instruments), et l'observation de la fiabilité de la simulation associée. L'erreur est alors censée être la part de l'hypothèse qui n'avait pas été prévue, et l'erreur en prévision l'écart entre les deux simulations.

8.7.4 L'INTERET DE CETTE TECHNIQUE

Elle a un coût et un avantage. Le coût est l'allongement de la durée des calculs.

Pour obtenir une mesure précise de la distribution, un minimum de 500 répétitions est jugé nécessaire, mais 1000 ou même 10000 est un meilleur chiffre. C'est clairement la raison principale de l'utilisation relativement récente de cette technique: il y a 30 ans, la résolution 1 000 fois d'un modèle même petit pouvait prendre plusieurs heures sur un ordinateur central. Aujourd'hui, l'utilisateur est prêt à payer le prix dans la plupart des cas: sur un micro-ordinateur de puissance moyenne, la simulation sur 20 périodes du modèle à 705 équations MacSim-CAN prend moins d'une seconde. Multiplié par 1 000, on obtient environ 15 minutes, un chiffre raisonnable qui pourrait cependant sembler élevé pour les modélisateurs actuels, en particulier si les simulations font partie d'un processus itératif à la recherche de la meilleure prévision.

8.7.5 RETOUR A NOTRE EXEMPLE

Nous allons résoudre le modèle `_fra_1` plus de 500 périodes, en commençant comme d'habitude en 2002S1.

Le seul changement nécessaire concerne la déclaration `solveopt`, où «`d = d`» est suivi par «`s = a`»:

En outre, l'option `solveopt` «`r = entier`» détermine le nombre de répétitions (par défaut: `r = 1000`), et «`b = nombre`» contrôle l'intervalle de confiance (par défaut: `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)
```

En fait, l'option «a» n'est pas la seule associée à des simulations stochastiques. C'est celle qui donne le plus de renseignements, y compris les intervalles de confiance, alors que «m» ne donne que la moyenne, «s» que l'écart type et «b» la moyenne et l'écart-type.

Dans notre cas, la solution va créer pour chaque variable dans le modèle:

- Une série de moyennes, ajoutant «m» au nom de la variable simulée. Pour le PIB et un suffixe actuel F, le nom de cette variable sera «GDP_FM».
- Une série d'écart-type, utilisant «S», donc appelée GDP_FS.
- Une série pour les limites supérieure et inférieure, utilisant «H» et «L», donc GDP_FH et GDP_FL.

8.7.6 LES RESULTATS

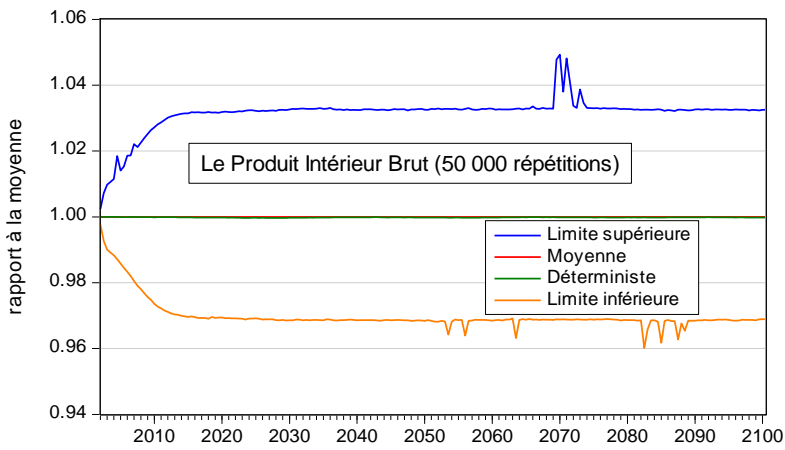
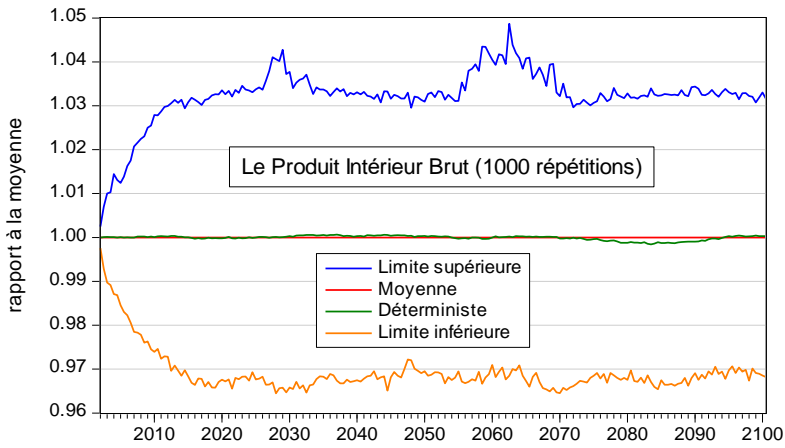
Nous allons effectuer une simulation stochastique sur notre petit modèle, dans les mêmes conditions que la prévision déterministe, ce qui nous permettra de mesurer l'erreur qu'elle introduit.

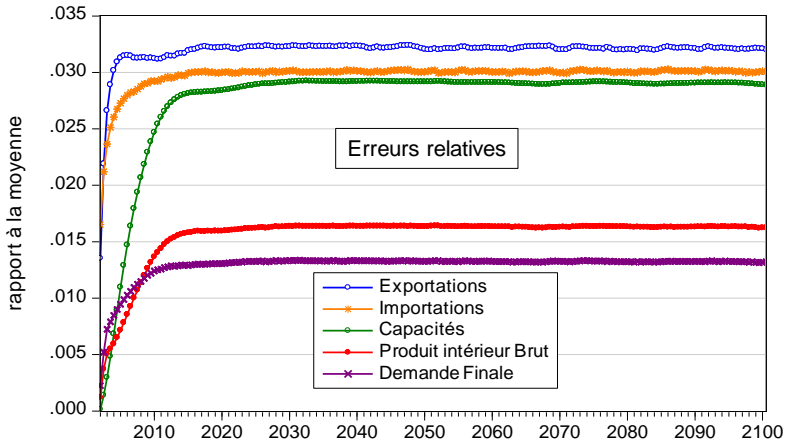
Le programme est très similaire au précédent, sauf pour la ligne:

```
_fra_1.solveopt(n=t m=1000,c=1e-6,o=g,d=s,a)
```

Les graphiques suivants montrent:

- Pour le PIB (Gross Domestic Product), l'évolution de sa valeur limite inférieure et supérieure (Lower / Higher stochastic value), et de la simulation déterministe Deterministic), rapportées à la valeur moyenne (rapport 1). Deux graphes sont fournis, avec 1000 et 50000 répétitions.
- Le rapport entre l'écart-type et la moyenne, pour un ensemble de variables. De haut en bas:
 - Produit intérieur brut,
 - Imports
 - Demande finale
 - Exports
 - Capital





En outre, le tableau suivant donne la différence moyenne par rapport à la moyenne, comparée à l'erreur obtenue lors de la simulation *a posteriori* sur le passé.

Variable	Erreur stochastique	Erreur ex post
CAP	2,75	1,69
CO	1,36	1,43
FD	1,29	1,36
I	4,37	2,85
K	2,77	2,70
LE	1,77	1,68
LED	1,61	1,36
LT	1,35	1,29
M	3,19	2,81
Q	1,56	1,36
RHI	1,36	1,30
UR	1,30	1,00
X	2,98	1,77

On peut observer:

- que l'erreur est relativement faible (?);
- que l'erreur croît dans un premier temps pour toutes les variables, avec l'augmentation de leurs sources;
- qu'elles se stabilisent à moyen terme, plus rapidement pour les éléments de commerce extérieur (qui utilisent un cadre de cointégration). Ceci est cohérent avec les propriétés de stationnarité;
- que l'évolution est beaucoup plus lisse pour le capital (ses valeurs sont fortement corrélées);
- que les éléments de commerce extérieur montrent la plus forte variabilité (et sont corrélés). Cette corrélation ne vient pas de Q;

en comparaison avec l'erreur *ex post*, on peut observer que l'erreur stochastique est généralement plus grande, en particulier pour les variables de capacité. Cela peut être dû au fait que la somme des résidus n'est plus nulle, ce qui avait permis à la solution de revenir vers les valeurs de référence à la fin.

8.8 ALLER PLUS LOIN: L'ETUDE DES PROPRIETES DU MODELE

En plus de l'utilisation des simulations directes, les fonctions mathématiques fournies par EViews permettront d'appliquer des techniques plus complexes.

8.8.1 L'ANALYSE DES VALEURS PROPRES

Cette technique peut aider à comprendre la dynamique de convergence ou de divergence du modèle, et de les interpréter en termes économiques.

Son application demande que le modèle soit linéarisé autour d'une solution de référence (maintenant Δ représente la différence entre les deux simulations):

$$\Delta y_t = A_t \cdot \Delta y_{t-1} + B_t \cdot \Delta x_t$$

Si A est censée être constante dans le temps, un choc Δx appliqué à l'instant t (uniquement) sur le vecteur x des exogènes se traduira à la période $t+k$ pour une variation de y :

$$\Delta y_{t+k} = A^k \cdot B \cdot \Delta x_t$$

et les conséquences à long terme dépendront donc de l'évolution des puissances de A^k , lorsque k tend vers l'infini. Trois cas peuvent se produire:

- A^k tend vers 0.
- A^k augmente vers l'infini (ou au moins l'un de ses termes le fait).
- A^k se stabilise à une valeur finie, non nulle.

Chacun de ces trois cas est associé à la mesure du rayon spectral de A (le module de sa plus grande valeur propre): ce rayon devra être respectivement inférieur, égal ou supérieur à l'unité. Ceci nous conduira à raisonner en valeurs propres, en déclarant:

$$A = V^{-1} \cdot \Lambda \cdot V$$

(Λ diagonale, contenant les valeurs propres de A , et V la base des vecteurs propres associés).

On obtient alors:

$$V \cdot \Delta y_{t+k} = V \cdot A^k \cdot \Delta y_t = \Lambda^k \cdot V \cdot \Delta y_t = \Lambda^k \cdot V \cdot B_t \cdot \Delta x_t$$

Par conséquent, pour un choc donné sur x , sa dynamique peut être étudiée par la décomposition des effets de première période sur la base des vecteurs propres de A , chaque élément de cette décomposition évoluant ensuite avec le temps proportionnellement à la valeur propre correspondante. Les évolutions peuvent être:

- divergentes, convergentes ou fixes, en fonction du module de la valeur propre (supérieur, inférieur ou égal à 1) ;
- monotone ou cyclique, en fonction de la nature réelle ou complexe de la valeur propre.

En fait, la stabilité de A au cours du temps est beaucoup mieux assurée si l'on considère les variations relatives:

$$\Delta y_t / y_t = A_t \cdot \Delta y_{t-1} / y_{t-1} + B_t \cdot \Delta x_t / x_t$$

Cette propriété a deux explications:

- La plupart des équations estimées sont formulées en logarithmes, avec des élasticités de long terme unitaires..
- En divisant la variation par la valeur, le résultat devrait être stable dans le temps.

En effet, cette modification est équivalente à un remplacement de chaque variable par son logarithme dans les équations du modèle.

Cela ne change pas fondamentalement les considérations ci-dessus, sauf que la convergence et la non-convergence seront associées maintenant à des variations relatives: une valeur propre supérieure à un signifiera que c'est rapportée à la variable que la différence à la valeur de base tend vers l'infini.

Comme la plupart des variables croissent avec le temps, ce critère de convergence sera moins restrictif que le premier. Mais la question suivante doit être prise en compte: doit-on associer la divergence à une variation absolue ou relative ? Par exemple, supposons que les ménages bénéficient d'une exonération d'impôt, à deux époques différentes dans le temps. Pour comparer les deux, tiennent-ils compte des niveaux à prix constants, ou en rapport à leur revenu actuel ? La vérité est certainement entre les deux.

8.8.1.1 Premier exemple: un modèle très simplifié

À titre d'exemple, nous allons utiliser une simplification extrême de notre modèle habituel.

$$I_t = a \cdot (Q_t - Q_{t-1}) + b$$

Les entreprises investissent pour adapter leurs capacités de production à l'évolution de la demande.

$$CO_t = c \cdot Q_t + d \cdot Q_{t-1} + e$$

Les ménages consomment une part des niveaux de production actuels et passés (introduire un retard rend l'exemple plus intéressant).

$$Q_t = I_t + CO_t + g_t$$

La production s'adapte immédiatement à la demande, composée de l'investissement, de la consommation et de la demande de l'État.

Raisonnant en variations, on obtient:

$$\begin{pmatrix} \Delta I \\ \Delta CO \\ \Delta Q \end{pmatrix}_t = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \Delta I \\ \Delta CO \\ \Delta Q \end{pmatrix}_t + \begin{pmatrix} 0 & 0 & -a \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta I \\ \Delta CO \\ \Delta Q \end{pmatrix}_{t-1} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \Delta g_t$$

ou

$$\begin{pmatrix} \Delta I \\ \Delta CO \\ \Delta Q \end{pmatrix}_t = 1/(1-a) \begin{pmatrix} 0 & 0 & (c-1) \\ 0 & 0 & c(c-a) \\ 0 & 0 & (c-a) \end{pmatrix} \begin{pmatrix} \Delta I \\ \Delta CO \\ \Delta Q \end{pmatrix}_{t-1} + 1/(1-a) \begin{pmatrix} a \\ 0 \\ 1 \end{pmatrix} \Delta g_t$$

Ce processus contient une seule variable retardée, et une seule valeur propre:

$$\lambda = (d - a)/(1 - a - c)$$

et le multiplicateur à la période k d'un changement unitaire de la demande de l'État à la période 0 est:

$$1/(1 - a - c) \cdot \lambda^k$$

À ce niveau de simplification, une interprétation formelle est encore possible. La valeur propre ajoute deux effets dynamiques décalés:

- la sensibilité de la consommation aux revenus des ménages passés: d ;
- la sensibilité des investissements au niveau précédent de la production, représentant les niveaux de capacité acquis auparavant: $-a$.

Ces deux effets sont amplifiés par le mécanisme multiplicateur, en raison de l'influence de l'investissement et de consommation dans la production actuelle: $1 / (1 - a, c)$.

On peut donc logiquement supposer:

- que a , c et d sont positifs ;
- que $0 < c + d < 1$: les ménages consomment une partie de leur revenu supplémentaire, elle-même inférieure à la production supplémentaire.

Notre but n'est pas de théoriser sur ces formulations. Cependant, nous pouvons identifier quelques cas particuliers:

- $d + c = 1$ (impact unitaire de Q sur la consommation), la valeur propre est 1, et le processus dynamique ne converge pas, ni ne diverge: l'augmentation de la consommation maintient le niveau de la production à la période suivante, ce qui stabilise l'investissement, qui n'a pas d'impact supplémentaire sur la croissance Q . Un choc de demande à la période 1 maintiendra ses effets indéfiniment.

En d'autres termes, l'effet multiplicateur: $1 / (1 - ac)$ est compensé à la prochaine période par l'effet réducteur: $d-a$ ou $(1-a-c) - (1-d-c)$ plus bas que $(1-a-c)$ si $d + c < 1$.

Ou encore: avec $c + d < 1$, l'effet multiplicateur sera étendu aux deux périodes, avec une intensité plus faible que le choc initial. Le processus va converger.

- $c = 0$ (consommation indépendante de Q), la valeur propre est $-a / (1 - a)$, et la dynamique converge si $a < 1/2$. Sinon, revenir au niveau initial de l'investissement représente plus de la moitié de la précédente augmentation de la production, et le multiplicateur est plus du double de ce changement, ce qui conduit à un autre processus divergent.
- $a = 0$, la valeur propre est c , ce qui conduit à une convergence générale si $|c| \in]0, 1[$
- Plus généralement, lorsque $0 < a < 1$ et $0 < c + d < 1$, le processus converge si:

$$c + d > 2a - 1$$

Condition moins restrictive que $a > 1/2$ (si $c = 0$, parce que l'effet positif de la consommation peut alors compenser la divergence alternative présentée plus haut).

Nous pouvons voir que, même pour un exemple très simple de l'interprétation formelle peut présenter une certaine complexité. Pour les cas plus réalistes, il faudra recourir à des techniques numériques, interprétant les résultats en termes de mécanismes économiques en les associant à des variables du modèle.

Pour cela on peut (selon une méthodologie introduite par Deleau et Malgrange (1978)) éliminer tour à tour pour chaque variable son influence retardée (en supprimant la colonne correspondante et la ligne de A , puis en recalculant les valeurs propres). Dans le cas le plus favorable, une valeur propre unique

disparaîtra, créant une association naturelle avec la variable éliminée. Dans d'autres cas la recherche sera plus difficile, et on aura peut-être besoin d'éliminer simultanément des groupes de variables, que l'on va associer à un groupe de valeurs propres de même dimension.

Notre exemple est trop simple pour présenter cette méthode, car elle ne contient qu'une seule influence retardée (et donc une seule valeur propre). Mais l'analyse que nous venons de faire donne une illustration de la méthode de base. Pour étendre les caractéristiques dynamiques du modèle ci-dessus, on aurait pu ajouter une équation pour le capital et la capacité de production.

8.8.1.2 Deuxième exemple: les effets de la règle de Taylor

Nous allons maintenant expliquer, sur un exemple très simple, le rôle de l'élément réel de la règle de Taylor dans notre type de modèle.

Rappelons que la règle de Taylor représente le comportement d'une banque centrale qui utilise le taux d'intérêt pour atteindre une cible d'inflation (et également réduire sa volatilité). Voir Taylor JB. (1993).

Elle va augmenter le taux si:

- l'inflation est supérieure à la cible (2% pour la Banque Centrale Européenne) ;
- des tensions apparaissent sur les capacités productives (l' «output gap») signalant une inflation potentielle.

Cette action devrait permettre de réduire la demande (investissement et consommation), l'activité et donc l'inflation.

La formule originale est la suivante:

$$IR_t = tx(p_t) + 0.5 \cdot (tx(p_t) - tx^*(p_t)) + 0.5 \cdot gap$$

avec

- IR le taux d'intérêt réel
- p l'indice des prix (indice implicite des prix de la consommation ?)
- $tx^*(p)$ la cible d'inflation
- gap l'écart de l'indicateur de production, normalement la différence relative entre les valeurs réelles et potentielles de celle-ci.

Considérons le modèle suivant, mesuré en termes réels, une simplification de notre exemple, sans commerce extérieur

$$\begin{aligned}
 i_t / k_t &= 0.65i_{t-1} / k_{t-1} + 0.10[(q_t - q_{t-1}) / q_{t-1} + (ur_t - ur^*) / ur^*] + d + a(ur_t - ur^*) \\
 c_t &= r_t q_t (1 - b(ur_t - ur^*)) \\
 ur_t &= q_t / (k_t p k_t) \\
 k_t &= k_{t-1} (1 - dd_t) + i_{t-1} \\
 q_t &= c_t + i_t + g_t
 \end{aligned}$$

avec les variables endogènes suivantes:

- i l'investissement productif
- c la consommation des ménages c
- k le capital
- q le PIB
- ur le taux d'utilisation des capacités de production

et les variables exogènes:

- dd le taux d'amortissement
- g la consommation des administrations publiques
- pk la productivité du capital
- r le ratio de la consommation au PIB
- ur^* la cible de taux d'utilisation

Pour la règle de Taylor, nous n'avons considéré que le rôle de l'écart de production, que nous avons représenté (ce qui peut être discuté) par le taux d'utilisation des capacités.

Donc, nous allons utiliser les paramètres suivants:

- a la sensibilité de l'investissement au taux d'intérêt
- b la sensibilité de la consommation au taux d'intérêt

Si nous calculons la dynamique de ce modèle en différenciant numériquement ses équations,

$$\Delta y_t / y_t = A \Delta y_t / y_t + B \Delta y_{t-1} / y_{t-1} + C \Delta x_t / x_t$$

où A , B et C sont des matrices jacobiniennes constantes avec le temps,

nous obtenons

$$\Delta y_t / y_t = (I - A)^{-1} (B \Delta y_{t-1} / y_{t-1} + C \Delta x_t / x_t)$$

Pour décrire le processus dynamique, il nous faut maintenant calculer les valeurs propres de $(I - A)^{-1} B$,

une matrice 6 x 6.

Mais bien sûr, indépendamment des valeurs de a et b , nous ne devrions avoir que deux valeurs propres non nulles, car seulement deux équations contiennent des éléments décalés (et trois de ces éléments).

Avec $a = b = 0$, on obtient deux valeurs propres complexes, avec

Module: 0,977 époque: 18,8

Maintenant, si nous attribuons à « a » la valeur -0,03, conformément à notre règle de Taylor (produit de -0,06 donné par nos estimations de l'équation d'investissement, par 0,5 coefficient théorique de la règle de Taylor elle-même), nous obtenons:

Module: 0,872 période: 20,6

La période n'est pas beaucoup affectée, mais la convergence est beaucoup plus rapide. Au bout de 10 périodes un choc initial est réduit de 75 % (au lieu de 21 %).

Concernant b (le rôle du taux d'intérêt de la consommation), nous obtenons seulement une faible augmentation de la période, et de la vitesse de convergence.

En conclusion: la dynamique converge beaucoup plus vite, et elle est presque entièrement expliquée par les éléments réels de la formation de capital, y compris le taux d'intérêt réel.

Les éléments permettant de produire l'étude ci-dessus peuvent être très simplement mis en œuvre en utilisant EViews.

8.8.1.3 Les sentiers de croissance équilibrée

Allons plus loin dans la détermination des conditions de l'existence solutions du modèle, tel que le fait que l'ensemble des variables croisse à un taux constant (un «chemin de croissance régulière»). L'application de ces conditions aux formulations du modèle permettra, en modifiant les équations, de produire un modèle de long terme, aux causalités potentiellement très différentes du modèle original. On étudiera également la stabilité de cette trajectoire, face à un choc ponctuel, et verrons si la dynamique de long terme ramène à la trajectoire ou en éloigne (encore une fois, en raisonnant en termes absolus ou relatifs).

Ces techniques peuvent être clarifiées grâce à notre modèle simple. Rappelons-en les formulations (en format EViews):

```

CAP = pk * K(-1)
Q + M = FD + X
UR = Q / CAP
IC = tc * Q
CI / Q(-1) = - 0.157 * @pch(Q) - 0.0464 * @pch(Q(-1)) + ec_ci
I / K(-1) = 0.825 * I(-1) / K(-2) + 0.0279 * UR + 0.152 * .25 * Q / Q(-4) - 0.0525 + ec_i
Log(PRLE_T) = 2.591 + 0.0279 * (t - 2002) + 0.0215 * (t - t1) * (t<t1) + 0.0145 * (t - t2) * (t<t2)
LED = Q / PRLE_T
DLog(LE) = 0.422 * DLog(LED) + 0.455 * Log(LED(-1) / LE(-1)) + 0.000731 + ec_le
LT = LE + Ig
RHI = wr * LT + r_rhiq * Q
IH = r_ih * RHI
CO = RHI * (1 - sr)
FD = CO + I + gd + CI + IH
TD = FD + IC
RES_M = Log(M / TD) + 1.322 * Log(UR) + 0.420 * Log(compm) + 0.0126 * (@TREND(60:1)
* (t<=2002) + @ELEM(@TREND(60:1) , "2002S2") * (t>2002))
DLog(M) = 1.200 * DLog(TD) + 0.282 * DLog(UR) - 0.212 * RES_M(-1) - 0.629 + ec_m

```

$$RES_X = \text{Log}(X / wd) + 0.686 * \text{Log}(UR) - 4.87 \text{ E-}05 * (@\text{TREND}(60:1) * (T \leq 2002) + @\text{ELEM}(@\text{TREND}(60:1), "2002S2") * (T > 2002))$$

$$D\text{Log}(X) = 0.940 * D\text{Log}(wd) - 0.0129 - 0.195 * RES_X(-1) + ec_x$$

$$K = K(-1) * (1 - dr) + I$$

Si nous appelons

- txq le taux de croissance des quantités
- txn le taux de croissance des populations.

Nous obtenons:

$$CAP = pk * K / (1 + txq)$$

$$Q + M = FD + X$$

$$UR = Q / CAP$$

$$IC = tc * Q$$

$$CI / Q / (1 + txq) = -0.157 * txq - 0.0464 * txq + ec_ci$$

$$I / K / (1 + txq) = 0.825 * I / K / (1 + txq) + 0.0279 * UR + 0.152 * .25 * ((1 + txq)^4 - 1) - 0.0525 + EC_I$$

$$\text{Log}(PRLE_T) = 2.591 + 0.0279 * (t - 2002) + 0.0215 * (t - t1) * (t < t1) + 0.0145 * (t - t2) * (t < t2)$$

$$LED = Q / PRLE_T$$

$$\text{Log}(1 + txn) = 0.422 * \text{Log}(1 + txn) + 0.455 * \text{Log}(LED / LE) + 0.000731 + ec_le$$

$$LT = LE + Ig$$

$$RHI = wr * LT + r_rhiq * Q$$

$$IH = r_ih * H * RHI$$

$$CO = RHI * (1 - sr)$$

$$FD = CO + I + gd + CI + IH$$

$$TD = FD + CI$$

$$RES_M = \text{Log}(M / TD) + 1.504 * \text{Log}(UR) + 0.483 * \text{Log}(COMPM) + 0.0124 * (@\text{TREND}(60:1) * (T \leq 2002) + @\text{ELEM}(@\text{TREND}(60:1), "2002S2") * (t > 2002))$$

$$\text{Log}(1 + txq) = 1.200 * \text{Log}(1 + txq) - 0.282 * RES_M - 0.629 + EC_M$$

$$\text{RES_X} = \text{Log}(X / \text{wd}) + 0.686 * \text{Log}(\text{UR}) - 4.87 \text{ E-}05 * (\text{@TREND}(60:1) * (\text{T} \leq 2002) + \text{@ELEM}(\text{@TREND}(60:1), "2002S2") * (\text{T} > 2002))$$

$$\text{Log}(1+\text{txq}) = 0.940 * \text{Log}(1+\text{txq}) - 0.0129 - 0.195 * \text{RES_X} + \text{ec_x}$$

$$K = K / (1+\text{txq}) * (1 - \text{dr}) + I$$

On peut d'abord observer que, dans chaque équation, à condition que tous les éléments explicatifs croissent au taux théorique, l'élément expliqué le fera également. En particulier, tous les termes contenus dans une somme devraient croître, en tenant compte des variables qui y participent, à ce même rythme.

Cela signifie qu'il existe une solution cohérente dans laquelle tous les éléments croissent au taux théorique. Comme le modèle ne possède qu'une seule solution, c'est la solution, et elle devrait être atteinte, à condition que les mécanismes de correction d'erreur mènent dans le bon sens (à la vue des coefficients, on peut supposer qu'ils le font).

Nous pouvons également construire le modèle à long terme, qui obéira à une logique quelque peu différente de la logique dynamique, une fois les déséquilibres stabilisés.

De (20) on obtient I / K comme le taux qui permet d'adapter le capital à la fois à la croissance et à la dépréciation:

$$I / K = 1 - 1 / (1 + \text{TXQ}) * (1 - \text{dr})$$

De (6) on obtient UR en fonction de la vitesse d'adaptation. Même si l'adaptation est immédiate (homogénéité dynamique) UR dépendra néanmoins de txq, comme une accélération de la croissance attendue appelle à un effort supplémentaire dans la création de capital, pour une utilisation à la période suivante:

$$I / K / (1 + \text{TXQ}) = 0,825 * I / K / (1 + \text{TXQ}) + 0.0279 * \text{UR} + 0,152 * .25 * ((1 + \text{TXQ}) ^ 4 - 1) - 0,0525 + \text{EC_I}$$

De (18) on obtient RES_X.

De (16) on obtient RES_M.

De (9) nous obtenons LED / LE.

Si l'adaptation était immédiate aucun des trois ne dépendrait de txq ou txn . Ce serait également le cas pour UR si les coefficients respectaient exactement le cadre de correction d'erreur.

Ici, l'écart va augmenter avec le taux de croissance, et diminuer avec la vitesse d'adaptation.

De RES_X, UR et (17) nous obtenons X / wd et X

De RES_M, UR et (15) nous obtenons $M / (FD + IC) = M / (FD + tc * Q)$

De UR, (1) et (3) nous obtenons Q / K ; d'après (6), nous obtenons l / Q

De (5) nous obtenons CI / Q

De (8) et (9) nous obtenons LE / Q

De (10) et (11) nous obtenons RHI comme un terme proportionnel à Q, plus $wr * lg$

De (12) et (13) nous obtenons CO et H comme proportionnel à RHI, donc Q et $wr * lg$

De (14) on obtient FD comme proportionnel à Q plus un terme proportionnel à $wr * lg$, plus gd

Cela donne M / Q

Cela donne Q de (2)

Et tous les autres éléments.

Enfin, tous les termes de l'offre - l'équilibre de la demande sont définis comme étant proportionnels à Q, à l'exception de:

- La consommation des ménages et l'investissement qui sont partiellement proportionnels à du gouvernement.
- La demande du gouvernement qui est exogène⁹⁹.
- La somme de ces éléments rend la demande partiellement exogène.
- Les exportations qui sont proportionnelles à la demande mondiale.

⁹⁹ Bien sûr, les hypothèses lui donneront des valeurs proportionnelles à Q.

Nous pouvons voir que la logique des causalités du modèle est tout à fait différente de celle du modèle dynamique. En particulier,

Le taux d'utilisation dépend du taux de croissance de l'économie et du taux d'amortissement.

L'écart entre les objectifs et les valeurs réelles aussi (avec UR comme intermédiaire éventuel).

Connaissant cet écart, nous pouvons définir tous les éléments réels en pourcentage du PIB, sauf pour les exportations et un élément exogène de la demande. Appelons le dx .

L'équilibre offre-demande

$$Q + M + X = FD$$

devient:

$$Q + a \cdot (b \cdot Q + dx) = b \cdot Q + dx + c \cdot wd$$

$$Q = ((1 - a) \cdot dx + wd) / (1 - (1 - a) \cdot b)$$

avec a et b éléments constants en fonction d'hypothèses exogènes et du taux de croissance txq (et aussi txn).

Cela signifie que si

- gd a été définie comme une proportion constante de la demande
- lg a été définie comme une proportion constante de LE (ou de $LT = LE + lg$)

(Ce qui semble logique et même peut-être une hypothèse nécessaire.)

Alors le niveau exogène de la demande mondiale définira tous les éléments de l'économie locale (comme l'inflation mondiale va définir l'inflation locale pour un modèle plus complexe).

8.8.2 LE CAS DES MODELES A CORRECTION D'ERREUR: UN EXEMPLE SIMPLE

Nous allons maintenant montrer comment une formulation à correction d'erreur rend l'interprétation de la dynamique du modèle à la fois plus facile et plus significative.

Nous allons établir le modèle suivant

$$\begin{aligned}
 CO_t &= a \cdot Q_t \\
 Q_t &= CO_t + I_t + g_t \\
 K_t &= K_{t-1} \cdot (1 - dr_t) + I_t \\
 \Delta K_t &= \alpha \cdot b \cdot \Delta Q_t + \beta \cdot (b \cdot Q_{t-1} - K_{t-1})
 \end{aligned}$$

Nous avons défini un modèle à correction d'erreur, où l'objectif consiste à rendre K proportionnel à Q (ce qui équivaut à un taux constant d'utilisation des capacités si l'on suppose constante la productivité du capital).

Étudions la dynamique de ce modèle, à l'instar de l'analyse des valeurs propres présentée plus tôt.

En éliminant CO et I, nous pouvons réduire le système en:

$$\begin{aligned}
 Q_t &= a \cdot Q_t + K_t - K_{t-1} \cdot (1 - dr_t) + g_t \\
 K_t &= \alpha \cdot b \cdot Q_t + (\beta - \alpha) \cdot b \cdot Q_{t-1} + (1 - \beta) \cdot K_{t-1}
 \end{aligned}$$

$$\begin{pmatrix} \Delta Q_t \\ \Delta K_t \end{pmatrix} = \begin{pmatrix} a & 1 \\ \alpha b & 0 \end{pmatrix} \begin{pmatrix} \Delta Q_t \\ \Delta K_t \end{pmatrix} + \begin{pmatrix} 0 & -(1-d) \\ b(\beta - \alpha) & 1 - \beta \end{pmatrix} \begin{pmatrix} \Delta Q_{t-1} \\ \Delta K_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} g_t$$

Nous pouvons appliquer à ce système la même technique que précédemment, avec normalement deux valeurs propres. Le processus peut paraître un peu trop complexe pour cette formulation élémentaire. Nous allons le simplifier, en supposant que la vitesse d'adaptation est la même pour un nouvel écart, dû à un changement de la valeur de la cible, et un écart préexistant:

$$\alpha = \beta$$

Le système devient:

$$\begin{pmatrix} \Delta Q_t \\ \Delta K_t \end{pmatrix} = \begin{pmatrix} 1-a & -1 \\ 1-\alpha b & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -(1-d) \\ 0 & 1-\alpha \end{pmatrix} \begin{pmatrix} \Delta Q_{t-1} \\ \Delta K_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} g_t$$

Avec une seule valeur propre

$$\lambda = 1 - \alpha + \alpha (d - \alpha) / (1 - a - b \alpha)$$

Interprétons cet élément.

- Si $d = \alpha$ la valeur propre est $(1 - \alpha)$. Cela signifie que l'écart initial entre la variable et sa cible se réduira dans une proportion $1 - \alpha$. La raison en est claire. Si l'on considère le système d'équations, le capital va se déprécier à la vitesse d , ce qui est exactement la vitesse à laquelle les entreprises «veulent» réduire un écart apparu au cours des périodes précédentes, à la suite par exemple d'un choc sur g . La variation ex ante sera identique à celle désirée, et aucun mécanisme supplémentaire n'apparaîtra.
- Au contraire, si d est supérieur à α par exemple, la diminution naturelle du capital sera plus élevée que la valeur désirée. Les entreprises devront investir la différence, ce qui conduit à une augmentation de capital rendue plus élevée par la présence du multiplicateur.
- Mais pour que le processus diverge, la différence entre d et α doit être grande, ou l'effet multiplicateur très important (ce qui conduirait à des problèmes avec n'importe quel modèle).

Il peut sembler étrange qu'une diminution ex ante du capital (par le biais d'un d supérieur) puisse conduire à une augmentation ex post. La raison en est que le comportement des entreprises concernant le capital est fixé par (2), en tenant compte de la variation de la production, mais pas le taux d'amortissement supplémentaire, qui est compensé automatiquement par le biais d'investissements plus élevés (donc la production et l'investissement à nouveau).

Considérons maintenant le modèle à long terme.

Il est maintenant clair que toutes les variables: CO , Q , I , g , CAP augmentent à la même vitesse sur le chemin de croissance en régime permanent. Ce taux (que nous appellerons q) est défini par « g ».

Le modèle à long terme est

$$CO = a \cdot Q$$

$$Q = CO + I + g$$

$$K = K \cdot (1 - dr) / (1 + q) + I$$

$$K \cdot (q + \beta) / (1 + q) = (\alpha \cdot q + \beta) \cdot b \cdot Q / (1 + q)$$

ou encore

$$CO = a \cdot Q$$

$$Q = CO + I + g$$

$$I / K = q + dr$$

$$K / Q = b \cdot (\alpha \cdot q + \beta) / (q + \beta)$$

Nous observons qu'aucune des explications causales du modèle original n'a changé. Le rapport de la production au capital est fixe, sa valeur définie par la cible dans la formulation d'origine. Et maintient cette contrainte, en conduisant à une évolution implicite de la capacité productive égale au taux de croissance (en tenant compte des amortissements).

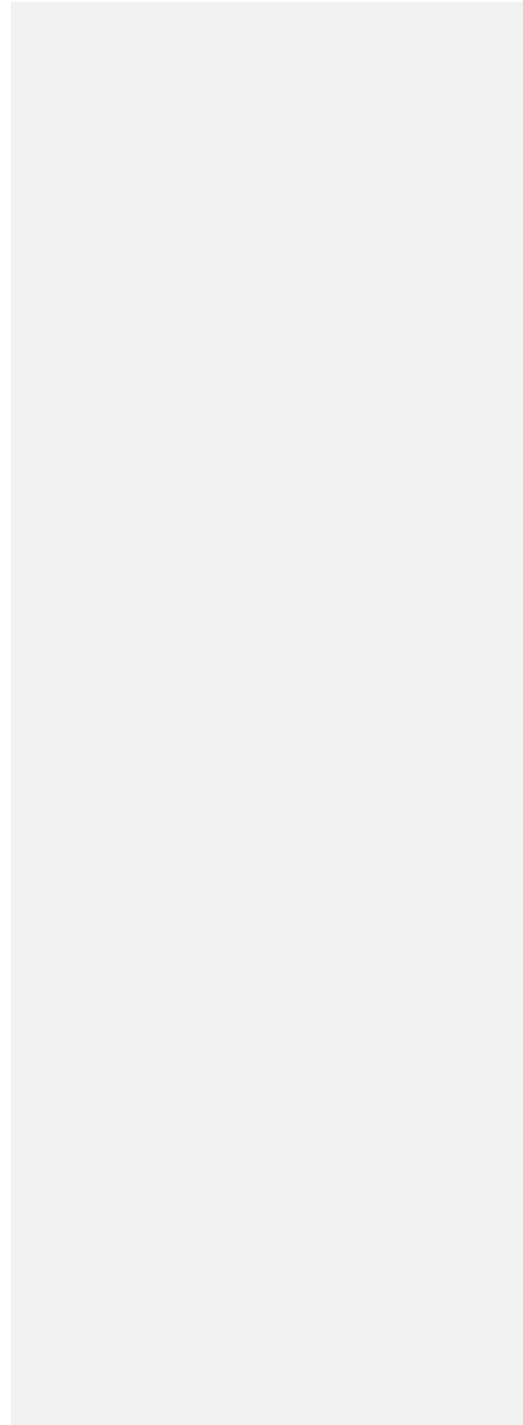
Mais on peut aussi voir que la cible n'est atteinte que dans des conditions particulières. Ceci est une caractéristique habituelle des modèles à correction d'erreur.

L'équation (4) montre que les cas sont les suivants:

$$\begin{array}{ll} \alpha = 1 & \text{(homogénéité dynamique, les nouveaux écarts sont immédiatement corrigés)} \\ q = 0 & \text{(la cible ne bouge pas)} \end{array}$$

Cela est vrai pour toutes les valeurs de départ, y compris celle qui répondent à $K / Q = b$: si q ou α ne sont pas nuls, un écart apparaît, qui ne disparaîtra jamais.

Nous nous arrêtons ici, car notre but était seulement d'illustrer le processus, et de montrer que les modèles de correction d'erreurs rendent la dynamique plus intéressante, et plus facile à étudier.



CHAPITRE 9: L'UTILISATION DES MODELES

Dans le chapitre 2, nous avons commencé à présenter les applications de la modélisation. Nous allons maintenant être plus précis, en nous concentrant sur les aspects pratiques.

9.1 LES DIAGNOSTICS OPERATIONNELS

La première application que nous avons citée, et la plus naturelle en apparence, consiste à demander au modèle un diagnostic réaliste sur des problèmes économiques. Ce diagnostic peut s'appuyer sur des scénarios ou des chocs.

L'auteur de ce type d'étude peut être:

- Un organisme gouvernemental, comme le ministère des Finances local, qui veut obtenir des informations sur l'évolution future de l'économie qu'il gère, ou l'efficacité des décisions politiques qu'il envisage.
- Une institution internationale, telle que le FMI ou l'OCDE, ou un institut non gouvernemental, qui souhaite évaluer l'avenir de l'économie mondiale.
- Une unité de recherche universitaire ou une personne qui veut étudier l'impact de décisions politiques ou de changements structurels, selon un modèle théorique donné.
- Une entreprise privée spécialisée dans les services économiques, qui fournit des prévisions et des études à ses abonnés, par une publication ou une étude distincte.

9.1.1 LES SCENARIOS ET LEURS DIFFERENTS TYPES

Deux types de scénarios peuvent être distingués.

- Les prévisions tendanciennes, qui s'appuient sur les hypothèses les plus probables (dans le cas d'un scénario unique), ou un ensemble d'hypothèses possibles, couvrant le champ des évolutions prévisibles.

Les hypothèses sont souvent développées par des experts extérieurs à l'équipe gérant le modèle. Prenons un exemple: les prévisions de MESANGE, un modèle macroéconomique français considérant les décisions de politique économique comme exogènes, et géré par le ministère français des Finances, dans deux de ses directions: à l'INSEE, l'Institut National de la Statistique et des Etudes Economiques, et au Trésor (Direction du Ministère des Finances).

Les variables décrivant l'environnement étranger sont évaluées au moyen d'un ensemble d'outils:

- MZE (Modèle Zone Euro) modèle du CEPII (une institution économique française spécialisée dans les études internationales), géré dans les mêmes établissements.
- Le modèle mondial NIGEM, géré et loué à l'institut britannique NIESR (Institut national de l'économie et de recherches sociales).
- Les résultats de prévisions économiques de l'OCDE, de l'ONU et du FMI.
- Des évaluations numériques établies par les experts en économie et commerce international, sans faire appel à un modèle.
- L'observation des évolutions internationales, comme la crise grecque, et l'évaluation des évolutions possibles.

Les décisions du gouvernement sont précisées à la suite de discussions avec le la Direction générale du Trésor et en particulier les personnes qui y produisent l'évaluation numérique du budget de l'année suivante.

En ce qui concerne les hypothèses sur l'évolution de la population et de sa structure, on utilisera les résultats des études réalisées par l'INSEE ou le ministère du Travail (INED).

Ce type de scénario donnera des indications soit sur l'évolution future la plus probable, soit sur la gamme des évolutions prévisibles, au mieux des connaissances des experts disponibles.

La production de la prévision finale réclamera un grand nombre de simulations, accompagnées de rencontres et de discussions. Elle durera généralement plusieurs semaines (disons 2 à 4). L'amélioration de la vitesse de calcul est utile, mais pas trop, car le temps est principalement utilisé pour discuter et améliorer la qualité des résultats.

- Des prévisions normatives, où les résultats sont invités à répondre à un certain nombre de conditions. Ces contraintes sont définies avant l'élaboration des hypothèses.

Ce type de scénario décrit une situation dans laquelle des efforts sont faits, par rapport à l'évolution prévisible de l'économie, pour atteindre certains objectifs¹ (comme répondre aux critères d'adhésion à l'Union Monétaire Européenne, généralement appelés les critères de Maastricht). Leur but essentiel sera de donner des informations sur l'ampleur des difficultés attendues, et les efforts requis, associés à un chemin possible.

¹ Toutefois, dans certains cas, les prévisionnistes peuvent également inclure dans l'ensemble potentiel des instruments des hypothèses qui ne sont pas gérables par l'État, comme un changement exogène sur le comportement des agents, ou des paramètres structurels tels que la productivité des facteurs. Ce n'est acceptable qu'à condition que l'option soit clairement indiquée dans la publication associée. La première question était alors: comment le comportement des agents, ou les paramètres structurels, doit-il changer afin d'atteindre un objectif spécifique ?

En général, le client de ce type d'étude sera également l'État ou un établissement public. En collaboration avec les gestionnaires du modèle, il fixera les contraintes et les éléments de politique économique utilisés pour y répondre. Mais les institutions non gouvernementales et les chercheurs peuvent aussi être intéressés par ce type de problème, peut-être plus que la production d'une prévision car elle implique un raisonnement plus économique.

Si des objectifs précis sont fixés (comme un déficit budgétaire de 3 points de PIB), on peut travailler soit par une suite d'approximations sur le modèle normal, ou en modifiant les spécifications du modèle par un renversement de causalité. Cette dernière technique présente deux problèmes:

- Une contrainte formelle: pour que le modèle conserve une solution unique, le nombre d'éléments endogènes qui sont maintenant fixés doit être égal au nombre d'instruments qui permettront de les atteindre. Par exemple, si l'on veut équilibrer à la fois la balance commerciale et le budget de l'État, il sera nécessaire de libérer deux variables de décision et deux seulement, pour éviter soit une impossibilité soit une indétermination².
- Une contrainte technique: il sera souvent trop complexe de produire un modèle identifié tenant compte de ce renversement explicite, et si l'algorithme de résolution réclame une telle formulation, il sera nécessaire de revenir à la première technique, la détermination de la solution par une suite d'approximations. Supposons par exemple que nous voulons équilibrer le commerce extérieur en augmentant les subventions à l'investissement. Pour transformer le modèle, il faudra identifier une chaîne de causalités menant des subventions au commerce extérieur (qui peut être assez longue) et d'inverser chacune à son tour, dans l'équation associée. Cela sera évidemment difficile, sinon impossible techniquement, et on peut préférer une technique numérique itérative: calculer en dehors du modèle une valeur approchée du montant des subventions nécessaires, l'utiliser pour résoudre le modèle standard, puis corriger le changement à partir de l'observation de l'erreur restante. Cette séquence d'approximations devrait faire diminuer l'écart jusqu'à une valeur acceptable.
- À la fin du livre, nous proposons une autre méthode, qui permet de résoudre le modèle pour des valeurs données d'un certain nombre de variables endogènes, à condition que le même nombre de variables exogènes soit libéré. Cette méthode ne fonctionne pas tout le temps, mais si c'est le cas, la solution est quasi immédiate, et peut être appliquée sous EViews sans accéder au code

² Sauf si le nombre d'instruments est plus grand que celui des objectifs, mais des limites sont fixées sur leurs variations, ainsi qu'un ordre de priorité. Ensuite, chaque instrument sera utilisé à son tour, en passant à un nouveau si la contrainte a été atteinte. Dans la pratique, nous nous retrouverons avec autant d'inconnues que de contraintes. On peut aussi créer un instrument synthétique en appliquant des pondérations prédéfinies à un ensemble plus vaste.

du modèle. Le programme lui-même est fourni, et peut être adapté en quelques minutes à n'importe quel modèle existant géré sous EViews.

Mais dans tous les cas, les résultats ne peuvent pas être pris pour plus de ce qu'ils sont, un équilibre économique associé à des hypothèses s'écartant de l'évolution normale, dont la probabilité d'occurrence ne peut pas être établie. Pour avoir le droit d'utiliser cette méthode, il faut afficher clairement et sans doute justifier les options qu'on a choisies.

Une fois ces restrictions posées, on n'est pas limité dans les éléments que l'on peut utiliser. Par exemple, on peut utiliser la modification du résidu de l'équation déterminant les investissements pour mesurer de combien les entreprises privées devraient les accroître (toutes choses égales par ailleurs) pour que la balance commerciale revienne à l'équilibre à long terme. Moins on contrôle les instruments, et moins la simulation doit être considérée comme le résultat d'une politique dirigiste. Mais cela ne rend pas les résultats (et l'observation des séquences causales) moins intéressants d'un point de vue théorique.

9.1.2 LA GESTION DES PREVISIONS: LE CIBLAGE DES SIMULATIONS

Dans le cas normatif, il faut manipuler les exogènes pour parvenir à une certaine solution. Mais cela se produit aussi dans les prévisions tendanciennes. On ne peut attendre du modèle qu'il donne spontanément une image de l'avenir compatible avec ce que nous attendons, en particulier à court terme, où nous avons déjà quelques informations (partielle) sur ce qui se passe.

Une procédure de ciblage doit être appliquée, qui peut être automatisée, par l'application de techniques spécifiques.

9.1.2.1 Sur le passé

Nous pourrions vouloir que la simulation de base reproduise exactement le passé connu. Cette opération n'est pas futile: il est conseillé de commencer les chocs d'analyse à partir de la valeur historique, ce qui élimine les erreurs de second ordre.

La technique est simple: nous appliquons aux équations de comportement leurs résidus d'estimation. Chaque équation individuelle fournira des valeurs historiques, et l'absence globale d'erreur permet aux valeurs historiques de représenter la solution du système.

9.1.2.2 En prévision

Pour les prévisions, cette méthode vise à obtenir des valeurs des variables appartenant à un certain intervalle, par le choix des hypothèses et des résidus d'estimation.

À court terme, l'objectif sera d'adapter les résultats aux valeurs déjà connues (ou à intervalles connus) de certains éléments du modèle. Ainsi, à la fin de l'année, nous connaissons déjà la croissance globale et certaines de ses composantes, l'inflation et la croissance mondiales, le chômage et la balance commerciale, et les prévisions devraient donc fournir leurs valeurs, si ce n'est que pour l'affichage. Face à une représentation incorrecte le passé déjà connu, le public mettra en doute les qualités de prévision du modèle.

À moyen terme, on cherchera à maintenir l'évolution générale dans des limites jugées raisonnables, en particulier respectant certaines conditions générales. Ainsi, on n'acceptera pas un écart trop important entre le pays décrit par le modèle et le reste du monde, en termes d'inflation et de croissance.

Mais le caractère fragmentaire des informations interdira l'utilisation de la méthode proposée dans le paragraphe précédent. Le constructeur du modèle travaillera en général par itérations, en assouplissant parfois certaines contraintes si elles s'avèrent difficiles à respecter. Ou on peut évidemment faire appel à l'algorithme spécifique que nous allons proposer.

9.1.3 LES DIFFERENTS TYPES DE CHOCS

Nous allons maintenant classer les différents types de chocs.

9.1.3.1 Chocs analytiques et chocs complexes

Nous avons déjà présenté les chocs analytiques, utilisés pour interpréter et valider les propriétés du modèle, par comparaison avec la théorie économique et les propriétés des autres modèles. Cette série de tests permet également l'interprétation des réponses du modèle aux chocs élémentaires, ce qui rend plus facile l'interprétation ultérieure des chocs plus complexes. En fait, les équipes de modélisation opérationnelles produisent toujours un «catalogue de chocs» donnant les réponses du modèle aux variations des principales hypothèses. Compte tenu de la quasi-linéarité numérique des modèles, on obtiendra une approximation assez réaliste des conséquences d'un ensemble complexe de modifications en combinant linéairement les résultats de ces chocs individuels, à l'aide d'un tableur, par exemple, ou plus simplement d'un programme EViews (combinant dans une seule page le contenu de plusieurs «chocs» simples enregistrés dans des pages différentes ayant la même structure).

Allant encore plus loin, quelques mois passés à travailler avec un modèle donneront à son utilisateur la capacité d'anticiper (au moins approximativement) ses réactions aux changements de données, et ce sans aucune simulation. Il aura déjà une base de «catalogues de chocs» disponible dans son esprit.

Mais si l'intensité des changements associés au choc varie avec le temps (par exemple pour tenir compte d'un relâchement progressif des quotas dans un accord commercial) nous rencontrons un problème plus vaste. Comme nous l'avons déjà dit, la plupart des modèles macroéconomiques sont à peu près linéaires

pour une période donnée, mais leur Jacobien n'est pas stable dans le temps, principalement en raison du rôle croissant du commerce extérieur. Bien sûr, plus le changement est important, plus le problème le sera. Mais dans le cas ci-dessus, le changement attendu dans la structure des biens exportés et des clients doit appeler à une prévision spécifique de toute façon.

On aura aussi recours à ce type de choc pour les études opérationnelles

- si les diagnostics demandés sont associés à des décisions simples,
- si l'on cherche à décomposer les effets de politiques plus complexes.

Mais pour obtenir les conséquences « exactes » de chocs complexes, avec un vrai but opérationnel, il faudra les appliquer simultanément à la simulation de base.

9.1.3.2 Chocs externes et chocs de politique économique

Tous ces chocs peuvent également se diviser en trois catégories :

- Les chocs sur l'environnement externe, qui tentent de mesurer les conséquences d'événements non-contrôlables. Pour le modèle à un seul pays, la question posée est donc: quelles sont les conséquences pour mon économie nationale d'un changement dans les conditions extérieures?

Deux sous-catégories peuvent être envisagées.

- Des chocs sur les hypothèses étrangères, associées pour les modèles à un seul pays avec la situation dans d'autres pays, ou des éléments définis au niveau mondial. Un choc typique représenterait une dépression en Asie, ou une augmentation du prix du pétrole, suivie éventuellement par une diminution de la demande mondiale et une augmentation de l'inflation mondiale. Techniquement, ces éléments ne devraient pas être exogènes, comme le pays ou l'ensemble de pays modélisés ont un certain impact sur l'économie du reste du monde. Mais le coût de la modélisation de cette rétroaction est généralement beaucoup trop élevé (il devrait normalement s'agir de la production d'un modèle mondial) par rapport à l'amélioration limitée des propriétés du modèle. Bien sûr, les rétroactions (et l'erreur associée) augmentent avec la taille du pays modélisé. S'il est raisonnable de considérer le monde en dehors du Danemark comme exogène à des événements de ce pays (nous avons pris volontairement un pays extérieur à la zone euro), ce n'est pas le cas de l'Allemagne (un grand pays qui partage une monnaie avec une part importante du monde) et encore moins pour l'ensemble de la zone euro elle-même.
- Les chocs sur les variables locales non contrôlées. Les exemples incluent le progrès technique, la population, ou le climat. Encore une fois, on peut dire que chacun de ces éléments n'est pas complètement exogène: la croissance économique accélère le progrès technique, les variations des taux de natalité influent sur la richesse et la santé des ménages, la croissance entraîne la

pollution et les changements climatiques. Mais les mécanismes associés sont difficiles à établir, formellement complexes et limités dans leur impact à court terme et à moyen terme (au moins pour un choc réduit). Par conséquent, le coût de leur prise en compte est généralement beaucoup plus élevé que les avantages, sauf dans des domaines très spécifiques.

La seule exception concerne les études à long terme. Mais s'il est vrai que nous avons résolu notre modèle sur une très longue période, l'objectif était seulement d'observer ses propriétés d'un point de vue technique, et non de produire une véritable prévision à cet horizon. Pour un authentique modèle de long terme, il faudrait un cadre assez différent, avec de nouvelles formulations et un cadre différent, s'appuyant davantage sur un raisonnement économique formel que sur des estimations réelles, car l'échantillon est généralement trop faible pour mettre en évidence statistiquement ce type d'évolutions.

- Les chocs de politique économique, où l'intérêt réside dans les conséquences des choix sur les instruments institutionnels, donc les variables de décision si l'on considère que l'État (ou un utilisateur se mettant à sa place) est le client de l'étude. La question est maintenant: que va-t-il se passer si je prends les décisions suivantes ? Et quand ces chocs se sont donné un objectif (par exemple: réduire le chômage par une marge donnée sans trop dégrader les autres critères), ils appellent à un processus itératif analogue à celui des scénarios normatifs.

Ces éléments ne sont pas indépendants du pays modélisé, car les décisions prises par le gouvernement sont en grande partie fonction de la situation économique locale. Mais le producteur de l'étude veut être autorisé à les gérer, et ne veut pas voir le modèle le faire à sa place. En outre, l'État a peu d'objectifs (chômage, inflation, déficits budgétaires et étrangers, croissance) et un grand nombre d'instruments potentiels. Mettre en évidence statistiquement le lien entre les décisions sur les instruments et l'état des objectifs serait très difficile, voire impossible. Les quelques études sur ce sujet ne donnent généralement pas des réponses fiables.

- Les deux approches peuvent également être combinées, si l'on recherche les décisions qui permettront de faire face à une évolution donnée des conditions extérieures (ou d'en profiter). On pourra, par exemple, déterminer comment l'État peut utiliser ses propres outils pour lutter contre la dégradation de la balance commerciale provenant d'un affaiblissement donné de la demande mondiale. Ou un pays producteur de pétrole peut décider de l'utilisation des recettes supplémentaires provenant de l'augmentation du prix du baril. Ce type d'exercice est à nouveau voisin des scénarios normatifs.

Il sera nécessaire, dans tous ces cas, de tenir compte de la présence de liens formels entre les variables exogènes. Par exemple, on ne peut pas supposer une augmentation du prix du pétrole sans tenir compte

de ses effets inflationnistes sur les prix à la production dans les pays non-producteurs³, ou de la dépression qu'elle fera subir à l'économie mondiale. De même, il semble difficile d'augmenter le nombre de fonctionnaires sans tenir compte de l'augmentation de la consommation de l'État (chauffage, éclairage, équipements divers) et de ses investissements (bureaux, hôpitaux, écoles), selon le type de postes créés.

Une question également très importante est l'impact des dépenses publiques (investissement, consommation ou embauche de fonctionnaires, pris ensemble ou séparément) sur l'offre locale. Bien sûr, elles modifient la demande (la création de postes de fonctionnaires augmente les revenus des ménages et réduit le chômage). Mais elles devraient aussi avoir un effet positif sur la production potentielle, pour de nombreuses raisons. Pour ne citer que deux exemples :

- La création ou l'amélioration du réseau routier, ou de l'éclairage des rues, augmente la productivité du secteur des transports, et de l'ensemble de l'économie (les produits agricoles peuvent être livrés frais à de plus longues distances, les ventes suivront plus rapidement la production, le nombre d'accidents diminuera ...). Certaines productions entièrement nouvelles deviendront possibles (l'accès à l'irrigation peut permettre de produire de nouveaux types de légumes).
- Acheter des manuels d'enseignement, ou embaucher davantage d'enseignants, permettra d'améliorer les compétences des élèves et leur productivité lorsqu'ils rejoindront le marché du travail.

Malheureusement, l'influence de ces éléments est très difficile à quantifier. Elle est laissée à la discrétion du modélisateur, qui le plus souvent la laisse nulle. Une option plus raisonnable serait certainement de supposer que l'effet sur la capacité de production globale serait le même que pour l'investissement productif des entreprises, mais avec un délai plus long dans son application. Cela voudrait dire que la principale différence de comportement entre les secteurs public et privé serait la moindre préoccupation pour des profits rapides, et un champ plus étendu des effets.

Enfin, on peut vouloir modifier les paramètres structurels, tels que la productivité des facteurs, ou le taux d'amortissement. On peut ranger dans la même catégorie les résidus de certaines équations de comportement: le rapport de la consommation des ménages aux revenus peut être modifié ex ante, ainsi que le partage de la valeur ajoutée entre les entreprises et les travailleurs. Selon la formulation, les

³ Sinon, le pays modélisé sera indûment soumis à des pertes de compétitivité, venant de l'augmentation de son taux d'inflation propre.

mécanismes de correction d'erreurs et les multiplicateurs peuvent entraîner une augmentation ou une diminution des effets ex post⁴.

Bien sûr, ces chocs entrent dans la première catégorie en ce qui concerne leur faisabilité, en tant qu'éléments incontrôlés externes.

9.1.4 LA PREVISION: ASPECTS TECHNIQUES

Il pourrait sembler qu'une fois qu'un modèle cohérent a été mis au point, avec une précision globale et des propriétés économiques validées par les tests précédents, les efforts restants pour produire une prévision correcte seront essentiellement liés à la définition des hypothèses. Ce n'est malheureusement pas le cas, et une bonne dose de savoir-faire sera nécessaire au constructeur de modèles pour produire des prévisions acceptables, portant notamment sur les points suivants:

- La définition des critères permettant de juger du caractère vraisemblable de la prévision.
- La recherche de l'intégration des évolutions des éléments endogènes, connus mais pas encore précisément quantifiés (en particulier en ce qui concerne les toutes premières périodes prévues). Cela va de valeurs partiellement connues (comme la récente évolution mensuelle du chômage pour une simulation trimestrielle) à des éléments plus vagues, tels que les anticipations fournies par des non-spécialistes de la modélisation. Au cours des périodes actuelles les prévisions de l'économie européenne appellent à un grand nombre de corrections conjoncturelles.
- Une fois corrigées ces anomalies concernant l'avenir proche, la manière dont ces éléments doivent être pris en compte pour les périodes suivantes. Par exemple, si on décide que les estimations des exportations doivent être diminuées à la première période par rapport à la valeur donnée par l'équation, cet écart doit-il être maintenu à moyen terme ?
- L'interprétation des résultats, au-delà de la simple observation des chiffres.
- En outre, certaines difficultés spécifiques peuvent apparaître lors de la génération des hypothèses:
 - En général, elles nécessitent l'intervention d'une expertise externe, car le constructeur du modèle macroéconomique ne peut pas être pleinement compétent dans l'ensemble des domaines concernés. Dans certains cas, il sera nécessaire de faire appel à un autre modèle pour obtenir des hypothèses cohérentes (par exemple un modèle mondial fournira des hypothèses sur l'environnement international).

⁴ La boucle prix-salaire va augmenter l'effet déflationniste d'une augmentation de salaire, mais l'impact de la demande additionnelle sur les importations sera réduit par l'adaptation des capacités.

- L'information externe, même quantifiée, n'est souvent présentée ni dans les unités, ni dans les catégories utilisées par le modèle (comme les accords commerciaux ou les résultats des négociations salariales).
- La transformation de certaines informations qualitatives (climat, grèves, les tensions, les attentes) en valeurs numériques est encore plus complexe.

9.1.5 LE CHANGEMENT DE SPECIFICATIONS

9.1.5.1 Les raisons

Normalement, le modélisateur doit réaliser ses prévisions avec le même modèle que celui qu'il a estimé. Pour procéder autrement, il devrait y avoir une raison acceptable.

On peut penser à quatre cas :

- Il est notoire que le comportement a changé, d'une certaine façon: par exemple, la Banque centrale a décidé de changer sa méthode de ciblage de l'inflation, ou elle vient de devenir indépendante et peut se prononcer sur le taux d'intérêt. Ou un nouveau gouvernement vient de décider de stabiliser le déficit de la sécurité sociale, en mettant le poids sur les contributions des entreprises et des travailleurs, à parts égales.
- Un comportement était présent dès le début, mais nous n'avions pas suffisamment de données, ou les variations de l'élément d'explication étaient trop faibles pour le mettre en évidence. Par exemple, le pays n'était confronté à aucun problème d'approvisionnement, mais les capacités ont moins augmenté que la demande, et le nouveau déséquilibre doit être pris en compte⁵.
- Un comportement est apparu vers la fin de la période d'estimation, et on ne dispose pas d'assez de périodes pour l'appuyer. Par exemple, une nouvelle loi a introduit le travail à temps partiel, avec un fort impact sur le processus d'embauche.
- On sait que l'économie future du pays suivra des comportements différents du passé. Cela peut s'appliquer à un pays en transition.

Dans le premier cas, le comportement est supposé connu. L'équation associée doit être introduite, par exemple une règle de Taylor.

⁵ Il était présent dès le début, mais n'a pas eu l'occasion de jouer un rôle.

Le deuxième cas est plus difficile à résoudre. Il s'applique par exemple à une période pendant laquelle l'inflation est restée stable, tandis que dans l'avenir on peut s'attendre à ce qu'elle varie beaucoup plus. Ou l'inflation locale était semblable à celle des pays partenaires (stabilisant la compétitivité, dont l'on ne pouvait identifier le rôle), tandis que des différences commencent à apparaître.

Dans ce cas, il n'y a pas d'autre moyen que d'établir des valeurs à partir de la théorie économique, de la mesure sur les autres pays et de l'observation des propriétés qui en découlent.

Les deux derniers cas peuvent généralement être associés:

- aux pays en développement,
- aux pays en transition (certains pays peuvent être les deux à la fois comme la Chine ou le Vietnam, ainsi peut-être que la Roumanie et la Bulgarie).

9.1.5.1.1 Identifier les pays en transition

Dans les pays en transition, le comportement de tous les agents a changé lorsque les mécanismes de l'économie de marché ont fait leur apparition. Mais bien que la transition technique puisse être assez rapide (comme la réunification allemande), le processus de transition a été généralement beaucoup plus lent.

On peut séparer deux cas:

- L'Europe centrale et orientale, avec l'ajout de la Russie et peut-être des anciennes républiques soviétiques.
- L'Asie: principalement la Chine et le Vietnam, peut-être le Laos et le Cambodge.

La différence est bien sûr que ces derniers pays sont encore gouvernés par des régimes communistes, et qu'ils peuvent toujours être considérés comme des pays en développement, ce qui est sans doute aussi le cas des éléments les moins avancés du premier ensemble, comme la Bulgarie, la Roumanie et l'Ukraine⁶.

En fait, à notre avis, si l'on considère la situation actuelle, le type de régime n'est pas le plus important, car le comportement de la Chine et du Vietnam se rapproche maintenant des économies de marché (peut-être plus que la Russie actuelle). Le seul problème est la date à laquelle la transition peut être considérée comme suffisamment avancée pour que l'économie soit modélisée de cette façon. La date officielle a été

⁶ Même si l'appréciation est très dépendante des critères (regardez la Chine par exemple).

1978 pour la Chine, et 1986 pour le Vietnam. Mais le processus a été très progressif, et l'on devrait envisager de retarder l'application d'un modèle structurel d'au moins 10 ans.

Pour le bloc de l'Est, la transition a commencé au début des années 90, et peut être considérée comme ayant été beaucoup plus rapide, en particulier parce que l'économie de ces pays avait suivi les standards occidentaux (en cadre et en niveau) 50 ans auparavant. Cette modélisation pourrait débuter en 1994-1995, avec des éléments supplémentaires:

- Le système statistique a pris un certain temps à adopter les méthodes nouvelles, en particulier le remplacement du système de matrice de comptabilité sociale par des comptes nationaux de type occidental (comme le SCN95 défini par l'ONU).

À l'heure actuelle, la quasi-totalité de ces pays fournissent des comptes trimestriels. Cela multiplie le nombre d'observations par quatre, ce qui est crucial ici (bien que le problème baisse avec le temps). Malheureusement, le gain en information est bien sûr beaucoup plus faible.

- Certains événements particuliers ont eu lieu au cours de la période, à l'instar de la partition de la Tchécoslovaquie en 1993, ou la décision de rejoindre l'Union Européenne (et même la zone euro).

9.1.5.1.2 En ce qui concerne les données

Plusieurs situations peuvent se présenter:

- Nous avons maintenant suffisamment de données pour estimer les équations. C'est par exemple le cas de la Pologne, où les données sont disponibles de 1995 à 2012 (74 observations), en utilisant la même méthodologie, et des comportements qui peuvent être considérés (avec quelques doutes quand même) comme représentatifs de l'économie de marché.

Mais nous devons également prendre en considération les deux derniers cas:

- Dans le cas 4 (cf. plus haut), l'équation doit être entièrement calibrée, avec des valeurs prises à partir de la théorie économique, des valeurs obtenues pour les autres pays et l'observation des propriétés qui en découlent.
- Dans le cas 3 également, certains renseignements supplémentaires peuvent être obtenus à partir des estimations. Mais ils ne devraient pas être utilisés seuls, même si toutes les statistiques semblent bonnes (sauf si nous obtenons de très bonnes statistiques sur un échantillon en principe trop petit).

Dans tous les cas, si l'influence d'un élément (comme la rentabilité sur le capital ou le chômage sur le taux de salaire) est considérée comme étant de plus en plus présente à la période actuelle, le coefficient

associé devrait également afficher une croissance pour les prévisions, probablement amortie avec le temps pour converger vers une constante. Ce processus (niveau initial de l'augmentation, vitesse de convergence) est de la responsabilité du modélisateur, aidé par l'observation des propriétés du modèle.

9.1.5.2 La technique

Nous pouvons utiliser exactement la même technique que pour tester les propriétés du modèle sur l'avenir, sauf que le rôle des changements dans les valeurs des paramètres et les formulations va augmenter.

9.1.6 LE CONTROLE OPTIMAL

Cette technique consiste à calculer les valeurs des instruments qui rendent les variables du modèle les plus proches possible d'une fonction objectif donnée, sous la contrainte du modèle. Par rapport à l'inversion des causalités, l'utilisateur du modèle fera face ici à une cible unique, normalement inaccessible, formalisée dans la plupart des cas par une fonction quadratique d'une ou de plusieurs variables. Il va alors appliquer un algorithme de maximisation sous contraintes (les équations du modèle ainsi que des limites sur l'évolution des instruments). Cela devra se faire en dehors du logiciel de modélisation (les feuilles de calcul proposent une version simplifiée de cette technique).

Cette technique a été très populaire dans les années cinquante et soixante, dans les pays occidentaux comme la France (ou elle était gérée par le Commissariat Général au Plan) et l'est restée dans une certaine mesure dans les pays socialistes⁷.

9.1.7 L'ENSEIGNEMENT PAR LES MODELES

Les modèles structurels permettent des applications éducatives ou scientifiques. Les diagnostics peuvent être observés pour eux-mêmes, comme l'illustration chiffrée de certaines théories. Mais leur petite taille les rend efficaces pour l'application des techniques de validation spécifiques, en particulier s'ils ont été conçus pour reproduire les propriétés d'un modèle opérationnel plus grand.

⁷ Il n'est pas surprenant que la Russie ait produit les meilleurs théoriciens dans ce domaine (comme Pontryagin) et dans le domaine connexe de la théorie des matrices (comme Gantmacher).

La simplicité de ces modèles leur donne le rôle naturel d'un instrument d'apprentissage. Selon le public et l'objectif, différents outils peuvent être utilisés.

- Des représentations extrêmement simplifiées, basées sur des exemples formels et de paramètres fixes, clarifiant les cadres théoriques (modèle de Solow, modèle IS-LM, modèle de Mundell-Fleming, de Fair ...). En les formalisant sur un exemple numérique et en les résolvant selon diverses hypothèses, on peut mieux apprécier leurs propriétés (la simplicité de ces modèles permet souvent un affichage graphique de leurs équations sous forme de courbes dynamiques). Les cours de macroéconomie sont maintenant souvent accompagnés par des programmes informatiques qui simulent ces représentations.
- D'autres modèles appliqués, qui reproduisent le plus fidèlement possible les propriétés des grands modèles utilisés pour les diagnostics opérationnels. Plutôt que d'illustrer une théorie économique particulière, le but ici est de réunir un nombre limité de mécanismes élémentaires, mettant en évidence les principales interactions qui régissent l'équilibre économique. Ce sera généralement réalisé grâce à des chocs, en se concentrant sur l'impact des différents instruments de politique économique et les ordres de grandeur de leurs effets.

Dans ce dernier cas, même si la précision de ces modèles reste limitée, la production de diagnostics acceptables nécessite de les appliquer à des cas réels. Par conséquent, ils seront généralement fondés sur des données observées, et souvent des estimations économétriques. Et pour rendre l'interprétation plus facile, on va généralement traiter des questions d'actualité: par exemple, mesurer les conséquences d'une diminution de la durée du travail, pour l'économie française de 1998, ou d'une réforme des retraites.

9.2 LA PRESENTATION DES RESULTATS

9.2.1 ELEMENTS GENERAUX

Le rôle joué dans l'efficacité du processus de simulation par la qualité de la présentation des résultats est souvent sous-estimé. Pour les chercheurs, il représente souvent une tâche mineure, à laquelle ils ne sont pas très bien préparés. Néanmoins, cet élément est souvent une condition nécessaire pour capter l'attention du public et assurer le succès d'une étude. Même si un travail médiocre mais agréable à l'œil rencontrera quelquefois un certain succès, un travail de qualité mal présenté ne parviendra généralement pas à atteindre son public, en décourageant ses auditeurs ou ses lecteurs potentiels, y compris les évaluateurs décidant de leur publication éventuelle.

Une bonne présentation doit être:

- sans erreur (ce qui est plus difficile qu'il n'y paraît). Une seule erreur visible jettera le doute sur l'ensemble des résultats ;

- explicite (la signification des chiffres présentés doit être évidente, en particulier leur nomenclature) ;
- pas trop longue ;
- agréable à l'œil (pas trop dense), mais pas trop commerciale (il ne faut pas donner à penser que l'auteur essaie d'occulter des problèmes, ou a détourné son temps de recherche) ;
- synthétisant clairement les principaux enseignements, dans le bon ordre logique ;
- autorisant un accès rapide à un élément particulier d'information ;
- adaptée à la fois à la nature du travail et à son public (pour des publics différents, des présentations différentes peuvent être produites pour le même travail⁸).

Il faut aussi garder à l'esprit que ces deux types de présentation, bien que complémentaires, devraient être conçus simultanément.

9.2.2 LES TABLEAUX

Il existe deux principaux types de tableaux, en fonction du public:

- Des tableaux de base, conçus pour le constructeur de modèles et de ses partenaires de travail, utilisant des formats simples résumant l'ensemble des informations sur le modèle. Ils permettent une interprétation rapide des principales caractéristiques des simulations et des chocs. Pour un modèle macroéconomique, le tableau contiendra généralement l'évolution des principales composantes de l'équilibre offre-demande (à prix constants), ainsi que l'évolution des prix, de l'emploi, des soldes extérieurs et budgétaires, et de certains ratios.

La définition des éléments présentés par le tableau sera limitée, parfois même à un nom représentatif du concept.

Le tableau sera plus simple à réaliser si des procédures adéquates ont été créées par l'utilisateur à l'intérieur du logiciel de création de modèle.

Pour des modèles plus détaillés, on n'aura pas à consulter l'ensemble des résultats, sauf s'il faut trouver les raisons de l'évolution anormale de certains éléments agrégés.

⁸ Ceci est facilité par les logiciels de présentation.

- Des tableaux plus élaborés, conçus pour des présentations à un public plus ou moins initié (le grand public, les clients d'une étude, les partenaires qui y ont collaboré directement, d'autres chercheurs, les non-spécialistes du sujet). Même dans le cas d'un document de travail, un certain niveau de qualité est nécessaire. Ceci est facilité par les logiciels de modélisation les plus sophistiqués, et leur production peut être réutilisée pour certains types de publications qui ne nécessitent pas une sophistication trop élevée. La qualité des tableaux EViews a augmenté de façon spectaculaire avec les dernières versions, et à notre avis, il n'est plus nécessaire de transférer des informations vers des logiciels de présentation, sauf dans des cas très spécifiques.

Dans la pratique, on observe que, même si le modèle est complètement assimilé, la production de tableaux relativement élaborés facilite le travail, et permet de mettre en évidence non seulement les propriétés, mais aussi les problèmes. Les produire dès que possible (même si le modèle n'est pas stabilisé) représente un bon investissement, dès la phase de développement. Une fois leur cadre précisé, ces tableaux peuvent être réutilisés indéfiniment sans coût supplémentaire.

9.2.3 LES GRAPHIQUES

L'avantage des graphiques est de clarifier les évolutions de la série par une présentation plus parlante. L'inconvénient est le manque de chiffres précis. Les graphiques doivent donc être utilisés:

- avec les tableaux, afin de compléter les résultats plus détaillés par une information synthétique et facilement accessible ;
- seuls, si le message transmis permet une synthèse simple.

On peut envisager plusieurs types de graphiques:

- Des points reliés ou non par des segments, ou plus rarement par des courbes (différenciés par une couleur, un type de ligne, et souvent un symbole).
- Des histogrammes, chaque observation étant associée à un ensemble de barres verticales (juxtaposées ou empilées) mesurant chacune la taille d'une variable.
- Des «camemberts» associés à une seule période, dont le cercle sera décomposé en tranches traduisant le rôle de chaque variable dans le total.

Cette liste n'est pas limitative, comme l'imagination peut aller loin dans la recherche de la synthèse visuelle. Par exemple, les barres des histogrammes sont souvent remplacées par les dessins appropriés (personnes, usines, produits ...).

En ce qui concerne les graphiques (dans l'ensemble des fonctions de présentation), les logiciels de modélisation ont réalisé des progrès considérables. Certains offrent des fonctions avancées par le biais d'un vaste ensemble de fonctions conviviales et EViews n'est pas le dernier à cet égard. Dès la fin de la

simulation, par exemple, on pourra afficher l'évolution des variables historiques et simulées, ou celle d'expressions choisies. Ce mode d'affichage sera utilisé en permanence dans le processus de développement d'un modèle, permettant une meilleure interprétation des résultats intermédiaires, et révélant des problèmes.

La qualité des graphiques proposés par EViews a aussi augmenté de façon spectaculaire avec les dernières versions. En fait, nous les préférons maintenant à ceux d'Excel, si l'on en reste à leur but précis: l'affichage d'un nombre limité de séries chronologiques, en couleur ou en monochrome, avec les définitions et titres associés.

Le langage utilisé peut être un peu ésotérique (bien qu'explicite) mais une fois qu'une présentation graphique a été conçue, les commandes peuvent être utilisées pour initialiser la suivante. Les fonctions graphiques d'EViews ont encore été améliorées dans la version 8. Vous pouvez maintenant:

- décider de la portée du graphique à l'aide d'une règle, disponible en bas de celui-ci ;
- ajouter des flèches personnalisées ;
- exporter le graphique dans un fichier PDF.
- Mais l'ajout le plus intéressant est la commande MAKEGRAPH, spécialement adaptée aux solutions d'un modèle. La syntaxe en est:

```
nom-du-modèle.makegraph (options) nom-des-séries nom-du-graphe
```

Si aucune option n'est spécifiée la solution de scénario courant est affichée.

Les principales options sont les suivantes:

- a: ajouter les valeurs observées
- c: comparer le scénario actif à la référence («baseline»)
- d: ajouter les écarts par rapport à la référence (comme un graphique supplémentaire dans le même cadre)
- n: ne pas inclure le scénario actif

Les résultats peuvent être présentés bruts ou transformés, et pour les simulations stochastiques un intervalle de confiance peut être affiché.

Une fois que le graphique est créé, les modificateurs habituels peuvent être appliqués (légende, type de ligne, couleurs).

Il faut noter que le graphique ne doit pas préexister. Si c'est le cas il faut d'abord le supprimer (s'il existe, le modificateur «noerr» permet d'éviter un message d'erreur).

```
delete(noerr) nom-du-graphe
```

CHAPITRE 10: L'APPLICATION DES PRINCIPES AUX CAS OPÉRATIONNELS

Dans les chapitres précédents, nous avons abordé tous les aspects techniques de la construction des modèles et ses applications, en les illustrant par un exemple simple. Même si les techniques proposées étaient tout à fait générales, elles étaient trop simples pour pouvoir suffire à la production de modèles opérationnels. Nous allons devoir les compléter maintenant.

Après avoir maîtrisé les techniques de production d'un très petit modèle, le professionnel devra faire face à deux types de difficultés:

- La structure du nouveau modèle doit être plus conforme à la théorie économique (celle qu'il a à l'esprit). Le détail des spécifications doit répondre à ses besoins, en termes d'hypothèses, de mécanismes et de résultats.
- La méthodologie de production et de gestion du modèle devra faire appel à des techniques légèrement différentes, adaptées à un problème de plus grande complexité.

Dans les pages suivantes, nous allons aborder ces deux questions. Comme nous ne proposons pas un cours de macroéconomie¹, nous décrivons uniquement les éléments qui, à notre avis, sont exigés pour la production d'un modèle opérationnel, les types de choix nécessaires, et nous décrivons les directions les plus prometteuses pour les autres options. Nous laissons au lecteur le choix des décisions effectives et l'introduction des fonctionnalités supplémentaires qu'il estime adaptées à son problème.

Il vous est donc demandé de ne pas critiquer nos explications comme étant trop basiques². Si cela est votre sentiment:

- vous êtes un assez bon économiste pour procéder par vous-même ;

¹ Nous n'avons pas donné non plus un cours d'économétrie.

² Certains lecteurs vont certainement décider qu'elles ne sont pas les bonnes. C'est un autre problème, mais la plupart des fonctionnalités que nous présentons ont été utilisées dans plusieurs modèles (ce n'est pas une preuve de leur validité, bien sûr).

- ou vous pouvez utiliser l'un des manuels plus avancés dans la liste que nous fournissons.

Encore une fois, nous rappelons que vous ne lisez pas un livre sur l'économétrie ou l'économie sophistiquées. Notre objectif est de montrer comment les deux domaines peuvent être fusionnés dans une stratégie qui fonctionne, en se concentrant sur les aspects techniques de la tâche.

Dans tous les cas, les éléments que nous décrirons devraient se révéler utiles, s'ils sont interprétés comme décrivant les options le plus fréquemment retenues par les principaux modèles opérationnels structurels (même si on pourra juger certaines d'entre elles comme démodées).

Pour cette raison, nous ne citerons pas une source dans laquelle les mécanismes que nous décrivons sont présentés. Au contraire, nous donnons une liste de livres sur la macroéconomie, et nous présenterons deux modèles complets, dans lequel la plupart d'entre eux seront trouvés dans un certain détail.

Nous allons être plus directifs sur les aspects EViews, bien sûr.

10.1 LE CADRE COMPTABLE

Avant de commencer à traiter la théorie macroéconomique, nous devons décrire le cadre et les concepts que nous allons utiliser. Nous avons déjà commencé, bien sûr, mais dans les cas opérationnels nous devons être beaucoup plus précis et complets.

En ce qui concerne l'économie, nous n'avons pas l'ambition de donner un cours de comptabilité nationale. À la fin du livre, on trouvera une liste des publications associées, allant des cours élémentaires à des manuels officiels indiquant les concepts officielles utilisées à l'heure actuelle par des organisations locales ou internationales.

À notre avis, la meilleure référence dans ce domaine est la suivante:

F. Lequiller, D. Lames (2006), *Comprendre la comptabilité nationale*, Éditions OCDE, téléchargeable à l'adresse: <http://www.oecd.org/dataoecd/37/12/38451313.pdf>

Pour définir le cadre d'un modèle macroéconomique, il faut définir:

- les agents impliqués,
- les opérations qu'ils partagent,
- les produits associés à ces opérations.

Selon le problème, et le modèle conçu pour les résoudre, ces éléments seront plus ou moins détaillés.

Nous allons commencer par un modèle à un seul pays (ou une série de modèles non liés), puis étendre cette classification. Les détails du processus de construction seront traités séparément.

10.1.1 LES AGENTS: UNE PREMIERE CLASSIFICATION

Les opérations comptables décrites dans un modèle seront gérées par des agents, caractérisés par leur rôle. Nous allons séparer:

- les entreprises, qui achètent et produisent des biens et services marchands, et distribuent des revenus ;
- les ménages, qui reçoivent des revenus et consomment une partie de celui-ci ;
- les institutions financières, qui gèrent les transactions financières: prêts, emprunts, assurances ;
- les administrations, qui reçoivent les impôts des agents ci-dessus, et répondent aux besoins sociaux par la redistribution et la production d'éléments associés ;
- le reste du monde, qui représente les agents étrangers négociant avec les quatre premiers.

10.1.2 LES OPERATIONS

Elles sont divisées en:

- Opérations sur biens et services: un bien (matériel, comme une bouteille de vin ou un téléviseur) ou un service (immatériel, comme le transport d'un endroit à un autre ou une leçon de musique) sont négociés entre deux agents, généralement avec échange d'argent. Mais il peut être échangé (un hébergement contre des services domestiques) ou fourni gratuitement, la plupart du temps par une administration (enseignement ou médicaments gratuits). Dans le dernier cas, la transaction est considérée comme «non-marchande».

- Transferts d'un agent à l'autre, qui peuvent venir en échange de quelque chose, indépendamment de l'opération ci-dessus. Par exemple, les salaires sont payés aux ménages par les entreprises, mais ils n'achètent pas en général la réalisation d'un bien, seulement une contribution à sa production. Alternativement, les retraites sont financées par les cotisations, mais pas nécessairement celles des bénéficiaires, et en tout cas avec un certain retard. Et les impôts ne sont pas liés à un service (même s'ils seront utilisés en partie pour le bien du contribuable, du moins en principe).

Les choses ne sont pas toujours aussi simples: un chauffeur de taxi indépendant vendra un service (et utilisera l'argent à titre de revenu). S'il appartient à une société, il va gagner un salaire (un transfert). La société va vendre le service et lui transférera une partie des recettes.

10.1.3 LES COMPTES ECONOMIQUES INTEGRES

Une fois que ces éléments ont été définis, ils peuvent être présentés dans un tableau, qui contient deux parties: gauche (ressources) et droite (emplois) avec une ligne par opération, et une colonne par agent. En fait, une colonne supplémentaire sera introduite pour la partie «biens et services».

Cette présentation est particulièrement utile pour les transferts: parfois un seul élément apparaîtra des deux côtés (sur la même ligne) comme pour l'impôt sur le revenu, l'emploi des ménages et les recettes de l'État. Mais si les salaires sont bien une recette des ménages, côté emplois ils sont répartis entre tous les agents, y compris les ménages (les deux sommes sont égales évidemment).

En outre, le tableau décrit les étapes du processus économique: les opérations seront décrites dans l'ordre séquentiel des sous-tableaux, chacun présentant un bilan sur la dernière ligne.

Nous rencontrerons successivement:

- Le compte de production, qui décrit toutes les opérations liées à la production, mais aussi les importations et les exportations. Il donne la valeur ajoutée. Cette partie exigera une colonne supplémentaire «Biens et Services», qui évitera un déséquilibre au niveau de la ligne (les exportations représentent une dépense pour le reste du monde, mais les ressources ne sont pas réparties entre les agents).
- Le compte de répartition primaire des revenus, qui introduit les dépenses immatérielles: les salaires, les impôts directs, les subventions (avec un positionnement inversé). Il donne l'excédent d'exploitation.

- Le compte de distribution secondaire du revenu, qui complète la distribution des revenus: les dividendes, les revenus de la propriété, ainsi que les transferts sociaux des administrations. Le solde est le revenu disponible.
- Le compte d'utilisation du revenu disponible, décrivant pour l'essentiel la consommation des ménages. Son solde est l'épargne.
- Le compte de capital, décrivant l'utilisation de l'épargne pour acquérir des biens. Il comprend diverses formes d'investissement et la variation des stocks. Son solde est la capacité de financement.
- Le compte financier, qui explique le financement des opérations précédentes. Il décrit les encours monétaires, les stocks et les crédits. Son solde est évidemment nul.

10.1.4 SECTEURS, BRANCHES ET PRODUITS

Nous pouvons maintenant séparer les «entreprises» ou plutôt «sociétés non financières» en catégories, en fonction de son type de production. Trois notions apparaissent ici:

- Le secteur: il catégorise une entreprise en fonction de son activité principale (construction automobile pour le constructeur automobile Peugeot).
- La branche: elle sépare l'entreprise en chacune de ses activités (automobiles, mais aussi production de pièces de rechange et ventes de voitures pour le constructeur automobile Peugeot).
- Le produit: il représente le bien ou le service effectivement produit par l'entreprise. Il est normalement équivalent à la branche, mais pas dans tous les cas: la différence peut provenir de «produits fatals» venant automatiquement d'un processus visant à créer un autre bien (comme l'hydrogène provenant du raffinage du pétrole). On peut également citer le bois, produit marchand, produit sans intervention par l'Etat dans les forêts domaniales.

Comme nous le verrons plus tard, cette décomposition appelle à la définition de la consommation intermédiaire, une variable à deux dimensions associée à la consommation d'un bien pour en produire un autre. Cela peut être vrai aussi pour l'investissement (en un produit par une branche) et les variations de stocks³.

³ Bien que dans ce cas, certains éléments sont identiquement nuls, comme l'investissement dans le produit énergie ou les stocks de services financiers.

10.1.5 LES SUBDIVISIONS DES AGENTS

Les comptes nationaux peuvent subdiviser les agents. Par exemple, on peut séparer:

- les ménages entre individuels (une famille) et collectifs (un ancien couvent) ;
- les institutions financières entre Banque centrale, autres banques et compagnies d'assurance ;
- les entreprises appartenant à l'État, les entreprises individuelles, les sociétés locales, ou les entreprises appartenant à l'étranger ;
- l'administration en gouvernement central, collectivités locales et organismes spécifiques, tels que l'organisme de sécurité sociale ;
- le reste du monde en pays et zones (même pour un modèle à un seul pays). Cela permettra de rendre les hypothèses du modèle plus explicites.

Bien sûr, d'autres classifications plus statistiques peuvent être utilisées, comme les ménages par niveau de revenu ou occupation du chef de famille.

10.1.6 UN MODELE A PLUSIEURS PAYS

Nous développerons plus tard les spécificités d'un modèle multi pays. Donnons simplement pour l'instant quelques éléments généraux:

Bien entendu, le système doit être reproduit pour chacun des pays décrits. Des éléments spécifiques peuvent être introduits dans certains cas (comme l'identification des hydrocarbures pour les pays producteurs).

Une cohérence doit être introduite entre les exportations de chaque pays et les importations des autres. Cela peut se faire soit par l'identification des flux individuels, ou en créant une variable de demande mondiale spécifique pour chaque pays, en fonction de la demande globale de chacun de ses clients.

Enfin, les catégories peuvent être des pays, des associations de pays (comme l'Union européenne) ou des zones liées (comme l'Afrique subsaharienne).

10.2 UN MODELE À UN SEUL PAYS ET UN SEUL BIEN

Nous allons commencer par les développements économiques nécessaires à notre modèle à atteindre le statut opérationnel.

10.2.1 LES ASPECTS ECONOMIQUES

Le très petit modèle que nous avons construit à titre d'exemple est bien sûr trop limité pour être utilisé pour les études opérationnelles. Il présente néanmoins:

- un lien entre la production et des revenus ;
- un commerce extérieur, fonction de la demande et de la capacité de production disponible ;
- une production cohérente avec la demande locale et les exportations nettes ;
- une fonction de production simple ;
- une description de l'écart de production, de ses conséquences et de la façon dont il est corrigé.

Que devons-nous ajouter, au minimum ?

- un système de prix, et ses liens avec le secteur réel (dans les deux sens) ;
- une meilleure description du comportement des entreprises et des ménages ;
- un secteur financier ;
- une description complète du budget de l'État et de ses instruments.

Ces éléments vont appeler aussi à une redéfinition du champ d'hypothèses.

10.2.1.1 Le processus productif

Cette partie du modèle (on parle souvent de « blocs ») ne définira pas la production, mais plutôt la production potentielle (ou capacité de production), en fonction de facteurs disponibles.

Pourquoi ne pas la production elle-même ? Il y a deux manières de considérer la production :

- La production réelle locale, contribuant avec les exportateurs étrangers à la satisfaction de la demande (à la fois locale et étrangère), le partage se faisant en fonction des prix relatifs et des capacités disponibles.
- La production potentielle, donnée par la fonction de production, en tenant compte du niveau des facteurs (capital et travail), eux-mêmes choisis par les entreprises en fonction de leurs coûts relatifs, de la demande anticipée, et de la profitabilité du processus.

Nous voulons que notre modèle suive la séquence causale la plus logique, qui est la suivante :

- La définition de la capacité cible en fonction des conditions de profitabilité et de la demande prévue.
- Le choix du niveau optimal de facteurs permettant cette capacité.
- L'adaptation partielle des niveaux réels des facteurs, ce qui donne le potentiel de production.
- La demande globale suivra, et sera partagée entre les producteurs locaux et étrangers pour donner la production réelle.

La connaissance imparfaite de la demande future, les inerties techniques et les dangers d'une adaptation trop rapide des facteurs contribueront à la création d'un écart entre les valeurs potentielles et réelles.

Cette comparaison entre la production réelle et potentielle jouera un rôle important dans plusieurs comportements.

Il s'agit ici de la séquence que décrit le modèle, la production effective étant obtenue tard dans le processus, une fois que les éléments de demande sont connus (comme dans le petit modèle).

Cette capacité de production sera mesurée :

- pour l'emploi, par des hommes-année ou des hommes-trimestre selon la périodicité du modèle ;
- pour le capital, par une valeur à prix constants, dans la monnaie du pays.

La fonction peut également inclure:

- la consommation d'énergie ;
- les biens intermédiaires (comme les matières premières).

En fait, les capacités sont généralement définies en termes de valeur ajoutée, une notion plus fiable comme nous l'expliquerons plus loin. Cela signifie que les deux derniers éléments ne sont pas pris en compte, ou plutôt que leur niveau sera défini automatiquement à partir de la valeur ajoutée elle-même.

La première question concerne le lien logique entre la capacité et les facteurs. Nous avons déjà présenté:

- les facteurs complémentaires. Pour une capacité donnée, il existe un seul processus optimal, qui utilise une combinaison donnée de travail et de capital. À partir d'une combinaison optimale, ajouter aux capacités une quantité d'un seul facteur n'augmente pas celle-ci, et ne permettra pas d'utiliser moins de l'autre facteur. Cette optimalité est évidemment indépendante des coûts relatifs. Mais en fait, la productivité du travail a généralement une certaine souplesse, et le capital est le seul facteur vraiment limitatif, car des hausses temporaires de la productivité du travail peuvent être réalisées (par exemple en augmentant le nombre d'heures travaillées).

Il s'agit de la solution la plus simple, dans sa formulation, son estimation et la compréhension de ses propriétés. Les modèles opérationnels utilisent des cadres en général plus sophistiqués:

- Cobb-Douglas. L'élasticité de substitution est unitaire. Cela signifie que si le rapport du coût de la main-d'œuvre au coût du capital varie de 1 %, le ratio optimal du capital au travail variera de 1 % aussi, pour une capacité désirée donnée.
- CES (élasticité de substitution constante). Cette fois, l'élasticité peut prendre n'importe quelle valeur fixe (avec le bon signe).

Bien sûr, l'option CES englobe les deux autres (avec des élasticités fixes de 0 et 1 respectivement).

Le cadre prévoit également:

10.2.1.1.1 Une définition du coût relatif.

Le coût relatif du travail et du capital n'est pas seulement mesuré par le rapport entre le taux de salaire et le déflateur de l'investissement. Il faut prendre également en compte:

- Les cotisations sociales des entreprises: elles contribuent au coût de la main-d'œuvre.
- Le taux d'intérêt: alors que le capital est acheté immédiatement, le travail peut être acheté au fur et à mesure, le moment venu (l'esclavage a été aboli depuis un certain temps). Ainsi, une entreprise qui a de l'argent peut l'épargner, et si elle n'en a pas elle n'a pas besoin d'emprunter⁴.
- Le taux d'amortissement: le capital s'épuise, tandis que lorsqu'un travailleur voit son efficacité diminuer avec le temps (physiquement ou par inadaptation aux nouvelles techniques), il partira et pourra être remplacé par un nouveau, sans dépense spécifique, si ce n'est le coût de formation de ce dernier (les retraites ont déjà été payées avec les salaires).
- L'évolution future des salaires: plus les salaires sont dans une phase de forte augmentation, et plus les entreprises peuvent s'attendre à ce que le travail devienne moins compétitif. Les bénéfices d'un transfert de production vers les pays en développement diminuent au fur et à mesure qu'ils comblent l'écart de salaire avec les pays développés. Ceci s'applique en particulier à la Chine.

10.2.1.1.2 Les éventuelles modifications de la technologie

- .La question ici est de savoir si la technologie est décidée définitivement au moment de l'investissement (qui définit les rôles de travail et de capital) ou peut changer plus tard.

Fondamentalement, les options sont les suivantes:

⁴ En fait certaines formes de capital (comme les bâtiments, les ordinateurs ou les brevets) peuvent être louées.

- Une seule technologie disponible (Clay-Clay).
- Une technologie choisie lors de l'installation, mais sans possibilité de modification ultérieure (Putty-Clay). Cela signifie essentiellement que l'option «facteurs complémentaires» s'applique une fois les facteurs installés.
- Une technologie avec possibilité permanente de changement (Putty-Putty). La faculté de substitution reste possible, à n'importe quelle période.

10.2.1.1.3 Un problème spécifique: la détermination statistique des capacités de production

Pour déterminer la capacité, nous avons plusieurs options, en fonction de l'information disponible

- Dans certains pays (comme la France), un sondage demande aux entreprises de combien elles pourraient augmenter leur production en utilisant les installations actuelles (mais en embauchant plus de personnes si nécessaire). Cela donne la capacité de production de l'entreprise. En utilisant les mêmes coefficients de pondération que pour le calcul de la production effective, on obtient une mesure comparable de la capacité, et le taux d'utilisation comme rapport des valeurs globales.

Ensuite, nous allons utiliser la série de capacité pour estimer les facteurs. Pour cela, nous pouvons spécifier le comportement réel des entreprises, et optimiser leurs profits sous une contrainte de capacité en utilisant la formule que nous voulons estimer. Ceci s'applique seulement lorsque les facteurs sont substituables (sinon la solution optimale est fixée dès le début, et ne dépend pas de coûts relatifs). Prendre successivement la dérivée de la fonction de production par rapport au travail et au capital donnera un ensemble de deux équations à coefficients communs, que l'on peut estimer comme un système. Cette méthode prend en compte pleinement et explicitement le rôle des coûts relatifs.

- Si nous ne connaissons que le niveau des facteurs (mais le capital manque parfois dans les statistiques du pays), on peut spécifier la fonction de production, et estimer ses paramètres en se basant sur les valeurs réelles de production. On peut supposer que la formule estimée donne la production normale, et que le résidu est l'écart de production. Encore une fois, le rapport des productions observée et «normale» donne le taux d'utilisation, mais ici à un facteur constant près (le taux moyen d'utilisation).
- Nous pouvons également (une meilleure solution à notre avis) appliquer la première méthode, en utilisant la production réelle au lieu de la capacité. Encore une fois, la capacité estimée (reconstruite en appliquant la fonction de production aux facteurs estimés, considérés comme optimaux) donnera un niveau normal de production, et la différence à la production réelle donnera l'écart de production.

- Si nous n'avons pas cette information, nous pouvons toujours lisser la production, et utiliser le résultat comme une «production normale» (associée à un taux normal d'utilisation des capacités). Pour cela, l'application aux données réelles d'un filtre Hodrick-Prescott est la technique la plus courante. Si nous supposons que le taux «normal» d'utilisation des capacités est constant au fil du temps, nous obtenons la capacité à un facteur multiplicatif inconnu.

Cette technique ne nécessite pas de choix de la fonction de production, ou de la disponibilité d'une série de capital (qui est souvent absente ou peu fiable). Pas plus qu'elle ne formalise l'offre, ce qui sera un problème pour la spécification du modèle.

10.2.1.2 Les variations de stocks

Nous ne voyons aucune raison particulière de modifier le cadre utilisé par le petit modèle. Des formulations plus sophistiquées peuvent être utilisées:

- Un cadre complet à correction d'erreurs, à condition de connaître le niveau des stocks.
- Une influence de la demande: si elle croît brusquement, une partie peut être satisfaite en puisant dans les stocks. Cet élément sera difficile à mettre en évidence, car il demande une influence négative, opposée à celle de la valeur ajoutée, alors que les deux éléments sont positivement corrélés. Cela signifie que la sur estimation d'un coefficient peut être compensée par une sur estimation du second.
- Une influence des prix: plus les biens stockés coûtent cher, moins de temps ils le resteront.

10.2.1.3 Le chômage

Il s'agit d'un nouveau concept, qui aurait pu être introduit facilement dans le petit modèle⁵.

⁵ Nous l'avons testé, et cela marche.

Nous allons considérer que les variations de l'emploi ne s'étendent pas comptablement au chômage. La création d'emplois va attirer sur le marché du travail des personnes précédemment inactives, qui vont occuper une partie des emplois proposés : la population active (emploi + chômage) va augmenter.

Par exemple, la création d'une entreprise dans une zone faiblement industrialisée permettra aux ménagères de concilier emploi et travail domestique⁶. Ou les employés licenciés par une fermeture d'usine ne vont pas nécessairement rester sur le marché du travail si leur qualification n'est pas demandée ailleurs.

Mais le niveau du taux de chômage devrait également influencer sa dynamique. S'il est élevé, l'incitation à rejoindre le marché du travail sera plus faible. Des perspectives d'emploi favorables conduiront les jeunes vivant avec leurs parents à commencer leur vie professionnelle. Au contraire, un marché du travail déprimé va persuader les travailleurs âgés de prendre leur retraite plus tôt (et ils y seront incités). Et une partie des chômeurs cessera de chercher un emploi, pour quitter le marché du travail.

En outre, un taux de chômage élevé augmentera la qualité des meilleurs chômeurs, ainsi que l'éventail de leurs compétences. Constatant la situation, les chômeurs s'estimant moins bons pronostiqueront une baisse de leurs chances de trouver un emploi, ce qui peut les amener à quitter le marché.

Au contraire, face à un faible taux de chômage, les chômeurs sentiront que leurs chances sont élevées par rapport à leurs concurrents, la plupart d'entre eux étant perçus soit comme inefficaces, soit comme n'étant pas vraiment à la recherche d'un emploi, ou travaillant dans des domaines peu recherchés.

Cela correspond évidemment à un cadre de correction d'erreur, conduisant à une cible de taux de chômage (et aussi de participation des travailleurs potentiels au marché du travail, comme nous le verrons plus tard).

⁶ Qui n'est pas considéré comme un emploi (peut-être parce qu'il n'est pas payé, et n'affecte pas le PIB, même si les travaux de ménages payants le font). Ainsi le fait qu'une femme au foyer fasse appel à une aide extérieure augmente le PIB (mais il s'agit alors d'un service marchand).

10.2.1.4 Le système de prix

Le rôle des prix dans un modèle est bien sûr essentiel. Mais il n'est pas si simple à mettre en place, même pour un modèle minimal comme celui que nous avons présenté ci-dessus. Dans ce cas, plusieurs déflateurs doivent être introduits simultanément, associés à des éléments de l'équilibre offre-demande:

- le PIB ;
- la demande finale ;
- les exportations ;
- les importations.

Et en plus:

- le taux de salaire (y compris éventuellement les cotisations de sécurité sociale) ;
- éventuellement, les déflateurs de chaque élément dans la décomposition de la demande: la consommation, l'investissement productif, l'investissement logement ... ;
- le prix de la monnaie étrangère (le taux de change) ;
- les prix des prêts et des emprunts (le taux d'intérêt).

Par ailleurs, les prix du commerce doivent être définis deux fois, en incluant et excluant les taxes. Cette distinction s'applique au commerce extérieur (pour la définition de la compétitivité et de la balance commerciale) et à demande locale (pour définir les consommations finales et intermédiaires).

Tous ces éléments n'ont pas à être estimés. Des comportements doivent être associés:

- au PIB (les entreprises décident du prix auquel elles vendent, une fois pris en compte le coût des intrants) ;
- aux exportations (les exportateurs locaux font de même) ;

- aux importations (nous considérons maintenant les exportateurs étrangers)⁷ ;
- aux salaires (le résultat d'une négociation entre les travailleurs et les gestionnaires de l'entreprise).

Le prix de la demande finale devra être utilisé pour équilibrer l'offre et la demande aux prix courants. Le modèle donne un ensemble équilibré de quatre éléments, à prix constants, et trois des déflateurs ont déjà été décidés. Le prix de la demande est le seul élément qui permet d'équilibrer l'ensemble:

$$Pfd \cdot FD + Px \cdot X = Pq \cdot Q + Pm \cdot M$$

ou

$$Pfd = (Pq \cdot Q + Pm \cdot M - Px \cdot X) / FD$$

Nous allons maintenant aborder les liens entre les prix

Dans le système, les déflateurs dépendront l'un de l'autre. Pour le moment, nous ne donnons que des indications. Un raisonnement plus détaillé viendra avec les estimations réelles.

- Le déflateur du PIB dépend du taux de salaire, ou plutôt du coût salarial.

Si les coûts salariaux augmentent, les entreprises devront augmenter leurs prix pour maintenir leurs marges.

⁷ Rappelez-vous, nous construisons un modèle à un seul pays. Le cadre de description sera différent avec plusieurs pays.

Elles n'ont pas à le faire immédiatement, et n'y sont pas obligées si elles veulent maintenir leur compétitivité sur les marchés locaux et étrangers (pour les entreprises exportatrices).

Et en fait, il pourrait être préférable d'utiliser le coût global, en y intégrant l'amortissement du capital.

- Le taux de salaire dépend du prix à la consommation, mais peut-être aussi du prix de la valeur ajoutée.

Si les prix augmentent, les travailleurs vont réclamer une augmentation de salaire pour maintenir leur pouvoir d'achat. Mais encore une fois, les entreprises sont moins susceptibles d'accepter celle-ci si elles ne sont pas en mesure d'augmenter leurs propres prix.

- Les prix du commerce extérieur dépendront du coût supporté par les exportateurs, et du prix fixé par ses concurrents. Cela signifie qu'ils doivent maintenir leurs marges et leur compétitivité en même temps.

Ce comportement est évidemment basé sur des prix à la production, le prix auquel ils vendent. Cela signifie que le coût des consommations intermédiaires doit être pris en compte. Par exemple, un pays ayant accès à du pétrole à bon marché sera en mesure d'exporter à des prix inférieurs, au même prix de valeur ajoutée (et les mêmes marges). Mais cela pose un problème, car jusqu'à présent la fonction de production nous a permis de mettre de côté la consommation intermédiaire, une variable difficile à gérer car sa valeur dépend de la nomenclature utilisée.

Le comportement doit également s'appliquer à des prix mesurés dans la même monnaie. Si le prix à l'exportation utilise la monnaie de l'exportateur, le prix de ses concurrents mesuré en monnaie étrangère doit être corrigé par le taux de change.

- Le prix de la demande dépend de celui auquel les producteurs et les exportateurs étrangers vendent sur le marché local.

Il utilise l'identité ci-dessus.

Une autre question importante concerne la séparation entre les prix auxquels les entreprises locales vendent sur les marchés locaux et étrangers (prix à l'exportation).

Deux comportements peuvent être envisagés:

- Les entreprises définissent les deux prix de vente séparément. Les entreprises locales commencent par définir un prix de vente sur le marché local, en utilisant le comportement décrit ci-dessus. Ensuite, le prix à l'exportation sera une moyenne de ce prix et de celui des concurrents.
- Les entreprises vont d'abord définir un prix de vente global, permettant d'atteindre un taux de marge global, puis ils choisissent une combinaison des deux prix qui répond à cet objectif. Cela signifie que la diminution du prix à l'exportation (peut-être destinée à rester compétitif face à la déflation étrangère) devra être compensée par une augmentation du prix de vente local.

Le choix aura un impact important sur le système de prix. En fait, la deuxième option va augmenter l'intensité de la boucle prix-salaires: si les dépenses locales augmentent, les entreprises refusent d'appliquer complètement ces coûts à l'exportation (car elles veulent maintenir la compétitivité), et le maintien des marges globales nécessite une plus grande augmentation du prix de vente local (ce qui n'est pas le cas si les objectifs sont définis séparément).

Cet équilibre est soumis à des influences extérieures, endogènes ou exogènes.

- Endogènes
 - si la productivité du travail augmente, les entreprises ont besoin de moins de travailleurs et peuvent les payer plus cher. Mais elles peuvent également baisser leurs prix ;
 - si les ventes sont trop faibles par rapport aux capacités, les entreprises peuvent commencer par baisser les prix pour accroître les débouchés (plus tard, elles pourront adapter leurs capacités) ;
 - si le chômage baisse, les travailleurs peuvent augmenter leurs exigences, avec moins de risque que les gestionnaires de l'entreprise cherchent à les remplacer.
- Exogènes

Les variables concernées sont les taux des impôts indirects.

On considère généralement:

- la TVA ;
- les autres taux d'impôts indirects;
- les droits de douane.

Et aussi

- le taux des cotisations sociales payées par les entreprises.

Si les taux des impôts indirects⁸ (comme la TVA, la taxe sur l'essence, l'alcool, les cigarettes, les cotisations sociales payées par les entreprises ...) augmentent, les entreprises doivent adapter leurs prix si elles veulent conserver leurs marges.

Deux éléments:

- Tout d'abord, il est très important de séparer ces taxes dans un modèle, pour la raison habituelle: leur base est différente, et leur impact sur l'économie aussi.
- La TVA s'applique uniquement à la valeur ajoutée. Mais la caractéristique la plus importante est qu'elle ne s'applique pas aux exportations (les exportateurs peuvent la déduire avant de vendre à l'étranger), mais par contre aux importations. La TVA sur les voitures étrangères est la même que sur les locales, et s'applique à la valeur totale. Et si la voiture utilise des équipements électroniques éventuellement importés, augmenter la TVA sur ce bien ne changera pas la décision sur son origine, car l'entreprise peut déduire la TVA de toute façon⁹.
- Au contraire, les autres impôts indirects ne s'appliquent qu'aux productions locales, même si l'impact de cette différence n'est pas si grand, car les marchandises importées sont généralement imposées au moment où elles sont vendues. Par exemple, la taxe sur le tabac s'applique aussi aux cigarettes importées, et la taxe sur l'alcool au whisky importé («produit» par le distributeur local au moment de la commercialisation).

⁸ Ces impôts sont appelés indirects parce qu'ils ne sont pas versés directement par le payeur final, contrairement à l'impôt sur le revenu, l'impôt sur les bénéfices des sociétés

⁹ Il pourrait ne pas appliquer la totalité de l'augmentation, cependant, s'il craint une baisse de ses ventes.

En ce qui concerne les droits de douane locaux, ils ne sont pas déductibles dans le cas général. Cela signifie qu'un changement de leur taux affectera directement la compétitivité des marchandises importées, sauf si l'importateur décide de compenser l'effet en adaptant ses marges.

En ce qui concerne les droits de douane appliqués aux produits locaux par les autres pays, ils affectent directement la compétitivité des exportations. Cela signifie qu'ils doivent être pris en considération, même si leur valeur statistique n'est pas directement tirée des comptes nationaux.

- Deuxièmement, la formalisation du rôle des impôts repose évidemment sur les taux, la variable décidée par l'État (ou les États étrangers pour les droits de douane sur les exportations locales). Ces taux aura une incidence sur les déflateurs, et permettra le calcul du montant de l'impôt une fois que la base sera connue.

Il devrait être clair que la bonne façon de formaliser ces taxes est de fixer le taux comme une variable de décision exogène, et non pas l'estimer comme certains modélisateurs pourraient être tentés de le faire. Cela permet de manipuler facilement la décision, à la fois dans les prévisions et l'analyse de chocs. Et sur le passé, la technique est assez simple: le montant de l'impôt et la base sont connus, ce qui permet le calcul exact du taux «apparent», utilisé comme un rapport exogène. L'identité associée ($\text{taxe} = \text{taux de base} \times x$) sera vérifiée. Nous verrons plus loin comment gérer ces taux sur l'avenir.

En pratique, le taux obtenu sera différent du taux légal (généralement plus faible pour des raisons évidentes).

Cette technique est conforme à l'approche générale: rendre les décisions gouvernementales exogènes, mais identifier d'abord ce que représente la vraie décision.

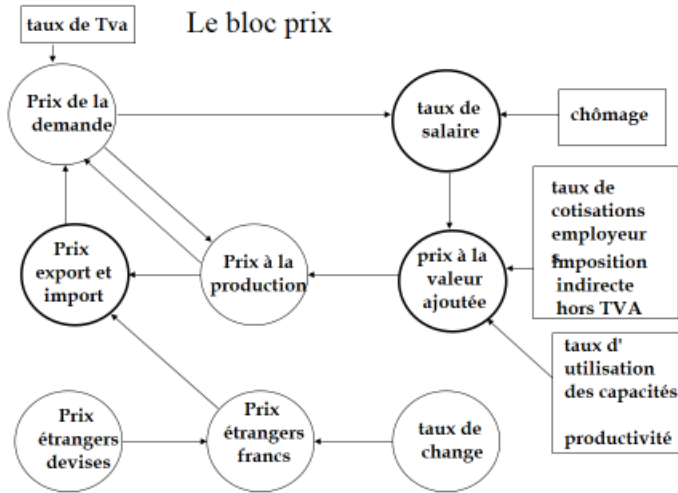
Les prix peuvent également influencer sur des éléments réels

- Le prix de vente des producteurs locaux détermine les quantités qu'ils vendent. Cela est également vrai des exportateurs, à travers la comparaison entre les prix à l'exportation et les prix sur le marché sur lequel ils vendent.
- Les coûts relatifs du travail et du capital influencent le choix des facteurs dans le processus productif.
- Plus généralement, les ratios de prix affectent des ratios d'éléments (ou leur part dans un total). Pour un niveau de consommation global donnée, la réduction du prix d'un bien augmente sa part.
- Une hausse de l'inflation réduit le pouvoir d'achat des économies réalisées jusqu'ici, appelant à un effort plus important de reconstitution.

- Et bien sûr les prix vont entrer dans la définition des variables à prix courants, quand elles sont séparées en volume et déflateur (éléments de la balance commerciale, salaires ...). Pour les éléments d'une somme, une évolution différente des déflateurs va modifier les parts à prix courants¹⁰.

Tout ceci est décrit par le graphique ci-dessous.

¹⁰ Par exemple, une augmentation de l'inflation locale va peser plus fortement sur les prix de la valeur ajoutée que sur le prix du capital (en grande partie importé). Cet effet va peser positivement sur le taux de profit.



10.2.2 LE COMPTE DES ENTREPRISES

Nous avons déjà traité de l'offre, la définition de l'adaptation des facteurs de production: le capital et l'emploi, à la capacité cible, ainsi que la décision sur les prix, sur la base généralement d'une maximisation à court terme des bénéfices.

Cela signifie que la plupart des équations restantes représenteront des définitions, décrivant le compte des entreprises sans faire appel à des éléments théoriques.

Il y a cependant deux exceptions:

- L'impôt sur les bénéfices, qui devrait être à nouveau calculé par application d'un taux à une base. La situation est cependant plus complexe que d'habitude, car:
 - le calcul des bénéfices dans un modèle est assez complexe, et tous les modèles ne sont en mesure de le faire. Parfois, il est nécessaire d'utiliser une variable approchée («proxy»), ce qui rend le taux apparent plus difficile à interpréter ;
 - le calendrier de l'impôt appelle à une équation dynamique, car il n'est pas généralement payé à la même période que les bénéfices associés (mais il peut y avoir une disposition demandant de payer immédiatement un impôt prévisionnel). Ainsi, une formule décrivant le mécanisme doit être mise en place ;
 - l'impôt sur des bénéfices négatifs n'est pas négatif, mais nul, introduisant un biais sur le taux apparent.
- Les dividendes versés par les entreprises, qui peuvent être estimés ou construits par une identité (en utilisant un taux de la même manière que les impôts). Encore une fois, il faut décider de la dynamique, car les dividendes suivent les bénéfices avec retard. En outre, le bénéficiaire des dividendes doit être identifié (le partage doit être fait entre les cinq agents habituels).

Bien sûr, la complexité des formulations (et même l'identification des éléments tels que les dividendes) dépend du rôle du modèle, selon qu'il est utilisé par des chercheurs tentant de répondre à des questions théoriques globales, ou par des conseillers en politique traitant en détail l'évolution du prochain budget de l'État.

10.2.2.1 Le compte des ménages

Au contraire, nous n'avons pas encore abordé le processus de décision des ménages.

Il comprend les éléments suivants:

- Les ménages tirent leurs revenus de plusieurs sources, dont les principales sont:
 - les salaires,
 - les revenus des entrepreneurs individuels,
 - les prestations sociales de toutes sortes,
 - les intérêts de prêts,
 - les dividendes,

- la location de logements à d'autres ménages (un service)¹¹.
- Ils utilisent ces recettes:
 - pour payer l'impôt sur le revenu,
 - pour consommer des biens et services divers,
 - pour épargner, en particulier en logement, mais aussi en dépôts, obligations, actions et biens (comme des tableaux ou des camping-cars).

Pour être considéré comme opérationnel, même un modèle à un seul produit doit utiliser un certain détail dans les revenus et les dépenses, car les processus économiques qui engendrent les premiers, et les conséquences des seconds, sont très différents les uns des autres.

Un autre principe de la modélisation: favoriser le détail qui permet d'identifier des comportements différents.

Fondamentalement

- Pour les revenus:
 - Les salaires versés par les entreprises sont le produit d'un taux de salaire moyen (venant du bloc de prix) par le nombre de travailleurs (venant du bloc de production).
 - Le nombre de fonctionnaires sera généralement exogène, mais pas leur taux de salaire, qui peut être différent de celui des entreprises.
 - Les salaires payés par les ménages (principalement aux domestiques) peuvent être identifiés ou non, selon le type de modèle.
 - Les prestations sociales sont généralement séparées en cinq catégories: la maladie, les allocations familiales, les indemnités de chômage, les indemnités d'invalidité et d'accidents de travail, les retraites. Il est clair que chacun de ces éléments dépend de l'inflation, mais à des degrés différents ;

¹¹ En fait, il serait étrange de considérer que si un ménage achète l'appartement qu'il louait, le service disparaît et le PIB diminue. Pour cette raison, les propriétaires de logements sont considérés par la comptabilité nationale comme se payant à eux-mêmes un loyer fictif.

- la plupart d'entre eux dépendent de la population, et souvent d'un type donné de population: par exemple le nombre d'enfants, de retraités, ou de chômeurs ;
- tous dépendent de l'activité économique, encore une fois d'une manière variable. Par exemple, les allocations de chômage diminuent avec le PIB, les accidents de travail augmentent, et le montant des retraites devrait être favorisé (en principe) par les recettes provenant des contributions qui les financent ;
- ils dépendent aussi évidemment d'une décision de l'État (le pouvoir d'achat de chaque retraite est maintenu).

Cela signifie qu'un modèle opérationnel devrait essayer de séparer ces éléments, afin de tenir compte de leurs différences de comportement.

De cette façon, le modèle montrera naturellement:

- la modification des prestations avec le nombre de bénéficiaires,
- la modification des prestations avec la décision.

Les intérêts seront décrits globalement, dans un paragraphe ultérieur. Soulignons seulement que pour les ménages les taux d'intérêt (prêts et emprunts) peuvent s'écarter des valeurs du marché grâce à l'intervention de l'État. En France, une quantité limitée de prestations d'épargne bénéficie d'un taux garanti et n'est pas imposable, et l'emprunt pour acheter un logement peut se faire à un taux réduit (0 % dans certains cas).

Les dividendes seront traités plus tard, avec le compte des entreprises.

En ce qui concerne les recettes venant du logement (loyers), leur rôle dans un modèle est limité, car il représente surtout un transfert de ménages à d'autres ménages. Pour les propriétaires de leur propre logement, c'est même un transfert fictif au sein d'un même ménage. Il faut cependant le prendre en compte, car il entre dans le PIB.

Il n'est pas nécessaire de tenir compte des éléments marginaux, tels que les gains de loterie, les héritages, les dons, les amendes ...

Enfin, on peut formaliser les transferts de l'étranger (ou vers l'étranger). Pour les pays en développement les envois de fonds peuvent représenter une part importante du revenu des ménages (plus d'un tiers du PIB du Tadjikistan). Pour un modèle à un seul pays ils devraient être exogènes, peut-être même en valeur (une exception notable à la règle générale).

Pour les dépenses:

- L'impôt sur le revenu doit être calculé comme un taux appliqué au revenu avant impôt, dont la valeur apparente sera calculée sur la période historique en divisant le montant par la base. Le modèle va ensuite obtenir l'impôt en prévision en appliquant ce taux exogène à la base simulée. Ceci pose le problème habituel de la dynamique temporelle: l'impôt peut être versé après que le revenu ait été obtenu (éventuellement en tenant compte d'un mécanisme de provision).
- En outre, l'application d'un taux moyen à l'ensemble des ménages peut être acceptable pour les prévisions (qui permettent à ce taux de changer avec le temps), mais moins pour les chocs adressés à une catégorie de ménages placée à une extrémité du spectre: dans un modèle macroéconomique traditionnel, une diminution de l'impôt sur les grandes fortunes ou une augmentation du Revenu Minimum de Solidarité (destiné aux catégories défavorisées), de même taille ex ante, aura les mêmes conséquences ex post¹². Pour éliminer cette erreur, une correction *ad hoc* doit être faite sur le taux d'épargne lui-même.

Ce problème apparaît dans la plupart des modèles, venant du fait que les outils permettant de le résoudre ne sont pas disponibles. Les comptes nationaux distinguent les entreprises selon les biens qu'elles produisent, mais pas les ménages selon une quelconque nomenclature, y compris le niveau de leurs revenus. Certaines enquêtes traitent le problème, et leurs enseignements pourraient être utilisés pour créer des données spécifiques. Cela signifie qu'une solution pourrait être trouvée, mais sans doute à un coût élevé. En fait, le même problème se pose si l'on veut séparer les entreprises non plus selon les secteurs, mais selon leur taille, sachant que les petites entreprises agissent différemment des grandes.

Une fois que le revenu disponible est connu, il reste à le séparer entre épargne et consommation, considérée dans son ensemble dans la plupart des modèles (pour les modèles multi-produits, la situation sera plus complexe).

La technique la plus courante consiste à calculer la consommation d'abord, comme un rapport au revenu, et l'épargne sous forme de résidu. Nous allons développer ceci par des estimations.

La consommation est généralement déterminée à prix constants (donc en pouvoir d'achat). Les déterminants habituels sont les suivants:

¹² Bien sûr, l'impact sur la consommation sera plus élevé si l'augmentation concerne les catégories défavorisées.

- Le niveau du revenu (mesuré également en pouvoir d'achat). La consommation augmente avec le revenu, mais sa part diminue (les catégories les plus défavorisées n'épargnent pas, et rappelez-vous que l'achat d'un logement est considéré comme de l'épargne).
- L'évolution récente du revenu. Les ménages mettent un certain temps à adapter leur comportement à une augmentation (ou diminution) de celui-ci. Et une hausse soudaine (surtout si elle est destinée à être permanente, comme une promotion) peut les amener à investir dans le logement, ce qui peut de fait réduire la consommation pendant un certain temps.
- L'inflation (l'effet «d'encaissements réelles»). Les économies réalisées jusqu'ici contiennent une part importante d'éléments monétaires (dépôts, obligations à taux fixes ...). L'inflation réduit leur pouvoir d'achat, ce qui oblige à les compléter par des économies supplémentaires. L'effort est proportionnel au niveau de l'inflation.
- Le taux de chômage. Pour les travailleurs salariés, une augmentation du risque de perdre leur emploi (mesuré plutôt par la variation courante que par la valeur¹³) conduit à épargner une plus grande part de leur revenu actuel, s'ils veulent optimiser leur utilité dans le temps.
- Le taux d'intérêt (à court terme): en général, les gens préfèrent satisfaire un besoin donné maintenant que plus tard. Mais cela a un coût, les intérêts qu'ils ont à payer. Plus le taux est bas, plus ils sont tentés par une consommation immédiate.

Cela est particulièrement vrai pour les biens durables: si un ménage veut regarder la télévision sur écran plat (et pense que, après son achat, dans toute sa vie il aura suffisamment de ressources pour se le permettre) la seule raison pour ne pas en acheter une dès maintenant (en augmentant sa satisfaction de façon permanente) est le coût actualisé, qui est réduit par une diminution des taux d'intérêt. Ce qu'il faut considérer n'est pas le coût du bien, mais le coût d'une acquisition plus précoce.

Si le bien est parfaitement durable, et peut être revendu à sa valeur d'origine à prix constants, les choses se passent exactement comme si le ménage était locataire du bien. Si le taux d'intérêt est divisé par 2, le «prix» du bien est divisé par 2.

Pour les biens non durables (on dit «fongibles»), la situation est différente. Le ménage a déjà optimisé sa consommation dans le temps. Si le taux d'intérêt change, il pourrait être tenté de consommer plus tôt, mais si l'utilité marginale du bien est en baisse rapide avec le

¹³ Bien que le taux réel joue également un rôle: une valeur plus élevée implique une rotation plus importante, et un risque élevé de participation à cette rotation.

niveau consommé, le mode de consommation ne sera pas beaucoup affecté. Une personne qui rêve de visiter les pyramides du Caire, et épargne à cet effet, pourra faire le voyage plus tôt, mais ne le fera pas.

Ce qui importe est le taux réel:

- Il permet de comparer les marchandises à prix constants.
- Si les ménages supposent que leurs revenus augmenteront avec l'inflation, leur volonté est d'optimiser en termes réels.

Une fois que la consommation est déterminée, l'épargne est calculée comme un résidu, et généralement comme un élément global. Cette option peut être discutée, car différents types d'épargne peuvent être supposés suivre des comportements différents.

En particulier, l'investissement immobilier est affecté négativement par le taux d'intérêt (il s'agit d'un taux spécifique, mais on peut supposer qu'il suit le taux global), tandis que l'épargne financière l'est positivement. L'achat d'un logement demande d'obtenir l'argent nécessaire auprès d'un autre agent, en contrepartie d'intérêts. L'achat d'une obligation représente un prêt à un autre agent qui pourra l'utiliser pour dépenser (peut-être acheter un bien durable), en échange d'intérêts, mais cette fois dans l'autre sens.

10.2.2.2 Le commerce extérieur

Dans un modèle à seul pays, le reste du monde est exogène.

Cela signifie que l'on considère seulement les influences du reste du monde sur le pays, et non l'inverse.

Bien sûr, ce n'est pas vraiment exact, même pour le plus petit des pays (ou à cet égard d'une région, d'une ville ou d'une personne): en augmentant votre consommation et votre production, vous créez une fraction d'un emploi, une petite quantité de revenu des ménages et donc une consommation.

Ce que nous considérons est que l'influence est assez faible pour que l'on puisse considérer que le coût de production (et d'utilisation) d'un modèle la décrivant soit trop élevé par rapport au gain sur l'exactitude des résultats. Ceci est certainement vrai pour des pays de taille faible ou moyenne comme la Lettonie ou la Bolivie, beaucoup moins pour les grands pays comme la France, et tout à fait faux pour les États-Unis ou l'Union européenne considérée comme un tout. Par exemple, lorsque nous avons utilisé (il y a quelques années) le modèle mondial MacSim pour une analyse de chocs, le multiplicateur keynésien pour la France pour l'an 2000 était de 1,3 si l'on considérait le modèle complet, mais seulement de 1,1 si l'on simulait le modèle français seul. Les rétroactions itératives des importations

allemandes en provenance de France, venant de l'augmentation des exportations allemandes, représentent la plus grande part de l'écart. Compte tenu de l'évolution du commerce mondial, la différence actuelle devrait être encore plus forte.

Cela signifie que les échanges du pays modélisé doivent être considérés du point de vue du pays lui-même:

- Les exportations sont la part de sa production de biens et services qui seront vendue par le pays au reste du monde.
- Les importations sont la part de la demande locale de biens et de services qui ne sera pas produite dans le pays, mais achetée au reste du monde.

Ces deux éléments seront calculés en utilisant la monnaie du pays. Cependant, l'utilisation des prix constants conduit à l'utilisation du taux de change de l'année de base, de sorte que la question de la monnaie n'est pas pertinente, étant traitée seulement par l'application un facteur constant¹⁴.

Toutefois, les éléments de commerce extérieur ayant le même caractère, leurs déterminants logiques seront les mêmes. La principale différence ne viendra de la taille relative des deux marchés (acheteur et vendeur) dans le fonctionnement de la négociation: l'importance du seul pays décrit (ou son PIB) sera toujours beaucoup plus faible que celle du reste du monde, même si c'est un peu moins évident si l'on modélise les États-Unis ou l'Union européenne dans son ensemble.

Ces éléments sont les suivants:

- La demande: pour qu'un pays vende un bien donné à un pays partenaire, la demande pour ce bien doit y être présente, une partie de cette demande doit être adressée au marché mondial, et la qualité des produits locaux doit être attractive pour le pays importateur. Par exemple, les exportations françaises de vin dépendront de la demande mondiale de vin, de la préférence naturelle des pays importateurs pour les vins étrangers (en commençant par leur statut de producteurs de vin), et pour le vin français en particulier.

¹⁴ Ce n'est vrai que si l'on considère un seul reste du monde, ou qu'on considère une monnaie unique. Nous reviendrons sur cet aspect plus tard.

Pour les importations, nous avons déjà vu que l'inclusion de la consommation intermédiaire dans l'équilibre offre-demande (avec la production d'un côté et la demande totale locale de l'autre) est un problème pour les modèles, car le niveau de cette consommation intermédiaire dépend du nombre d'étapes dans le processus de production. L'adoption d'un produit unique a jusqu'à maintenant éliminé la nécessité de considérer celle-ci. Mais les importations contiennent des biens intermédiaires, qu'ils représentent l'énergie (pétrole, gaz, et même électricité) ou des biens primaires (des bois non traités aux composants électroniques). Et ces biens intermédiaires sont nécessaires pour produire les biens exportés.

Une solution simple est de considérer comme fixe le ratio de la consommation intermédiaire à la valeur ajoutée. En observant les chiffres, on peut en effet observer que les «coefficients techniques», le nombre d'unités nécessaires pour produire une unité de PIB ou de valeur ajoutée, sont assez constants.

Nous pouvons donc considérer une demande composite, comme la somme de la demande finale elle-même, et d'une consommation intermédiaire fonction du PIB (ou plutôt de la valeur ajoutée, car les consommations intermédiaires sont déterminées hors TVA).

- La compétitivité prix: au moment de décider d'acheter ou non un produit à un producteur, local ou producteur étranger, un pays va comparer les prix associés. Et pour choisir entre les vendeurs potentiels, le pays importateur prendra en considération leur prix correspondant à une qualité donnée (rappelez-vous que les déflateurs considèrent des marchandises au même niveau de qualité, une augmentation de la qualité augmentant la valeur à prix constants).

Les critères statistiques montrent généralement qu'au niveau global le prix relatif joue un moins grand rôle lorsque l'acheteur doit choisir entre acheter des produits locaux ou étrangers, que quand il a le choix entre différents vendeurs étrangers. Ceci est assez logique: les produits locaux sont censés être conçus pour le marché local, et certains produits ne sont pas ou difficilement substituables (billets d'autobus locaux ou journaux).

Cela signifie dans notre cas que la sensibilité des exportations à la compétitivité prix devrait être supérieure à celle des importations. Les exportations dépendent de la demande sur le marché mondial, et une fois qu'un pays a décidé d'importer, le prix jouera un rôle plus important dans le choix du pays où il achète.

Enfin, bien sûr la mesure de la compétitivité doit utiliser des indices définis dans la même monnaie. Peu importe laquelle, car l'application du même taux de change aux deux éléments du rapport ne changera pas le rapport. Dans le cas des exportations, cela signifie que la mesure de leur indice de prix en monnaie locale demande d'utiliser pour le prix étranger dans les mêmes unités. Une fois le taux de change identifié, ce prix étranger sera donc endogène, comme le produit de deux hypothèses: le prix en devises étrangères, et le taux de

change (ou plutôt son déflateur). Il serait peut-être plus logique (et équivalent en pratique) d'examiner les deux prix en monnaie étrangère, le prix local étant corrigé par un taux de change symétrique.

- Les capacités disponibles

Le troisième élément est la possibilité de répondre à une demande supplémentaire, et donc la présence de capacités de production disponibles. La variable pertinente est naturellement le taux d'utilisation des capacités, indépendamment de la taille de l'économie.

Le choix de cette option n'est pas si simple, cependant. On pourrait faire valoir que, tant que le taux est inférieur à l'unité, la demande supplémentaire peut être satisfaite. Nous avons déjà montré que ce n'est pas vrai: la demande porte sur une large gamme de produits, que l'on ne peut en général pas remplacer par un autre, en particulier à court terme. Et certains produits peuvent voir leur capacité complètement saturée.

Laissez-nous vous expliquer le processus à nouveau, cette fois d'une manière tournée vers l'importation.

En fait, le taux moyen d'utilisation est basé sur une distribution de valeurs de taux, de zéro (espérons-le seulement dans quelques cas) à un (probablement en nombre important). Lorsque la demande mondiale augmente, elle s'adresse à un éventail d'entreprises, en intensité variable. Certaines de ces demandes seront adressées aux entreprises déjà incapables de fournir plus, et pour quelques autres elles vont leur faire atteindre ce niveau. La proportion des entreprises qui travaillent à pleine capacité va augmenter. Bien sûr, dans certains cas, un autre produit disponible peut représenter un substitut. Mais l'option la plus simple consiste à importer le même produit, car le produit manquant devrait être disponible quelque part dans le monde (peut-être à un prix plus élevé, mais cela sera traité par la compétitivité prix).

La part «manquante» de la demande augmente avec la part des entreprises qui ne peuvent pas augmenter leur production, après avoir atteint leur capacité.

Bien sûr, ce phénomène s'applique essentiellement à court terme, car les entreprises vont réagir en investissant, ce qui augmentera les capacités et réduira l'écart de production avec le temps. Mais ce processus peut être lent, même si une adaptation complète doit être obtenue dans le long terme.

Mais si on suit le raisonnement ci-dessus, on observe:

- Que plus grand est le pays, et plus faible sera la probabilité qu'une augmentation donnée absolue, mais aussi relative devra faire face à des problèmes d'approvisionnement locaux. Cette augmentation de la demande sera plus diversifiée, et les capacités disponibles seront plus diversifiées aussi¹⁵.
- Que, dans notre cas, le reste du monde n'est pas censé faire face à des problèmes d'approvisionnement, ce qui signifie que pour les importations et les exportations de notre pays seul le taux d'utilisation local doit être pris en compte.
- Et une dernière condition peut apparaître pour le pays exportateur. Si le monde a besoin d'un bien donné, les caractéristiques de sa production dans le pays doivent également être adaptées à la demande, qui peut être tout à fait spécifique. Par exemple, face à une augmentation de la demande étrangère d'automobiles, un pays pourra en proposer à un prix compétitif, mais le type de voitures dont il dispose ne conviendra pas. Cela peut concerner la taille, la sophistication, l'énergie consommée, les dispositifs de sécurité ... Bien sûr, on ne peut pas envisager d'entrer dans un tel détail dans un modèle, même si les données étaient disponibles (ce qui n'est pas le cas).

Malheureusement, trouver un élément statistique décrivant cette fonctionnalité est moins simple que pour les autres, en particulier pour un modèle à produit unique. L'idée la plus simple est d'utiliser l'âge du capital, en supposant que les processus de production récents seront mieux adaptés à la satisfaction de la demande¹⁶. Par exemple, une usine de construction automobile récente pourra suivre les tendances du marché dans la production de voitures plus petites, ou plus économes en énergie. L'âge du capital peut être déduit de la seule chronologie des investissements et des amortissements, si l'on considère que cette dépréciation s'applique identiquement à toutes les générations de capital, et que le capital disparaît brutalement un certain nombre d'années après sa mise en œuvre. Une autre hypothèse conduit à des formules plus complexes, mais gérables.

¹⁵ Ce ne serait pas le cas si la demande supplémentaire se faisait dans un bien spécifique.

¹⁶ En particulier de la demande étrangère car son rôle est croissant avec le temps, de sorte que son influence sur la nature de l'investissement sera de plus en plus élevée.

10.2.2.3 Le budget de l'État

Dans les modèles opérationnels, décrire complètement le budget de l'État est une exigence absolue.

Cela est vrai même si le modèle ne va pas être utilisé par les conseillers du gouvernement, mais plutôt par des experts en économie appliquée. L'objectif général de ces chercheurs est d'évaluer les conséquences pour l'économie des décisions du Gouvernement, des événements externes, ou des changements structurels, sous les conditions les plus probables ou un éventail de possibilités, et peut-être selon différentes formulations du modèle (par exemple différentes options sur les taux d'intérêt). L'approche peut être plus ou moins appliquée (les conseillers peuvent chercher à produire une image du prochain budget, qui sera présenté au Parlement, et les scientifiques vont essayer de voir comment l'adoption par la Banque centrale d'une règle de Taylor stabilisera l'économie), mais l'outil nécessaire est assez similaire.

Comme nous l'avons indiqué ci-dessus, la meilleure façon de définir les équations associées est de construire des identités, calculant un revenu ou une dépense endogène comme le produit d'une base endogène par un taux exogène. Les équations sont valables sur le passé, et le modélisateur sera responsable de (et autorisé à) l'établissement des hypothèses futures sur les taux. Bien sûr, il n'est pas nécessaire de maintenir ce taux constant, et on peut établir comme base de départ une tendance estimée. Mais la décision finale sera la sienne.

Cette technique répond à l'objection suivante: si l'on considère la TVA, même avec des taux légaux constants, le taux apparent va changer (positivement) avec la richesse des ménages, amenés à augmenter la part des produits fortement taxés dans leur consommation. Une solution consiste à établir une tendance, utilisée comme valeur de base, la décision consistant à s'écarter de cette tendance.

Si ces principes sont respectés, il sera possible de produire un tableau montrant l'évolution de tous les éléments du budget, en termes courants et en points de PIB, les deux présentations étant évidemment nécessaires.

Un autre principe important de la modélisation: si vous avez des difficultés à choisir entre des présentations possibles pour un concept donné (valeur à prix constants, aux prix courants, taux de croissance, ratio par rapport à une autre variable) il suffit de regarder comment ce concept est présenté dans les publications économiques (en privilégiant celles conçues pour le grand public). Ou d'attendre que vous ayez à utiliser les chiffres dans vos propres présentations, puis mesurer votre réaction et celle du public.

10.2.2.4 Les éléments financiers et monétaires

Dans tout modèle, il s'agit de la partie la plus variable et controversée. Les premiers modèles comportaient peu ou pas d'équations financières. Même à ce stade, le bloc financier peut être limité à la définition de quelques taux, et à leur impact sur le secteur réel (ces taux peuvent même être exogènes, généralement en termes réels). Au contraire, ce bloc peut être si développé que le but de la définition d'un secteur réel sera considéré comme le moyen de décrire les liens entre les éléments financiers, par exemple la création et l'origine de prêts supplémentaires si une baisse des taux d'intérêt tire les investissements vers le haut.

À notre avis, même un modèle orienté côté réel devrait inclure:

- un taux d'intérêt de base fixé par la Banque centrale du pays ;
- des taux à court et à long terme dans la monnaie du pays ;
- un taux moyen sur l'endettement net de la période ;
- un taux sur la dette actuelle, calculé à partir de la chronologie des taux passés, peut-être comme une fonction autorégressive ;
- un ou plusieurs taux étrangers, appliqués à des emprunts en devises à la fois dans le pays et dans le reste du monde ;
- les intérêts nets versés par l'ensemble des (cinq) agents, dans les deux monnaies pour les intérêts versés au reste du monde.

Un exemple de ce cadre sera présenté bientôt.

À partir de cette option de base, on peut considérer des sophistications:

- identifier la dette des agents (ou leurs avoirs financiers) ;
- la séparer en monnaies locales (dollars américains, peut-être euros pour les pays non UEM) ;
- la séparer par échéances, entre court terme et à long terme ;
- en outre, une ou plusieurs formes de la masse monétaire peuvent être formalisées.

La plupart de ces équations vont être établies comme des identités, basées sur les données disponibles ou des hypothèses. Des exceptions peuvent concerner:

- le taux de la Banque centrale, qui peut suivre une règle de Taylor, mais pas nécessairement. En fait, le même modèle devrait permettre plusieurs options (on peut établir une équation unique, utilisant un opérateur booléen) ;

- les taux à court et long terme peuvent inclure une prime de risque, en fonction par exemple du déficit budgétaire actuel ou de ses valeurs les plus récentes ;
- l'écart entre long et court terme pourrait dépendre des anticipations de croissance (ceci est plus vrai si les anticipations sont partiellement ou totalement rationnelles) et de la santé de l'économie locale.

Nous nous arrêterons ici, car le traitement questions financières est un objectif mineur pour cette publication.

10.2.3 LES PROGRAMMES EViews

Nous allons maintenant présenter des programmes EViews produisant les données et le cadre pour un modèle du type que nous venons de décrire. Nous utiliserons encore le cas français, mais nous allons passer à des données trimestrielles, ce qui rendra notre exemple plus opérationnel.

Remarque: tous les programmes présentés ici sont accessibles directement sur CD ou sur le site du modèle.

Tout comme dans l'exemple précédent, nous allons commencer par établir complètement les identités, tout en limitant la définition des équations de comportement à une déclaration du type de fonction que nous avons l'intention d'estimer (ce qui sera fait plus tard).

Comme nous l'avons indiqué précédemment, la production des données et la définition des équations peuvent être faites dans n'importe quel ordre, mais la création des groupes et le contrôle résiduel doivent attendre que les deux tâches soient accomplies.

Nous allons cependant séparer le programme qui crée les données de celui qui crée les équations (et produit la vérification résiduelle).

Les raisons en seront clarifiées plus tard, mais nous pouvons déjà affirmer la principale: pendant le processus de construction du modèle, les deux tâches vont évoluer, mais séparément la plupart du temps. La modification des spécifications du modèle sera beaucoup plus fréquente que le changement des données, et nous voulons éviter de répéter des opérations inutiles.

Il nous a semblé que notre présentation serait plus claire si nous commençons par la production du modèle. Cela pourrait être également la séquence la plus naturelle pour le constructeur: d'abord observer globalement quelles sont les données disponibles (sans s'intéresser en détail aux séries individuelles¹⁷), puis écrire les équations dans l'ordre, en observant plus précisément si chaque série est disponible (directement ou par l'intermédiaire d'une transformation), puis en créant le programme de construction des données. Il suffit de n'utiliser dans les spécifications que des séries dont l'utilisateur sait disposer, ou qu'il sait pouvoir construire.

¹⁷ On peut partir par exemple de la liste des variables présentées à la fin du livre.

10.2.3.1 La production du modèle

Voyons maintenant comment les éléments économiques que nous avons décrits peuvent être combinés dans un programme EViews, en suivant les éléments du cas le plus simple.

Dans ce programme, nous allons étendre les principes de la méthode que nous avons décrits plus haut, pour créer le cadre d'un petit modèle opérationnel. Comme dans l'exemple, nous nous arrêterons avant toute estimation. Ce que nous devons réaliser est un ensemble d'équations dans lesquelles:

- les identités sont entièrement définies ;
- les équations de comportement sont définies comme des, qui présenteront, en utilisant les notations spécifiques, la variable expliquée et ses éléments explicatifs.

Le texte contiendra essentiellement les déclarations EViews, et les commentaires présents dans le fichier lui-même (un guillemet apparaît au début de la ligne). Les commentaires supplémentaires apparaîtront sans guillemet, et utiliseront une police différente. Ils sont généralement associés à des fonctions réclamant de plus longues explications, qui auraient pris trop de place dans le programme.

```
'Nous commençons par la définition habituelle du répertoire  
cd "c:\program files\evIEWS8\book\fra_0"
```

```
' Nous décidons que les résultats des tests seront exportés  
' vers un fichier Rich Text Format nommé _mod_1.rtf
```

```
output(r) mod_1
```

```
' Nous fermons les fichiers d'entrée et de sortie  
' dans le cas où une version est déjà ouverte
```

```
close data_1  
close mod_1  
open data_1
```



```
wfsave mod_1
```

- ' Nous définissons la période échantillon comme le maximum disponible
- ' Les restrictions apparaîtront avec des estimations réelles

```
smpl 1960Q1 2005Q4
```

- ' Nous devons maintenant créer un nouveau modèle (vide)
- ' Astuce
- ' Nous donnons au modèle un nom commençant par «_»
- ' «_» est le premier élément apparaissant dans n'importe quel ordre alphabétique
- ' Comme les éléments de l'espace de travail sont affichés par ordre alphabétique
- ' Cela signifie qu'il va apparaître en premier dans l'affichage
- ' (Bien sûr parmi les éléments en utilisant la même astuce)
- ' Ceci est très utile pour les éléments fréquemment utilisés
- ' Dans le cas où le modèle existe déjà, on le détruit

```
delete(noerr) _mod_1
```

Cela signifie que nous repartons à zéro

Si nous ne le faisons pas, les équations précédentes seront conservées, même si les nouvelles équations redéfinissent les mêmes variables (y compris en utilisant des formulations identiques). Cela signifie que les variables seront définies deux fois, ce qui sera bien évidemment refusé par EViews, conduisant à la production d'un modèle erroné. Cela se produit parce que les identités sont ajoutées («appended») en tant que texte à la spécification du modèle. Jusqu'à EViews 8 (cf. plus haut), il n'était pas possible de supprimer directement une identité dans un modèle, bien que des techniques artificielles puissent réaliser cette tâche.

Nous ne les présenterons pas, car nous considérons beaucoup plus efficace de créer le modèle complet dans un programme unique. Nous avons déjà présenté les raisons de cette méthode: la clarté du texte du modèle, une meilleure identification des erreurs, des transferts facilités avec d'autres utilisateurs, une gestion des programmes plus aisée.

Cela n'est pas vrai pour les équations estimées, comme nous le verrons plus tard

```
model _mod_1
' on définit le scalaire f, pour la production des identités «de comportement»
scalar f
```

Nous commençons par le bloc de production

```
'-----
' Le bloc de production
'-----
' Le taux d'utilisation des capacités est le rapport entre le PIB effectif et le PIB potentiel
_mod_1.append UR = Q / CAP
' Le PIB équilibre l'offre et la demande
_mod_1.append GDP + M = FD + X
```

Pour obtenir la valeur ajoutée nous soustrayons le TVA à prix constants, produit du taux de TVA de l'année de base par la demande finale hors TVA, à prix constants.

Identifier la valeur ajoutée est nécessaire pour calculer les marges des entreprises et les déflateurs hors TVA.

Dans un modèle à prix constants, le taux de TVA sera plus ou moins constant, et la TVA proportionnelle au PIB. Mais ici, l'utilisation de la TVA aux prix courants, et d'un déflateur hors TVA, demande de l'enlever d'un dénominateur mesuré à prix constants. Nous utiliserons la valeur ajoutée Q dans la fonction de production (même si le PIB fournirait une explication similaire).

```
_mod_1.append Q=GDPM-r_vat0*FD/(1+r_vat0)
```

L'investissement dépend du PIB, du taux d'utilisation, de la rentabilité, du niveau précédent du capital, et du taux d'intérêt réel à long terme. Si l'on considère la substitution capital-travail, il dépendra également du coût relatif des facteurs.

Nous voyons que nous pouvons déjà présenter des décalages (pour le niveau précédent du capital) et des éléments de formulations (pour le taux d'intérêt réel).

```
_mod_1.append I=f*(Q+UR+RPROF+K(-1)+(IRL-100*@pchy(PC))+RELC)
```

Nous avons choisi de mesurer le capital à la fin de la période.

Il s'agit de la somme du capital précédent non déprécié et de l'investissement réalisé au cours de la période.

En fait, ce que nous considérons ici est l'investissement acheté au cours de la période. Il n'y a aucune garantie qu'il puisse encore être utilisé pour produire.

```
_mod_1.append K=K (-1)*(1-rdep)+I
```

‘ La variation des stocks dépendra du PIB

```
_mod_1.append IC=f*(Q)
```

‘ L'emploi par les entreprises aussi, et peut-être aussi du coût relatif des facteurs.

```
_mod_1.append LF=f*(Q+RELC)
```

‘ La productivité du travail est le ratio de la valeur ajoutée à l'emploi des entreprises

```
_mod_1.append PL=Q/LF
```

‘ La capacité de production dépend de l'emploi des entreprises et du niveau initial du capital

```
_mod_1.append CAP=f*(LE+K(-1))
```

‘ L'emploi total prend en compte les fonctionnaires

‘ En pratique nous devrions aussi identifier les salariés des ménages

‘ Comme les domestiques et les concierges

```
_mod_1.append LT=LF+lg
```

‘ La population active dépend de l'emploi et de la main-d'œuvre potentielle, dans la pratique la population en âge de travailler.

```

_mod_1.append POPAC=f*(LT+pop65)
' Les chômeurs sont définis par la population active sans emploi
_mod_1.append UN = POPAC-LT
' Le taux de chômage
_mod_1.append UN=POPAC-LT
'-----
' Le bloc de prix
'-----

' Le coût salarial unitaire est représenté par le salaire
' (plus les contributions des entreprises)
' nécessaire pour produire une unité de valeur ajoutée
_mod_1.append UWC=WR*(1+r_scf)/PL
' Le déflateur de la valeur ajoutée dépendra d'UWC
' et du taux d'utilisation
_mod_1.append PQ=f*(UWC+UR)
' Le prix à la production utilise le prix à la valeur ajoutée
' et le prix des consommations intermédiaires
' pondérés par leur influence
' Il sera nécessaire pour définir les prix des échanges extérieurs
_mod_1.append PP=(PQ+tc*PFDXT)/(1+tc)
' Le déflateur de la demande finale est le ratio de la demande
' à prix courants et constants
' Il solde l'offre-l'équilibre de la demande aux prix courants
' Tous les autres éléments sont ou seront calculés ailleurs

_mod_1.append PFD=(GDPMVAL+MVAL-XVAL)/(GDPM+M-X)

```

- ' Nous calculons également le déflateur hors TVA
- ' en inversant la relation avec PFD

_mod_1.append $PFDXT = PFD * (1 + r_vat0) / (1 + r_vat)$

- ' Les déflateurs des éléments de la demande dépendront de la valeur globale
- ' La relation peut être estimée, ou on peut appliquer un taux exogène

_mod_1.append $PC = f * (PFD)$

_mod_1.append $PI = f * (PFD)$

_mod_1.append $PIG = f * (PFD)$

- ' Le taux de salaire dépend de déflateurs
- ' PC à du point de vue des travailleurs
- ' PQ du point de vue des gestionnaires de l'entreprise
- ' et aussi de la productivité du travail et du chômage

_mod_1.append $WR = f * (PC + LP + UNR + PQ)$

- ' Les déflateurs des échanges en monnaie locale dépendent des prix à la production locaux et étrangers
- ' qui doivent être convertis en monnaie locale par le biais du taux de change

_mod_1.append $PX = f * (PP + ppx + ER)$

_mod_1.append $PM = f * (PP + ppx + ER)$

- ' Le taux de change peut être exogène
- ' ou dépendre de l'écart d'inflation (hypothèse Parité de pouvoir d'achat)

_mod_1.append $ER = f * (PP + ppx + ERX)$

- ' Le taux d'intérêt à court terme peut être exogène
- ' en termes nominaux ou réels
- ' ou dépendre de l'inflation et de l'écart entre production effective et potentielle
- ' (règle de Taylor)

```
_mod_1.append IRS=f*(IRS+(IRSR+100*@pchy(PC))+ (150*&pchy(PC)+50*(UR-urd)/urd))
```

- ' Le taux d'intérêt à long terme dépend d'une structure de retards de taux à court terme
- ' Peut-être avec un écart

```
_mod_1.append IRL=f*(IRS+spread)
```

- ' Le taux des nouveaux emprunts est une moyenne des taux à court et long terme

```
_mod_1.append IRL=f*(IRS+spread)
```

- ' Le taux sur les emprunts antérieurs dépend de la valeur précédente
- ' et du taux actuel
- ' en fonction de la vitesse de remboursement

```
_mod_1.append IR=f*(IRS+IRL)
```

'-----

- ' Le bloc ménages

'-----

- ' Les entreprises paient des salaires

_mod_1.append $WF=WR*LF$

- ' Le salaire des fonctionnaires
- ' pourrait être différent en moyenne de celui des entreprises

_mod_1.append $WG=WR*Ig$

- ' Le total des salaires

_mod_1.append $W=WF+WG$

- ' Les prestations sociales sont exogènes en pouvoir d'achat par tête
- ' Elles pourraient être séparées par risque

_mod_1.append $SOCB=socbr*PC*popt$

- ' Les revenus supplémentaires sont liés à la valeur ajoutée

_mod_1.append $REVQ=r_{revq}*QVAL$

- ' ou exogènes en pouvoir d'achat

_mod_1.append $REVX=r_{revx}*PFD$

- ' Des formulations plus complexes peuvent être utilisées
- ' Les cotisations sociales utilisent un taux exogène

_mod_1.append $SCW=r_{scw}*W$

' Le revenu des ménages est défini comme la somme de ses éléments

```
_mod_1.append HI=W-SCW+REVQ+REVM+SOCB
```

' L'impôt sur le revenu s'applique aux revenus de l'année précédente

' Encore une fois, le mécanisme peut être rendu plus complexe

' ou utiliser la valeur courante des revenus

' selon les règles appliquées par le pays

```
_mod_1.append ICT=r_ict*HI(-1)
```

' Maintenant, nous calculons le revenu disponible

```
_mod_1.append HDI=HI-ICT
```

' et également en pouvoir d'achat

```
_mod_1.append HRDI=HDI/PC
```

' La consommation des ménages dépend:

' * du revenu disponible réel

' * de l'inflation

' * du chômage

' * du taux d'intérêt à court terme

```
' _mod_1.append HCO=f*(HRDI+HCO(-1)/HRDI(-1)+PC+d(UNR)+(IRS-100*@pch(PC)))
```


' Une relation plus explicite pourrait être utilisée

' `_mod_1.append HCO=f*(HRDI+HCO(-1)/HRDI(-1)+PC+d(UNR)+(IRS-100*@pch(PC)))`

ou encore:

' `_mod_1.append Dlog(HCO) =f*(Dlog(HRDI)+Log(HCO(-1)/HRDI(-1)+Dlog(PC)+d(UNR) +(IRS-100*@pch(PC)))`

' Cela ne changera pas le diagnostic sur la structure du modèle

' mais indiquera le type de relation que l'on veut estimer

' à titre de rappel personnel ou pour une présentation à des partenaires

'-----

' Le bloc entreprises

'-----

' La valeur ajoutée à prix courants

`_mod_1.append QVAL=PQ*Q`

' La taxe à la valeur ajoutée

`_mod_1.append VAT=r_vat*PFDXT*FD/(1+r_vat0)`

' Le Produit intérieur brut

`_mod_1.append GDPMVAL=QVAL+VAT`

' et son déflateur

```
_mod_1.append PGDPM=GDPVAL/GDPM
```

' Les subventions sont proportionnelles à la valeur ajoutée

```
_mod_1.append SUBS=r_subs*QVAL
```

' Les marges sont calculées comme suit:

' Valeur ajoutée plus subventions moins autres impôts indirects

' et les salaires et cotisations de sécurité sociale des entreprises

```
_mod_1.append MARG=QVAL*(1+r_subs-r_oit)-WR*LF*(1+r_scf)
```

' Le taux de marges

```
_mod_1.append RMARG=MARG/QVAL
```

' L'impôt sur les bénéfices est basé sur les bénéfices passés

' Une règle plus complexe peut être appliquée

```
_mod_1.append IFP=(PROF(-1)+IFP(-1))*r_ifp
```

' Des bénéfices sont exclus:

' * Le revenu des ménages tiré de la production

' * L'impôt sur les bénéfices

' * Les intérêts nets versés

```
_mod_1.append PROF=MARG-REVQ-IFP-NIF
```

' Le taux de profit s'applique au capital au coût de renouvellement

_mod_1.append RPROF=PROF/(PI*K(-1))

' On peut aussi calculer le ratio des marges au capital

' Un concept plus stable et plus fiable

_mod_1.append RPROB=MARG/(PI*K(-1))

' La capacité de financement exclut les dépenses en

' * Investissement productif

' * Variation des stocks

' Car les marges comprennent la valeur ajoutée invendue

_mod_1.append FCAPF=PROF-PI*I-PFD*IC

' Les intérêts nets versés dépendent

' * de leur valeur antérieure

' * du taux sur les dettes passées

' * du taux sur les dettes nouvelles

' * de la capacité

_mod_1.append NIF=f*(NIF(-1)+IRM+IR+FCAPF)I

'-----

' Le bloc du commerce extérieur

'-----

' Le prix à l'importation, y compris les droits de douane

```
_mod_1.append PMT=PM*(1+r_tar)/(1+r_tar0)
```

- ' La compétitivité des prix à l'importation compare
- ' * Le prix à la production locale
- ' * Le prix à l'importation, y compris les droits de douane

```
_mod_1.append COMPM=PMT/PP
```

- ' La demande finale à prix constants est la somme de ses composantes
- ' y compris un résidu proportionnel au PIB résiduelle
- ' (une hypothèse raisonnable)
- ' L'investissement en logements pourrait être identifié

```
_mod_1.append FD=COH+I+IC+CG+IG+fdxr*Q
```

- ' Les importations dépendent
- ' * de la demande finale et la demande intermédiaire
- ' (proportionnelle à la valeur ajoutée)
- ' * du taux d'utilisation des capacités
- ' * de la compétitivité prix

```
_mod_1.append M=f*(FD+Q+UR+COMPM)
```

- ' La compétitivité-prix des exportations compare
- ' * Le prix de la production à l'étranger
- ' * Le prix à l'exportation local, y compris les droits de douane
- ' Le taux de change corrige l'écart de change

```
_mod_1.append COMPX=PX*(1+r_tarx)/(1+r_tarx0)/(PPX*ER)
```

- ' Les exportations dépendent
- ' * de la demande mondiale (à la fois finale et intermédiaire)
- ' * du taux d'utilisation des capacités
- ' * de la compétitivité prix

```
_mod_1.append X=f*(WD+UR+COMPX)
```

- ' Les flux commerciaux sont calculés en termes courants

```
_mod_1.append MVAL=PM*M
```

```
_mod_1.append XVAL=PX*X
```

- ' Les taux d'importation-exportation sont calculés:
- ' * à prix courants
- ' * à prix constants
- ' * pour les déflateurs

```
_mod_1.append RCVAL=XVAL/MVAL
```

```
_mod_1.append RCVOL=X/M
```

```
_mod_1.append TTRAD=PX/PM
```

- ' La balance commerciale

```
_mod_1.append TRB=XVAL-MVAL
```

- ' Les intérêts nets versés au reste du monde

```
' sont séparés selon la monnaie.  
' Ils dépendent:  
' * de leur valeur antérieure  
' * du taux sur les dettes passées  
' * du taux sur les dettes nouvelles  
' Le taux international est utilisé pour des intérêts  
' en monnaie étrangère  
' * de la balance commerciale  
' * du taux de change de la dette en monnaie étrangère  
  
_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  
  
' La capacité de financement  
  
_mod_1.append FCAPX=TRB-NIX  
  
'-----  
' Le budget de l'État  
'-----  
' La plupart de ses éléments ont déjà été calculés  
' comme transferts entre l'État et un autre agent  
' Les cotisations sociales payées par les entreprises  
  
_mod_1.append SCF=r_scf*Wf  
  
' Les autres impôts indirects
```

_mod_1.append OIT=r_oit*QVAL

‘ Les droits de douane

_mod_1.append TAR=r_tar*MVAL

Les cotisations sociales payées par l'État

_mod_1.append SCG=R_SCG*WG

‘ Le total des revenus

_mod_1.append REVG=SCF+SCG+SCW+OIT+IFP+ICT+VAT+TAR+r_rev*QVAL

‘ L'investissement public à prix courants

_mod_1.append IGV=IG*PIG

‘ La consommation publique à prix courants

_mod_1.append CGV=CG*PFD

‘ La demande du gouvernement à prix courants

_mod_1.append FDGV=CGV+IGV

Les Intérêts nets versés par l'État dépendent

- ' * de leur valeur antérieure
- ' * du taux sur les dettes passées
- ' * du taux sur les dettes nouvelles
- ' * de la capacité de financement

_mod_1.append $NIG=NIG(-1)*IRM/IRM(-1)-IR/100*FCAPG$

- ' Les dépenses totales

_mod_1.append $EXPG=FDGV+WG+SUBS+SOCB+NIG+SCG+r_expg*QVAL$

- ' Le solde budgétaire

_mod_1.append $FCAPG=REVG-EXPG$

- ' Le solde budgétaire en points de PIB

_mod_1.append $FCAPGP=100*FCAPG/GDPMVAL$

- ' Le PIB total
- ' La somme du PIB marchand et non marchand
- ' Assimilé au coût salarial total de l'État

_mod_1.append $GDPVAL = GDPMVAL+WG+SCG$

'-----

- ' Fin de la spécification du modèle

'-----

10.2.3.2 La production des données: l'exemple de la base OCDE

Nous allons maintenant présenter un programme créant les données nécessaires au modèle ci-dessus. Nous avons utilisé un cas très simple, dans lequel le constructeur du modèle a accès aux «Perspectives économiques» de l'OCDE, un fichier contenant environ 5 000 séries trimestrielles décrivant l'économie mondiale. Chaque pays de l'OCDE y est décrit individuellement, et d'autres pays importants (comme la Chine) ou des zones (comme l'Amérique latine) en utilisant moins de détails. Cependant, les définitions varient légèrement d'un pays à l'autre, et bien que le même ensemble de séries est toujours techniquement présent, certaines d'entre elles ne contiennent aucune valeur, ou très peu. Par exemple, la notion d'épargne n'est pas toujours la même, et le stock de capital n'est pas toujours disponible.

Comme nous le verrons, cette base contient pour la France toutes les données dont nous avons besoin, à une ou deux exceptions près.

De toute évidence, ce programme est tout à fait spécifique, et un utilisateur partant d'une autre base (comme celle du FMI) devra en construire un différent. Mais:

- l'utilisateur peut partir de cette base OCDE, une source très fréquemment utilisée ;
- certaines des tâches (création du fichier de travail, production d'une tendance ...) seront présentes dans tous les cas ;

et en fait, le programme ne changera pas beaucoup. La plupart des variables accessibles dans la base de l'OCDE le seront aussi dans n'importe quelle base de ce type, et il suffira simplement de remplacer les noms «type OCDE», en modifiant quelques-uns des concepts si nécessaire.

Les seuls problèmes supplémentaires pour l'utilisateur potentiel du programme proviendront:

- de données non disponibles qui doivent être estimées, ou inventées ;
- de données additionnelles requises par l'introduction de nouveaux éléments dans le modèle.

```
'=====
```

```
' Un exemple de transfert de données
```

```
'=====
```

```
' Ce programme partira de données originales françaises
```

```
' fournie par les «Perspectives Economiques» de l'OCDE
' et nommées fra_*
' le préfixe utilisé par l'OCDE pour identifier les statistiques françaises
' Nous décidons du répertoire
' Cet ordre n'est généralement pas nécessaire
' sauf si l'on travaille sur plusieurs projets
' ou gère plusieurs répertoires pour le même projet
' Il garantit que des versions d'essai ne détruisent pas les informations officielles

cd "c:\program files\Eviews5\Book\fra_cf"

' Il va créer les données de notre modèle
' avec le préfixe f_*
'-----
' Cette technique peut être utilisée avec n'importe quelle source
' Pour la série d'origine, utilisez le même préfixe
' Les résultats seront créés avec n'importe quel préfixe différent
' Dans le meilleur des cas, si l'ensemble disponible est le même (ou plus important) que l'ensemble de l'OCDE
' Il suffit de remplacer les noms de l'OCDE par les nouveaux
'-----
' Nous fermons le fichier original fra_1
' contenant les seules données fra_*
' et quelques séries mondiales de l'OCDE nommées OECD_*
' Nous fermons également le fichier qui recevra les données françaises
' dans le cas où il est déjà ouvert
' Avoir deux versions d'un même fichier ouvertes en mémoire est très dangereux ...

close fra_1
```

```
close data_1
```

- ' Nous ouvrons le fichier d'origine (actuellement fermé)
- ' et l'enregistrons sous le nom data_1 pour les données françaises

```
open fra_1
```

```
save data_1
```

- ' Maintenant, le fichier ne doit contenir que des données originales
- ' appelées fra_*
- ' Nous supprimons toute série existante f_*
- ' au cas où elles existeraient, ce qui ne devrait pas arriver ...

```
delete f_*
```

- ' Nous devons faire une hypothèse sur la répartition des impôts indirects
- ' entre la TVA et les autres impôts indirects
- ' comme l'OCDE ne fournit qu'une variable globale
- ' P_oit = hypothèse sur la part des oit dans les impôts indirects

```
scalar f_p_oit=0.2
```

- ' Nous créons une tendance temporelle
- ' avec la valeur de l'année pour le premier trimestre
- ' à laquelle on ajoute 0,25 pour chaque trimestre suivant de l'année
- ' 1994: 1994T1
- ' 1994,25: 1994T2
- ' 1994,50: 1994T3

' 1994,75: 1994T4
' 1995,00: 1995T1
' Ceci sera très utile pour:
' Créer des tendances temporelles annuelles
' Remplacer les variables muettes réelles par des expressions logiques en utilisant des conditions
' Ceci sera beaucoup plus facile à gérer
' comme nous le verrons

```
smpl 1962Q1 1962Q1  
genr f_t=1962  
smpl 1962Q2 2004Q4  
genr f_t=f_t(-1)+0.25
```

' Maintenant, nous commençons par l'équilibre offre-demande

```
smpl 1962Q1 2004Q4  
genr f_gdpval=fra_gdp
```

' Le Produit intérieur brut à prix courants

```
genr f_gdp=fra_GDPV
```

'= Produit intérieur brut à prix constants

```
»-----  
' L'offre  
»-----
```

' Le modèle sépare PIB marchand et non marchand

genr f_gdpmval=fra_gdp-fra_cgw

'= Produit intérieur brut à prix courants – les coûts salariaux de l'État à prix courants

genr f_pcog=fra_cg/fra_cgw

'= Consommation totale du gouvernement à prix courants / = consommation totale du gouvernement à prix constants

genr f_gdpm=fra_GDPV-fra_CGw/f_pcog

genr f_pgdpm=f_gdpmval/f_gdpm

' Le commerce à prix courants et constants, les déflateurs associés

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

' Le déflateur de la demande finale (y compris la TVA et autres impôts indirects)

genr f_pfd = (f_GDPmVAL + f_MVAL-f_XVAL) / (+ f_GDPm f_m-f_X)

genr f_fd = fra_TDDV-fra_CGw/f_pcog

' Identification des impôts indirects

' Cette séparation est importante, car la TVA s'applique seulement à la demande finale (y compris importée)

```

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)
scalar f_r_vat0=@elem(f_r_vat,"1995")
genr f_r_oit=f_oit/(f_gdpmval-f_vat-f_oit)
scalar f_r_oit0=@elem(f_r_oit,"1995")
genr f_pfdxt=f_pfd*(1+f_r_vat0)/(1+f_r_vat)

```

' La valeur ajoutée hors TVA (mais y compris 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, taux d'utilisation, capacité

```

genr f_k=fra_KBV
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

```

' La demande

' Les consommations intermédiaires
' non identifiées dans la base de données de l'OCDE

' nous fixons le taux à l'unité (une valeur courante)

genr f_tc=1

genr f_ic=f_tc*f_q

' Le prix à la production

genr f_pp=(f_qval+f_ic*f_pfdxt)/(f_q+f_ic)

' Les éléments de la demande et leurs déflateurs

' La consommation des ménages '

genr f_coh=fra_CPV

genr f_pcoh=fra_cp/fra_cpv

' L'investissement et son taux d'amortissement

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)

' L'investissement en logements

genr f_hih=fra_IHV

' L'investissement public (à l'exception des entreprises d'État)

genr f_ig=fra_IGV

genr f_igv=fra_IG
genr f_pig=f_igv/f_ig

‘ La consommation des administrations publiques (à l'exception des entreprises d'État)

genr f_cogv=fra_cg-fra_cgw
genr f_cog=(fra_cg-fra_cgw)/f_pcog

‘ La demande du gouvernement

genr f_fdgv=f_cogv+f_igv
genr f_fdg=f_cog+f_ig

‘ Les variation de stocks

genr f_ci=fra_iskv

‘ Les ratios donnant les déflateurs individuels de la demande

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

Le taux de salaire

genr f_wr=fra_WAGE/(fra_ET-fra_ES)
genr f_r_scf = fra_WSSS/fra_WAGE -1

' Un résidu montrant si la décomposition de la demande finale est correcte

$\text{genr } f_{\text{dxr}} = (f_{\text{fd}} - f_{\text{coh}} - f_{\text{i}} - f_{\text{hih}} - f_{\text{ic}} - f_{\text{fdg}}) / f_{\text{q}}$

' L'emploi, le chômage, la population

' L'emploi

$\text{genr } f_{\text{lt}} = \text{fra}_{\text{ET}}$
 $\text{genr } f_{\text{lg}} = \text{fra}_{\text{EG}}$
 $\text{genr } f_{\text{lf}} = \text{fra}_{\text{ETB}}$

' La productivité du travail

$\text{genr } f_{\text{pl}} = f_{\text{q}} / f_{\text{lf}}$

' Les salaires et coûts salariaux

$\text{genr } f_{\text{w}} = f_{\text{wr}} * f_{\text{lt}}$
 $\text{genr } f_{\text{wf}} = f_{\text{lf}} * f_{\text{wr}}$
 $\text{genr } f_{\text{uwc}} = f_{\text{wr}} * (1 + f_{\text{r_scf}}) / f_{\text{pl}}$

' Le taux de chômage

$\text{genr } f_{\text{un}} = \text{fra}_{\text{UN}}$
 $\text{genr } f_{\text{unr}} = \text{fra}_{\text{UN}} / (\text{fra}_{\text{ET}} + \text{fra}_{\text{UN}})$

‘ Les populations

genr f_pop=fra_POPT
genr f_pop65=fra_POPT
genr f_popac=f_lt+f_un

‘ Éléments financiers

‘ Le taux de change

genr f_er=1/(fra_EXCHEB/@elem(fra_EXCHEB,"1995"))

‘ Les taux d'intérêt nominaux

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

‘ Les taux d'intérêt réels et les suppléments exogènes

‘ par rapport à la formule adoptée

‘ Ici, nous prenons en compte les formules utilisées pour le court terme

- ' et le taux d'intérêt moyen sur les dettes passées
- ' Leur logique sera expliquée plus tard

```

genr f_irr=f_ir-100*@pchy(f_pcoh)
genr f_irl_ec=0
genr f_irsr=f_irs-100*@pchy(f_pcoh)
genr fIRST=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)

```

- ' Le coût relatif (pour la substitution entre travail et 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

```

- ' Les ménages

- ' Éléments globaux

```

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)

```

- ' Éléments individuels

```
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
```

```
'-----  
' Le commerce extérieur  
'-----
```

```
' Les droits de douane
```

```
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")
```

```
' Les éléments étrangers
```

```
genr f_wd = fra_XMVMKT  
genr f_ppx=OECD_PGDP
```

‘ La compétitivité

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)

‘ Les ratios et les soldes

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

‘ Les intérêts payés

‘ Dans les données de l'OCDE, les intérêts versés à l'entreprise ne sont pas identifiés

‘ Nous appliquons les formules utilisées par le modèle

‘ Elles seront expliquées plus tard

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_fcpx=f_trb-f_nix  
simpl 1962Q1 2004Q4
```

```
'-----  
' Le budget de l'État  
'-----
```

```
' Le revenu
```

```
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_revfg=f_ict+f_oit+f_vat+f_scf+f_scg+f_tar+f_scw+f_ifp  
genr f_r_revfg=0  
genr f_recg=0  
genr f_r_recg=0
```

```
' Les dépenses
```

```
genr f_socbr = fra_TRRH /f_pcoh/f_pop  
genr f_subs=fra_TSUB  
genr f_r_subs = fra_TSUB / (f_qval)
```

```
' Les taux d'intérêt et les soldes
```

```
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_revq-f_fcapg
genr f_r_expg=(f_expg-(f_fdg+f_wg+f_scg+f_nig+f_socb+f_subs))/f_qval

```

‘ Le solde en points de PIB

```
genr f_fcagp=100*f_fcapg/f_gdpval
```

‘ Le compte des entreprises

‘ Les marges

```

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

```

Les intérêts payés, les bénéfices et les soldes

```

genr f_prof=fra_PROF
smpl 1978Q1 2004Q4
genr f_fcagf=-fra_NLB
smpl 1978Q1 1978Q1
genr f_nif=f_ir*f_fcagf/11
smpl 1978Q2 2004Q4
genr f_nif=f_nif(-1)*f_irm/f_irm(-1)-f_ir/400*f_fcagf
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_fcacf_er=(f_fcacf-(f_prof-f_pi*f_i-f_pfd*f_ci))/f_qval

```

```

-----
'  Les facteurs de relaxation
-----

```

```

genr f_relax_q=1
genr f_relax_pfd=1
save data_1

```

10.2.3.3 La création du modèle et la vérification de la cohérence équations - données

Ces deux tâches peuvent être lancées que lorsque le modèle a été défini, et les données créées.

La création des groupes nécessite non seulement le modèle, mais aussi les données, car les groupes EVIEWS ne peuvent être construits qu'à partir de séries (existantes).

Ce programme peut être exécuté après les deux tâches, dans la même session. C'est pourquoi nous n'allons pas introduire les déclarations habituelles spécifiant le répertoire et le fichier de travail à ouvrir.

Le programme est un modèle indépendant. Il

- crée une liste de variables endogènes et exogènes ;
- résout deux fois les équations du modèle séparément les unes des autres, pour différentes valeurs de la grandeur scalaire «f» ;
- sépare les endogènes en identités et comportements en utilisant le fait que ces dernières contiennent le scalaire «f» ;

- calcule les différences en niveau et relatives entre les valeurs historiques et les résultats de l'une des simulations ci-dessus.

```

=====
' Création des groupes
=====
' Nous créons des groupes pour les endogènes et exogènes
' La déclaration «makegroup»
' * S'applique au modèle «_mod_1»
' * Crée un groupe d'éléments endogènes (@endog) ou exogènes (@exog)
' * Avec comme noms _g_vendo et _g_vexog
' Les modificateurs spécifient les éléments considérés lors de la création des groupes
' «n» signifie que nous ne voulons pas des éléments suffixés
' (par le suffixe courant, en fait «_c»)
' «a» signifie que nous voulons les noms réels (sans suffixe)
' Cette opération remplace les étranges options par défaut d'EViews

_mod_1.makegroup (a,n) _g_vendo @endog
_mod_1.makegroup (a,n) _g_vexog @exog

' Maintenant, nous voulons séparer les identités et les variables comportementales
' Nous allons utiliser une astuce
' Les futures équations de comportement utilisent actuellement le scalaire «f»
' Nous résolvons le modèle à deux reprises pour des valeurs différentes de «f»
' En utilisant des suffixes différents pour la solution
' Nous avons besoin de mettre à jour le modèle
' Pour que la modification du paramètre soit prise en compte

```

```
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
```

‘ Nous créons deux groupes vides

```
group _g_vbeha
group _g_viden
```

‘ Ensuite, nous examinons tour à tour chaque variable endogène

‘ `_g_vendo.@count` est le nombre d'éléments de `_g_vendo`

‘ `_g_vendo.@Seriesname (!i)` est le nom de l'élément `i` dans le groupe

‘ Nous considérons la variable construite à partir de ce nom et les deux suffixes " `_c` " et " `_d` "

‘ Si les valeurs sont différentes, alors la variable est comportementale

‘ Nous testons la condition sur une période donnée

‘ (Mais pas l'année de référence car il y a une chance que la valeur de droite donne zéro

‘ Dans ce cas, la méthode ne fonctionne pas)

‘ Si la condition est vraie nous ajoutons l'élément à la liste des éléments comportementaux

‘ sinon à la liste des éléments d'identité

‘ «endif» ferme la condition

‘ «next» ferme la boucle

‘ Les crochets «{» et «}» délimitent le paramètre

‘ Ils sont éliminés après le remplacement

```

for li=1 to _g_vendo.@count
%1=_g_vendo.@seriesname(li)
scalar {%1}_eq=(@elem{%1}_c,"2000Q1")<>@elem{%1}_d,"2000Q1")
if {%1}_eq=1 then
_g_vbeha.add {%1}
else
_g_viden.add {%1}
endif
next

```

```

'=====

```

```

' Le calcul des résidus

```

```

'=====

```

```

' Maintenant, nous vérifions les résidus

```

```

' Nous définissons le suffixe comme « _c »

```

```

_mod_1.append assign @all _c

```

```

' Nous réduisons l'échantillon aux périodes

```

```

' qui peuvent être calculées pour chaque équation

```

```

' compte tenu de la présence de retards

```

```

smpl 1995Q2 2004Q4

```

```

' Mais seulement pour les identités

```

```

' Toutefois, nous avons pu calculer les comportements

```

```

' Au cours de la période spécifiée

```

```

' Cela signifie que nous serons en mesure de les estimer

```

- ' Les $dc_{\{ \%1 \}}$ sont les différences absolues
- ' Les $pc_{\{ \%1 \}}$ les différences relatives
- ' Pour éviter une division par zéro, nous utilisons une astuce
- ' Nous ajoutons à la variable une valeur booléenne
- ' qui teste si la variable est égale à zéro
- ' Si elle est vraie nous obtenons 0 divisé par 1 = 0
- ' Si elle est fausse le calcul n'est pas affecté
- ' Cela signifie que nous avons simplement modifié le résultat des divisions,
- ' de sorte que la division de zéro par zéro donne maintenant zéro

```

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

```

10.2.3.4 L'estimation des équations: exemples

Nous allons maintenant appliquer les principes ci-dessus à l'estimation des équations. Fondamentalement, cela signifie que nous remplacerons dans le programme précédent les identités déclarant les intentions du modélisateur par de véritables équations de comportement.

Nous ne présentons pas ici le nouveau programme de création de modèle, associé à ce processus. Le programme est disponible sous forme de fichier programme («prg») sur le site du modèle. Il comprend un résumé des observations que nous allons faire.

Bien sûr, les résultats que nous obtiendrons sont spécifiques au cas français. Ils ne doivent être considérés que comme un exemple (qui fonctionne), et vont probablement se révéler beaucoup moins directement utiles que les programmes précédents.

On observera que nous ne respectons pas toujours une méthodologie pure et propre. Notre opinion est que nous nous comportons plutôt mieux que la plupart des modélisateurs. Et de toute façon, plutôt qu'un objet parfait, ce que nous proposons ici représente un exemple

de la méthodologie générale utilisée par les constructeurs de modèles, y compris les écarts par rapport au chemin «politiquement correct».

Par exemple, nous n'allons utiliser la cointégration qu'une seule fois, et construire la plupart des modèles à correction d'erreur en une seule étape. Vous pouvez deviner (à juste titre) que nous n'avons pas pu trouver une formulation respectant les conditions requises, soit parce que l'un des éléments était stationnaire, que la cointégration n'était pas présente, ou que les coefficients de l'équation de cointégration n'étaient pas significatifs ou satisfaisants.

Nous allons traiter les équations dans l'ordre, en considérant deux cas pour la fonction de production: facteurs complémentaires et Cobb-Douglas. Le choix aura une incidence sur d'autres éléments, car l'estimation d'une capacité de production différente modifiera le taux d'utilisation, une variable explicative dans plusieurs équations.

10.2.3.4.1 Une fonction de production à facteurs complémentaires

Traisons d'abord le cas des facteurs complémentaires. Le cadre est très similaire à celui que nous avons vu précédemment. L'investissement et l'emploi sont estimés séparément.

10.2.3.4.1.1 ESTIMATION DE L'INVESTISSEMENT

Comme dans le modèle réduit, la capacité sera définie par le capital.

Mais maintenant, nous allons estimer le capital comme variable de décision, et déduire les investissements en inversant l'équation comptable définissant précédemment le capital, selon le niveau qui permet d'atteindre une valeur donnée compte tenu de la dépréciation:

$$I_t = K_t - K_{t-1} \cdot (1 - dr_t)$$

L'équation du capital est la suivante:

$$\Delta \text{Log}(K_t) = a \cdot \Delta \text{Log}(K_{t-1}) + b \cdot (0.5 \cdot \text{Log}(Q_t / Q_{t-4}) - \text{Log}(UR_t)) + c \cdot 1/3 \cdot \sum_{i=0}^2 RPROB_{t-i} + d \cdot (t - 2005) + e$$

En termes EViews:

```
equation _eq_k.ls(p) dlog(k)=c_k(1)*dlog(k(-1))+0.50*c_k(2)*(q-q(-4))/q(-4)
-c_k(2)*(1-ur)/ur+c_k(4)*@movav(rprob,3)+c_k(5)*((t-2004)*(t<=2004))+c_k(6)+k_ec
```

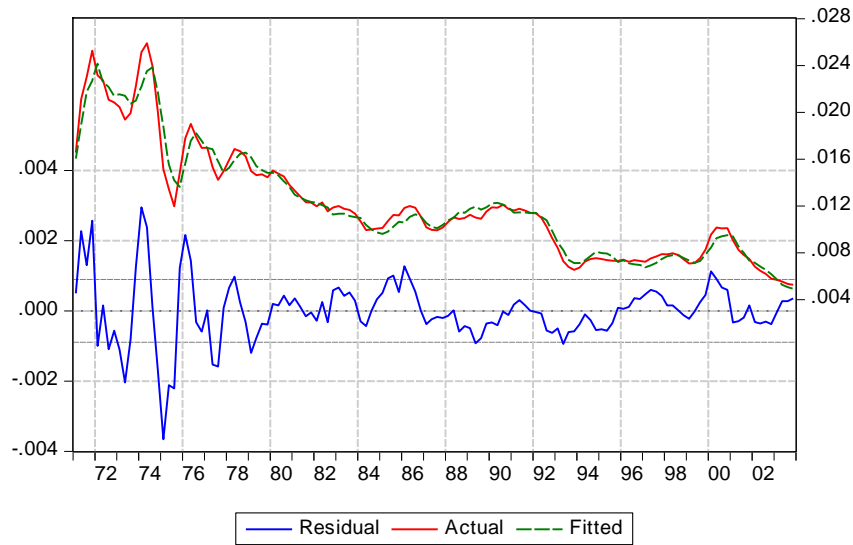
Dans notre équation, la variation du capital proviendra de:

- la variation de la valeur ajoutée au cours de la dernière année, qui représentera les anticipations de croissance future des entreprises ;
- l'adaptation du taux d'utilisation à sa cible (unitaire) ;
- la rentabilité attendue du capital, représentée par la moyenne des taux de profit des trois derniers trimestres.

Nous avons également dû introduire une tendance temporelle, que nous neutraliserons à partir de la fin de la période échantillon.

Dependent Variable: DLOG(K)
 Method: Least Squares
 Date: 11/07/12 Time: 18:51
 Sample (adjusted): 1971Q1 2003Q4
 Included observations: 132 after adjustments
 DLOG(K)=C_K(1)*DLOG(K(-1))+0.50*C_K(2)*(Q-Q(-4))/Q(-4)-C_K(2)*(URD-UR)/UR+C_K(4)*@MOVAV(RPROB,3)+C_K(5)*((T-2005)*(T<=2005))+C_K(6)

	Coefficient	Std. Error	t-Statistic	Prob.
C_K(1)	0.602302	0.046208	13.03471	0.0000
C_K(2)	0.019733	0.004273	4.618419	0.0000
C_K(4)	0.011318	0.002696	4.198355	0.0001
C_K(5)	-0.000117	1.85E-05	-6.338253	0.0000
C_K(6)	0.000221	0.000460	0.480779	0.6315
R-squared	0.966687	Mean dependent var		0.012048
Adjusted R-squared	0.965638	S.D. dependent var		0.004818
S.E. of regression	0.000893	Akaike info criterion		-11.16664
Sum squared resid	0.000101	Schwarz criterion		-11.05745
Log likelihood	741.9985	Hannan-Quinn criter.		-11.12227
F-statistic	921.3380	Durbin-Watson stat		0.788728
Prob(F-statistic)	0.000000			



Les résultats sont satisfaisants même si nous avons dû égaliser les coefficients de l'accélérateur et du taux d'utilisation.

10.2.3.4.1.2 L'ESTIMATION DE L'EMPLOI

Encore une fois, nous allons utiliser le même cadre que dans le petit modèle. Nous commençons par définir l'évolution de la productivité du travail:

```
smpl 1962Q1 2003Q4
coef(10) c_plt
```

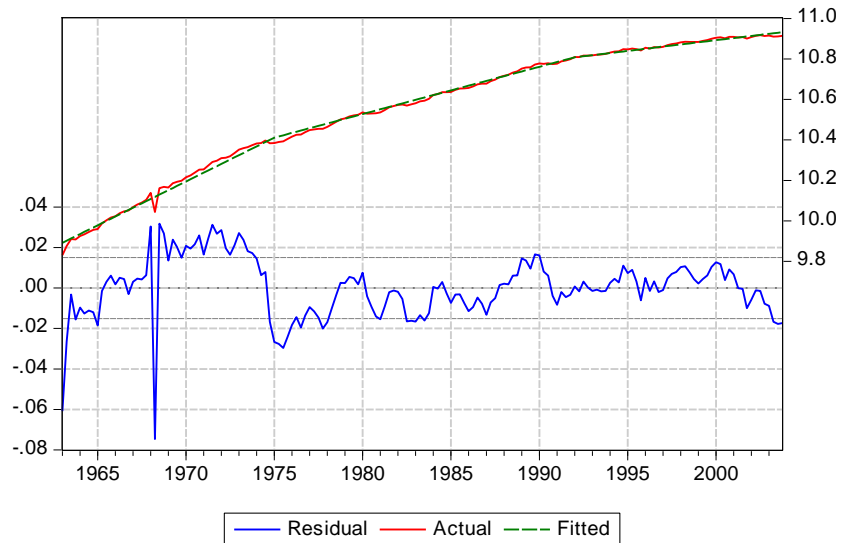


```
equation _eq_plt.ls(p) LOG(Q/LF) = c_plt(1) + (t-2004)*(c_plt(2)*(t<=2004)
+4*log((1+txq)/(1+txn))*(t>2004))+c_plt(3)*((T<=1975)*(T-1975)) + c_plt(4)*((T<=1992)*(T-1992))
_eq_plt.resids(p)
close _eq_plt
smpl 1962Q1 2004Q4
genr log(plt) = c_plt(1) + (t-2004)*(c_plt(2)*(t<=2004)+4*log((1+txq)/(1+txn))*(t>2004))
+c_plt(3)*((T<=1975)*(T-1975)) + c_plt(4)*((T<=1992)*(T-1992))
genr lfd=q/plt
```

Dependent Variable: LOG(Q/LF)
 Method: Least Squares
 Date: 11/07/12 Time: 18:53
 Sample (adjusted): 1963Q1 2003Q4
 Included observations: 164 after adjustments

$$\text{LOG(Q/LF)} = C_PLT(1) + (T-2005) * (C_PLT(2) * (T \leq 2005) + 4 * \text{LOG}((1+TXQ) / (1+TXN)) * (T > 2005)) + C_PLT(3) * (T \leq 1975) * (T-1975) + C_PLT(4) * (T \leq 1992) * (T-1992)$$

	Coefficient	Std. Error	t-Statistic	Prob.
C_PLT(1)	10.94511	0.004502	2431.041	0.0000
C_PLT(2)	0.010535	0.000469	22.47874	0.0000
C_PLT(3)	0.019946	0.000646	30.89304	0.0000
C_PLT(4)	0.012821	0.000658	19.49688	0.0000
R-squared	0.997580	Mean dependent var	10.54892	
Adjusted R-squared	0.997534	S.D. dependent var	0.303824	
S.E. of regression	0.015086	Akaike info criterion	-5.525963	
Sum squared resid	0.036416	Schwarz criterion	-5.450357	
Log likelihood	457.1290	Hannan-Quinn criter.	-5.495270	
F-statistic	21983.12	Durbin-Watson stat	0.812272	
Prob(F-statistic)	0.000000			



Les deux ruptures sont identifiées comme ayant lieu en 1972 et 1992 (premiers trimestres dans chaque cas).

Comme dans le cas précédent, on constate que la productivité du travail observée est stationnaire autour de la tendance.

Cela nous donne la cible d'emploi, à partir de laquelle nous construisons le modèle à correction d'erreur:

```
coef(10) c_lf
smpl 1962Q1 2004Q4
genr lf_ec=0
smpl 1962Q1 2003Q4
```

```

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

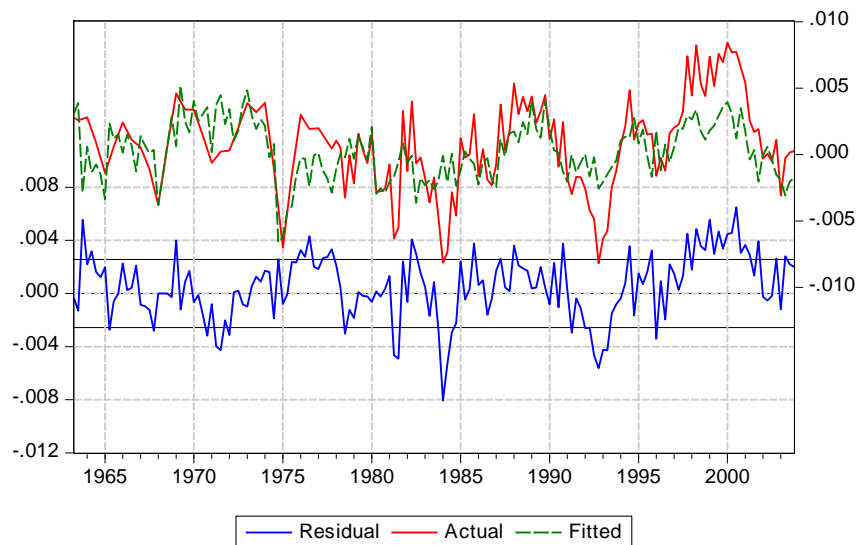
Date: 11/07/12 Time: 18:54

Sample (adjusted): 1963Q2 2003Q4

Included observations: 163 after adjustments

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)

	Coefficient	Std. Error	t-Statistic	Prob.
C_LF(1)	0.260651	0.028089	9.279331	0.0000
C_LF(2)	0.107863	0.015245	7.075301	0.0000
C_LF(3)	0.022898	0.003906	5.861669	0.0000
C_LF(4)	-0.019260	0.003948	-4.877970	0.0000
C_LF(5)	-0.009769	0.002640	-3.700988	0.0003
R-squared	0.395867	Mean dependent var		0.000753
Adjusted R-squared	0.380573	S.D. dependent var		0.003267
S.E. of regression	0.002571	Akaike info criterion		-9.058593
Sum squared resid	0.001045	Schwarz criterion		-8.963693
Log likelihood	743.2754	Hannan-Quinn criter.		-9.020065
Durbin-Watson stat	0.907945			



Cette fois, nous identifions trois variables muettes pour 1968, une année très particulière en France en particulier pour le marché du travail. Répétons notre explication du problème.

Au cours du premier trimestre et tout le second, une «révolution étudiante» (dite de Mai 68) a effectivement bloqué l'économie, en particulier les transports. Le PIB a diminué fortement, mais l'emploi n'a pas vraiment suivi, car les entrepreneurs ont estimé (à juste titre) que cette situation était transitoire. Ceci réclame une variable muette pour ce trimestre (avec un coefficient positif). Au premier trimestre, le PIB avait déjà augmenté sans création d'emploi. Mais surtout, au troisième, la forte croissance du PIB a pu être atteinte sans embauche, grâce aux personnes non licenciées plus tôt.

En fait, ce processus de transition pourrait être exprimé par une variable unique avec des valeurs de -0,5 et -0,5, 1 au cours des trois trimestres pertinents. La somme nulle rendrait l'impact transitoire.

Dependent Variable: DLOG(LF)

Method: Least Squares

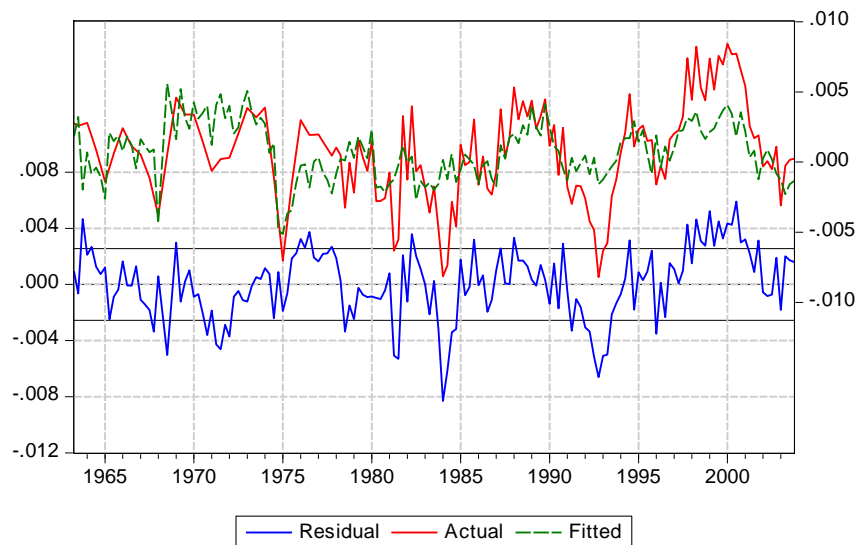
Date: 11/07/12 Time: 18:58

Sample (adjusted): 1963Q2 2004Q4

Included observations: 167 after adjustments

$$\text{DLOG(LF)} = \text{C_LF(1)} * \text{DLOG(LFD)} + \text{C_LF(2)} * \text{LOG(LFD(-1)/LF(-1))} + \text{C_LF(3)} * (-0.5 * (\text{T}=1968) + 1 * (\text{T}=1968.25) - 0.5 * (\text{T}=1968.50)) + \text{C_LF(4)}$$

	Coefficient	Std. Error	t-Statistic	Prob.
C_LF(1)	0.220392	0.024174	9.116813	0.0000
C_LF(2)	0.111036	0.014601	7.604636	0.0000
C_LF(3)	0.020235	0.003322	6.091305	0.0000
C_LF(4)	0.000546	0.000198	2.761393	0.0064
R-squared	0.394013	Mean dependent var		0.000756
Adjusted R-squared	0.382860	S.D. dependent var		0.003228
S.E. of regression	0.002536	Akaike info criterion		-9.093091
Sum squared resid	0.001048	Schwarz criterion		-9.018408
Log likelihood	763.2731	Hannan-Quinn criter.		-9.062779
F-statistic	35.32752	Durbin-Watson stat		0.822083
Prob(F-statistic)	0.000000			



10.2.3.4.2 Une fonction de type Cobb-Douglas

Dans le cadre d'un modèle opérationnel, le choix de facteurs complémentaires est assez discutable et limité. Mais il est aussi très aisé à mettre en œuvre, et conduit à une interprétation plus facile des propriétés. Le passage à une spécification de type Cobb-Douglas peut paraître simple à première vue. Par exemple, nous pourrions envisager de modifier légèrement le cadre des facteurs complémentaires, en définissant l'emploi et le capital séparément en fonction de la capacité cible, et en construisant la capacité de production en fonction de ces éléments, à l'aide de la fonction Cobb-Douglas elle-même.

Mais alors nous perdrons la propriété la plus intéressante de ce cadre: la prise en compte de la sensibilité (endogène) des facteurs travail et capital au rapport de leurs coûts relatifs.

Cela signifie qu'un choc produisant une augmentation de 1 % de la capacité cible conduirait à long terme à la même augmentation relative de ces deux facteurs, même si le choc a modifié le rapport des coûts de la main-d'œuvre et du capital.

Cela paraît trop simpliste, et nos formules tiendront compte de cet effet. Intégrées au modèle, elles montreront, par exemple, qu'une baisse des cotisations sociales des entreprises sera particulièrement propice à l'emploi, et que l'effet d'un choc de demande sur le chômage et les salaires favorisera le capital.

Cette hypothèse simple conduit donc à un cadre assez compliqué. Considérons les éléments à leur tour.

10.2.3.4.2.1 LE CADRE THEORIQUE

Nous allons maintenant décrire en détail le cadre en allant plus loin que nous ne l'avons fait plus haut.

10.2.3.4.2.1.1 LA MAXIMISATION DES MARGES

Dans ce cadre, les entreprises cherchent à maximiser leurs marges

$$pq_t \cdot Q_t - w \text{cost}_t \cdot LE_t - k \text{cost}_t \cdot K_t$$

sous la contrainte de la fonction de production:

$$\text{Log}(CAP_t) = \alpha \cdot \text{Log}(LE_t) + (1 - \alpha) \cdot \text{Log}(K_t) + \beta \cdot t + \gamma$$

ce qui nous amène à maximiser

$$pq_t \cdot Q_t - wcost_t \cdot LE_t - kcost_t \cdot K_t$$

$$(3) - \lambda \cdot (CAP_t - \exp(\beta \cdot t + \gamma) \cdot LE_t^\alpha K_t^{1-\alpha})$$

par rapport à la fois LE_t et K_{t-1} .

La dérivation de (3) donne:

$$-wcost_t - \lambda \cdot \alpha \cdot \exp(\beta \cdot t + \gamma) \cdot LE_t^{\alpha-1} K_t^{1-\alpha} = 0$$

ou

$$-wcost_t - \lambda \cdot \alpha \cdot \exp(\beta \cdot t + \gamma) \cdot (K_t / LE_t)^{1-\alpha} = 0$$

et de façon équivalente

$$-kcost_t - \lambda \cdot (1 - \alpha) \cdot \exp(\beta \cdot t + \gamma) \cdot (K_t / LE_t)^\alpha = 0$$

La division de (4) par (5) des deux côtés donne:

$$(6) \quad w \text{ cost}_t / k \text{ cost}_t = \alpha / (1 - \alpha) \cdot (K_t / LE_t)$$

ce qui montre bien l'élasticité du rapport des facteurs au rapport des coûts.

De (2) et (6) nous obtenons:

$$\begin{aligned} \text{Log}(LE_t / CAP_t) &= (1 - \alpha) \cdot \text{Log}(LE_t / K_t) - \beta \cdot t - \gamma \\ &= (1 - \alpha) \cdot \text{Log}(\alpha / (1 - \alpha) k \text{ cost}_t / w \text{ cost}_t) - \beta \cdot t - \gamma \\ \text{Log}(K_t / CAP_t) &= -\alpha \cdot \text{Log}(LE_t / K_t) - \beta \cdot t - \gamma \\ &= -\alpha \cdot \text{Log}(\alpha / (1 - \alpha) k \text{ cost}_t / w \text{ cost}_t) - \beta \cdot t - \gamma \end{aligned}$$

Mais pour appliquer ce cadre à un modèle complet, nous devons prendre en compte plusieurs éléments.

10.2.3.4.2.1.2 CIBLES ET VALEURS OBSERVEES

La présentation ci-dessus s'applique à des objectifs: connaissant les coûts relatifs, les entreprises vont estimer une cible de capacités, et les niveaux cibles des facteurs qui permettent d'atteindre cette capacité.

Cela signifie que nous devons définir cette capacité, puis calculer le niveau cible de deux facteurs, puis définir le processus qui mène des valeurs cibles aux valeurs réelles.

10.2.3.4.2.1.3 LA CAPACITE DE PRODUCTION

Pour définir la capacité, nous supposerons que les entreprises ont une cible de taux d'utilisation, constante sur la période. Cela signifie que la capacité cible est proportionnelle à la production (ou plutôt à la valeur ajoutée). Dans l'estimation des équations (7) et (8), nous remplacerons le rapport des facteurs de capacité par le rapport des facteurs de production, sans perte de généralité (le taux cible d'utilisation sera absorbé par le terme constant).

10.2.3.4.2.1.4 LE FACTEUR TEMPS

Dans le cadre ci-dessus, nous avons utilisé uniquement les éléments instantanés. Mais nous devons tenir compte de la nature de nos variables.

- Le capital et la capacité sont mesurés en un point donné du temps
- L'emploi, dans notre définition, est un niveau moyen sur une période (un trimestre).

Nous supposerons que les cibles de capital et d'emploi (et la cible de capacité implicite) sont données par le système, en utilisant le niveau réel de la production, mais que la capacité réelle de la période est donnée par l'emploi effectif et le niveau initial du capital. Comme le capital est mesuré en fin de période, nous allons utiliser la valeur retardée.

Selon une méthode courante, nous identifierons la valeur estimée des facteurs à la cible, et supposerons que le résidu représente l'écart à la décision effectivement appliquée.

10.2.3.4.2.1.5 L'INERTIE DES FACTEURS

Nous allons également supposer que les décisions optimales ne sont pas mises en œuvre immédiatement. Comme expliqué précédemment, les raisons sont à la fois techniques (la longueur du processus d'investissement) et psychologiques (aversion au risque). Ainsi, les entreprises vont seulement faire une partie du chemin vers la cible, à partir du niveau de décision précédent. Nous allons essayer d'estimer le facteur d'inertie, avec des valeurs différentes pour le travail et le capital.

Le premier facteur devrait être moins inerte, car le coût des erreurs y est plus faible (si l'on compare des salaires trimestriels au coût de l'investissement correspondant), et la gestion de leurs conséquences est également plus facile (licencier des travailleurs est plus facile techniquement¹⁸ que vendre des équipements inutiles).

10.2.3.4.2.1.6 LES COÛTS RELATIFS

Bien sûr, il faut comparer le prix du capital (en fait l'investissement¹⁹) et le prix du travail (le taux de salaire). En fait, les choses sont un peu plus complexes.

- Le taux de salaire devrait inclure les cotisations sociales payées par les entreprises.
- Une fois acheté, le capital peut être utilisé tant qu'il n'est pas détruit ou obsolète, alors que l'emploi est acheté pour une seule période²⁰.
- Le prix du capital devrait tenir compte du fait qu'il doit être acheté tout de suite, tandis que le facteur de remplacement, le travail, est payé au moment où il est utilisé, ou même plus tard. Ce délai devrait réclamer l'introduction du taux d'intérêt.
- Le capital se déprécie au fil du temps. Pour les travailleurs, l'efficacité augmente puis diminue généralement, selon le produit. Mais les entreprises peuvent toujours remplacer les travailleurs âgés par de nouveaux, à un coût minime (le financement de la retraite est généralement inclus dans le coût salarial)²¹.

Nous allons effectivement comparer le coût salarial annuel avec le prix de l'investissement, qui équivaut au prix du capital au coût de renouvellement. Il faut en effet considérer que l'augmentation de l'efficacité du capital est incluse dans la variable à prix constants, et

¹⁸ Nous ne considérons pas évidemment les aspects sociaux, que les modèles ne traitent que rarement, sans doute faute d'informations. Cela ne péjore pas des qualités humaines du modélisateur.

¹⁹ Rappelons-nous que l'évolution de la qualité du capital s'applique à sa valeur à prix constants, et non pas à son déflateur, mesuré à une qualité (ou efficacité) donnée. Cela signifie que le déflateur actuel s'applique également aux valeurs passées du capital, comme l'efficacité d'une grandeur est indépendante du moment où elle a été achetée. Appliqué à un niveau de capital donné, il vous donnera la valeur de l'investissement nécessaire pour remplacer ce capital, avec la même efficacité.

²¹ Même remarque que plus haut.

non pas dans le déflateur (c'est ce qu'on appelle l'«effet qualité»). Pour répartir le coût du capital sur sa durée d'utilisation, nous allons diviser son déflateur par un facteur estimé, ce qui devrait s'interpréter comme le nombre de périodes de sa vie productive.

$$relc = (wf \cdot (1 + r_{scf}) / pi) \cdot f$$

$$f = 1 / ((ir - \dot{pc}) - 4 \cdot \text{Log}(1 - rdep))$$

Quelques explications sont nécessaires pour le dernier terme. L'augmentation du coût du capital pour une période donnée correspond à la somme du taux d'intérêt et de l'amortissement annuel:

$$z = (ir - \dot{pc}) + 4 \cdot \text{Log}(1 - rdep)$$

Cela signifie que le rapport entre le coût du travail et du capital donnera au premier à la période k un avantage de:

$$1 / (1 + z^k)$$

ou pour toute la période de

$$1 / (1 + z) + 1 / (1 + z)^2 + 1 / (1 + z)^3 + \dots = 1 / z$$

10.2.3.4.2.2 L'ESTIMATION

Voyons comment nous pouvons appliquer cette technique dans notre modèle, sous EViews. Nous avons vu que nous devons estimer deux équations, avec des coefficients communs. Cela se fera grâce à un «système», une caractéristique que nous avons maintenant l'occasion de présenter.

Voyons d'abord comment calculer le coût relatif, comme expliqué ci-dessus:

```
smp1 1962Q1 2004Q4
genr relc=wr*(1+r_scf)/pi/(ir/100-@pchy(pc)-4*log(1-rdep))
genr k_ec=0
genr lf_ec=0
```

Maintenant, nous devons estimer un système de deux équations à coefficients communs.

- Comme d'habitude, nous créons deux vecteurs: les coefficients et paramètres.
- Nous détruisons tout système préexistant du même nom.
- Nous initialisons les valeurs par une hypothèse.

Ceci est important car la convergence de ce système est un peu délicate sous EViews. En effet nous allons prendre le logarithme d'un coefficient, et comme le vecteur est initialisé à zéro, le processus peut s'interrompre dès le départ.

```
coef(10) c_cd
vector(10) p_cd
delete cd
smp1 1962Q1 2003Q4
system cd
c_cd(4)=0.1
```

```
c_cd(1)=100
c_cd (3) = 0,02 'la croissance annuelle de la productivité totale des facteurs: initialisé à 2 %
c_cd (2) = 0,65 ' la part de l'emploi dans le processus, initialisée à 0,6522
p_cd(8)=0.0
p_cd(9)=0.0
cd.append log(k/q(-1))=-c_cd(3)*(t-2005)-c_cd(7)+c_cd(2)*log(rele)
cd.append log(lf/q(-1))=-c_cd(3)*(t-2005)-c_cd(1)+(c_cd(2)-1)*log(rele)
```

Nous envisagerons deux méthodes: le maximum de vraisemblance à information complète (FIML) et la régression sans relation apparente (SUR, Zellner). Nous testerons les deux mais nous avons choisi la dernière.

System: CD

Estimation Method: Seemingly Unrelated Regression

Date: 11/07/12 Time: 19:18

Sample: 1968Q1 2005Q4

Included observations: 152

Total system (balanced) observations 304

Linear estimation after one-step weighting matrix

	Coefficient	Std. Error	t-Statistic	Prob.
C_CD(3)	0.005050	0.000587	8.605400	0.0000
C_CD(7)	6.454626	0.166179	38.84134	0.0000
C_CD(2)	0.637878	0.014030	45.46507	0.0000
C_CD(4)	0.015137	0.002106	7.187983	0.0000
C_CD(1)	6.908534	0.166183	41.57180	0.0000
Determinant residual covariance	7.51E-05			

Equation: $\text{LOG}(K/Q(-1)) = -C_CD(3)*(T-2005) - C_CD(7) + C_CD(2) * \text{LOG}(\text{RELC}) - C_CD(4)*(T \leq 1982)*(T-1982)$			
Observations: 152			
R-squared	0.154620	Mean dependent var	0.480164
Adjusted R-squared	0.137484	S.D. dependent var	0.277316
S.E. of regression	0.257548	Sum squared resid	9.816973
Durbin-Watson stat	0.010576		
Equation: $\text{LOG}(LF/Q(-1)) = -C_CD(3)*(T-2005) - C_CD(1) + (C_CD(2)-1) * \text{LOG}(\text{RELC}) - C_CD(4)*(T \leq 1982)*(T-1982)$			
Observations: 152			
R-squared	-0.210476	Mean dependent var	-10.63960
Adjusted R-squared	-0.235013	S.D. dependent var	0.232456
S.E. of regression	0.258331	Sum squared resid	9.876747
Durbin-Watson stat	0.005369		

L'estimation donne un coefficient de 0,64 pour le travail (une valeur très habituelle) et une tendance positive significative de la productivité globale des facteurs.

Nous utiliserons l'estimation pour déterminer la série des valeurs cibles («souhaitées»).

```
smpl 1970Q4 2004Q4
```

```
genr log(kd/q(-1))=-c_cd(3)*(t-2005)-c_cd(7)+c_cd(2)*log(relc)
```

```
genr log(lfd/q(-1))=-c_cd(3)*(t-2005)-c_cd(1)+(c_cd(2)-1)*log(relc)
```

Nous estimons que les résultats réels comme une moyenne pondérée du résultat estimé et de la valeur réelle précédente, avec une tendance additionnelle.

En fait, nous avons fixé les facteurs d'inertie, à des niveaux qui donnent au modèle des propriétés raisonnables²³.

```

smp1 1970Q4 2003Q4
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-2005)*(t<=2005)+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-2005)*(t<=2005)+lf_ec
genr lf_ec=resid

```

Enfin, nous générons la capacité de production en appliquant la formule estimée aux valeurs observées des facteurs, et nous calculons le taux d'utilisation des capacités.

```

genr log(cap)=c_cd(7)*(1-c_cd(2))+c_cd(1)*c_cd(2)+c_cd(3)*(t-2005)*(t<=2005) +c_cd(2)*4*log(((1+txq)/(1+txn))^(t-
2005)*(t>2005)+c_cd(2)*log(lf)+(1-c_cd(2))*log(k(-1))
genr ur=q/cap

```

Comme on pouvait s'y attendre, ce taux d'utilisation est un peu différent de celui que nous avons obtenu avec la fonction à facteurs complémentaire. Cela signifie que les équations dans lesquelles il intervient devront être estimées deux fois: il s'agit du déflateur de la valeur ajoutée, des exportations et des importations.

10.2.3.4.3 Les variations de stocks

Nous allons utiliser la même explication que le modèle simple, sauf pour un rôle (négatif) des décisions antérieures, et une tendance se terminant en 1985.

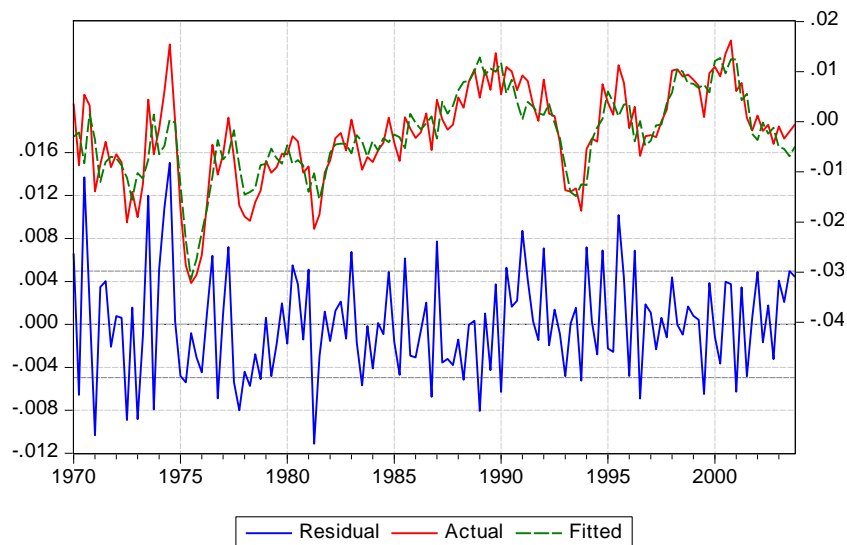
```

coef(10) c_ci
smp1 1962Q1 2004Q4
genr ci_ec=0
smp1 1963Q1 2003Q4
equation _eq_ci.ls(p) ci/q(-1)=c_ci(1)*@pchy(q)+c_ci(2)+c_ci(3)*ci(-1)/q(-2)
+c_ci(4)*(t-1985)*(t<=1985)+ci_ec
_eq_ci.resids(p)
close _eq_ci
genr ci_ec=resid

```

Dependent Variable: CI/Q(-1)
 Method: Least Squares
 Date: 11/07/12 Time: 19:24
 Sample: 1970Q1 2003Q4
 Included observations: 136
 $CI/Q(-1) = C_CI(1) * PCHY(Q) + C_CI(2) + C_CI(3) * CI(-1)/Q(-2) + C_CI(4) * (T - 1985) * (T \leq 1985)$

	Coefficient	Std. Error	t-Statistic	Prob.
C_CI(1)	0.195389	0.026646	7.332822	0.0000
C_CI(2)	-0.003401	0.000717	-4.744389	0.0000
C_CI(3)	0.584125	0.053030	11.01501	0.0000
C_CI(4)	0.000754	0.000119	6.332001	0.0000
R-squared	0.742572	Mean dependent var		-0.002669
Adjusted R-squared	0.736722	S.D. dependent var		0.009661
S.E. of regression	0.004957	Akaike info criterion		-7.747087
Sum squared resid	0.003243	Schwarz criterion		-7.661420
Log likelihood	530.8019	Hannan-Quinn criter.		-7.712274
F-statistic	126.9217	Durbin-Watson stat		2.260290
Prob(F-statistic)	0.000000			



10.2.3.4.4 Le chômage

Il s'agit d'une nouvelle équation. En fait, nous l'utiliserons pour modéliser la population active POPAC (emploi + chômage) plutôt que ce dernier. Comme nous la ferons dépendre de l'emploi, la qualité de l'estimation sera exactement la même (seul le R2 va changer). Le coefficient de POPAC sera augmenté de 1 par rapport à une formulation utilisant le chômage UN.

Par ailleurs, l'équation suivant pleinement le cadre défini plus haut.

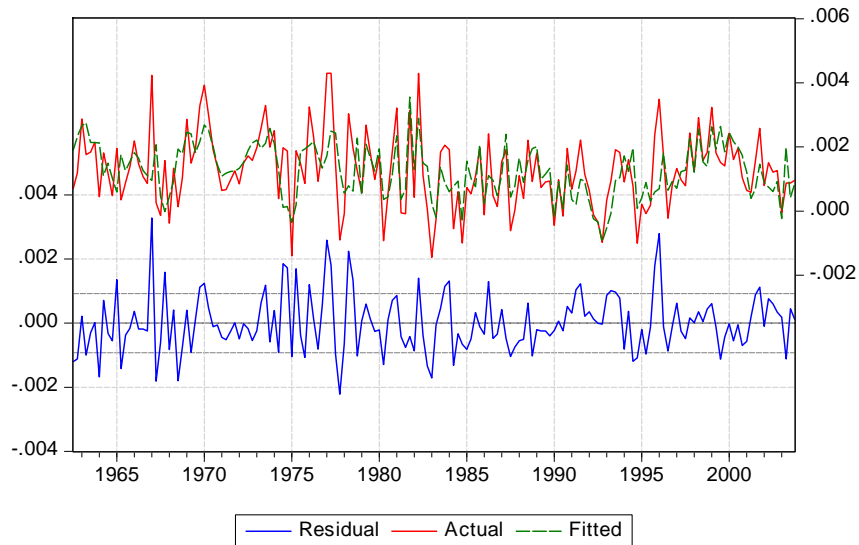
Les sensibilités à court terme sont raisonnables, et à long terme aussi, avec une valeur légèrement inférieure²⁴. Le coefficient de la force de travail potentielle n'est pas significatif, mais sa valeur est raisonnable. Nous pouvons observer une convergence très lente vers l'équilibre de long terme (avec une intensité non fiable), et une sensibilité plus forte à long terme de la population active à l'emploi (donc un impact plus faible sur le chômage). Il est logique de s'attendre à ce que l'amélioration permanente du marché du travail attire de plus en plus de demandeurs d'emploi.

```
coef(10) c_popac
smpl 1962Q1 2002Q4
genr popac_ec=0
equation _eq_popac.ls(pd(popac)/pop65(-1)=c_popac(1)*d(lt)/pop65(-1)      +c_popac(2)*d(pop65)/pop65(-1)-
c_popac(3)*(popac(-1)/pop65(-1)-c_popac(4)*lt(-1)/pop65(-1)-  c_popac(5))+c_popac(6)*(t-2003)+popac_ec
_eq_popac.resids(p)
genr popac_ec=resid
close _eq_popac
```

Dependent Variable: D(POPAC)/POP65(-1)
 Method: Least Squares
 Date: 11/07/12 Time: 19:40
 Sample (adjusted): 1962Q3 2003Q4
 Included observations: 166 after adjustments
 Convergence achieved after 121 iterations

$$D(POPAC)/POP65(-1) = C_POPAC(1) * D(LT)/POP65(-1) + C_POPAC(2) * D(POP65)/POP65(-1) - C_POPAC(3) * (POPAC(-1)/POP65(-1)) - C_POPAC(4) * LT(-1)/POP65(-1) - C_POPAC(5) + [AR(1) = C_POPAC(6)]$$

	Coefficient	Std. Error	t-Statistic	Prob.
C_POPAC(1)	0.698524	0.073991	9.440704	0.0000
C_POPAC(2)	0.394457	0.203308	1.940193	0.0541
C_POPAC(3)	0.621335	0.074271	8.365737	0.0000
C_POPAC(4)	0.432522	0.073996	5.845199	0.0000
C_POPAC(5)	1.731148	83.13383	0.020824	0.9834
C_POPAC(6)	0.999840	0.010059	99.39601	0.0000
R-squared	0.362344	Mean dependent var		0.001278
Adjusted R-squared	0.342417	S.D. dependent var		0.001138
S.E. of regression	0.000923	Akaike info criterion		-11.10322
Sum squared resid	0.000136	Schwarz criterion		-10.99074
Log likelihood	927.5673	Hannan-Quinn criter.		-11.05756
Durbin-Watson stat	1.812878			
Inverted AR Roots	1.00			



10.2.3.4.5 Le déflateur de la valeur ajoutée

La formule va changer selon le type de fonction de production: la mesure du taux d'utilisation sera différente, et dans le cas Cobb-Douglas le coût comprendra logiquement l'amortissement du coût du capital.

Dans les deux cas, l'équation suivra un modèle à correction d'erreur: une relation à long terme entre le taux d'utilisation et le taux de marge (ou plutôt la part des coûts dans la valeur ajoutée), et une équation dynamique décrivant l'évolution vers cette cible, en libérant l'élasticité au déflateur du coût.

Présentons d'abord le cas des facteurs complémentaires.

Nous avons contraint à la même valeur la sensibilité à court et à long terme aux taux d'utilisation.

```
smpl 1962Q1 2004Q4
coef(10) c_pq
genr pq_ec=0
smpl 1963Q1 2003Q4
equation _eq_pq.ls(p) dlog(pq)=c_pq(1)*dlog(uwc)+c_pq(2)*dlog(ur)+c_pq(3)*log(pq(-1)/uwc(-1))-c_pq(2)*c_pq(3)*log(ur(-1))+c_pq(5)+pq_ec
_eq_pq.resids(p)
close _eq_pq
```

Dependent Variable: DLOG(PQ)

Method: Least Squares

Date: 11/07/12 Time: 19:49

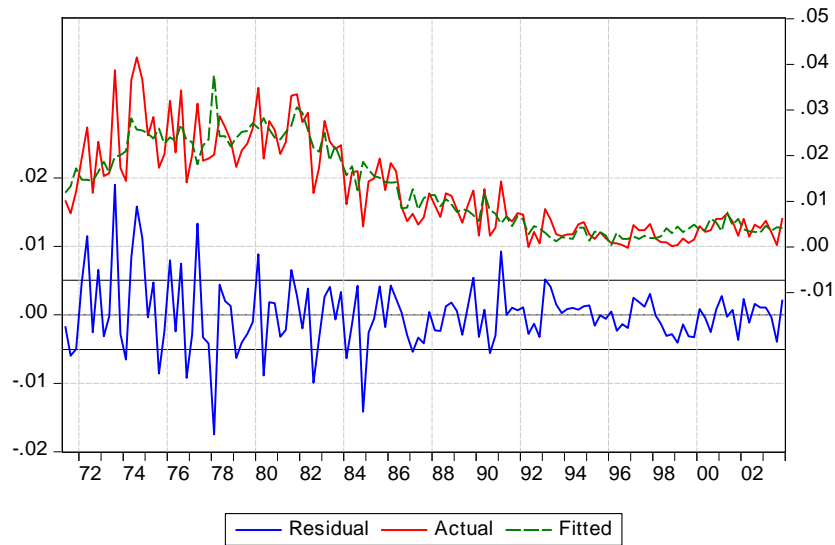
Sample (adjusted): 1971Q2 2003Q4

Included observations: 131 after adjustments

Convergence achieved after 1 iteration

$$\text{DLOG(PQ)} = \text{C_PQ(1)} * \text{DLOG(UWC)} + \text{C_PQ(2)} * \text{DLOG(UR)} + \text{C_PQ(3)} * \text{LOG(PQ(-1)/UWC(-1))} - \text{C_PQ(2)} * \text{C_PQ(3)} * \text{LOG(UR(-1))} + \text{C_PQ(5)}$$

	Coefficient	Std. Error	t-Statistic	Prob.
C_PQ(1)	0.415570	0.045536	9.126161	0.0000
C_PQ(2)	0.354049	0.086013	4.116222	0.0001
C_PQ(3)	-0.065965	0.007525	-8.765824	0.0000
C_PQ(5)	0.037465	0.003775	9.923875	0.0000
R-squared	0.773324	Mean dependent var		0.012717
Adjusted R-squared	0.767969	S.D. dependent var		0.010493
S.E. of regression	0.005054	Akaike info criterion		-7.707104
Sum squared resid	0.003244	Schwarz criterion		-7.619311
Log likelihood	508.8153	Hannan-Quinn criter.		-7.671430
F-statistic	144.4234	Durbin-Watson stat		2.015563
Prob(F-statistic)	0.000000			



Dans le cas Cobb-Douglas, il nous faut considérer le coût total: salaire et capital.

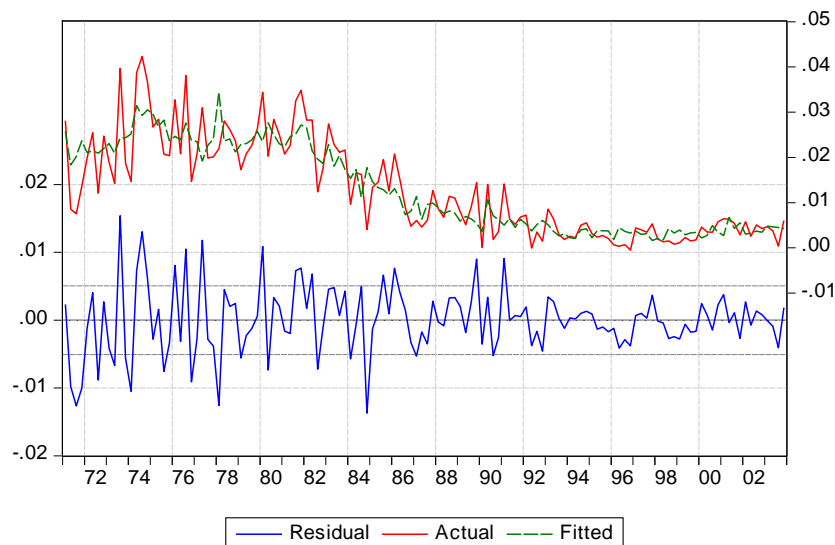
$$\text{genr cost} = (w \cdot l^{\alpha} (1 + r_{scf}) + 0.05 \cdot p \cdot k^{1-\alpha}) / q$$

Nous répartissons le coût du capital sur 20 périodes (5 ans). Réduire cette valeur ne changerait pas beaucoup les propriétés.

Les coefficients sont assez différents du cas précédent: effet plus dynamique, correction plus lente.

Dependent Variable: DLOG(PQ)
Method: Least Squares
Date: 11/07/12 Time: 19:18
Sample (adjusted): 1971Q1 2003Q4
Included observations: 132 after adjustments
Convergence achieved after 5 iterations
DLOG(PQ)=C_PQ(1)*DLOG(COST)+C_PQ(2)*DLOG(UR)+C_PQ(3)
*LOG(PQ(-1)/COST(-1))-C_PQ(2)*C_PQ(3)*LOG(UR(-1))+C_PQ(5)
+PQ_EC

	Coefficient	Std. Error	t-Statistic	Prob.
C_PQ(1)	0.441903	0.077407	5.708801	0.0000
C_PQ(2)	0.295159	0.120499	2.449460	0.0157
C_PQ(3)	-0.052352	0.008027	-6.522219	0.0000
C_PQ(5)	0.022348	0.002811	7.949870	0.0000
R-squared	0.783973	Mean dependent var		0.012987
Adjusted R-squared	0.778910	S.D. dependent var		0.010789
S.E. of regression	0.005073	Akaike info criterion		-7.699873
Sum squared resid	0.003294	Schwarz criterion		-7.612515
Log likelihood	512.1916	Hannan-Quinn criter.		-7.664375
F-statistic	154.8399	Durbin-Watson stat		2.056374
Prob(F-statistic)	0.000000			



10.2.3.4.6 Le taux de salaire

L'estimation du taux de salaire sera notre seule occasion d'utiliser la cointégration, et donc de nous conformer aux principes reconnus de l'économétrie.

Le processus d'estimation est exactement le même que pour le petit modèle.

Nous supposons que le coût salarial est indexé sur le long terme: 50 % sur le déflateur de la valeur ajoutée (les entreprises veulent stabiliser la part des salaires dans celle-ci) et pour 50 % sur le déflateur de la consommation (les travailleurs veulent que leur pouvoir d'achat suive leurs gains de productivité). La pondération pourrait utiliser un paramètre, éventuellement estimé.

Testons d'abord la stationnarité du coût réel du travail `_luwc`:

```
genr _luwc=log(uwc)-0.5*log(pq)-(1-0.5)*log(pcoh)
uroot(p) _luwc
```

Null Hypothesis: _LUWC has a unit root
 Exogenous: Constant
 Lag Length: 0 (Automatic - based on SIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.209301	0.9332
Test critical values:		
1% level	-3.482453	
5% level	-2.884291	
10% level	-2.578981	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(_LUWC)
 Method: Least Squares
 Date: 11/07/12 Time: 19:59
 Sample (adjusted): 1972Q2 2003Q4
 Included observations: 127 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
_LUWC(-1)	-0.001769	0.008454	-0.209301	0.8346
C	-0.001650	0.003746	-0.440503	0.6603

R-squared	0.000350	Mean dependent var	-0.000881
Adjusted R-squared	-0.007647	S.D. dependent var	0.008313
S.E. of regression	0.008345	Akaike info criterion	-6.718639
Sum squared resid	0.008705	Schwarz criterion	-6.673848
Log likelihood	428.6336	Hannan-Quinn criter.	-6.700441
F-statistic	0.043807	Durbin-Watson stat	1.855823
Prob(F-statistic)	0.834554		

ainsi que celle du taux de chômage lui-même.

uroot(p) unr

Null Hypothesis: UNR has a unit root				
Exogenous: Constant				
Lag Length: 1 (Automatic - based on SIC, maxlag=13)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic				
Test critical values:			-1.390809	0.5858
	1% level		-3.468980	
	5% level		-2.878413	
	10% level		-2.575844	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(UNR)				
Method: Least Squares				
Date: 11/07/12 Time: 20:01				
Sample (adjusted): 1962Q3 2004Q4				
Included observations: 170 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
UNR(-1)	-0.004118	0.002961	-1.390809	0.1661
D(UNR(-1))	0.674893	0.056522	11.94040	0.0000
C	0.000450	0.000238	1.888933	0.0606
R-squared	0.469104	Mean dependent var		0.000495
Adjusted R-squared	0.462746	S.D. dependent var		0.001945
S.E. of regression	0.001426	Akaike info criterion		-10.25100
Sum squared resid	0.000339	Schwarz criterion		-10.19566
Log likelihood	874.3347	Hannan-Quinn criter.		-10.22854
F-statistic	73.78123	Durbin-Watson stat		1.881471
Prob(F-statistic)	0.000000			

Les deux échouent.

Ensuite, nous testons la cointégration des deux éléments, qui fonctionne.

```
coint(b,p) _luwc unr
```

Date: 11/07/12 Time: 20:03
 Sample (adjusted): 1973Q2 2003Q4
 Included observations: 123 after adjustments
 Trend assumption: No deterministic trend (restricted constant)
 Series: _LUWC UNR
 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 *	0.153423	27.57871	20.26184	0.0041
At most 1	0.056032	7.092588	9.164546	0.1216

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 *	0.153423	20.48612	15.89210	0.0088
At most 1	0.056032	7.092588	9.164546	0.1216

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
-13.09680	-61.22113	-0.735028
9.578167	-2.900417	5.183698
Unrestricted Adjustment Coefficients (alpha):		
D(_LUWC)	0.003119	-0.000294
D(UNR)	-1.40E-05	0.000290
1 Cointegrating Equation(s):	Log likelihood	1080.812
Normalized cointegrating coefficients (standard error in parentheses)		
_LUWC	UNR	C
1.000000	4.674509	0.056123
	(0.72130)	(0.06394)
Adjustment coefficients (standard error in parentheses)		
D(_LUWC)	-0.040853	
	(0.00912)	
D(UNR)	0.000183	
	(0.00150)	

Le coefficient de UNR dans l'équation de cointégration est significatif avec le bon signe. Il reste à voir si cette équation donne de bonnes propriétés au modèle complet.

Pour l'équation dynamique, nous avons dû forcer l'indexation à une valeur globalement unitaire (homogénéité dynamique) avec une structure de retards. Tous les éléments sont très significatifs, mais la correction d'erreur est assez lente (une caractéristique que nous avons retrouvée dans d'autres circonstances pour la même équation).

Dependent Variable: DLOG(WR)

Method: Least Squares

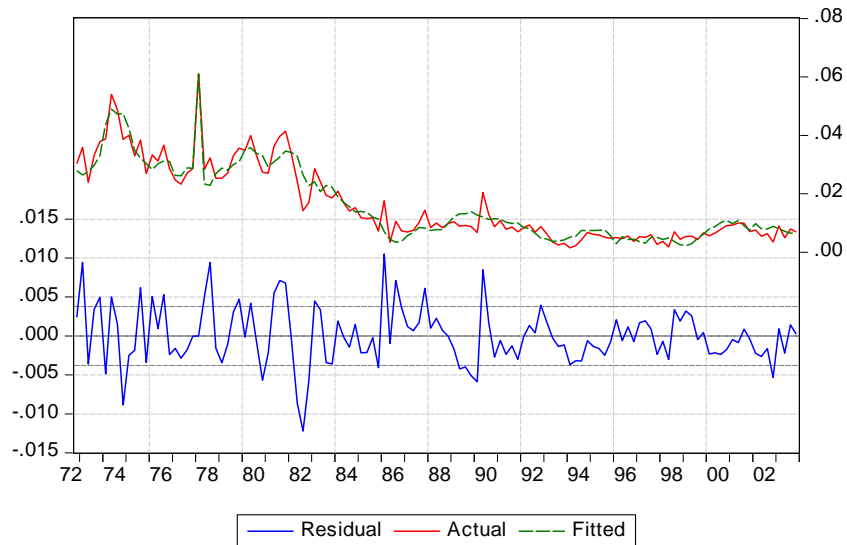
Date: 11/07/12 Time: 20:13

Sample (adjusted): 1972Q4 2003Q4

Included observations: 125 after adjustments

DLOG(WR)=C_WR(1)*@MOVAV(DLOG(PCOH),4)+C_WR(2)*D(UNR)
+C_WR(3)*RES_WR(-1)+C_WR(7)*(T=1978)+C_WR(8)*T

	Coefficient	Std. Error	t-Statistic	Prob.
C_WR(1)	1.148306	0.036142	31.77180	0.0000
C_WR(2)	-0.694329	0.180518	-3.846317	0.0002
C_WR(3)	-0.038498	0.004146	-9.285473	0.0000
C_WR(7)	0.033340	0.003844	8.673487	0.0000
C_WR(8)	1.29E-06	2.97E-07	4.330384	0.0000
R-squared	0.920773	Mean dependent var		0.016535
Adjusted R-squared	0.918132	S.D. dependent var		0.013225
S.E. of regression	0.003784	Akaike info criterion		-8.276957
Sum squared resid	0.001718	Schwarz criterion		-8.163824
Log likelihood	522.3098	Hannan-Quinn criter.		-8.230997
Durbin-Watson stat	1.594350			



10.2.3.4.7 Les déflateurs des échanges extérieurs

Pour cette équation, nous n'utiliserons pas un format à correction d'erreur, mais plutôt une indexation homogène. Cela signifie que nous ne séparerons pas les sensibilités à court et à long terme.

Comme d'habitude pour ce type d'équation, les exportateurs montrent une plus grande attention à leurs coûts qu'au prix de leurs concurrents. Cela est particulièrement vrai pour les exportateurs vers la France, où le rôle des prix français n'est pas tout à fait significatif.

Ceci freinera la dynamique de la boucle prix-salaires. L'impact des échanges sur le prix de la demande peut être expliqué de la manière suivante.

- Les Importations représentent une part de la demande globale. Elles sont achetées au prix à l'importation. Plus sa sensibilité aux coûts en devises est élevée, plus la différence avec le prix de la production locale sera forte, réduisant d'autant l'impact de la valeur des importations sur le prix de la demande globale.
- Si les producteurs locaux ont décidé de leur prix de vente sur les marchés locaux et étrangers de façon indépendante (un comportement possible que nous ne considérerons pas), l'indexation du prix de la demande aux prix locaux à la production sera donc nettement inférieure à l'unité.
- Mais dans notre cadre, le prix de production est décidé au niveau global, et la sensibilité réduite du prix à l'exportation aux prix à la production français devra être compensée par une plus grande sensibilité du prix auquel ils vendent sur le marché local. Plus l'impact des coûts locaux sur le prix à l'exportation sera élevé, plus la compensation nécessaire sera faible.

On peut voir que dans le passage du prix à la production au prix de la demande, plus le rôle du coût de production dans le prix fixé par l'exportateur (local ou étranger) est élevé, et plus le premier effet (négatif) l'emporte sur le second (positif)²⁵.

À l'extrême, si tous les exportateurs ne prennent en compte que leurs coûts, les prix à l'importation ne seront pas affectés, et comme le prix à l'exportation va changer comme le prix à la production local, aucune compensation ne sera nécessaire. L'effet d'amortissement sera maximal.

La tendance supplémentaire est négative et très significative. Elle représente probablement un changement structurel mais permanent dans les biens échangés, s'orientant vers ceux qui présentent le prix le plus bas.

Pour atteindre un équilibre à long terme, ces tendances doivent être supprimées après un certain temps. Ici, nous l'avons fait immédiatement, mais de véritables prévisions devraient prendre en compte une diminution progressive.

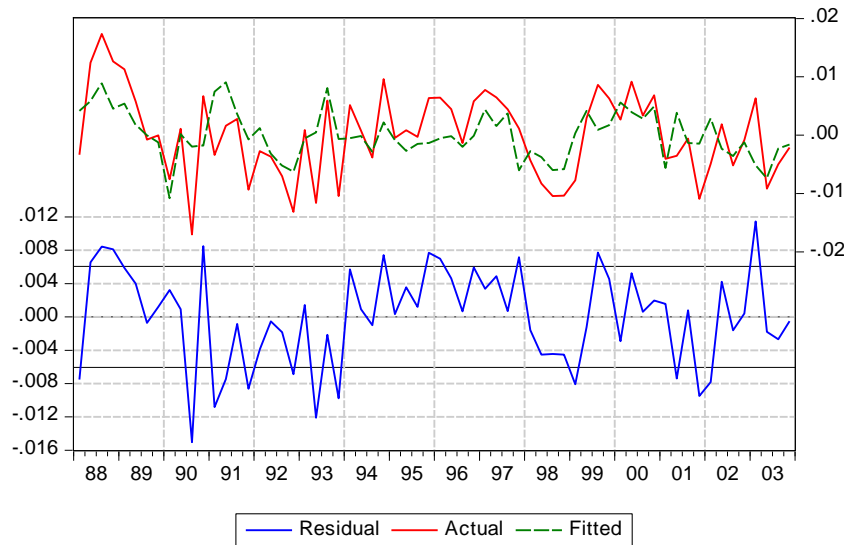
```
coef(10) c_px
smpl 1962Q1 2004Q4
genr px_ec=0
smpl 1962Q1 2003Q4
equation _eq_px.ls(p) log(px)=c_px(4)*log(pp)+(1-c_px(4))*log(pp*er)+c_px(6)+c_px(5)*(t-
2005)*(t<=2005)+[ar(1)=c_px(7)]+px_ec
_eq_px.resids(p)
close _eq_px
```

Dependent Variable: LOG(PX)
 Method: Least Squares
 Date: 11/07/12 Time: 20:17
 Sample (adjusted): 1970Q2 2003Q4
 Included observations: 135 after adjustments
 Convergence achieved after 1 iteration
 LOG(PX)=C_PX(4)*LOG(PP)+(1-C_PX(4))*LOG(PPX*ER)+C_PX(6)
 +C_PX(5)*(T-2005)*(T<=2005)+[AR(1)=C_PX(7)]

	Coefficient	Std. Error	t-Statistic	Prob.
C_PX(4)	0.614657	0.046654	13.17491	0.0000
C_PX(6)	-0.172383	0.035436	-4.864641	0.0000
C_PX(5)	-0.022531	0.003025	-7.447879	0.0000
C_PX(7)	0.963293	0.018031	53.42319	0.0000

R-squared	0.999545	Mean dependent var	-0.287933
Adjusted R-squared	0.999535	S.D. dependent var	0.445430
S.E. of regression	0.009606	Akaike info criterion	-6.423668
Sum squared resid	0.012088	Schwarz criterion	-6.337586
Log likelihood	437.5976	Hannan-Quinn criter.	-6.388687
F-statistic	95996.91	Durbin-Watson stat	1.682038
Prob(F-statistic)	0.000000		

Inverted AR Roots	.96
-------------------	-----



```
coef(10) c_pm
smpl 1962Q1 2004Q4
genr pm_ec=0
smpl 1962Q1 2003Q4
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-
2005)*(t<=2005)+[ar(1)=c_pm(7)]+pm_ec
_eq_pm.resids(p)
close _eq_pm
```

Dependent Variable: LOG(PM)

Method: Least Squares

Date: 11/07/12 Time: 20:23

Sample (adjusted): 1970Q2 2003Q4

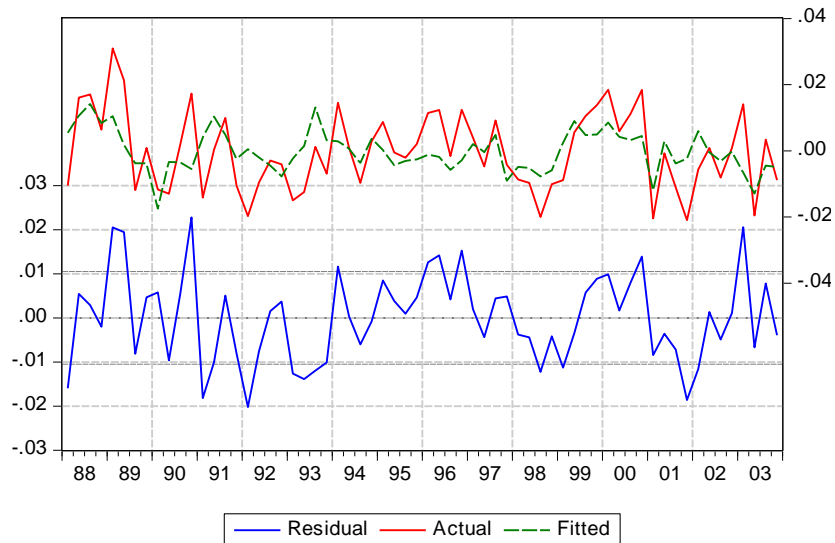
Included observations: 135 after adjustments

Convergence achieved after 1 iteration

$$\text{LOG(PM)} = \text{C_PM(4)} * \text{LOG(PP)} + (1 - \text{C_PM(4)}) * \text{LOG(PPX*ER)} + \text{C_PM(6)}$$

$$+ \text{C_PM(5)} * (\text{T} - 2005) * (\text{T} \leq 2005) + [\text{AR}(1) = \text{C_PM(7)}]$$

	Coefficient	Std. Error	t-Statistic	Prob.
C_PM(4)	0.118328	0.107132	1.104509	0.2714
C_PM(6)	-0.270806	0.078177	-3.464005	0.0007
C_PM(5)	-0.034081	0.006262	-5.442496	0.0000
C_PM(7)	0.960815	0.018142	52.95997	0.0000
R-squared	0.997835	Mean dependent var		-0.264537
Adjusted R-squared	0.997785	S.D. dependent var		0.469326
S.E. of regression	0.022088	Akaike info criterion		-4.758370
Sum squared resid	0.063913	Schwarz criterion		-4.672288
Log likelihood	325.1900	Hannan-Quinn criter.		-4.723389
F-statistic	20122.07	Durbin-Watson stat		1.417189
Prob(F-statistic)	0.000000			
Inverted AR Roots	.96			



10.2.3.4.8 La consommation des ménages

Notre équation suivra comme d'habitude une spécification à correction d'erreur (estimation en une seule étape !). Elle utilisera presque complètement le cadre présenté plus tôt.

L'évolution de la consommation dépendra:

- de la variation du revenu réel (par rapport à l'année dernière) ;
- de la variation du taux de chômage ;
- de l'inflation (l'effet «d'encasmes réelles») ;
- d'un terme de correction d'erreur ;

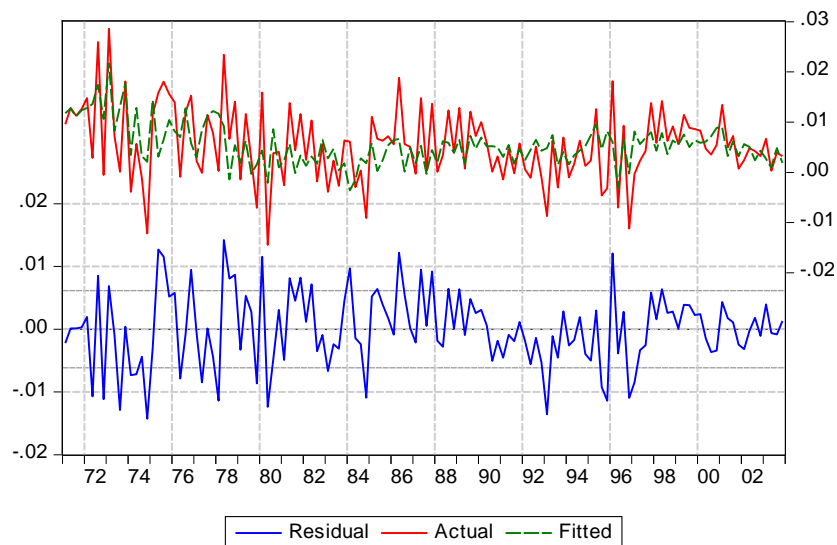
- d'une tendance temporelle négative, qui représente l'augmentation de la richesse des ménages et de leur potentiel d'épargne (traduisant en particulier leur accession à la propriété d'un logement).

La seule influence que nous n'avons pas pu mettre en évidence est celle du taux réel d'intérêt à court terme.

```
coef(10) c_coh
smpl 1962Q1 2004Q4
genr coh_ec=0
smpl 1962Q1 2003Q4
equation _eq_coh.ls(p) dlog(coh)=c_coh(1)*.25*log(hr di/hr di(-4)) +c_coh(2)*dlog(unr)+c_coh(3)*log(pcoh/pcoh(-
4))+c_coh(5)*dlog(coh(-1)) +c_coh(6)+c_coh(7)*log(coh(-1)/hr di(-1))+c_coh(8)*(t-2005)*(t<=2005)+coh_ec
_eq_coh.resids(p)
```

Dependent Variable: DLOG(COH)
 Method: Least Squares
 Date: 11/07/12 Time: 20:45
 Sample (adjusted): 1971Q1 2003Q4
 Included observations: 132 after adjustments
 DLOG(COH)=C_COH(1)*.25*LOG(HRDI/HRDI(-4))+C_COH(2)
 *DLOG(UNR)+C_COH(3)*LOG(PCOH/PCOH(-4))+C_COH(5)
 *DLOG(COH(-1))+C_COH(6)+C_COH(7)*LOG(COH(-1)/HRDI(-1))
 +C_COH(8)*(T-2005)*(T<=2005)

	Coefficient	Std. Error	t-Statistic	Prob.
C_COH(1)	0.445908	0.158975	2.804900	0.0058
C_COH(2)	-0.043832	0.019937	-2.198509	0.0298
C_COH(3)	-0.064211	0.027487	-2.336033	0.0211
C_COH(5)	-0.330953	0.079638	-4.155725	0.0001
C_COH(6)	-0.003384	0.002005	-1.688118	0.0939
C_COH(7)	-0.066734	0.023394	-2.852604	0.0051
C_COH(8)	-0.000269	0.000109	-2.461291	0.0152
R-squared	0.328381	Mean dependent var		0.005510
Adjusted R-squared	0.296144	S.D. dependent var		0.007319
S.E. of regression	0.006140	Akaike info criterion		-7.296357
Sum squared resid	0.004713	Schwarz criterion		-7.143482
Log likelihood	488.5596	Hannan-Quinn criter.		-7.234236
F-statistic	10.18626	Durbin-Watson stat		1.932245
Prob(F-statistic)	0.000000			



10.2.3.4.9 Les exportations

Pour les exportations, nous utiliserons à nouveau un cadre à correction d'erreur, estimé en une seule étape.

L'effet de substitution apparaît à travers la moyenne des taux d'utilisation au cours des deux derniers trimestres (les effets dynamiques et à long terme ne sont pas séparés). La compétitivité des prix utilise la même technique.

Une tendance importante a dû être introduite et sera abandonnée dans le futur. Il est possible que la mesure de la demande mondiale soit biaisée.

```
genr x_ec=0
smp1 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

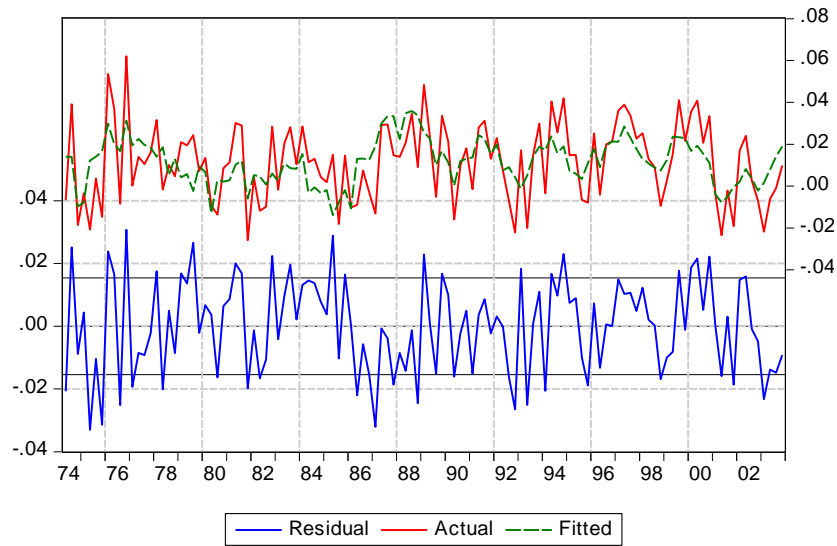
Date: 11/07/12 Time: 20:46

Sample (adjusted): 1974Q2 2003Q4

Included observations: 119 after adjustments

$$\text{DLOG(X)} = \text{C_X(1)} * \text{DLOG(WD)} + \text{C_X(2)} * \text{LOG(X(-1)/WD(-1))} + \text{C_X(3)} * 0.5 * (\text{LOG(UR)} + \text{LOG(UR(-1)))) + \text{C_X(4)} * 0.5 * (\text{LOG(COMPX(-1))} + \text{LOG(COMPX)}) + \text{C_X(5)} + \text{C_X(6)} * (\text{T}-2005) * (\text{T} \leq 2005)$$

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.385320	0.083551	4.611796	0.0000
C_X(2)	-0.230013	0.050940	-4.515338	0.0000
C_X(3)	-0.232284	0.087825	-2.644856	0.0093
C_X(4)	-0.183650	0.052253	-3.514653	0.0006
C_X(5)	-0.034008	0.016122	-2.109421	0.0371
C_X(6)	-0.007193	0.001669	-4.309415	0.0000
R-squared	0.340963	Mean dependent var		0.011474
Adjusted R-squared	0.311802	S.D. dependent var		0.018593
S.E. of regression	0.015424	Akaike info criterion		-5.456639
Sum squared resid	0.026884	Schwarz criterion		-5.316515
Log likelihood	330.6700	Hannan-Quinn criter.		-5.399739
F-statistic	11.69246	Durbin-Watson stat		2.024301
Prob(F-statistic)	0.000000			



Le cas Cobb-Douglas donne des résultats similaires avec un autre UR:

Dependent Variable: DLOG(X)

Method: Least Squares

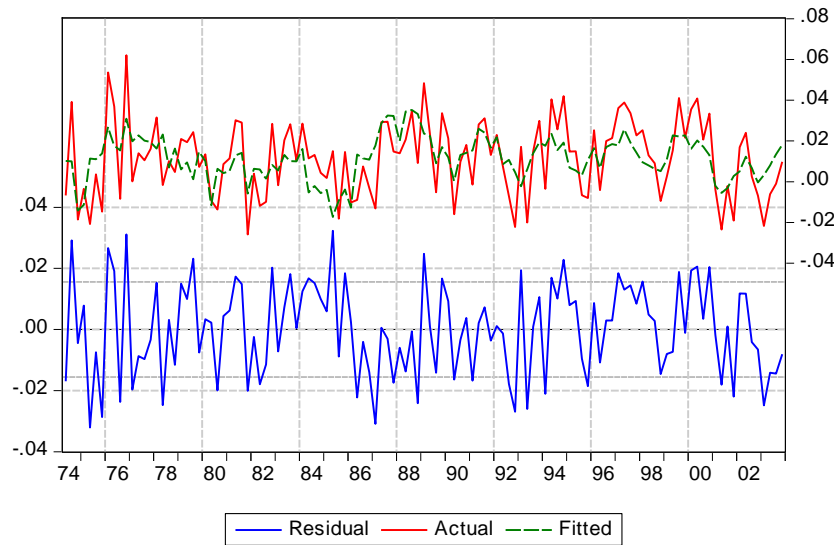
Date: 11/07/12 Time: 19:18

Sample (adjusted): 1974Q2 2003Q4

Included observations: 119 after adjustments

$$DLOG(X) = C_X(1) * DLOG(WD) + C_X(2) * LOG(X(-1)/WD(-1)) + C_X(3) * 0.5 * (LOG(UR) + LOG(UR(-1))) + C_X(4) * 0.5 * (LOG(COMPX(-1)) + LOG(COMPX)) + C_X(5) + C_X(6) * (T - 2005) * (T \leq 2005) + X_EC$$

	Coefficient	Std. Error	t-Statistic	Prob.
C_X(1)	0.416462	0.084406	4.934023	0.0000
C_X(2)	-0.188997	0.048520	-3.895210	0.0002
C_X(3)	-0.221336	0.136816	-1.617768	0.1085
C_X(4)	-0.199217	0.065871	-3.024365	0.0031
C_X(5)	-0.046719	0.023643	-1.976001	0.0506
C_X(6)	-0.007443	0.002100	-3.544910	0.0006
R-squared	0.316007	Mean dependent var		0.011474
Adjusted R-squared	0.285742	S.D. dependent var		0.018593
S.E. of regression	0.015714	Akaike info criterion		-5.419471
Sum squared resid	0.027902	Schwarz criterion		-5.279347
Log likelihood	328.4585	Hannan-Quinn criter.		-5.362571
F-statistic	10.44128	Durbin-Watson stat		2.071681
Prob(F-statistic)	0.000000			



10.2.3.4.10 Les importations

Pour les importations, nous utilisons le même cadre, mais:

- nous avons dû fixer le coefficient dynamique de la demande à l'unité ;
- le taux d'utilisation n'est pas retardé ;
- la compétitivité des prix est mesurée sur 6 périodes.

coef(10) c_m

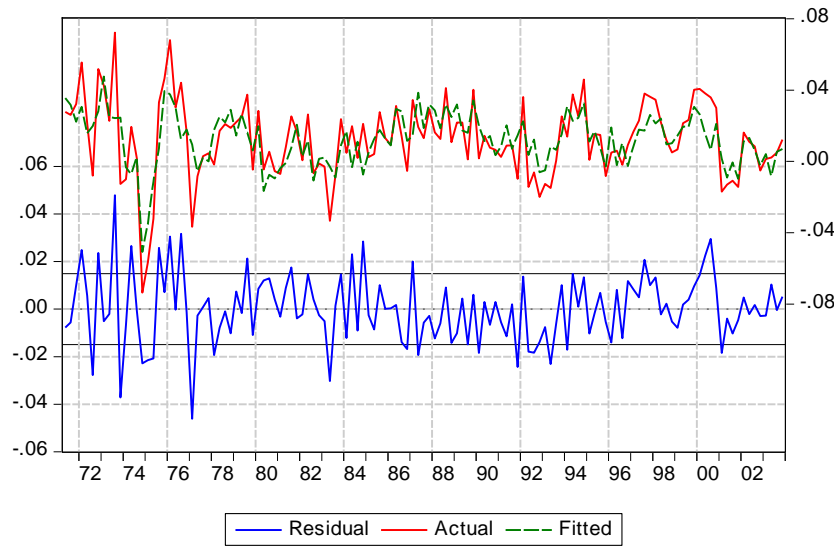
```

smp1 1962Q1 2004Q4
genr m_ec=0
smp1 1962Q1 2003Q4
equation _eq_m.ls(p) dlog(m)=dlog(fd+tc*q)+c_m(2)*log(ur)+c_m(3)*log(@movav(compm,6))+c_m(4)
+c_m(5)*(t-2005)*(t<=2005)+[ar(1)=c_m(6)]+c_m(7)*log(m(-1)/(fd(-1)+tc(-1)*q(-1)))+m_ec
_eq_m.resids(p)
close _eq_m
genr resid_m=resid

```

Dependent Variable: DLOG(M)
Method: Least Squares
Date: 11/07/12 Time: 20:49
Sample (adjusted): 1971Q2 2003Q4
Included observations: 131 after adjustments
Convergence achieved after 1 iteration
DLOG(M)=DLOG(FD+TC*Q)+C_M(2)*LOG(UR)+C_M(3)
LOG(@MOVAV(COMP6,6))+C_M(4)+C_M(5)(T-2005)*(T<=2005)
+[AR(1)=C_M(6)]+C_M(7)*LOG(M(-1)/(FD(-1)+TC(-1)*Q(-1)))

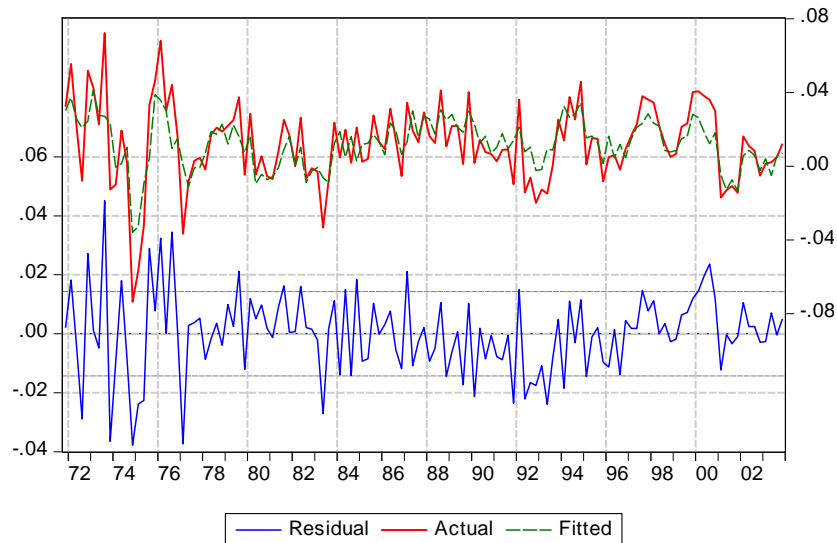
	Coefficient	Std. Error	t-Statistic	Prob.
C_M(2)	0.854879	0.259322	3.296598	0.0013
C_M(3)	-0.219006	0.101418	-2.159443	0.0327
C_M(4)	-1.364308	0.175134	-7.790062	0.0000
C_M(5)	0.017528	0.002816	6.225037	0.0000
C_M(7)	-0.782949	0.097880	-7.999036	0.0000
C_M(6)	0.889459	0.052156	17.05367	0.0000
R-squared	0.545765	Mean dependent var		0.012263
Adjusted R-squared	0.527596	S.D. dependent var		0.021698
S.E. of regression	0.014914	Akaike info criterion		-5.528350
Sum squared resid	0.027802	Schwarz criterion		-5.396661
Log likelihood	368.1069	Hannan-Quinn criter.		-5.474839
Durbin-Watson stat	2.113970			
Inverted AR Roots	.89			



Le cas Cobb-Douglas donne des résultats différents, et moins significatifs.

Dependent Variable: DLOG(M)
Method: Least Squares
Date: 11/07/12 Time: 19:18
Sample (adjusted): 1971Q4 2003Q4
Included observations: 129 after adjustments
Convergence achieved after 11 iterations
DLOG(M)=DLOG(FD+0.5*X)+C_M(2)*LOG(UR)+C_M(3)
LOG(@MOVAV(COMP,6))+C_M(4)+C_M(5)(T-2005)*(T<=2005)
+[AR(1)=C_M(6)]+C_M(7)*LOG(M(-1)/(FD(-1)+0.5*X(-1)))+M_EC

	Coefficient	Std. Error	t-Statistic	Prob.
C_M(2)	0.109425	0.154100	0.710092	0.4790
C_M(3)	-0.056667	0.031988	-1.771487	0.0790
C_M(4)	-0.250145	0.093434	-2.677242	0.0084
C_M(5)	0.004009	0.001571	2.551585	0.0119
C_M(7)	-0.201193	0.071715	-2.805436	0.0058
C_M(6)	0.390800	0.121756	3.209690	0.0017
R-squared	0.575925	Mean dependent var		0.012036
Adjusted R-squared	0.558686	S.D. dependent var		0.021789
S.E. of regression	0.014475	Akaike info criterion		-5.587442
Sum squared resid	0.025770	Schwarz criterion		-5.454428
Log likelihood	366.3900	Hannan-Quinn criter.		-5.533396
Durbin-Watson stat	2.089100			
Inverted AR Roots	.39			



10.2.3.5 Une nouvelle vérification de la cohérence équations - données

Maintenant que les équations ont été estimées, nous devons vérifier à nouveau la cohérence des données et des équations, cette fois sur toutes celles-ci.

Les identités doivent rester vraies, et en ajoutant les résidus d'estimation à la fin des équations de comportement elles doivent aussi donner les valeurs historiques.

Donc, ce processus pourrait sembler inutile. Ce sera généralement le cas si toutes les estimations ont été établies juste avant le contrôle, ce qui signifie qu'elles ont été regroupées, ainsi que la définition des identités, dans le même programme. C'est en fait la meilleure

technique, car elle permet de produire un modèle en une seule fois, et non à travers une séquence (qui nécessite une organisation et peut conduire à des erreurs).

Mais si les estimations datent quelque peu, et si l'on n'est pas sûr que les données n'ont pas été modifiées depuis, des erreurs ont pu apparaître. Et de toute façon, cette vérification est presque gratuite (contrôler que tous les résidus sont négligeables peut être fait par un seul programme, avec une seule réponse globale). Il n'y a donc aucune raison de s'en passer.

Bien sûr, notre exemple donne des résultats acceptables²⁶.

smpl 1995Q2 2004Q4

```
' Maintenant, nous calculons les résidus
' pour toutes les endogènes
' Le processus est le même qu'auparavant
' Les dc_ {%1} sont les différences absolues
' Les pc_ {%1} celles relatives
' Pour éviter une division par zéro, nous utilisons une astuce
' Nous ajoutons à la variable une valeur booléenne
' Nous testerons si la variable est égale à zéro
' Si c'est vrai que nous obtenons 0 divisé par 1 = 0
' Sinon le calcul n'est pas affecté
```

²⁶ Mais obtenir ce résultat a pris un certain temps ...

```

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

```

10.2.3.6 La résolution du modèle sur le futur

Nous arrivons maintenant à la partie cruciale du test du modèle, l'observation de la manière dont il se comporte sur le terrain où il sera utilisé: le futur.

Globalement, la technique est la même que celle expliquée précédemment. Les différences sont minimes.

Nous commençons par les déclarations habituelles: changement de répertoire, création d'un fichier spécial étendu appelé «proj_1».

```
cd "c:\Program Files\evIEWS5\book\fra_cf"
```

```
' Nous avons transféré les informations vers proj_1, trimestriel de 1962 à 2100
```

```
close proj_1.wf1
```

```
close mod_1.wf1
```

```
open mod_1.wf1
```

```
save proj_1
```

```
pagestruct(end=2100Q4) *
```

Maintenant, nous utilisons des déflateurs. Pour les hypothèses à long terme, nous devons inclure leur taux de croissance:

```
' Nous définissons des taux de croissance à long terme
scalar txq=0.05
scalar txn=0.02
scalar txp=0.006
```

Il est évident que l'ensemble des variables exogènes, et des hypothèses à définir, est plus important.

Ce programme simule le modèle sur le futur

```
' Nous étendons les séries pour la projection
' Commençons par les variables exogènes
' Croissance nulle
smpl 2005Q1 2100Q4
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_IGT 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
genr {%1}={%1}(-1)
next
' Croissance comme le PIB
for %1 COG IG WD R_REVX HIH
genr {%1}={%1}(-1)*(1+txq)
```

```

next
‘ Croissance comme la population
for %1 LG POP65 POP
  genr {%1}={%1}(-1)*(1+txn)
next
‘ Croissance comme les déflateurs
for %1 PPX
  genr {%1}={%1}(-1)*(1+txp)
next
‘ Croissance spéciale
for %1 SOCBR
  genr {%1}={%1}(-1)*(1+txq)/(1+txn)
next
genr t=t(-1)+0.25

```

Il sera aussi utile (mais pas indispensable) de créer des valeurs pour les endogènes:

```

smpl 2005Q1 2100Q4
‘ Nous étendons les séries pour la projection: on initialise les variables endogènes
‘ Croissance nulle
for %1 COMPM COMPX ER FCAPGP IR IRL IRM IRS RCVAL RCVOL RES_WR RMARG RPROB RPROF TTRAD
UNR UR
  genr {%1}={%1}(-1)
next

```

```

' Croissance comme le PIB
for %1 CAP COH FD HRDI I CI K M Q X GDPM IC
  genr {%1}={%1}(-1)*(1+txq)
  next
' Croissance comme la population
for %1 LFD LF LT POPAC UN popt
  genr {%1}={%1}(-1)*(1+txn)
  next
' Croissance comme les déflateurs
for %1 PCOH PFD PFDXT PI PIG PM PMT PP PQ PX UWC
  genr {%1}={%1}(-1)*(1+txp)
  next
  next
' Croissance comme le PIB en valeur
  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
  genr {%1}={%1}(-1)*(1+txp)*(1+txq)
  next
' Croissance comme la productivité du travail
for %1 PL PLT
  genr {%1}={%1}(-1)*(1+txq)/(1+txn)
  next
' Croissance comme la productivité du travail à prix courants

```



```

for %1 WR
genr {%1}={%1}(-1)*(1+txp)*(1+txq)/(1+txn)
next

```

Mais la question essentielle est la garantie qu'un sentier stationnaire sera obtenu.

10.2.3.6.1 Éléments généraux

Les éléments développés ci-dessus s'appliquent aussi au cas actuel.

Mais la plus grande complexité du problème rend les méthodes proposées précédemment plus efficaces sinon nécessaires:

- faire des chocs sur les principales hypothèses et en observer les conséquences ;
- exclure certaines variables, ou modifier certains comportements.

10.2.3.6.2 Favoriser la convergence à long terme: de nouveaux éléments

Revenons aux contraintes que le modèle doit suivre, et voyons comment les principes ci-dessus s'appliquent à un cas plus complexe.

Tout d'abord, il faut prendre soin d'éliminer toute tendance restante.

Comme nous l'avons dit précédemment, toutes les élasticités devraient être unitaires dans les équations estimées, sauf bien sûr pour les éléments sans dimension (ratios). La méthode la plus simple est évidemment d'utiliser un cadre à correction d'erreurs: les contraintes peuvent être définies dans les équations de cointégration (les estimer en un seul passage ne modifie pas cette propriété, juste la validité de la méthode). Ils peuvent être libérés dans le VAR (l'équation dynamique), sans frais pour la convergence, en principe.

On observera que presque tous nos équations estimées, même si la cointégration n'a été appliquée qu'une seule fois, contiennent une expression de long terme liant des rapports sans dimension. Les seules exceptions sont les suivantes:

- Les prix du commerce où aucune différence n'est faite entre court terme et long terme, mais l'homogénéité est obtenue par des contraintes sur les coefficients.
- Pour les investissements:
 - dans le cas des facteurs complémentaires, la plupart des éléments représentent des dérivées (sans dimension). Les seules exceptions sont le taux d'utilisation et le taux de profits, deux ratios qui devraient se stabiliser sur le long terme. Toutefois, le taux d'utilisation n'est plus fixé comme dans le cas antérieur ;
 - dans le cas Cobb-Douglas, le ratio capital- valeur ajoutée est relié à un ratio de prix (du travail et du capital).
- Le seul vrai problème réside dans l'emploi, qui revêt une dimension particulière.

Nous avons déjà traité le cas des facteurs complémentaires. L'évolution de la productivité du travail y est estimée. À long terme, les deux tendances de l'emploi (une population) et de la valeur ajoutée (une variable à prix constants) sont fixées, donc l'évolution de la productivité du travail, avec une valeur qui doit remplacer celle estimée²⁷.

Pour le cas Cobb-Douglas, trois éléments sont pris en compte pour définir la capacité de production:

- l'évolution de la productivité globale des facteurs, estimée jusqu'ici, croissant au taux a
- la contribution du travail: $\alpha \cdot txn$
- la contribution du capital: $(1 - \alpha) \cdot txq$

Fondamentalement, le taux de croissance total de la capacité devrait être le même que pour toutes les variables à prix constants, ce qui donne:

$$txq = a + \alpha \cdot txn + (1 - \alpha) \cdot txq$$

²⁷ Ou l'une des tendances théoriques être fixée à la valeur estimée.

$$a = \alpha \cdot (txq - txn)$$

Cette formule est facile à interpréter: les unités de capital auront une productivité constante (rappelez-vous que la valeur à prix constants intègre l'augmentation de la qualité). Mais pas les unités de travail. Pour obtenir leur part dans l'augmentation de la capacité, leur productivité doit augmenter comme la différence entre la croissance des quantités et la croissance de la main-d'œuvre.

De cette façon, la croissance totale de la capacité va respecter la contrainte.

Pour présenter la question de manière plus claire, il pourrait être préférable de considérer, non pas l'emploi, mais des «unités efficaces» pour lesquelles chaque travailleur sera mesuré par sa contribution potentielle à la production, tout comme le capital. La valeur globale de ces unités à prix constants augmentera à la même vitesse que la production, et il n'y aura pas besoin d'introduire une tendance de productivité supplémentaire.

Une fois que ces conditions sont définies, le modèle doit avoir une solution à long terme, et les formulations (en particulier les comportements de correction d'erreur) doivent faire converger le modèle vers celle-ci.

10.2.3.6.3 Faire converger le modèle à court terme

Les problèmes sont les mêmes que dans le cas le plus simple, mais

- Le nombre d'éléments et de mécanismes perturbateurs augmente, et donc en particulier la probabilité de cycles éventuellement parasites.
- Les déflateurs sont généralement plus volatiles que les quantités (ils sont moins accrochés aux équilibres réels).
- Le risque de commettre une erreur sur des hypothèses est plus élevé, car le processus de définition est plus complexe: chercher à produire une prévision précise sur un futur partiellement connu va nous amener à définir explicitement certains éléments détaillés.
- Les tendances qui doivent être bloquées sont plus nombreuses, et elles peuvent apparaître dans des équations plus complexes.
- Enfin, il est possible que la gestion des erreurs des périodes de prévision initiales doive traiter de fortes variations, et en particulier des cycles importants.

10.2.3.6.4 Faire converger le modèle à moyen et long terme

Encore une fois, le principal danger réside dans les cycles, avec une probabilité plus élevée.

Mais le fait que le modèle a convergé sur les périodes précédentes nous fournit des informations importantes, ainsi que la possibilité d'effectuer des tests supplémentaires, comme nous l'avons déjà vu.

10.2.3.6.5 Éléments spécifiques aux prévisions

Maintenant que nous avons fait converger le modèle sur toute la période, on peut tester ses propriétés par le biais des réponses aux chocs. Ensuite, nous pourrons passer aux véritables prévisions.

10.2.3.7 La production d'une prévision

Nous vérifions que les taux de croissance convergent vers les valeurs théoriques.

La technique est la même que d'habitude. Ici, nous utilisons l'algorithme de Gauss-Seidel.

```
' Nous résolvons le modèle dans l'avenir (sans choc)
smpl 2005Q1 2100Q4
' La lettre associée aux simulations de base est b
_mod_1.append assign @all _b
Nous utilisons la méthode de Gauss-Seidel
_mod_1.solveopt(n=t m=1001,c=1e-6,o=g,d=d)
_mod_1.scenario "scenario 1"
' Nous nous assurons que le modèle complet est résolu

_mod_1.exclude
```

```

smpl 2005Q1 2100Q4
solve(n=t m=1001 c=1e-6 o=g d=d) _mod_1
' Nous calculons les taux de croissance
for !i=1 to _g_vendo.@count
%st1=_g_vendo.@seriesname(!i)
smpl 2005Q1 2100Q4
genr {%st1}_tc= 100*({%st1}_1-{%st1}_1(-1))/({%st1}_1(-1)+({%st1}_1(-1)=0))
genr {%st1}_tc0= 100*({%st1}-{%st1}(-1))/({%st1}(-1)+({%st1}_1(-1)=0))
smpl 2005Q1 2100Q4
series {%st1}_DTC=100*({%st1}_tc-{%st1}_tc0)
next

```

10.2.3.8 La production des chocs d'hypothèse

Par rapport à l'exemple précédent, le processus sera un peu plus complexe. En particulier, nous allons considérer sept chocs.

Les commentaires sont insérés dans le programme.

```
' Maintenant, nous allons produire un ensemble de chocs (7 dans le cas présent)
```

```
smpl 2000 2200
```

```
' Le groupe appelé shocks_v contiendra la liste des 7 variables choquées
```

```
group shocks_v ig erx r_vat r_tar r_tarx wd m_ec
```

- ' Le groupe a appelé shocks_l contient 7 lettres associées à chaque choc
- ' Mais jusqu'à EViews 6 les lettres doivent être connues comme des séries
- ' car les groupes EViews ne pouvaient contenir que des séries
- ' Nous créons une série artificielle
- ' Mais seulement pour les éléments absents

```
for %z g r t f y w n j
if @isobject(%z) =0 then
genr {%z}=na
endif
next
```

- ' Nous créons le groupe

```
group shocks_l g r t f y w n j
```

- ' Maintenant, nous calculons le changement supplémentaire pour chaque hypothèse
 - ' en utilisant le nom de la variable
 - ' Le choc ne commencera qu'en 2006
 - ' laissant une période non choquée
 - ' pour vérifier que la différence ne vient que du choc
 - ' Le nom de la variable choquée combinera celui de la variable originale et la lettre
- '1 - chocs g: +1 point de PIB sur la demande du gouvernement

genr ig_g=ig+.01*gdpm_b*(t>=2006)

'2 – Choc r: dévaluation de 1 % de l'euro

genr erx_r=erx*(1+.01*(t>=2006))

'3 – Shock t: -1 point de taux de TVA

genr r_vat_t=r_vat-.01*(t>=2006)

'4 – Choc f: -1 point sur le taux de droits de douane local

genr r_tar_f=r_tar-.01*(t>=2006)

'5 – Choc y: -1 point sur le taux de droits de douane étranger

genr r_tarx_y=r_tarx-.01*(t>=2006)

'6 - Choc w: +1 % sur la demande mondiale adressée à la France

genr wd_w=wd*(1+.01*(t>=2006))

'7 – Choc n: +1 % sur les importations ex-ante

```
genr m_ec_n=m_ec+.01*(t>=2006)
```

'La boucle sur les chocs

```
for !j=1 to shocks_v.@count
```

```
smp1 2005Q1 2100Q4
```

‘ Nous obtenons le nom de la variable de choc et la lettre associée

```
%2=shocks_v.@seriesname(!j)
```

```
%3=shocks_!j.@seriesname(!j)
```

‘ Nous avons établi le suffixe solution en utilisant la lettre

‘ Nous contrôlons qu'aucune endogène n'est exclue de la simulation

‘ Nous remplaçons la variable associée au choc

‘ Nous résolvons le modèle

```
_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)
```

‘ Nous calculons la différence par rapport à la simulation de base, en termes absolus et relatifs


```

for !i=1 to _g_vendo.@count
%st1=_g_vendo.@seriesname(!i)
series d{%3}_{%st1}={%st1}_{%3}-{%st1}_b
series p{%3}_{%st1}=100* d{%3}_{%st1}/({%st1}_b+ ({%st1}_b=0))
series dv_{%st1}=d{%3}_{%st1}
series pv_{%st1}=p{%3}_{%st1}
next

```

- ' Nous créons des groupes de variations pour une liste spécifique de variables importantes
- ' Le nombre est limité, afin d'établir un tableau lisible

```

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

```

- ' ou très importantes
- ' affichées dans un graphique
- ' Le nombre est encore restreint pour rendre le graphe lisible

```

group g_w{%3} P{%3}_FD P{%3}_gdpm P{%3}_X P{%3}_M P{%3}_PQ

```

- ' Nous créons un groupe de variations pour toutes les variables

```

group g_v2 pV_* dV_*

```

Nous conservons tous les chocs dans les fichiers Excel (horizons 2025 et 2100)

' en utilisant la lettre associée au choc

```
smp1 2005Q1 2025Q4  
write(t=xls) v2_{%3}.xls g_v2  
smp1 2005Q1 2100Q4  
write(t=xls) w2_{%3}.xls g_v2  
next
```

10.2.4 L'APPLICATION DES PROGRAMMES

Nous allons maintenant présenter les résultats obtenus à partir des programmes précédents.

10.2.4.1 La production du cadre du modèle

Fondamentalement, le programme crée un objet appelé MOD_1 dans le fichier de travail. Bien sûr, ce modèle ne peut être résolu à l'heure actuelle. Toutefois, l'accès à cet élément donne

- la liste des variables,
- la liste des équations,
- le code source original,
- la structure par blocs.

La liste des équations et le code répètent le modèle en l'état, et sont de peu d'intérêt. La liste des variables, comme indiqué précédemment, permet de localiser les problèmes tels que des variables endogènes dont l'équation est absente (elles apparaissent comme exogènes) ou des fautes de frappe dans les noms de variables (elles créent également un élément exogène).

Le dernier élément est le plus intéressant. Montrons le résultat pour notre modèle. D'abord pour la version à facteurs complémentaires.

Number of equations: 85

Number of independent blocks: 5

Number of simultaneous blocks: 2

Number of recursive blocks: 3

Block 1: 2 Recursive Equations

ict(40) ifp(51)

Block 2: 55 Simultaneous Equations

gdp(1)

i(4)

pl(8)

popac(11)

uwc(14)

pdf(17)

pi(21)

pm(25)

irl(28)

relc(31)

w(34)

revx(37)

hdi(41)

qval(44)

marg(49)

q(2)

ci(6)

lt(9)

un(12)

pq(15)

pdfxt(18)

wr(23)

er(26)

ir(29)

wf(32)

socb(35)

scw(38)

hrdi(42)

vat(45)

nif(52)

ur(3)

lf(7)

cap(10)

unr(13)

pp(16)

pcoh(19)

px(24)

irs(27)

irm(30)

wg(33)

revq(36)

hi(39)

coh(43)

gdpmval(46)

prof(53)

rprof (54)	fcapf (56)	pmt (57)
compm (58)	fd (59)	m (60)
compx (61)	x (62)	mval (63)
xval (64)		

Block 3: 23 Recursive Equations

k (5)	pcog (20)	pig (22)
pgdpm (47)	subs (48)	rmarg (50)
rprob (55)	rcval (65)	rcvol (66)
ttrad (67)	trb (68)	nixl (69)
nixx (70)	nix (71)	fcapx (72)
scf (73)	oit (74)	tar (75)
scg (76)	revg (77)	igv (78)
cogv (79)	fdgv (80)	

Block 4: 3 Simultaneous Equations

nig (81)	expg (82)	fcapg (83)
----------	-----------	------------

Block 5: 2 Recursive Equations

fcapp (84)	gdpval (85)
------------	-------------

Ce qui précède montre que les 85 équations de notre modèle peuvent être décomposées en une séquence de cinq «blocs». Trois d'entre eux sont définis comme récursifs, deux comme simultanés.

Nous allons définir ces notions, qui sont associées au processus de résolution du modèle.

- Les blocs récursifs peuvent être ordonnés de telle manière qu'aucune variable n'apparaît dans une équation avant celle où elle est calculée. Il est évident que ce mode de calcul donnera à chaque variable sa valeur exacte, et après le calcul de chaque variable aucune autre action n'est nécessaire.
- Dans les blocs (bloc), chaque variable de l'ensemble dépend d'une variable calculée ultérieurement, soit directement, soit par l'intermédiaire d'une séquence d'influences. Par conséquent, cette variable ne prend pas sa valeur exacte immédiatement, sauf si la variable responsable de la non-récursivité a pris sa valeur exacte au début.

À partir de n'importe quel ordre, EViews est capable de construire un ensemble récursif initial, comme contenant des variables qui ne dépendent que de variables exogènes, retardées, ou de variables qui apparaissent auparavant dans le même ensemble. Nous appellerons cet ensemble un «prologue».

Il est également capable de construire un ensemble final récursif, contenant les variables qui n'influencent aucune variable dans le reste de l'ensemble, ou seulement des variables qui ont déjà été incluses dans l'ensemble. Nous appellerons cet ensemble un «épilogue».

Une fois cette opération réalisée, le reste du modèle pourrait être considéré comme simultané. Toutefois, cet ensemble peut éventuellement être séparé en deux blocs consécutifs, si aucune variable du deuxième ensemble n'influe sur n'importe quelle variable du premier. Le premier ensemble peut aussi avoir un épilogue, le second un prologue, qui peuvent être assemblés en un bloc intermédiaire récursif. Et le procédé peut être appliqué de nouveau aux deux blocs simultanés, jusqu'à ce qu'aucune partition ne soit plus possible.

Dans notre cas, nous observons:

- Un petit bloc initial récursif contenant ICT (impôt sur le revenu) et IFP (impôt sur les bénéfices des entreprises). Ils dépendent évidemment d'un taux d'imposition exogène et des revenus provenant de la période précédente.
- Un grand bloc (55 équations) contenant deux boucles interconnectées: la boucle keynésienne équilibrant l'offre et la demande, et la boucle prix-salaires. Si l'on prend deux variables quelconques de l'ensemble, on peut relier la première à la seconde, et la seconde à la première.
- Un grand bloc récursif (22 équations), contenant des éléments descriptifs (comme les termes de l'échange), les éléments de budget de l'État (montant des droits de douane) et des variables influençant seulement l'avenir (le capital).

- Un petit bloc non récursif de 3 équations, équilibrant pour le Gouvernement la liaison entre les dépenses (qui comprennent les intérêts versés), le solde budgétaire, et les intérêts eux-mêmes qui en dépendent. Ces éléments dépendent du reste du modèle, mais n'ont aucune influence, sauf bien sûr sur le solde en points de PIB, qui apparaît dans le dernier bloc (simultané), ainsi que le PIB global à prix courants qui est un élément purement descriptif.
- Ceci définit un bloc final de deux équations.

On peut s'interroger sur l'absence d'un bloc de 3 équations pour les entreprises, semblable à celui obtenu pour le gouvernement. Il est en fait intégré dans le bloc à 55 équations, car les profits influent sur leur comportement.

Ensuite, la version Cobb-Douglas

```

Number of equations: 88
Number of independent blocks: 7
Number of simultaneous blocks: 3
Number of recursive blocks: 4
Block 1: 1 Recursive Equations
ict(43)

Block 2: 51 Simultaneous Equations
gdpm(1)                q(2)                ur(3)
  kd(4)                lfd(5)                lf(6)
  ic(8)                lt(9)                id(10)
  i(11)                cap(13)                popac(14)
  un(15)                unr(16)                pq(18)
    
```

pp (19)	pfd (20)	pfdxt (21)
pc (22)	pi (23)	wr (26)
cost (27)	px (28)	pm (29)
er (30)	irs (31)	irl (32)
ir (33)	relc (35)	revq (36)
revx (37)	socb (38)	wg (39)
w (40)	scw (41)	hi (42)
hdi (44)	hrdi (45)	coh (46)
qval (47)	gdpmval (48)	wf (50)
pmt (60)	compm (61)	fd (62)
m (63)	compX (64)	x (65)
mval (66)	xval (67)	vat (76)

Block 3: 11 Recursive Equations

pl (7)	k (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

nif (55)	prof (56)	fcapf (59)
----------	-----------	------------

Block 5: 18 Recursive Equations

rprof (57)	rprob (58)	rcval (68)
rcvol (69)	ttrad (70)	trb (71)

634 LA MODELISATION ECONOMETRIQUE, PRINCIPES ET TECHNIQUES

```
nix1 (72)          nixx (73)          nix (74)
fcapx (75)         scf (77)           oit (78)
tar (79)           scg (80)          revg (81)
igv (82)           cgv (83)          fdgv (84)
```

Block 6: 3 Simultaneous Equations

```
nig (85)           expg (86)          fcapg (87)
```

Block 7: 1 Recursive Equations

```
fcappg (88)
```

Le nombre de blocs est maintenant de 7. Cela vient du fait que les bénéfices n'influencent plus les investissements comme une variable. Pour les entreprises, la boucle: les intérêts – profits – capacités de financement est maintenant déconnectée de la boucle principale, à l'instar du gouvernement.

10.2.4.2 La production des données

Nous espérons que les commentaires introduits dans ce programme parlent d'eux-mêmes. Fondamentalement, les données de l'OCDE sont utilisées pour créer les séries du modèle dans l'ordre, et les résultats sont stockés. Bien sûr, les éléments utilisés dans tout calcul doivent avoir été créés dans une précédente instruction (si la séquence avait défini un modèle, il devrait être récursif).

Pour adapter ce programme à un autre modèle en utilisant les mêmes concepts, nous vous conseillons de remplacer toutes les références à l'original «FRA_» par le nom spécifique dans le fichier original. Cela devrait produire une version presque correcte, à quelques exceptions près:

- les variables non disponibles qui devront être créées (un exemple fréquent est le capital) ;

- les variables connues en utilisant une définition différente. Ceci va changer la logique des instructions. Par exemple l'OCDE fournit deux définitions des salaires, comprenant ou non les cotisations sociales, mais n'individualise pas les cotisations elles-mêmes, qui devront donc être calculées par différence. Si elles sont disponibles, un transfert direct est possible.

10.2.4.3 La production des groupes et la vérification de la cohérence modèle-données

Ce programme est également auto-explicatif. Les éléments suivants sont produits:

- Des groupes pour les endogènes (séparées automatiquement entre identités et comportements) et les exogènes.

Comme il est expliqué dans le programme, le fait que les équations de comportement, et seulement elles, contiennent l'élément «f» est utilisé pour les séparer des identités.

Le programme va exécuter deux simulations résiduelles avec des valeurs différentes de «f», et identifier comme comportementales les variables pour lesquelles le résultat va changer.

- Un ensemble de différences absolues et relatives entre les séries chronologiques, et le résultat donné par l'équation associée.

Commentons un peu la vérification que le résidu est nul.

En fait, la valeur de la valeur résiduelle est rarement exactement nulle. En raison de la précision relativement faible de EViews, nous allons généralement obtenir une valeur du type 10^{-6} (pour cent, de sorte que la précision est de 8 chiffres). Une fois que le modèle atteint une certaine taille, il devient difficile de vérifier visuellement que la valeur est suffisamment faible (l'erreur peut concerner un nombre limité de périodes, ou même une seule).

Il y a au moins deux façons de traiter le problème, l'une immédiate mais limitée, l'autre plus difficile à implémenter mais plus efficace et plus exacte.

- La première consiste à produire un graphe EViews. Si tous les résidus sont négligeables, nous devrions obtenir un ensemble de mouvements browniens, avec de très faibles valeurs maximales supérieure et inférieure. Nous savons donc immédiatement s'il n'y a pas d'erreurs. Mais il est difficile d'identifier les coupables par un graphe (EViews limite le nombre de séries affichées).

- La seconde consiste à exporter les erreurs relatives vers une feuille Excel, trier le fichier pour une année donnée (autre que l'année de base qui est moins sujette à erreurs), et se concentrer sur les valeurs supérieures et inférieures. Le processus est répété jusqu'à ce que seules des erreurs négligeables subsistent.

10.2.4.4 L'estimation des équations

Nous avons déjà présenté ses éléments en détail.

10.2.4.5 La résolution du modèle sur le futur

Le processus de résolution donne les mêmes résultats que d'habitude. Si la convergence est atteinte (c'est notre cas), aucun message n'est produit. Sinon, nous avons décrit précédemment les techniques que l'on peut utiliser pour résoudre le problème.

10.2.4.6 La production d'une prévision

Dans notre cas, le but n'est pas de produire une prévision proprement dite, mais une simulation qui donne des résultats suffisamment acceptables pour être utilisés comme point de départ pour les chocs qui vont suivre.

Cette propriété peut être contrôlée en affichant les taux de croissance des principales variables à court et moyen termes.

Dans notre cas, le modèle converge normalement en utilisant Gauss-Seidel, vers des valeurs assez proches des historiques pour les premières périodes (nous avons seulement présenté la convergence des taux d'utilisation UR et du déflateur de la valeur ajoutée PQ).

Model: _MOD_1

Date: 06/12/11 Time: 16:00

Sample: 2005Q1 2020Q4

Solve Options:

Dynamic-Deterministic Simulation

Solver: Gauss-Seidel

Max iterations = 5000, Convergence = 1e-008

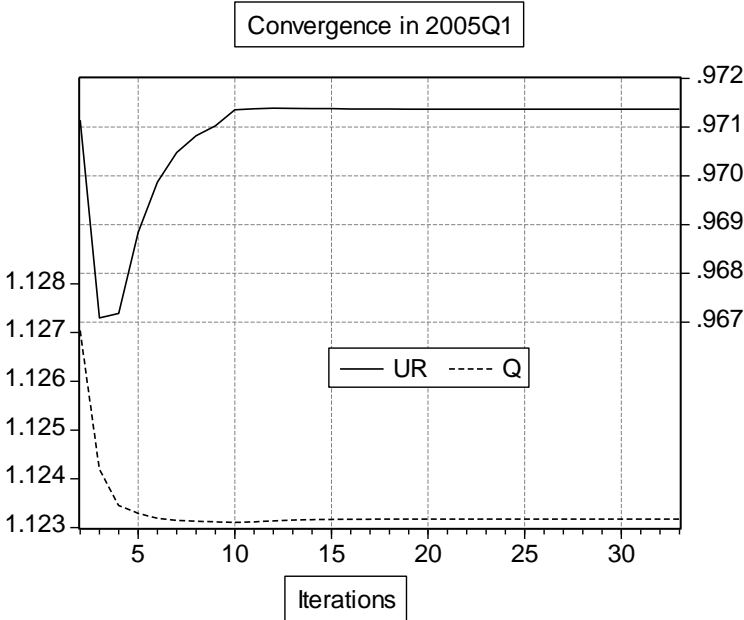
Scenario: Scenario 1

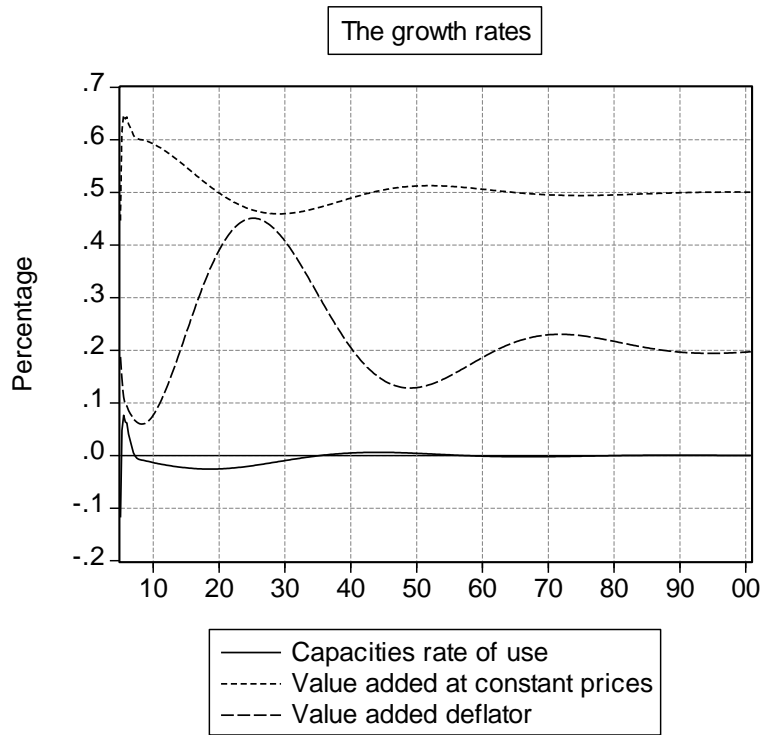
Solve begin 16:00:24

2005Q1	Block 1 - 1 eqns	Solved (recursive block)
2005Q1	Block 2 - 51 eqns	Convergence after 41 iterations
2005Q1	Block 3 - 11 eqns	Solved (recursive block)
2005Q1	Block 4 - 3 eqns	Convergence after 7 iterations
2005Q1	Block 5 - 18 eqns	Solved (recursive block)
2005Q1	Block 6 - 3 eqns	Convergence after 7 iterations
2005Q1	Block 7 - 1 eqns	Solved (recursive block)
2005Q2	Block 1 - 1 eqns	Solved (recursive block)
2005Q2	Block 2 - 51 eqns	Convergence after 33 iterations
2005Q2	Block 3 - 11 eqns	Solved (recursive block)
2005Q2	Block 4 - 3 eqns	Convergence after 7 iterations
2005Q2	Block 5 - 18 eqns	Solved (recursive block)
2005Q2	Block 6 - 3 eqns	Convergence after 8 iterations
2005Q2	Block 7 - 1 eqns	Solved (recursive block)
2005Q3	Block 1 - 1 eqns	Solved (recursive block)
2005Q3	Block 2 - 51 eqns	Convergence after 34 iterations
2005Q3	Block 3 - 11 eqns	Solved (recursive block)
2005Q3	Block 4 - 3 eqns	Convergence after 7 iterations
2005Q3	Block 5 - 18 eqns	Solved (recursive block)

2005Q3	Block 6 - 3 eqns	Convergence after 7 iterations
2005Q3	Block 7 - 1 eqns	Solved (recursive block)
2020Q1	Block 1 - 1 eqns	Solved (recursive block)
2020Q1	Block 2 - 51 eqns	Convergence after 43 iterations
2020Q1	Block 3 - 11 eqns	Solved (recursive block)
2020Q1	Block 4 - 3 eqns	Convergence after 9 iterations
2020Q1	Block 5 - 18 eqns	Solved (recursive block)
2020Q1	Block 6 - 3 eqns	Convergence after 8 iterations
2020Q1	Block 7 - 1 eqns	Solved (recursive block)
2020Q2	Block 1 - 1 eqns	Solved (recursive block)
2020Q2	Block 2 - 51 eqns	Convergence after 44 iterations
2020Q2	Block 3 - 11 eqns	Solved (recursive block)
2020Q2	Block 4 - 3 eqns	Convergence after 9 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 44 iterations
2020Q3	Block 3 - 11 eqns	Solved (recursive block)
2020Q3	Block 4 - 3 eqns	Convergence after 10 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 44 iterations
2020Q4	Block 3 - 11 eqns	Solved (recursive block)
2020Q4	Block 4 - 3 eqns	Convergence after 9 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)





Le dernier graphique présente la convergence des taux de croissance du taux d'utilisation (Rate of use), de la valeur ajoutée (value added) à prix constants, et de son déflateur.

Nous pouvons voir que ces derniers se stabilisent après 100 ans. Mais les quantités et l'écart production-capacités le font assez rapidement, tandis que le déflateur conserve ses oscillations pendant une longue période.

10.2.4.7 Les chocs d'hypothèses: le cas des facteurs complémentaires

Voyons maintenant comment nos modèles répondent aux chocs sur leurs hypothèses. Nous serons brefs, et présenterons et commenterons les résultats pour seulement quatre d'entre eux: notre objectif est surtout de montrer que le cadre actuel et les estimations associées peuvent fournir un modèle aux propriétés acceptables. Avec un peu de chance, vous devriez être capable de faire la même chose, ou peut-être mieux. Les modèles que nous vous présentons sont très loin d'être parfaits, et nous n'avons pas cherché à les rendre tels²⁸.

Comme présenté précédemment, nous envisageons deux fonctions de production: facteurs complémentaires et Cobb-Douglas. En utilisant les deux modèles associés, nous verrons que leurs propriétés diffèrent de façon substantielle, en raison principalement (mais pas seulement) de cette fonction.

Nous allons commencer par le premier cas.

Tous les chocs débiteront au premier trimestre de 2006, et seront maintenus pendant toute la période, jusqu'en 2100. Pour faciliter l'interprétation, nous allons limiter les périodes aux quinze premières années (pour montrer la dynamique à court et moyen terme) et aux trois dernières (pour les effets à long terme).

²⁸ En fait, on peut identifier quelques coefficients qui, avec des valeurs différentes, permettront d'améliorer légèrement les propriétés du modèle. Bien sûr, nous ne pourrons pas appliquer ces modifications.

10.2.4.7.1 Une augmentation de la demande de l'État

Dans ce choc, nous augmenterons la demande du gouvernement de 1 % du PIB de référence. C'est le premier choc qu'on a dans l'esprit, car il définit le multiplicateur keynésien, montrant comment les mécanismes économiques internes au pays modifient les conséquences d'un choc externe sur la demande. Techniquement, le multiplicateur sera le rapport des variations du PIB évaluées ex post et ex ante, cette dernière étant représentée par le changement de l'un des éléments exogènes de la demande.

Étudiant le rôle de la demande du gouvernement (investissement IG ou la consommation CG), nous pouvons voir qu'il ne touche d'abord que:

- la demande finale, et le PIB par l'équilibre offre- demande ;
- les dépenses du gouvernement et le déficit.

Comme indiqué précédemment, nous ne considérons pas l'augmentation de la productivité globale provenant de l'investissement du gouvernement.

Concentrons-nous sur l'équilibre offre-demande:

$$GDP + M = FD + X$$

avec:

$$FD = COH + IC + cg + ig + fdxr * Q$$

En termes de PIB, l'impact ex-ante du choc est bien sûr de 1 % (ce qui est la raison du choix de sa taille). L'évolution de la valeur ex post provient des éléments endogènes:

- éléments de commerce extérieur: les exportations et les importations;
- éléments de demande: la consommation, l'investissement, les variations de stocks.

Évidemment, nous aurons:

- Un effet positif sur la demande, car la nécessité de capacités supplémentaires accroît les investissements et crée des emplois, qui génèrent des salaires, des revenus et de la consommation des ménages.
- Mais les importations jouent un rôle négatif, car une part de de cette demande supplémentaire devra être importée.

Deux autres effets doivent être pris en compte:

- À court terme et moyen terme, la demande va faire face à une saturation des capacités locales pour certains produits, qui devront être importés. Et comme les producteurs locaux vont augmenter la part de leurs capacités satisfaisant la demande locale, ils seront moins actifs dans la recherche de marchés à l'exportation.

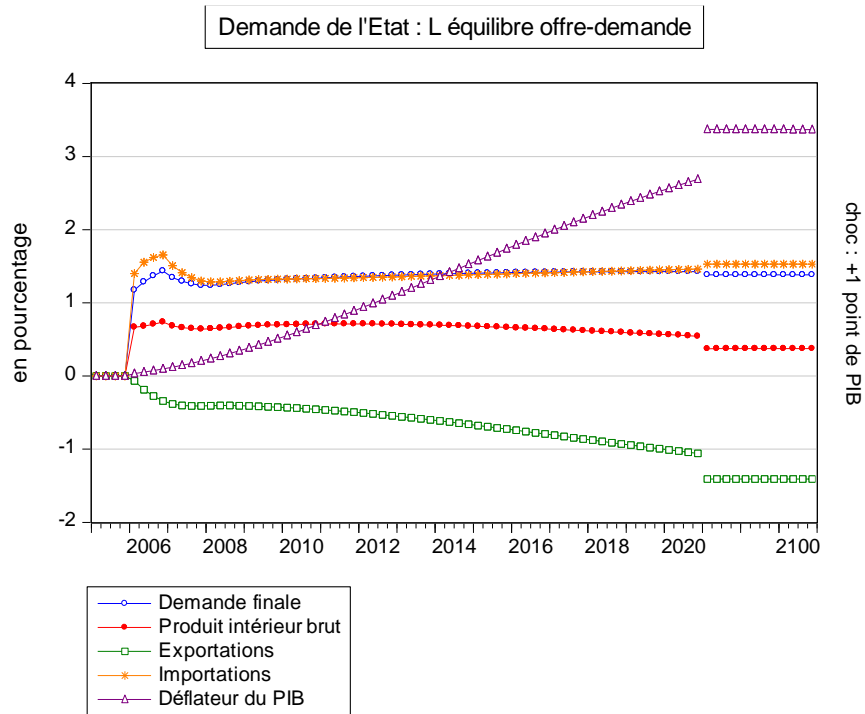
Cet effet disparaîtra progressivement à mesure que les producteurs locaux vont adapter leurs capacités (par l'investissement).

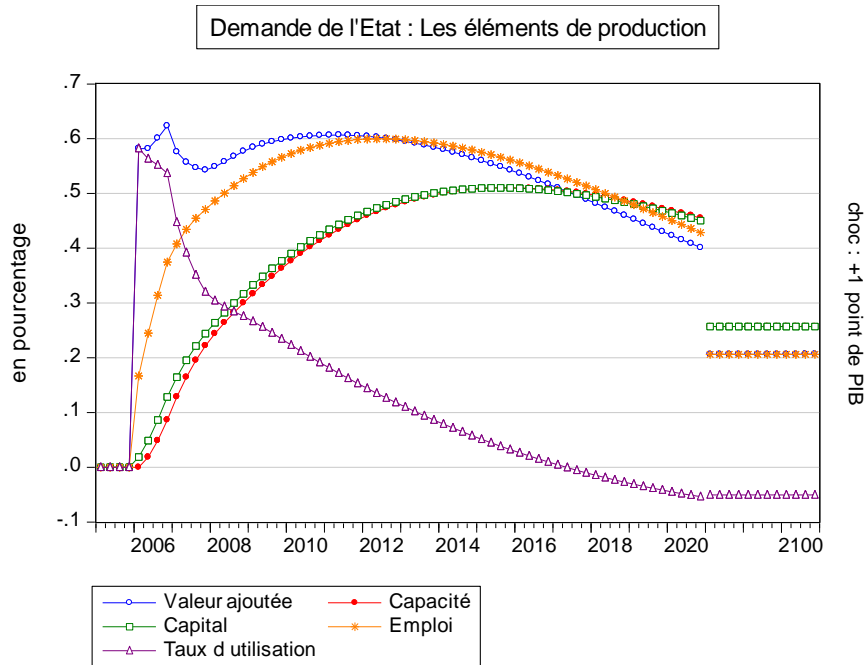
- Le taux d'inflation va évoluer, en raison de déséquilibres:
 - sur les capacités, car face à une augmentation des débouchés (par rapport aux capacités, par le biais du taux d'utilisation) le comportement d'optimisation des profits conduira les entreprises à augmenter leurs prix. Cet effet disparaîtra dans le long terme ;
 - sur les salaires, car la diminution du taux de chômage améliorera le pouvoir de négociation des travailleurs. Tant que le chômage restera réduit, cet effet persistera.

L'inflation réduira la compétitivité des producteurs locaux, sur les marchés étrangers et locaux.

Le graphique suivant illustre nos commentaires. Ex post la demande augmente de 1,5 % (contre 1 % ex ante, le solde commercial française étant à peu près à l'équilibre). Le multiplicateur est affecté dans à court et moyen termes par les goulots d'étranglement, dans le long

terme par des pertes de compétitivité. C'est à moyen terme que la perte est la plus faible (l'inflation est encore limitée, les capacités sont largement adaptées).



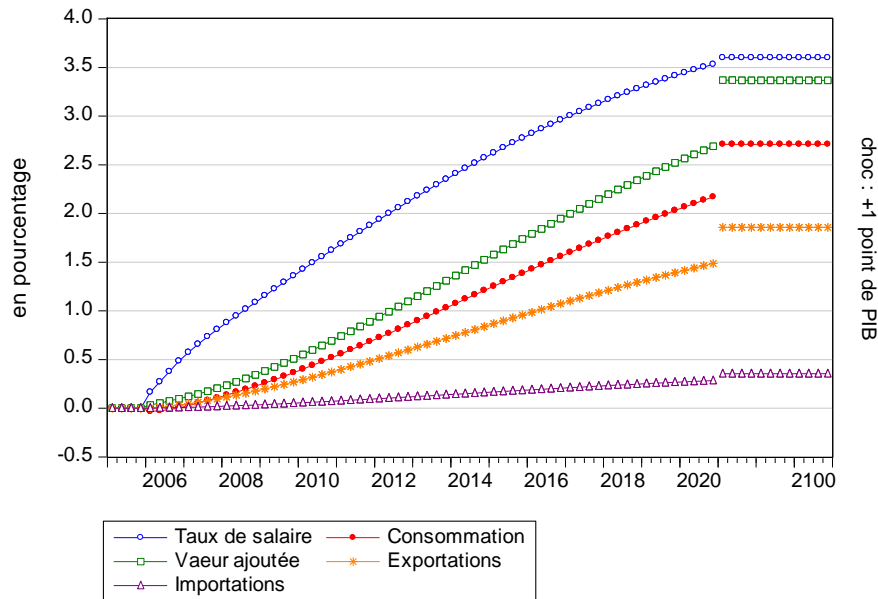


Les deux graphiques montrent l'adaptation des facteurs au nouveau niveau de production (avec des inerties différentes). Nous pouvons constater que:

- L'emploi s'adapte beaucoup plus rapidement (de façon cohérente avec les coefficients estimés).
- À long terme, le taux d'utilisation ne revient pas à la valeur de référence. Cela n'est pas trop difficile à expliquer. Comme le prix à l'exportation est plus sensible aux coûts locaux que le prix à l'importation, un changement dans le déflateur local à la valeur ajoutée freinera son impact sur le prix de la demande locale, défini par une identité.

$$PFD = (GDPMVAL + MVAL - XVAL) / (GDPM + M - X)$$

Demande de l'Etat : Les prix



L'augmentation modérée du prix de l'investissement, par rapport aux recettes qui le financent, va accroître la rentabilité et autorisera pour les entreprises un taux d'utilisation plus faible, suivant le mécanisme présenté dans l'estimation du déflateur de la valeur ajoutée.

Les faibles niveaux d'activité à long terme sont bien sûr dus à l'inflation, qui prend beaucoup de temps pour arriver à sa valeur d'équilibre (on peut constater que les coefficients de correction de l'erreur sont assez faibles).

En exécutant les programmes associés, le lecteur sera en mesure d'observer d'autres éléments, tels que la réduction ex post du coût budgétaire à court terme (en raison des divers revenus fiscaux supplémentaires), suivie par une augmentation (le gouvernement doit payer des intérêts sur la dette accumulée). Mais il peut également produire son propre modèle, et observer comment ses modifications (limitées ou étendues) vont modifier les propriétés du modèle.

10.2.4.7.2 une dévaluation (du franc français !)

Bien sûr, ce choc n'est pas réaliste, car la France partage avec d'autres pays sa monnaie actuelle (pour l'instant).

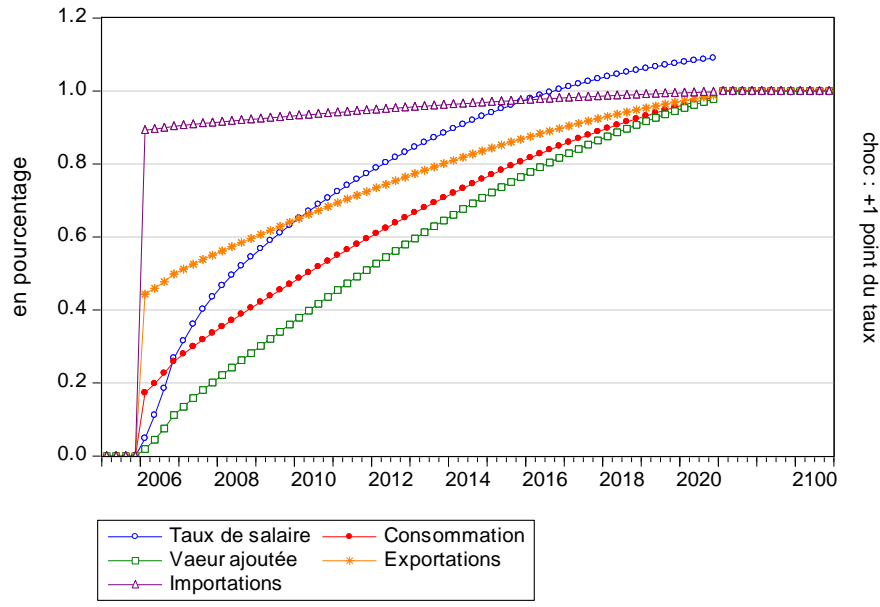
Nous devons donc supposer que la France est sortie de la zone Euro (ce qui conduirait certainement à un équilibre économique différent)

Nous allons montrer tout de même les résultats, pour illustrer les conséquences générales d'une dévaluation, en utilisant le cadre de modèle que nous proposons.

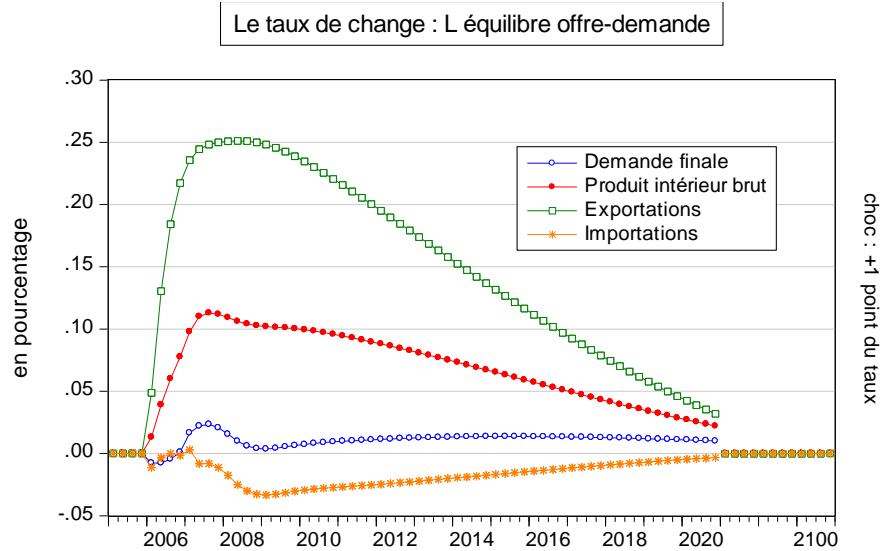
Nous allons déprécier la monnaie française (le franc ?) de 1 %.

Nos deux premiers graphiques présentent les prix et les quantités.

Le taux de change : Les prix



choc : +1 point du taux



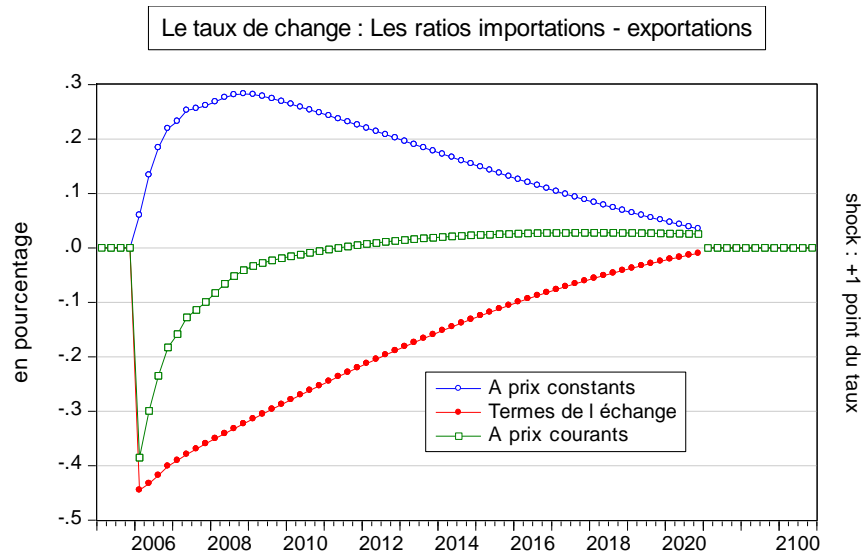
Sur le premier, nous voyons que, dans le long terme, tous les prix augmentent de 1%. Mais l'augmentation est lente, et présente quelques suradaptations à moyen terme (en raison de l'inertie provenant de l'erreur faible des coefficients de correction). Logiquement, les déflateurs des importations (essentiellement) et des exportations sont les plus rapides à prendre la décision en compte, et donc présentent le plus faible dépassement.

En ce qui concerne les quantités, les exportations vont profiter au maximum du choc, améliorant le PIB. Mais les importations vont augmenter un peu, les gains de compétitivités étant plus que compensés par le surcroît de demande (pas seulement de l'investissement et de la consommation locale, mais aussi des exportations, qui appellent à l'importation de biens intermédiaires).

À long terme, l'adaptation complète de l'inflation fait disparaître tous les effets, après une dégradation temporaire due à la suradaptation des déflateurs locaux.

Le dernier graphique présente l'évolution du commerce extérieur. On voit que les gains de compétitivité (associés à une perte sur les termes de l'échange) sont réduits progressivement à partir de la première période, mais que les gains sur les échanges à prix constants prendront un certain temps à atteindre leur maximum (en raison des goulets d'étranglement initiaux et de l'augmentation de la demande).

Un gain très limité aux prix courants n'apparaît que dans le moyen terme.



10.2.4.7.3 une diminution des droits de douane étrangers

Comme vous l'avez probablement deviné, ce choc est essentiellement axé sur la demande.

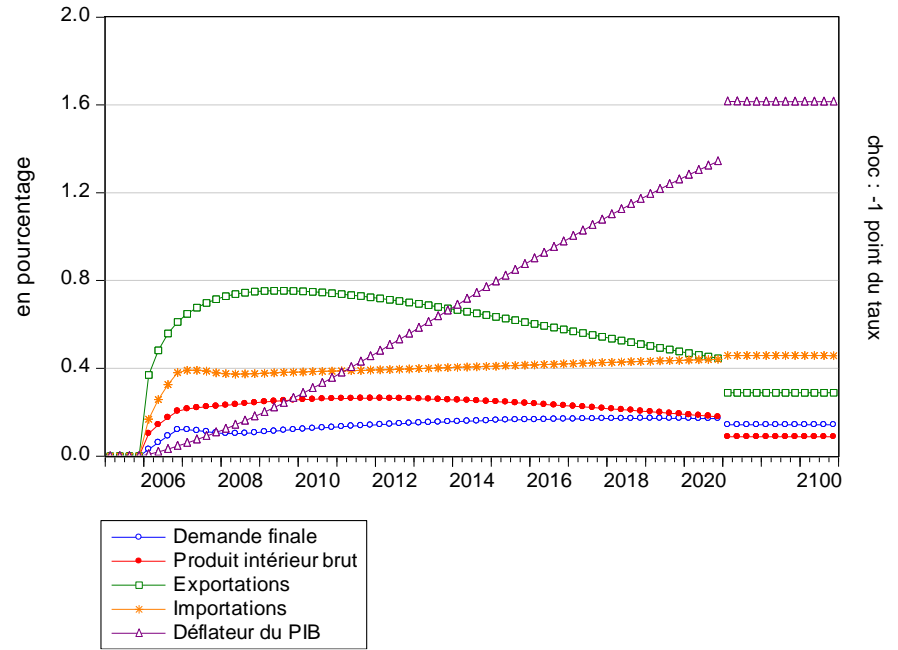
L'amélioration de la compétitivité va accroître la demande adressée à la France, avec essentiellement les mêmes effets sur le PIB que la demande du gouvernement, sur l'équilibre offre-demande.

Bien sûr, la baisse des prix à l'importation des produits français va réduire l'inflation dans le reste du monde, ce qui devrait influencer sur les prix des importations françaises. Le PIB mondial devrait aussi changer (d'une manière incertaine, positive grâce à la désinflation, négative à travers la croissance de la part de la France dans le commerce mondial). Mais ces deux effets peuvent être considérés comme négligeables, surtout si l'on considère le coût de leur prise en compte: construire un modèle mondial fiable.

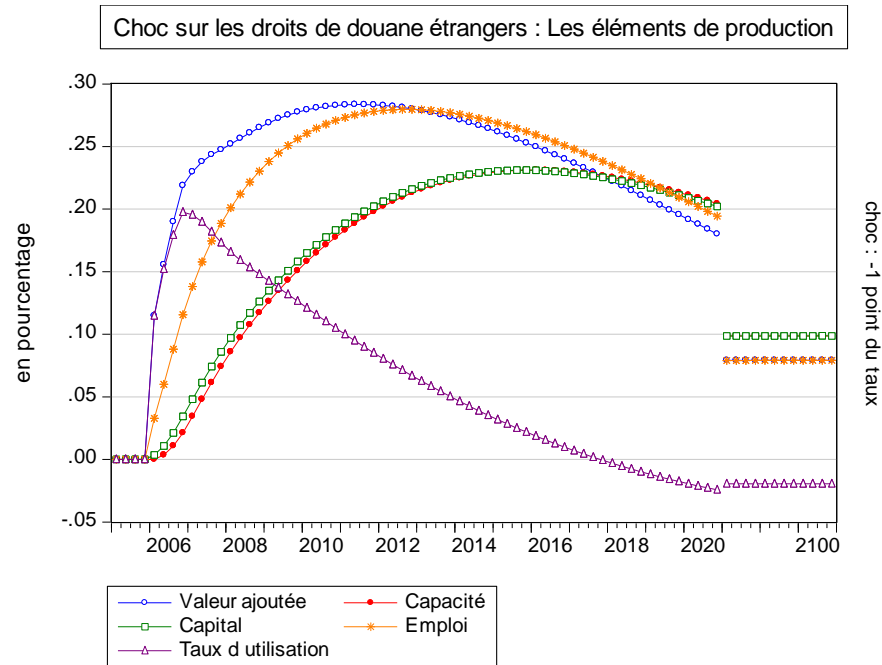
Il est intéressant de constater que l'augmentation de l'activité locale inverse le gain sur les échanges à prix constants à long terme, mais pas autant qu'à prix courants, cet écart étant dû à l'inflation locale, qui «améliore» les termes de l'échange.

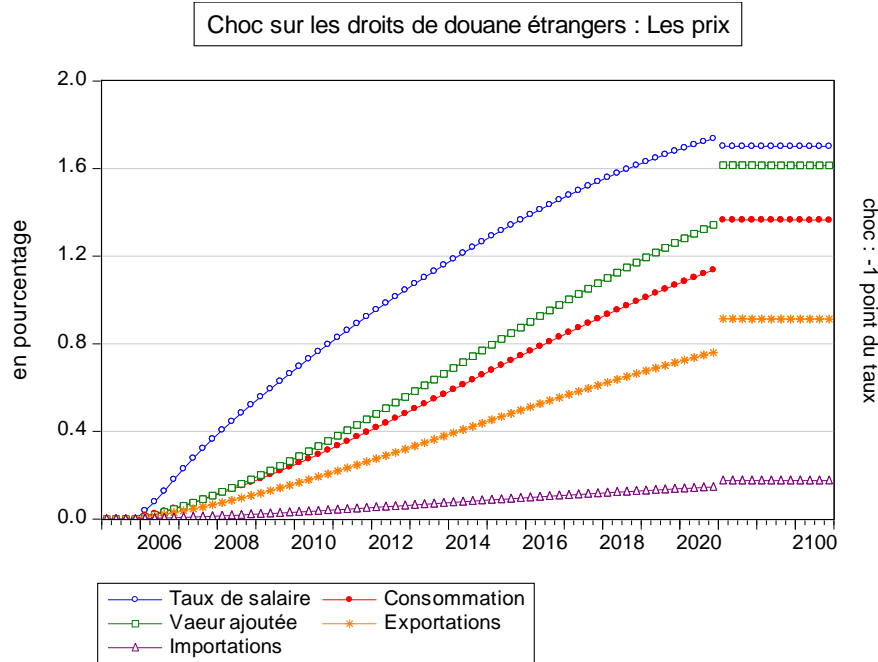
Dans l'ensemble, le gain à prix courants se stabilise après un certain temps, lorsque les capacités se sont adaptées au nouveau niveau de la demande.

Choc sur les droits de douane étrangers : L'équilibre offre-demande



choc : -1 point du taux





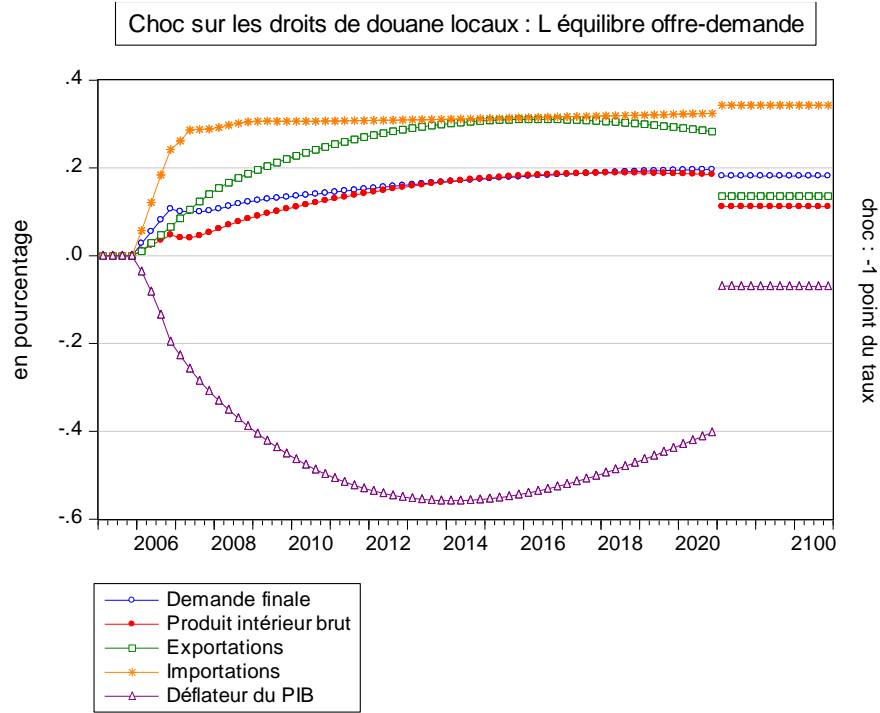
10.2.4.7.4 une diminution des droits de douane locaux

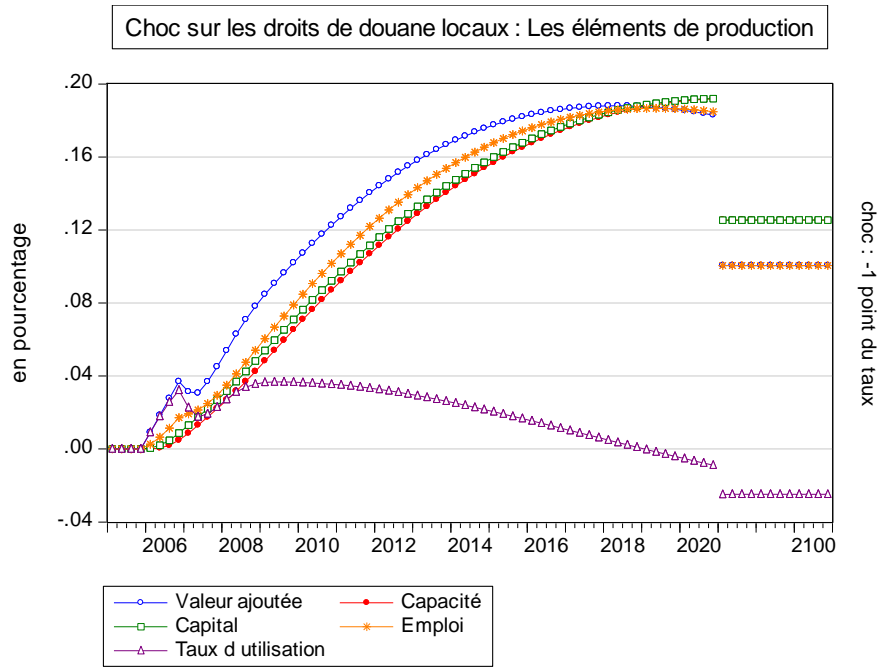
Ce choc réunit traditionnellement les mécanismes les plus complexes, et ses conséquences sont très volatiles d'un modèle à l'autre. Deux canaux principaux doivent être pris en compte:

- L'amélioration de la compétitivité des importations augmente leur part dans la demande locale, réduisant ainsi la production locale, avec les conséquences traditionnelles d'un choc de demande, cette fois négatif.
- La baisse des prix à l'importation se transmet aux prix intérieurs, en particulier de la demande, mais aussi de la valeur ajoutée grâce à une baisse des salaires (indexés sur les prix à la consommation) et du coût des investissements. Cette désinflation permet aux entreprises locales de concurrencer les producteurs étrangers sur les marchés locaux et étrangers, réduisant l'écart initial sur le premier.

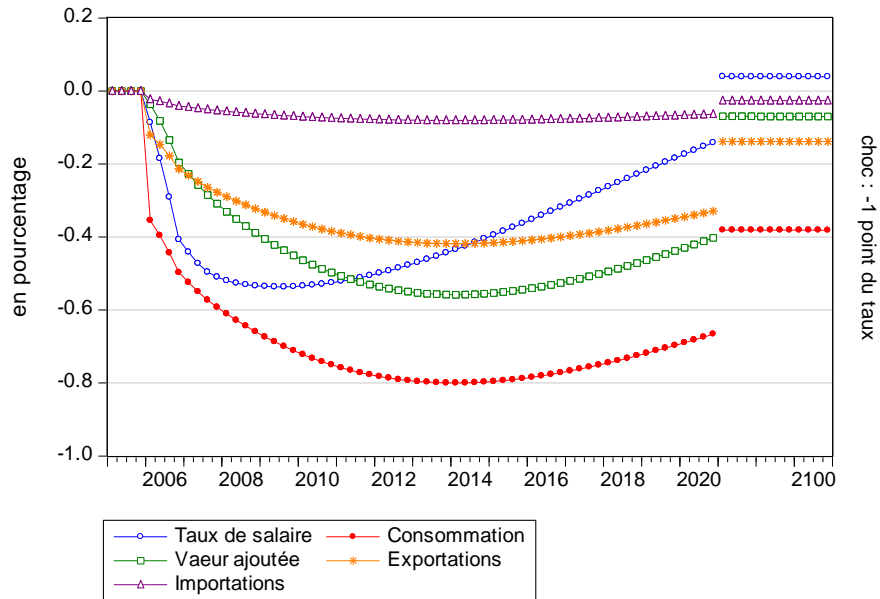
Le coût du capital (en partie importé) diminue plus que le prix de la valeur ajoutée. Sa rentabilité augmente et favorise les investissements et des capacités de production, indépendamment de la demande. Les entreprises qui travaillent à pleine capacité verront la contrainte relâchée²⁹, et les producteurs locaux dans leur ensemble gagneront des parts de marché sur les marchés locaux et étrangers. Dans le même temps, pour améliorer la demande pour ces nouvelles capacités, ils vont baisser leurs prix.

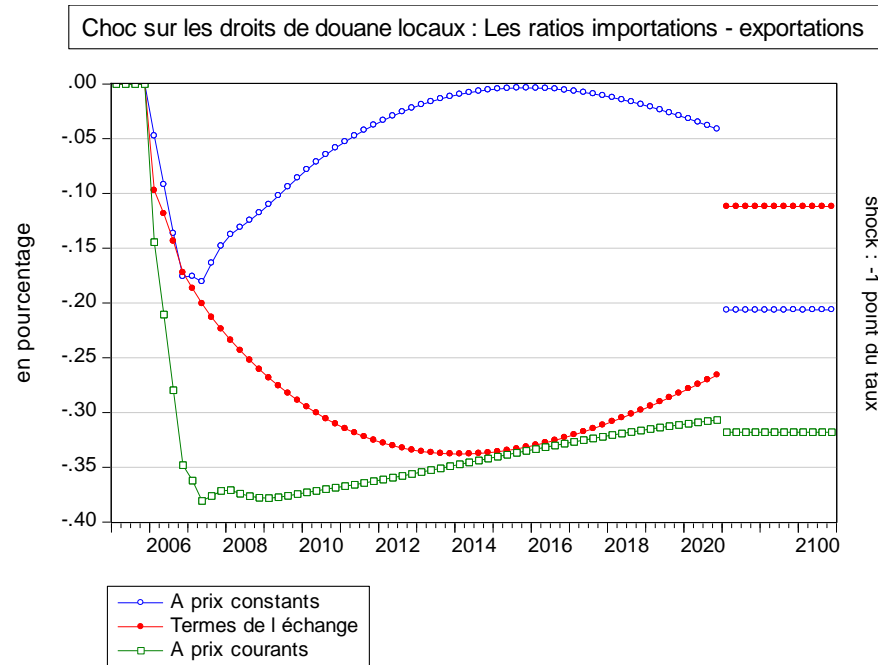
²⁹ Et éventuellement réapparaître à un niveau de production plus élevé.





Choc sur les droits de douane locaux : Les prix





On peut voir:

- Que les importations affichent la plus forte augmentation, car ex ante le choc exerce seulement une influence positive. Mais l'amélioration de la compétitivité locale permet de réduire cet effet, et également d'augmenter les exportations. L'écart initial n'est pas compensé, mais le PIB augmente en volume. Ce serait étrange si les exportations nettes étaient le seul déterminant du PIB, mais nous observons également une augmentation significative de la demande finale locale, provenant d'une consommation et d'investissements moins coûteux (et plus profitables pour ces derniers).

- Que la baisse de l'inflation suit un ordre logique (le même que dans la légende). Nous rencontrerons dans l'ordre:
 - le prix à l'importation hors droits de douane, lié essentiellement au prix mondial ;
 - le prix à l'exportation, moins sensible au prix mondial ;
 - le prix à la valeur ajoutée, qui n'est pas touché directement.
 - le prix de la demande, une moyenne pondérée des importations et des prix locaux (avec un faible impact des exportations) ;
 - le taux de salaire, indexé sur les prix à la consommation. Mais quand l'activité s'améliore, la baisse du chômage lui fait gagner du pouvoir d'achat.

Nous n'avons pas présenté le prix à l'importation taxes incluses, qui baisse d'un point de plus que le prix hors taxes.

Enfin, la balance commerciale est constamment dégradée, à un niveau relativement constant, d'abord par l'augmentation des importations à prix constants, puis par la perte sur les termes de l'échange³⁰. Si la balance commerciale en termes réels se rétablit, c'est grâce à la dégradation des prix relatifs.

10.2.4.8 Le cas Cobb-Douglas

Voyons maintenant comment l'utilisation d'une fonction de production Cobb-Douglas modifie le diagnostic. Nous commencerons comme d'habitude par un choc sur la demande du gouvernement.

Pour rendre les résultats comparables au cas des facteurs complémentaires, nous allons calibrer l'équation d'importations, en donnant à la compétitivité-prix la même élasticité.

En ré estimant les autres coefficients, on observe que les valeurs ne changent pas trop. Nous avons décidé de garder les valeurs originales.

³⁰ Qui utilise évidemment des prix à l'importation hors taxes.

Cela se fait par:

‘ Estimation des importations

```
vector(10) p_m
```

```
smpl 1962Q1 2004Q4
```

```
genr m_ec=0
```

```
smpl 1962Q1 2003Q4
```

```
equation _eq_m1.ls(p) dlog(m)=1*dlog(fd+q)+c_m(2)*log(ur)+c_m(3)*log(@movav(compm,6))+c_m(4)+c_m(5)*(t-2005)*(t<=2005)+[ar(1)=c_m(6)]+c_m(6)*log(m(-1)/(fd(-1)+q(-1)))+m_ec
```

```
p_m(1)=1
```

```
p_m(2)=c_m(2)
```

```
p_m(3)=-0.22
```

```
p_m(5)=c_m(5)
```

```
p_m(6)=c_m(6)
```

```
equation _eq_m.ls(p) dlog(m)=p_m(1)*dlog(fd+tc*q)+p_m(2)*log(ur)+p_m(3)*log(@movav(compm,6))+c_m(4)+p_m(5)*(t-2005)*(t<=2005)+[ar(1)=c_m(6)]+p_m(6)*log(m(-1)/(fd(-1)+q(-1)))+m_ec
```

```
_mod_1.merge _eq_m
```

‘ La réestimation a été faite par (rappelons que l'élasticité de la demande est calibrée dans tous les cas):

```

vector(10) p_m
simpl 1962Q1 2004Q4
genr m_ec=0
simpl 1962Q1 2003Q4
p_m(1)=1
p_m(3)=-0.22
equation _eq_m.ls(p) dlog(m)=p_m(1)*dlog(fd+tc*q)+c_m(2)*log(ur)+p_m(3)*log(@movav(compm,6))+c_m(4)+c_m(5)*(t-
2005)*(t<=2005)+[ar(1)=c_m(6)]+c_m(6)*log(m(-1)/(fd(-1)+q(-1)))+m_ec
_mod_1.merge_eq_m

```

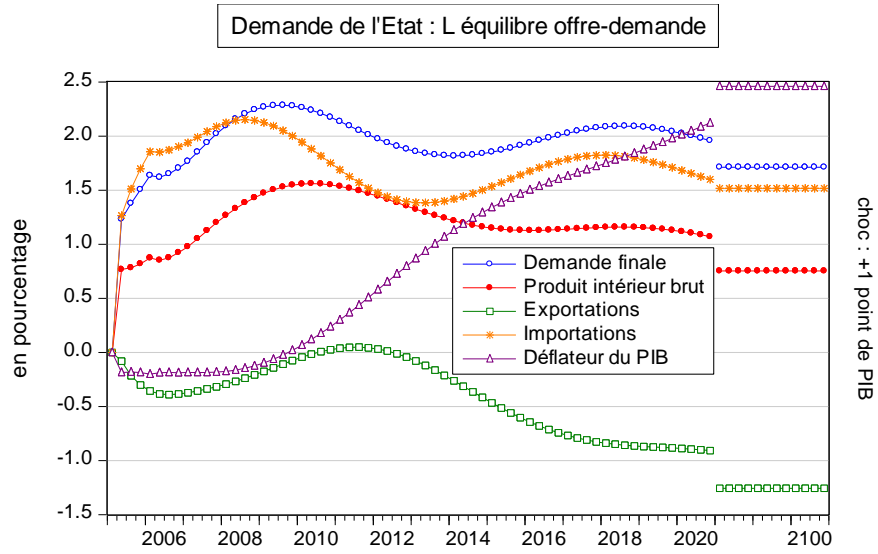
10.2.4.8.1 une augmentation de la demande publique

Encore une fois, nous augmentons la demande du gouvernement de 1 % du PIB de la simulation de référence. Par rapport au cas précédent:

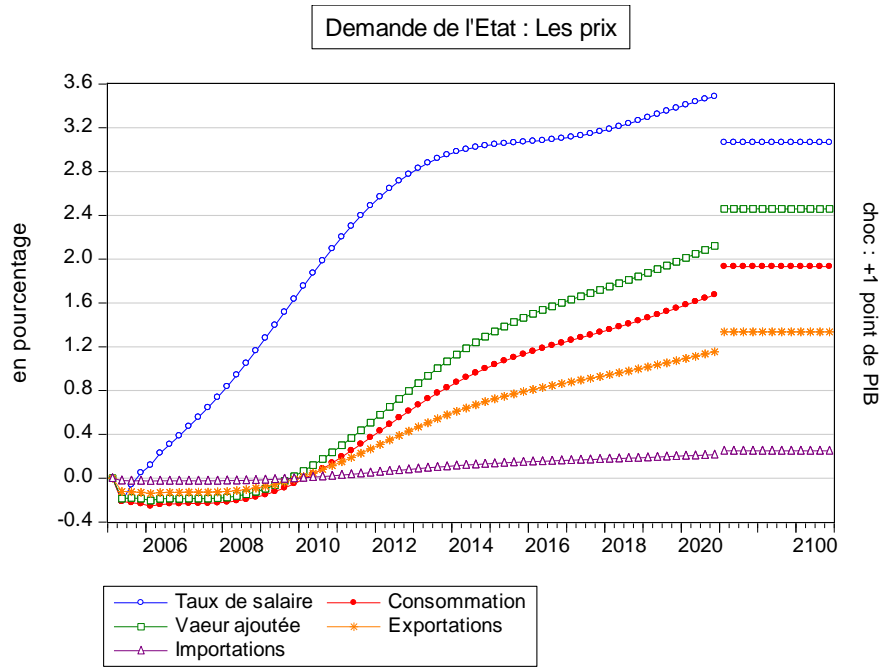
L'effet inflationniste est plus faible, avec une dynamique similaire. Cela vient de la variable de coût utilisé par les entreprises pour cibler leurs marges: au salaire est maintenant ajouté le coût du capital. Les deux éléments sont indexés sur le prix de la demande, qui augmente moins que le prix à la valeur ajoutée. Mais sur les salaires cet impact est plus que compensé par les gains provenant de la baisse du chômage. Cela signifie que, par rapport au déflateur de la valeur ajoutée, le coût salarial augmente, le coût du capital diminue.

Le multiplicateur est plus élevé (conséquence logique). La perte sur les exportations est plus faible, les importations augmentent moins que la demande finale (mais n'oubliez pas qu'elles dépendent également des exportations).

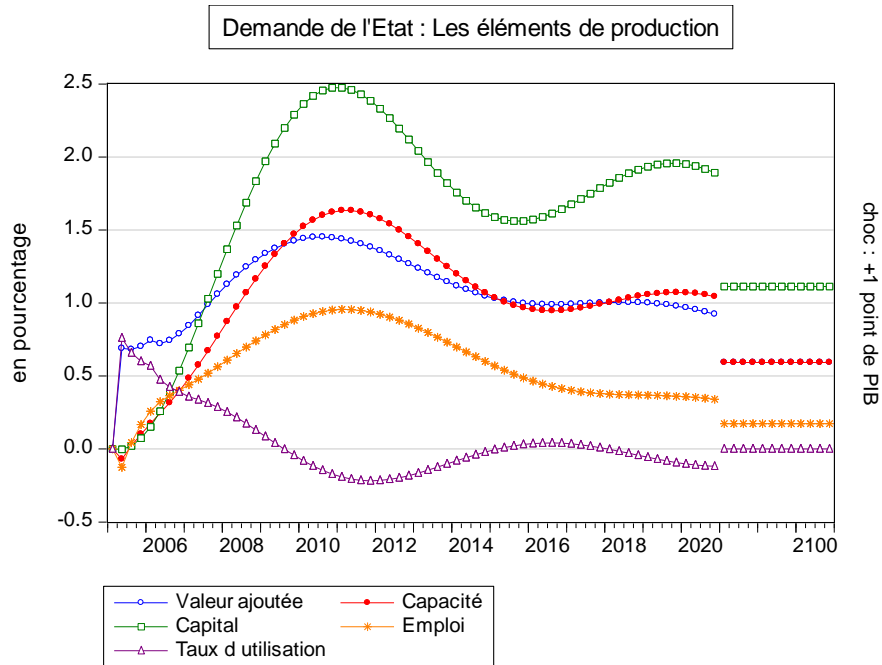
Le court terme présente un cycle mineur.



Le graphique ci-dessous présente les éléments de prix.



Le graphique suivant présente le processus de production, ainsi que l'évolution relative des facteurs et de leur coût.



Nous constatons que:

- le coût du travail finit par augmenter par rapport au capital (à cause de la baisse du taux de chômage) ;
- à long terme, cette augmentation favorise le capital par rapport au travail ;

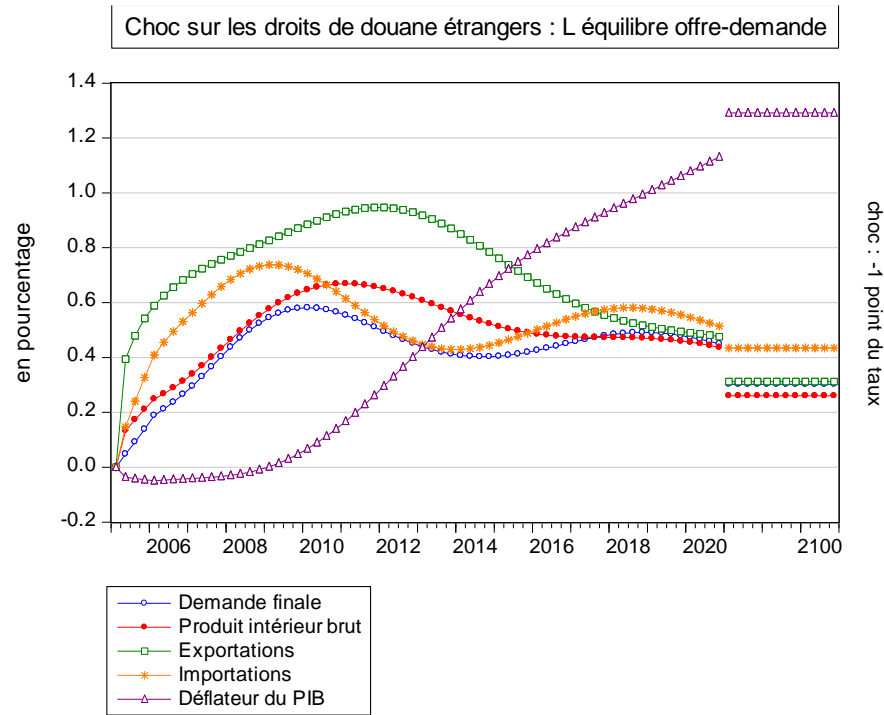
- mais que cet effet met du temps à s'instaurer, principalement à cause de l'inertie sur l'investissement.

10.2.4.8.2 une dévaluation (du franc français)

Le cas dévaluation ne montre pas de nouveauté intéressante.

10.2.4.8.3 une diminution des droits de douane étrangers

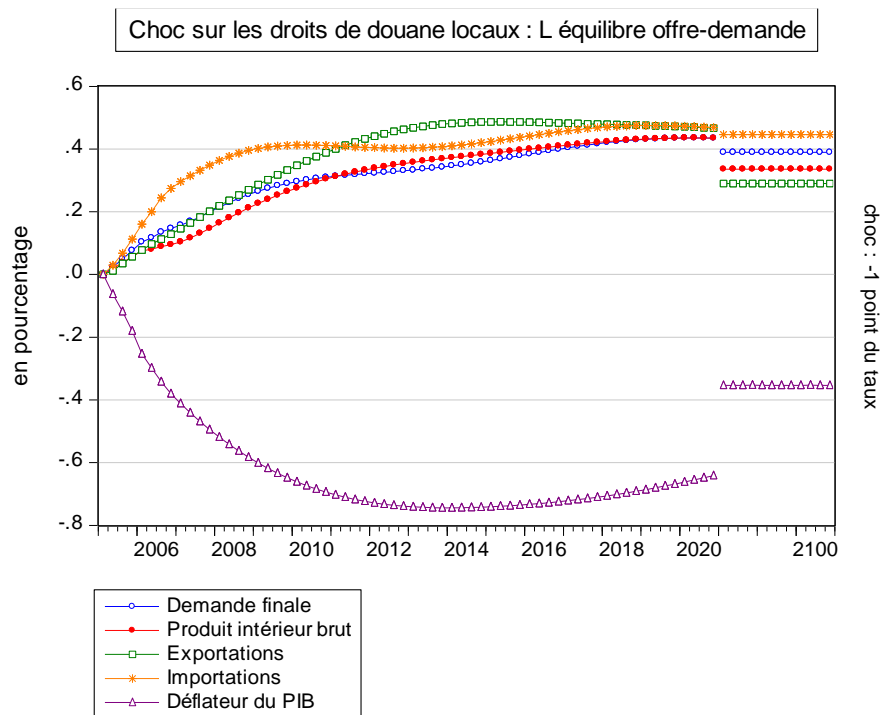
Pour ce choc de demande d'autre part, les changements ne viennent que de la baisse de l'inflation.



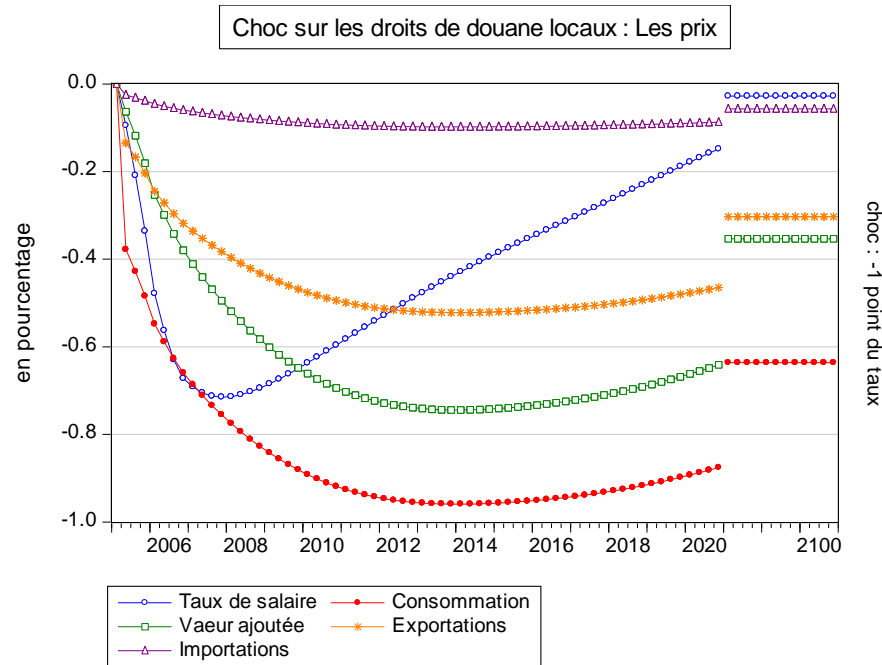
10.2.4.8.4 une diminution des droits de douane locaux

Le principal impact de la modification de la formulation est tout à fait logique: comme la nouvelle définition du coût comprend principalement des biens d'équipement (en grande partie importés), il se réduit plus, conduisant à plus de désinflation et une amélioration

plus forte des exportations. Bien sûr, cette augmentation conduit à un accroissement des importations et le PIB ne s'améliore que modérément.

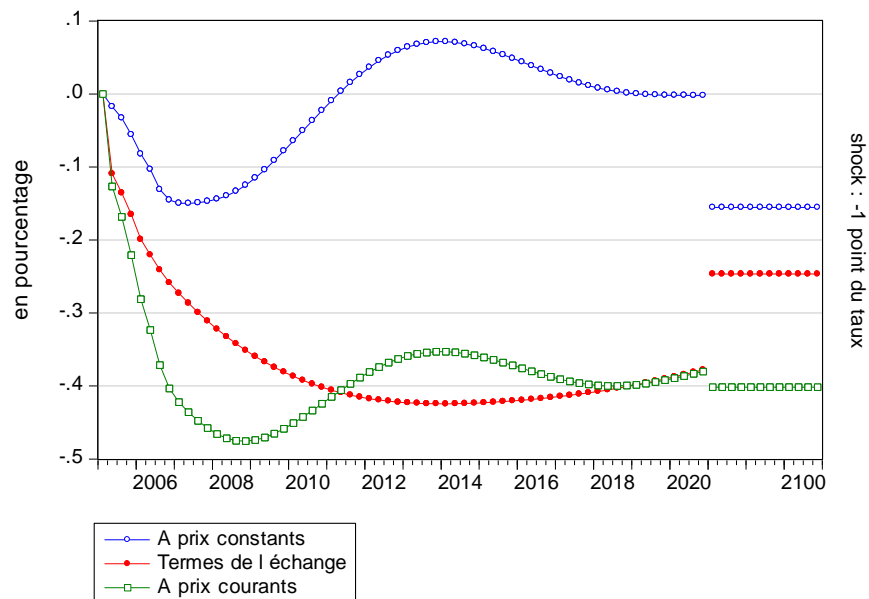


En ce qui concerne les prix, l'évolution est beaucoup plus régulière, et la convergence à la baisse à long terme est monotone.



Quant à la balance commerciale, on peut voir encore une fois que, à prix courants, une évolution plus favorable de la balance à prix constants est compensée (et également justifiée) par une augmentation de la perte sur les termes de l'échange. Mais quant à l'évolution des prix, elle est beaucoup plus régulière que dans le cas des facteurs complémentaires.

Choc sur les droits de douane locaux : Les ratios importations - exportations



10.3 UN MODELE MONO PAYS, MULTI PRODUITS

Considérons maintenant la séparation des produits de notre modèle en plusieurs catégories. Nous examinerons tour à tour:

- les raisons de la décomposition en produits,

- les caractéristiques spécifiques introduites par la décomposition en produits, et la façon dont elle peut être traitée.
- Les principales raisons de la décomposition par produits

On peut trouver trois raisons principales pour la décomposition du produit:

- Elle permet de prendre en compte des différences dans les valeurs des éléments structurels.

Cela se produit lorsque les paramètres structurels prennent des valeurs différentes d'un produit à l'autre, ce qui implique qu'ils vont réagir différemment aux influences extérieures (exogènes ou endogènes), ou à des hypothèses exogènes concernant le produit.

- Elle permet de prendre en compte des différences dans la sensibilité à certains éléments explicatifs.
- Pour certains éléments ou domaines économiques, les liens entre les variables peuvent suivre des mécanismes différents, appelant à des spécifications différentes qui ne peuvent être résumées par agrégation mathématique.

10.3.1 LES DIFFERENCES STRUCTURELLES

Elles peuvent être facilement identifiées dans le cadre de notre modèle. Les principaux éléments sont les suivants:

- La productivité des facteurs: la quantité de travail et / ou de capital nécessaire pour produire une unité de bien peut être différente. Cela est particulièrement vrai pour l'agriculture, où la productivité du travail est particulièrement faible³¹. Et aussi pour les services, où le besoin en capital est généralement limité.
- Le taux de salaire: les salaires moyens obtenus par les travailleurs peuvent être très différents (souvent en liaison avec la productivité du travail, ce qui peut compenser l'effet lorsque l'on considère le coût unitaire).
- Les parts différentes dans chaque élément de la demande, ce qui permet de présenter des sensibilités différentes à une augmentation donnée de la consommation ou de l'investissement.

³¹ Surtout dans les unités familiales.

- Le montant des consommations intermédiaires de chaque bien nécessaire pour produire une unité d'un bien donné.
- Le partage de la destination de la production entre les marchés locaux et étrangers (exportations).
- Le partage de la demande locale entre les producteurs locaux et étrangers (importations).
- La séparation entre les entreprises, les ménages, l'État et des agents étrangers dans le processus de production locale, avec des conséquences différentes sur les recettes et leur utilisation.
- La taxation des biens (TVA, autres impôts indirects, droits de douane sur les importations et sur les exportations locales).

Tous ces éléments seront pris en compte, sans coût, en séparant les éléments dans le modèle.

10.3.2 LES DIFFERENCES DE SENSIBILITE

Cela peut s'appliquer à tous les paramètres des équations estimées. Bien sûr, on s'intéressera aux plus importants, et la théorie économique peut indiquer pour lesquels des différences significatives peuvent être attendues. Par exemple, la nécessité de la formation de stocks sera plus importante dans le secteur manufacturier, le rôle du chômage sera limité dans la définition du salaire agricole, ou la sensibilité du commerce extérieur à la compétitivité-prix peut être faible pour les services.

10.3.3 LES DIFFERENCES DE COMPORTEMENT LOGIQUE

Cette fois, la formulation même d'un comportement ou sa logique causale sera différente.

Ceci s'applique essentiellement à deux domaines.

10.3.3.1 Le processus de production

Dans le secteur agricole, on peut supposer que les producteurs maximisent leur production. Pour ce faire, ils peuvent utiliser des plantations (arbres et champs), des animaux, ou une combinaison des deux (comme l'élevage des truites dans les étangs). Ils ont aussi besoin d'une certaine quantité de travail, d'outils et de biens de consommation intermédiaires (comme du fourrage ou des engrais). L'augmentation de ces niveaux peut améliorer le rendement, de même que de meilleures infrastructures (installations d'irrigation, de

stockage et de transport³²). Mais une fois que ces éléments sont définis, la production suit, corrigée fortement par des conditions climatiques aléatoires. Ceci est également vrai dans une grande mesure pour la pêche. L'impact d'une augmentation de la demande est limitée au moins dans le court terme (même si peut-être que plus de bovins seront abattus, ou plus de poissons seront pêchés).

Dans le secteur manufacturier, on peut supposer l'existence de plusieurs processus, entre lesquels le producteur a choisi en fonction des coûts relatifs du capital et du travail. Supposer une élasticité constante et unitaire conduit à une formulation de type Cobb-Douglas. Nous pouvons aussi choisir une simple fonction à facteurs complémentaires, ainsi que d'une fonction plus générale de type CES. En fait, l'un des avantages de l'identification par produits est d'associer une fonction sophistiquée aux produits qui le méritent réellement. Dans le seul cas des modèles à un seul bien, les éléments de l'alternative: utiliser une fonction simple ou associer une fonction Cobb-Douglas à l'ensemble du spectre économique nuisent tous deux à la qualité du modèle.

Mais le principal avantage de cette identification est de définir une capacité de production et un taux d'utilisation associés. Cet élément jouera un rôle important (comme nous l'avons déjà vu) pour trois éléments: le rôle des producteurs locaux dans la satisfaction des besoins locaux et étrangers, le comportement d'investissement, et le choix à court terme par les entreprises de la marge appliquée au coût de production.

Enfin, dans le secteur des services, le choix des facteurs peut dépendre du coût relatif (des ordinateurs remplaceront des personnes), mais le rôle de la capacité est moins clair, car dans de nombreux cas, les producteurs ont une influence considérable sur la production, indépendamment du capital installé et même de l'emploi: une augmentation soudaine et forte de la demande pour les voyages touristiques peut généralement être satisfaite par les agences de voyages, si les clients sont prêts à adapter leurs choix à l'offre. La qualité du service va probablement diminuer, mais pas son coût, l'élément par lequel la production est mesurée. L'insatisfaction des clients ne réduit pas la valeur d'un bien acheté, sauf si un remboursement doit être effectué.

³² Rappelons-nous que la production ne comprend que les biens qui ont effectivement été achetés (pensez à l'équilibre offre-demande) ou les variations de stocks, qui ne s'appliquent pas aux produits agricoles (les produits transformés sont considérés comme manufacturés). Cela signifie que les produits gâchés ne sont pas pris en considération, et l'amélioration des transports augmentera la production agricole, toutes choses étant égales par ailleurs.

10.3.3.2 Le commerce extérieur

Le commerce des produits agricoles, des biens manufacturés et des services obéit à des règles bien évidemment différentes.

Pour l'agriculture, le prix à l'exportation est en général fixé au niveau mondial pour une qualité donnée, et un pourcentage donné de la production locale est proposé sur le marché étranger, selon le type et la qualité des produits, la politique locale, et les contrats préétablis.

Pour les produits manufacturés, en plus de la demande, la compétitivité des prix est essentielle, ainsi que la capacité disponible qui peut créer des goulots d'étranglement temporaires sur la fourniture d'éléments spécifiques.

Pour les services, les déterminants sont complètement différents et beaucoup moins clairs, mais les montants échangés sont moins importants (sauf pour le tourisme qui est très difficile à modéliser).

10.3.3.3 Les conséquences pour la décomposition par produits

Les remarques ci-dessus ont conduit à préciser que la décomposition minimale devrait introduire trois catégories: primaire, secondaire et tertiaire des produits.

Allant plus loin, on peut considérer :

- l'extension de la décomposition à d'autres produits (voir plus loin).
- L'idée la plus immédiate concerne l'énergie, en particulier pour les pays producteurs d'hydrocarbures. Il est clair que le niveau de production peut changer considérablement selon la volonté du producteur, que les coûts variables de production sont relativement faibles, et que toute la production ne devrait avoir aucun problème pour être vendue, à un prix très variable mais défini par le marché mondial.
- Cela est également vrai en général, pour l'exploitation des mines, qui devrait donc être regroupée avec les biens secondaires, avec l'énergie, ou placée dans sa propre catégorie.
- Le bâtiment peut également être identifié, comme déterminé par d'autres productions (entrepôts, usines), et par les revenus des ménages. Le processus de production est relativement simple pour un type donné, et il n'est ni importé ni exporté, même si ses intrants peuvent l'être (comme le bois et l'acier).

- l'introduction du type de propriétaire. Les entreprises peuvent être divisées en publiques et privées, et les secondes en entreprises individuelles et sociétés. Les entreprises étrangères peuvent être identifiées (avec les investissements directs étrangers, et le rapatriement des profits).

10.3.4 INTRODUIRE LES CONSOMMATIONS INTERMEDIAIRES

Un développement très important est associé aux consommations intermédiaires.

Tout modèle, pour un produit, doit respecter l'équilibre:

$$\text{Production} + \text{importations} = \text{demande totale} + \text{exportations}$$

Cependant, au niveau global, on peut soustraire les consommations intermédiaires de chaque côté, ce qui donne:

$$\text{PIB} + \text{importations} = \text{demande finale} + \text{exportations}$$

Cela libère le modèle de la définition des consommations intermédiaires, un concept difficile à gérer, car il dépend de la décomposition du produit lui-même, plus précisément du nombre d'étapes dans le processus de production.

Par exemple un processus qui séparerait dans la production de vêtements l'impression des motifs et la découpe des pièces générerait moins de production qu'un processus intégré, car les vêtements imprimés et non découpés ne seraient plus pris en compte.

Le seul rôle qui lui reste est le fait que dans les équations d'échanges extérieurs, où la compétitivité des prix se fonde sur les prix de production: la préférence pour les produits originaires d'un pays est évidemment basée sur le prix auquel ils sont vendus.

Cela signifie que les prix des échanges (exportations et importations) sont des prix de production, et que la compétitivité compare le coût global de l'exportateur et le prix auquel ses concurrents vendent (compte tenu de celui des consommations intermédiaires). Cela signifie par exemple qu'un pays qui a accès au pétrole ou de gaz à bon marché (comme les États-Unis, la Russie ou l'Algérie) dispose d'un avantage comparatif par rapport à l'Allemagne ou la Chine, face à la fois aux exportateurs vers son marché et à ses concurrents sur les marchés étrangers³³.

L'équation ci-dessus

$$\text{Production} + \text{importations} = \text{demande totale} + \text{exportations}$$

peut être écrite comme:

$$\text{Valeur ajoutée} + \text{consommation intermédiaire par le produit} + \text{importations} = \text{consommation finale locale} + \text{consommation intermédiaire du produit} + \text{exportations}$$

Cela conduit à l'identification d'une matrice des consommations intermédiaires, carrée avec la dimension de la décomposition.

³³ Toutefois, comme nous utilisons des déflateurs, cela signifie que les prix du commerce de ces pays seront moins sensibles au coût de l'énergie, qui représente une faible part du coût total. Ils vont gagner en compétitivité si le prix monte, mais perdre si elle se réduit (comme ils ne profitent pas du choc).

En règle générale, la définition de chacun de ces éléments à prix constants supposera que la production de chaque unité d'un bien donné nécessite une quantité donnée de chaque bien (y compris lui-même).

A prix courants, on pourra appliquer le déflateur du produit utilisé (hors TVA, car elle ne s'applique pas ici) pour la demande globale du bien, ou pour la consommation intermédiaire globale. La disponibilité de la matrice à prix courants améliore la situation en permettant de définir un déflateur individuel pour chaque case.

10.3.5 LES ASPECTS SECTORIELS SPECIFIQUES

Maintenant que nous connaissons les raisons de la décomposition par produits, nous allons voir comment cela affecte la structure du modèle.

Le changement le plus immédiat est bien sûr la multiplication des catégories. Selon l'élément, l'extension peut s'appliquer à un critère, voire à deux critères en même temps.

Une liste précise des objets appartenant à chaque catégorie sera fournie en annexe.

On peut d'abord considérer les produits:

- Produit 1: produits agricoles (tels que les aliments non transformés)
- Produit 2: produits manufacturés (tels que les appareils électroménagers)
- Produit 3: services (comme les transports).

Cette décomposition s'applique naturellement aux éléments échangés, tels que la consommation et les exportations.

Mais la décomposition apparaîtra aussi pour les unités de production, en fonction de leur production. Des fermes, des usines et des agences de voyages seront classées dans chacune de nos trois catégories, dans l'ordre.

Et ici, une distinction doit être faite, car une unité donnée peut produire des biens différents. Dans ce cas:

- l'ensemble sera classé dans la branche de son activité principale ;
- ou il sera fractionné en produits, chaque partie étant allouée au secteur qu'elle représente.

Dans notre modèle, nous ne séparerons pas les branches et secteurs, et nous allons utiliser alternativement les deux termes. Mais branches et produits peuvent jouer des rôles différents, et nécessiteront parfois pour la même variable l'introduction d'une double indexation.

En outre, il peut paraître intéressant d'aller plus loin dans la décomposition, dans deux cas:

- Pour l'agriculture (et la pêche), le modèle bénéficiera d'une séparation en unités artisanales et industrielles:
 - Les unités artisanales utilisent moins de capital (machines) et beaucoup plus de travail, avec une productivité plus basse et une inertie très forte de l'emploi.
 - L'emploi artisanal ne sera pas très sensible à la production (les unités familiales comprennent souvent une main-d'œuvre généralement inactive, à laquelle on fera appel si nécessaire).
 - Elles utilisent aussi moins de biens intermédiaires (tels que les engrais et le fourrage).
 - Leur chiffre d'affaires est entièrement approprié par les ménages.
 - Les exportations représentent une part moins importante de leur production.

Ceci est particulièrement intéressant dans les pays en développement, où le rôle de l'agriculture est encore important, au moins socialement.

- Le secteur manufacturier comprend l'énergie et les mines. Le pétrole, le gaz et les minerais sont négociés et produits selon un processus différent des autres produits:
 - Le prix international est fixe, et la compétitivité joue un rôle très limité (même pour les importations).
 - La capacité ne joue également aucun rôle dans le court terme: les quantités exportées sont décidées par l'exportateur, et les importations sont directement liées à la demande³⁴.

Cela signifie que nous allons séparer dans notre exemple:

- La production agricole et la pêche en artisanale et industrielle.

³⁴ Mais si la production locale diminue, les importations peuvent croître.

- La production, les exportations et les importations en produits énergétiques et non énergétiques. Nous n'identifierons pas les mines, même si elles jouent un rôle de premier plan dans l'économie de certains pays (Chili, Maroc, Australie ...).

Ceci est un exemple du cas (fréquent) où des décompositions peuvent conduire à d'autres décompositions:

- c'est seulement pour l'agriculture que la décomposition artisanale / industrielle est vraiment efficace,
- et la séparation de l'énergie des biens manufacturés provient du fait qu'elle se comporte plus comme le produit primaire. Cela appelle une correction locale de la décomposition générale, si nous voulons que le modèle présente des équations d'échanges fiables.

10.3.6 LES DONNEES

Pour des données détaillées, très souvent, l'information devient plus rare. Les problèmes suivants peuvent se produire:

- Le détail sur les biens et services à prix courants est connu seulement du côté de la production, et non pas du côté de la demande.
- Pour les séries détaillées qui sont disponibles, les déflateurs sont connus seulement à un niveau plus global, ce qui rend le calcul des variables à prix constants approximatif.
- L'investissement (et le capital) ne sont disponibles que par produit, ou par branche, et pas dans un détail à deux dimensions. Le même problème peut apparaître pour les variations de stocks.

Heureusement l'investissement est réalisé essentiellement dans les produits manufacturés et le bâtiment (plus le cheptel et les plantations pour l'agriculture). Si le modèle utilise un produit secondaire global, la dimension des produits n'est pas nécessaire³⁵. Dans tous les cas, de nombreuses cellules de la matrice contiennent des valeurs nulles.

³⁵ Bien que l'investissement comporte une part des services (production de plans de bâtiments, brevets).

- La consommation intermédiaire par produit et utilisation est disponible plus souvent, mais pas toujours (cela signifie que la matrice d'entrée-sortie n'est pas entièrement connue, ou peut-être seulement pour certaines périodes, nécessitant une interpolation).
- L'emploi n'est pas détaillé par branche ou par catégorie de propriétaire de l'entreprise.
- La partie du compte d'entreprises donnant les transferts (salaires, subventions ...) n'est pas détaillée par branche.
- La période pour laquelle les données sont connues est plus courte, ou plus difficile à collecter. Par exemple, les tableaux peuvent être séparés année par année.

10.3.7 CONSEQUENCES POUR LES ESTIMATIONS

Si l'échantillon est trop court, on peut toujours chercher à estimer les équations. Mais il ne faut alors pas trop se fier aux résultats, même favorables. Nous devons simplement les utiliser comme une indication sur le comportement associé, complétant la théorie économique, les observations sur la façon dont fonctionne l'économie locale, et les estimations au niveau mondial.

Mais très souvent, les problèmes ci-dessus ne permettent pas d'estimer des équations individuelles au niveau détaillé.

On peut alors:

- calibrer les équations en utilisant les coefficients théoriques ;
- utiliser les résultats estimés au niveau agrégé ;
- prendre les coefficients dans d'autres modèles, décrivant un pays similaire (la Thaïlande pour le Vietnam par exemple, ou l'Autriche pour la République Tchèque). Cette option présente un aspect favorable surprenant: un pays plus avancé a généralement un système statistique plus élaboré, et ses estimations sont basées sur une période qui représente l'avenir de notre pays modélisé, la période réelle sur laquelle les études économiques sera effectuée.

Et en fait, cette difficulté a un bon côté: partiellement exempt de restrictions économétriques, on peut appliquer les formulations qui sont conformes à la théorie économique, dans la spécification et la valeur des coefficients. En particulier, on peut appliquer un cadre de correction d'erreur dans chaque formule, conduisant à un modèle qui:

- fournit une solution à long terme sur une période future, sans aucun changement aux spécifications (avec quelques restrictions sur les hypothèses) ;

- sépare complètement la spécification à long terme de la dynamique qui y conduit ;
- permet d'interpréter complètement et facilement les coefficients (en utilisant des valeurs qui sont conformes à la théorie).

10.3.8 LA FONCTION DE PRODUCTION

Comme nous l'avons indiqué ci-dessus, on peut utiliser des solutions différentes pour chaque branche.

10.3.8.1 Le produit primaire

Il s'agit d'un domaine assez difficile à gérer. La capacité peut dépendre:

- du capital en produit 1: terrain prêt à l'emploi, plantations et bétail ;
- du capital en produit 2: machines, plus ou moins sophistiquées ;
- du capital en services: transport, stockage ;
- du capital public: disponibilité de l'eau et de l'électricité, réseau routier ;
- des intrants: fourrage, engrais ;
- des conditions climatiques (exogènes, bien sûr).

Les formulations peuvent être plus ou moins complexes, allant d'une simple fonction de production (complémentaire ou Cobb-Douglas) à un système logique, comprenant des conditions et des fonctions non continues. La production pourrait également être décomposée en plus de catégories (agriculture / pêche / sylviculture) ou en produits (riz et café pour le Vietnam, fruits pour l'Amérique centrale ...). Dans ce dernier cas, des quantités physiques pourraient être utilisées (tonnes, litres, quantités).

Mais fondamentalement, la définition à court terme des quantités produites ne devrait pas trop dépendre de la demande, mais plutôt du potentiel de production, affecté par les facteurs climatiques (dans le long terme, bien sûr, la rentabilité et la taille du marché auront une incidence sur la création de capacités).

Cela a un impact important sur les propriétés du modèle. Comme l'équilibre offre-demande doit toujours être appliqué, il faut le solder par un élément nouveau, qui peut être les importations, les exportations, la demande globale ou l'une de ses composantes. La solution la plus simple est d'utiliser les importations, ce qui signifie que:

- la demande locale doit être satisfaite³⁶ ;
- les exportations sont déterminées par la demande mondiale et la compétitivité (avec un impact limité car les prix agricoles à l'exportation sont essentiellement déterminés au niveau mondial) ;
- la part de la demande mondiale qui ne peut être satisfaite localement doit être importée³⁷.

D'autres cadres peuvent être envisagés, tels que:

- la demande locale doit être satisfaite ;
- les importations sont contrôlées par les agents locaux ou l'État ;
- ce qui reste de la production locale est exporté ;

ou:

- les exportations sont contrôlées par les agents locaux ou l'État, en fonction des conditions du marché ;
- les agents locaux sont autorisés à acheter la part non exportée de la production locale ;
- les importations sont également contrôlées ;

cela donne par solde la demande locale.

Il est tout à fait possible d'envisager plusieurs de ces cadres en même temps, mais seulement si deux ou trois catégories de biens sont définies, chacune avec ses quatre éléments qui vérifient un équilibre particulier.

On peut s'attendre à des conséquences pour les propriétés du modèle, pour lequel ce choix est tout à fait important.

³⁶ Mais n'oubliez pas qu'elle est endogène, car elle suit l'activité mondiale.

³⁷ Mais si le prix des marchandises importées est trop élevé, sa consommation va diminuer et se déplacer également vers d'autres biens.

10.3.8.2 Le produit industriel

L'un des objectifs principaux de la décomposition est d'identifier une fonction de production sophistiquée pour le secteur manufacturier, en laissant de côté les autres branches pour lesquelles ce choix est beaucoup moins naturel, ce qui aurait un impact négatif sur la qualité d'une estimation globale.

Une option modérément complexe est la fonction Cobb-Douglas, que nous avons présentée plus tôt.

10.3.9 LE CHOMAGE

Le taux de chômage suit la même logique que le modèle simple. Cependant on pourrait définir des sensibilités différentes à l'emploi selon la branche. Par exemple, dans le secteur primaire, la plus faible proportion de travailleurs qualifiés peut rendre les chômeurs immédiatement productifs. Et pour l'agriculture artisanale, un supplément de production peut souvent être réalisé sans création d'emplois, par appel à des membres de la famille inactifs.

10.3.10 LES VARIATIONS DE STOCKS

L'élément nouveau est que nous devrions développer l'équation dans deux directions: les biens stockés et la branche qui les utilise. Cela signifie que, en principe, nous aurons au moins 9 variables (sans compter les totaux). Cependant les services ne sont pas stockés.

Pour définir les coefficients, on peut considérer la théorie, des conseils d'experts ou le rapport de la variation des stocks à la variation totale de la valeur ajoutée.

10.3.11 LA CONSOMMATION DES MENAGES

La décomposition apporte une nouvelle fonctionnalité: la consommation doit être divisée en produits.

Nous calculerons d'abord la consommation globale comme s'il y avait un produit unique.

Puis nous séparerons la consommation en produits, en utilisant un cadre à correction d'erreur, avec une cible répartissant les biens en fonction de leurs prix relatifs, à partir d'une décomposition de base (à prix identiques). Après l'application de la décomposition, il faut contrôler que la somme des valeurs obtenues correspond à la valeur globale (cela peut se faire à l'aide d'un système ou par application d'un facteur de correction).

La situation est plus simple si l'on considère deux produits, car nous avons seulement besoin d'un seul ratio, puisque la consommation globale est connue.

10.3.12 LE COMMERCE EXTERIEUR

Il s'agit du deuxième champ dans lequel la décomposition du produit permet d'introduire des différences dans les formules selon les produits.

Cette originalité viendra de:

- Différents poids: la part des exportations dans la production et des importations dans la demande est différente d'un produit à l'autre.
- Des sensibilités variables: l'estimation du rôle de la compétitivité des prix peut donner des réponses différentes. Si l'équation doit être calibrée, la théorie économique pourrait également conduire à utiliser des valeurs différentes.
- Différentes formulations: le rôle des tensions sur les capacités peut être limité aux produits manufacturés.
- Et nous avons vu que l'équilibrage de l'offre et de la demande peut donner un rôle différent aux éléments de commerce extérieur.

10.3.13 LES SALAIRES

Pour les salaires, on partira du même cadre (théorique) que celui du modèle simple.

Toutefois, le processus d'indexation est un peu plus complexe. Nous avons encore le choix (en particulier sur le long terme) entre une indexation sur le prix à la consommation et le déflateur de la valeur ajoutée. Mais si le premier élément est mesuré à l'échelle globale, le second est associé à la branche. Cela signifie que nous pouvons observer les conséquences des chocs dissymétriques (comme une

augmentation des subventions à un secteur donné), ou de chocs ayant des impacts sectoriels (comme une augmentation des marges des entreprises exportatrices, principalement industrielles).

En outre, la sensibilité au chômage peut être très différente, conduisant à des propriétés inflationnistes dissymétriques.

Enfin, l'identification des branches artisanales et industrielles de l'agriculture permettra de tenir compte des productivités et des salaires différents.

10.3.14 LES PRIX

10.3.14.1 Le déflateur de la valeur ajoutée

Avec des modes de production différents, le coût variable utilisé dans l'objectif à long terme va changer. Et le rôle du taux d'utilisation devrait être plus important dans le secteur manufacturier.

10.3.14.2 Le prix à la production

C'est un domaine dans lequel la complexité augmente: la consommation intermédiaire d'une branche est la somme des différents éléments bidimensionnels, chacun évalué à son prix (qui doit être défini).

10.3.14.3 Le prix des échanges

Une fois que les prix de production sont connus, nous pouvons passer aux prix des échanges. Pour définir le prix des concurrents, nous avons besoin d'identifier un prix moyen de production étranger pour chacun de nos produits, qui dépendra de la structure des biens échangés, ainsi que des clients et fournisseurs.

10.3.14.4 Le prix de la demande

Les déflateurs ci-dessus nous permettent de calculer la valeur ajoutée et le commerce à prix courants. Comme nous connaissons déjà les valeurs à prix constants, les déflateurs de la demande peuvent être définis par des identités.

Cependant, il faut s'assurer que, en prix courants, les données vérifient toujours l'équilibre entre la demande et l'offre. Une petite erreur peut être acceptée, et soit

- vous la négligez: c'est dangereux car le modèle ne pourra jamais obtenir un ajustement historique parfait, une fois que les équations estimées auront été alimentées par leurs résidus ;
- soit vous la considérez comme un résidu (multiplicatif). Nous avons vu que c'est une bien meilleure option.

10.3.14.5 Le budget de l'État

Comme dans le modèle simple, nous devons définir un budget détaillé du gouvernement. Les éléments associés aux biens et services (demande, taxes et subventions) devront être séparés. Mais d'autres non (transferts aux ménages en particulier).

10.3.14.6 Le programme EViews

Nous pouvons maintenant construire un modèle décomposé, basé sur le même cadre que notre produit unique, et utilisant les mêmes concepts.

Nous avons décidé que le problème était trop complexe pour être présenté ici. Nous allons cependant présenter un tableau utilisant le niveau de décomposition appliquée à ces concepts. Selon la variable, aucune, une ou deux dimensions seront ajoutées.

10.3.14.7 Une liste d'éléments

Nous présentons ici une liste d'éléments, venant en fait d'un modèle opérationnel à trois produits construit pour le Vietnam. Il contient la décomposition de l'agriculture en artisanale et industrielle, que nous avons présentée plus tôt.

Dans les «catégories» de colonne,

- T représente le total,
- 1, 2 et 3 les produits associés (des branches ou des secteurs),
- 1a et 1i la décomposition de l'agriculture entre artisanale et industrielle,
- 2e et 2m la décomposition du produit 2 entre l'énergie et le reste (pour la production et le commerce extérieur).

Nom	type	définition	Unités	Type	Catégories
CAP	Endog	Capacité de production	Monnaie locale constante	produit	1,2,3
CI	Endog	Variation des stocks	Monnaie locale constante	produit	T, 1,1a,1i,2,3
COG	Exog	Consommation des administrations	Monnaie locale constante	produit	1,2,3
COGV	Endog	Consommation des administrations	:Monnaie locale courante	produit	T
COH	Endog	Consommation des ménages	Monnaie locale constante	produit	T,1,2,3
COMPM	Endog	Compétitivité-prix des importations (prix à l'importation / prix de la production locale)		produit	1,2i,3
COMPX	Endog	Compétitivité des prix des exportations (prix à l'exportation / prix étranger)		produit	1,2e,2i,3
COST	Endog	Coût des salaires et des capitaux	Monnaie locale constante	branche	2
CPI	Endog	Déflateur du prix à la consommation	Déflateur année de référence = 1	produit	T
CRV	Endog	Rapport exportations-importations à prix courants	Rapport	produit	T
CRVOL	Endog	Rapport exportations-importations à prix constants	Rapport	produit	T
ER	Endog	Taux de change (la valeur de la monnaie locale)	Déflateur année de référence = 1		T
EXPG	Endog	Dépenses des administrations	:Monnaie locale courante		T
FBAL	Endog	Capacité de financement des entreprises	:Monnaie locale courante	produit	1i,2,3
FD	Endog	Demande finale locale	Monnaie locale constante	produit	T,1,2,3
FDGV	Endog	Demande finale des administrations à prix courants	:Monnaie locale courante	total	T
FDI	Endog	Investissement étranger direct à prix constants	:Monnaie locale courante	produit	1,2,3
FDV	Endog	Demande finale locale	:Monnaie locale courante	produit	T,1,2,3
FOBAL	Endog	Balance commerciale	:Monnaie locale courante	produit	T

GBAL	Endog	Capacité de financement des administrations	:Monnaie locale courante	produit	T
GDPM	Endog	Produit intérieur brut marchand	Monnaie locale constante	produit	T,1,2,3
GDPMV	Endog	Produit intérieur brut marchand aux prix courants	:Monnaie locale courante	produit	T,1,2,3
GDPV	Endog	Produit intérieur brut total à prix courants	:Monnaie locale courante	produit	T
HDI	Endog	Revenu disponible des ménages	:Monnaie locale courante	produit	T
HI	Endog	Revenu des ménages avant impôt	:Monnaie locale courante	produit	T
HRDI	Endog	Revenu disponible des ménages en pouvoir d'achat	Monnaie locale constante	produit	T
IC	Endog	Consommation intermédiaire à prix constants	Monnaie locale constante	Produit x Direction	(1,2,3)*(1,2,3)
ICV	Endog	Consommation intermédiaire à prix courants	:Monnaie locale courante	Produit x Direction	(1,2,3)*(1,2,3)
IG	Exog	Investissements publics à prix constants	Monnaie locale constante	produit	1,2,3
IGV	Endog	Investissements publics à prix courants	:Monnaie locale courante	total	T,1,2,3
IHH	Endog	Investissement logement des ménages	Monnaie locale constante	total	T,2,3
INCT	Endog	Impôt sur le revenu	:Monnaie locale courante	total	T
IP	Endog	Investissement productif	Monnaie locale constante	total	(1,2,3)*(T,1a,1i,1,2,3),T*
IPD	Endog	Cible d'investissement productif	Monnaie locale constante	Produit x branche	2*2
IR	Endog	Taux d'intérêt global sur les nouveaux prêts	Points		T
IRL	Endog	Taux d'intérêt à long terme sur les nouveaux prêts	Points		T
IRM	Endog	Taux d'intérêt moyen sur les dettes	Points		T
IRS	Endog	Taux d'intérêt à court terme sur les nouveaux prêts	Points		T
K	Endog	Capital productif	Monnaie locale constante	Produit x branche	(1,2,3)*(T,1a,1i,1,2,3),T*

KD	Endog	Cible de capital productif	Monnaie locale constante	Produit x branche	2*2
KV	Endog	Capital productif à prix courants	:Monnaie locale courante	branche	1a,1i,1,2,3
L	Endog	Emploi total	Personnes Thsds	branche	T,1a,1i,1,2,3
L_G	Exog	Emploi des administrations	Personnes Thsds		
L_M	Endog	Emploi des ménages	Personnes Thsds		
LD	Endog	Cible d'emploi des entreprises	Personnes Thsds	branche	1a,1i,2,3
LF	Endog	Emploi des entreprises	Personnes Thsds	branche	1a,1i,2,3
LH	Exog	Emploi des ménages	Personnes Thsds	branche	1a,1i,2,3
LP	Endog	Productivité du travail	Personnes Thsds	branche	T,1,2,3
LPTA	Exog	Tendance de productivité du travail	Personnes Milliers	branche	1a,1i,3
LS	Endog	Salaires	Personnes Milliers	total	1,1a,1i,2,3
LS_F	Endog	Salaires versés par les entreprises	Personnes Milliers	total	T,1a,1i,2,3
LS_H	Endog	Salariés versés par les ménages	Personnes Milliers	total	T,1a,1i,2,3
M	Endog	Importations	Monnaie locale constante	total	T,1,2e,2i,3
MARG	Endog	Marges des entreprises	Monnaie locale courante	branche	1,1a,1i,2,3
MV	Endog	Importations à prix courants	Monnaie locale courante	total	T,1,2e,2i,3
NIF	Endog	Intérêts nets versés par les entreprises	Monnaie locale courante	branche	1i,2,3
NIG	Endog	Intérêts nets versés par le gouvernement	Monnaie locale courante		
NIX	Endog	Intérêts nets versés au reste du monde	Monnaie locale courante		
OIT	Endog	Autres impôts indirects	Monnaie locale courante	produit	T,1,1a,1i,1,2,3
PCOG	Endog	Déflateur de la consommation des administrations	Déflateur année de référence = 1	produit	1,2,3
PENS	Endog	Retraites	Monnaie locale courante		
PENSR	Exog	Taux des retraites	Monnaie locale constante		
PEX	Endog	Déflateur des exportations hors droits de douane	Déflateur année de référence = 1	produit	T,1,2,2e,2i,3
PEXT	Endog	Déflateur des exportations, y compris les droits de douane	Déflateur année de base = 1, USD	produit	1,2e,2i,3

PFD	Endog	Déflateur de la demande locale	Déflateur année de référence = 1	produit	T,1,2,3
PFDG	Endog	Déflateur de la demande du gouvernement	Déflateur année de référence = 1	produit	T
PFDXT	Endog	Déflateur de la demande locale hors TVA	Déflateur année de référence = 1	produit	1,2,3
PGDPM	Endog	Déflateur du PIB marchand	Déflateur année de référence = 1	produit	T
PIC	Endog	Déflateur de la consommation intermédiaire	Déflateur année de référence = 1	Produit x branche	(1,2,3)*(1,2,3)
PIG	Endog	Déflateur des investissements du gouvernement	Déflateur année de référence = 1	produit	2,3
PIHH	Endog	Déflateur de l'investissement logement des ménages	Déflateur année de référence = 1	produit	T
PIM	Endog	Déflateur des importations hors droits de douane	Déflateur année de référence = 1	produit	T,1,2e,2i,3
PIMT	Endog	Déflateur des importations, y compris les droits de douane	Déflateur année de référence = 1	produit	1,2e,2i,3
PIMU	Exog	Déflateur des importations d'énergie en dollars américains	Déflateur année de référence = 1	produit	2e
PIP	Endog	Déflateur des investissements productifs	Déflateur année de référence = 1	produit	T,1,2,3
POP	Exog	Population	Personnes Milliers		
POPAG	Exog	Population en âge de travailler	Personnes Milliers		
PP	Endog	Déflateur des prix de production	Déflateur année de référence = 1	produit	1,2,3
PQ	Endog	Déflateur de la valeur ajoutée	Déflateur année de référence = 1	produit	1,2,3
PROF	Endog	Bénéfices des entreprises	Monnaie locale courante	branche	1i,2,3
PT	Endog	Impôt sur les bénéfices des entreprises	Monnaie locale courante	branche	T,1i,2,3
PWM	Exog	Déflateur mondial pondération importations	Déflateur Année de base = 1, USD	produit	2i,3

PWX	Endog	Déflateur mondial pondération exportations	Déflateur année de référence = 1	produit	T,1,2e,2i,3
Q	Endog	Valeur ajoutée	Monnaie locale constante	produit	T,1,1a,1i,2,3
QA	Endog	Valeur ajoutée cible	Monnaie locale constante	produit	'2
QV	Endog	Valeur ajoutée à prix courants	Monnaie locale courante	produit	T,1,1a,1i,2,3
R_COGV	Exog	Ratio de la consommation du gouvernement hors salaires au PIB	Unités		
R_CPI	Exog	Ratio de CPI au déflateur de la demande finale	Monnaie locale constante		
R_EXPG	Exog	Ratio des dépenses non identifiées des administrations au PIB	Monnaie locale constante		
R_IHH	Exog	Ratio de l'investissement logement des ménages à leur revenu	Monnaie locale constante	branche	2,3
R_INCT	Exog	Taux de l'impôt sur le revenu	Monnaie locale courante		
R_IP	Exog	Ratio de l'investissement productif au PIB	Monnaie locale constante	Produit x Direction	3*(1,1a,1i,2,3)
R_LFA	Exog	Part de l'emploi des entreprises dans le total marchand		branche	1a,1i,2,3
R_OIT	Endog	Taux des autres impôts indirects (que la TVA)		branche	1,1a,1i,2,3
R_PCOG	Exog	Ratio du déflateur de la consommation de l'État à celui de la demande finale	Déflateur année de référence = 1	branche	1,2,3
R_PFDG	Exog	Ratio du déflateur la demande de l'État à celui de la demande finale	Déflateur année de référence = 1		
R_PIC	Exog	Ratio du déflateur de la consommation intermédiaire à celui de la demande (hors taxes)	Rapport	Produit x Direction	(1,2,3)*(1,2,3)
R_PIG	Exog	Ratio du déflateur de l'investissement du Gouvernement à celui de la demande finale	Rapport	produit	2,3
R_PIIH	Exog	Ratio du déflateur de l'investissement logement au déflateur de la demande finale	Déflateur année de référence = 1		

R_PIP	Exog	Ratio du déflateur de l'investissement productif au déflateur de la demande finale	Monnaie locale courante	produit	1,2,3
R_PTI	Exog	Taux de l'impôt sur les bénéfices	Monnaie locale courante	branche	1i,2,3
R_REVG	Exog	Ratio des recettes au PIB Administrations autre	Monnaie locale constante		
R_REVQ	Exog	Ratio des revenus non salariaux des ménages issus de la production, en part du PIB	Monnaie locale courante	branche	1i,2,3
R_REVX	Exog	Revenus non salariaux des ménages non issus de la production, en pouvoir d'achat	Monnaie locale courante		
R_SCF	Endog	Taux des cotisations sociales payées par les entreprises	Monnaie locale courante	branche	1,1a,1i,2,3
R_SCF_G	Exog	Taux des cotisations sociales payées par le Gouvernement	Monnaie locale courante		
R_SCF_H	Exog	Taux des cotisations sociales payées par les ménages	Monnaie locale courante		
R_SCW_G	Exog	Taux des cotisations sociales payées par les fonctionnaires	Monnaie locale constante	gouvernement	
R_SCWA	Exog	Taux des cotisations sociales par les salariés du privé	Monnaie locale courante	branche	1a,1i,2,3
R_SUBS	Exog	Taux de subvention (à la valeur ajoutée à prix courants)	Monnaie locale courante	branche	1,1a,1i,2,3
R_TARIFF	Exog	Taux de droits de douane locaux	Monnaie locale courante	branche	1,2,2e,2i,3
R_TARX	Exog	Taux de droits de douane étrangers	Monnaie locale courante	branche	1,2e,2i,3
R_UN	Endog	Taux de chômage	Monnaie locale constante		
R_VAT	Exog	Taux de TVA	Monnaie locale constante	produit	1,2,3
R_WG	Exog	Rapport du taux de salaire du Gouvernement au taux du marché	Monnaie locale courante		
R1_A	Exog	Part de l'agriculture artisanale	Monnaie locale constante		
R1_I	Exog	Part de l'agriculture industrielle	Monnaie locale courante		

RDEP	Exog	Taux de dépréciation du capital	Monnaie locale constante	Produit x Direction	(1,2,3)*(1a,1i,2,3)
RELC	Endog	Rapport coût du travail – coût du capital	Monnaie locale courante	branche	2
REVG	Endog	Revenus des administrations	Monnaie locale constante		Eq364
REVQ	Endog	Revenu des ménages, issu de la production, autres que les salaires	Monnaie locale constante	branche	T,1i,2,3
REVX	Endog	Revenus non salariaux des ménages, non issus de la production	Points		
RFDI	Exog	Part des investissements directs étrangers dans l'investissement	Points	produit	1,2,3
RMARG	Endog	Taux de marges	Points	branche	1,1a,1i,2,3
RPROB	Endog	Ratio des marges au capital au coût de renouvellement	Points	branche	1i,2,3
RPRO	Endog	Taux de profit	Monnaie locale constante	branche	1i,2,3
RS_F	Exog	Part des entreprises dans l'emploi salarié	Monnaie locale constante	branche	1a,1i,2,3
RS_H	Exog	Part des ménages dans l'emploi salarié	Monnaie locale courante	branche	1a,1i,2,3
SCF	Endog	Cotisations sociales versées par les entreprises	Personnes Milliers	branche	T,1,1a,1i,2,3
SCF_G	Endog	Cotisations sociales versées par l'État	Personnes Milliers		
SCF_H	Endog	Cotisations sociales versées par les ménages employeurs	Personnes Milliers		
SCW	Endog	Cotisations sociales versées par les salariés	Personnes Milliers		
SOCB	Endog	Prestations sociales	Personnes Milliers		
SOCBR	Exog	Prestations sociales en pouvoir d'achat	Personnes Milliers		
SUBS	Endog	Subventions aux entreprises	Personnes Milliers	branche	T,1,1a,1i,2,3
TARIFF	Endog	Droits de douane locaux	Personnes Milliers	produit	T,1,2,2e,2i,3
TC	Exog	Coefficient technique	Personnes Milliers	Produit x Direction	(1,2,3)*(1,2,3)
TRBAL	Endog	Balance commerciale	Personnes Milliers	produit	T

TTRAD	Endog	Termes de l'échange (rapport prix à l'exportation - à l'importation)	Personnes Milliers		
UN	Endog	Chômage	Monnaie locale constante		
UN_R	Exog	Chômage rural	Monnaie locale courante		
UN_U	Endog	Chômage urbain	Monnaie locale courante		
UR	Endog	Taux d'utilisation des capacités	Monnaie locale courante	branche	1,2,3
VAT	Endog	Taxe à la valeur ajoutée	Monnaie locale courante	total	T,1,2,3
W	Endog	Taux de salaire marchand	Monnaie locale courante	branche	1,1a,1i,2,3
W_G	Endog	Taux de salaire du Gouvernement	Monnaie locale courante		
WAGE	Endog	Total des salaires	Déflateur année de référence = 1	branche	T,1,1a,1i,2,3
WAGE_F	Endog	Salaires versés par les entreprises	Monnaie locale courante	total	T,2,3
WAGE_G	Endog	Salaires versés par les administrations	Monnaie locale constante		Eq271
WAGE_H	Endog	Salaires versés par les ménages	Déflateur année de référence = 1	total	T,2,3
WCOST	Endog	Coût salarial unitaire	Déflateur Année de base = 1, USD	branche	1,2,3
WD	Exog	Demande mondiale en prix constants	Déflateur année de référence = 1	produit	1,2e,2i,3
WF	Endog	Taux de rémunération versé par les entreprises	Déflateur année de référence = 1	branche	2,3
WH	Endog	Taux de salaire payé par les ménages	Déflateur année de référence = 1	branche	2,3
X	Endog	Exportations à prix constants	Déflateur année de référence = 1	total	T,1,2,2e,2i,3
XV	Endog	Les exportations à prix courants	Déflateur année de référence = 1	total	T,1,2,2e,2i,3

10.4 UN MODELE MULTI PAYS, MONO PRODUIT

Nous allons maintenant aborder la question de la construction, de la gestion et de l'utilisation d'un modèle décrivant l'économie de plusieurs pays. Nous allons également examiner la question similaire d'un modèle à un seul pays, mais plusieurs régions.

10.4.1 PREMIER SUJET: LES MODELES

Pour produire un modèle multi pays, nous devons d'abord produire ceux des pays individuels. Deux questions doivent être prises en compte:

- le degré de similitude des modèles des pays ;
- la description du Reste du Monde.

10.4.1.1 Les modèles pays

Nous supposons que les modèles par pays suivent les principes introduits ci-dessus. Ce qui importe maintenant, c'est la différence autorisée dans les spécifications, d'un pays à l'autre. Plusieurs cas peuvent être envisagés:

A - Les modèles sont autorisés à être complètement différents. Il en va de même des concepts utilisés, tels que la décomposition en produits. La seule condition est que, une fois identifiés les éléments reliant les modèles (essentiellement les variables d'échanges à prix constants et courants), ils peuvent être convertis de manière à ce que les éléments produits par un modèle puissent être transférés comme entrée pour les autres modèles. Mais, par exemple, des modèles à produit unique et multi-produits peuvent coexister. Le processus de transfert comportera alors une agrégation ou une désagrégation selon les besoins. Dans ce cas, les noms des variables seront sans doute également différents. Il convient également de tenir compte des éléments communs à un ensemble, comme la présence d'une monnaie ou d'un taux d'intérêt directeur communs.

B - Les modèles utilisent les mêmes concepts, mais le contexte économique est différent. Par exemple, un modèle peut utiliser une fonction de production Cobb-Douglas, l'autre une CES.

C - Les modèles utilisent les mêmes équations, mais les coefficients sont différents. Certaines exceptions peuvent apparaître, par exemple une variable retardée supplémentaire, mais la logique reste la même. En général, ces différences devraient être dues à l'estimation, ce qui signifie qu'un ensemble commun de données doit être disponible.

D - Les modèles utilisent les mêmes équations et les mêmes coefficients. Cela signifie que les formules ont été estimées «en données de panel» (en utilisant les données de tous les pays simultanément) ou calibrées.

On peut envisager un cas intermédiaire où les formules sont estimées en même temps, mais seulement certains coefficients prennent des valeurs communes. La mise en œuvre de cette technique est très aisée sous EViews.

Le choix de l'option dépend du cas:

- L'option A ne peut être envisagée que si les modèles existent déjà, ce qui simplifie le processus. La justification est encore plus forte si les modèles doivent continuer à vivre comme des outils indépendants.

C'est le cas par exemple pour le projet LINK, qui relie dans une version multi pays, sous la direction de l'ONU et de l'Université de Toronto, des modèles gérés comme outils de prévision indépendants par les institutions des pays participants.

Un autre exemple correspond au greffage d'un modèle national préexistant dans un environnement international, lui aussi préexistant. Par exemple, le modèle MODUX géré par l'institut statistique du Luxembourg (STATEC) a été intégré dans l'environnement MacSim (voir plus loin) afin d'enrichir ses projections et surtout ses analyses de chocs. Une telle possibilité est proposée par les gestionnaires de MacSim (et donc par l'auteur de ce livre).

L'inconvénient est bien sûr que les modèles peuvent présenter des propriétés économiques différentes, ainsi que des différences dans les mécanismes, conduisant à des sensibilités différentes aux chocs, sans que cela soit nécessairement justifié par des spécificités nationales fiables (si elles le sont, l'option A peut représenter un bon choix). Comme indiqué précédemment, l'économétrie basée sur les mêmes données peut justifier des formulations aux propriétés très différentes, même par pur hasard (deux économistes partageant la même philosophie économique peuvent atteindre des formulations différentes dans les méandres des estimations successives).

- L'option B est l'option A, sans modèles initiaux. Ensuite, les modèles seront plus faciles à gérer, mais les problèmes ci-dessus demeurent, sauf si les originalités restent limitées.
- L'option C est encore plus simple à gérer et ses propriétés faciles à interpréter. Le seul inconvénient vient des contraintes sur les formulations du modèle, qui interdit en principe l'introduction de formulations spécifiques à chaque pays. Mais bien sûr, des différences minimes peuvent être acceptées, comme une structure de retards plus longue ou l'indexation des salaires à long terme sur les déflateurs différents.

- L'option D limite les spécificités du pays voire plus. Ce qui subsiste est seulement l'effet de la taille du pays, la structure de ses partenaires (importations et exportations) et des paramètres structurels tels que la productivité du travail et du capital, le poids du secteur public, la structure fiscale, le partage des revenus de la production entre les entreprises et les travailleurs.

La seule justification est tout simplement celle-ci: que la seule différence dans les propriétés vient d'éléments indiscutables, dont ils décrivent fidèlement les conséquences.

À partir des éléments ci-dessus, nous pouvons faire la proposition suivante: si les modèles ne préexistent pas, l'option C est clairement la meilleure, avec quelques éléments de l'option B s'ils sont réellement justifiés, notamment par l'observation des caractéristiques du pays.

Par exemple, pour un pays producteur de pétrole le commerce en éléments énergétiques pourrait être individualisé. Pour le taux d'intérêt, l'identification de la politique de la banque centrale devrait conduire à l'introduction de la règle associée. Et peut-être que s'il existe une preuve suffisante que la substitution des facteurs suive une élasticité nulle ou unitaire, la fonction de production associée (facteurs complémentaires ou Cobb-Douglas) doit être utilisée.

10.4.2 DEUXIEME SUJET: LES ECHANGES ENTRE PAYS

Une fois les modèles mono pays définis, ils doivent être fusionnés en une seule entité.

Le problème principal est de rendre cohérentes les importations et les exportations de chaque pays, à prix constants et courants.

La solution évidente est d'établir des comportements cohérents avec les processus de décision réels:

- Les importations sont décidées par les pays importateurs, et les exportations s'y adaptent.
- Les prix des exportations sont décidés par les exportateurs et les prix à l'importation s'en déduisent par des identités, en tenant compte du taux de change relatif.

Une fois les importations d'un pays décidées, elles doivent être partagées entre les pays qui exportent vers lui. On peut considérer deux options:

- Identifier les importations individuelles par leur source, et les convertir en exportations individuelles du pays fournisseur. C'est évidemment la solution la plus logique. Elle présente l'avantage essentiel de rendre les exportations et les importations cohérentes au niveau mondial.
- Toutefois, certains modèles utilisent une technique plus discutable: ils calculent les exportations mondiales potentielles d'un pays en appliquant des poids constants aux importations globales de ses partenaires.

Bien sûr, les exportations totales et les exportations auront alors des valeurs différentes, et une correction doit être appliquée. Le principal avantage de cette méthode est de permettre une estimation des exportations basées sur les valeurs globales réelles. En outre, elle permet d'éviter l'identification d'un grand nombre de flux commerciaux (croissant presque comme le carré du nombre de pays).

Nous favorisons fortement la première option.

- Elle identifie plus de variables, mais l'information associée est importante. L'interprétation des conséquences de chocs (sur le commerce extérieur ou les variables associées) est beaucoup plus claire et plus informative.
- Elle ne appelle pas de corrections, donnant automatiquement des résultats cohérents.
- Elle ne nécessite pas de plus amples informations: les poids utilisés par l'autre méthode peuvent être utilisés pour calculer les flux commerciaux.
- Mais le plus important à notre sens est qu'elle localise le processus décisionnel au bon endroit: les exportations ne sont pas décidées par les exportateurs, mais par les pays importateurs qui, une fois qu'ils ont décidé d'importer, doivent choisir parmi les fournisseurs potentiels, en tenant compte essentiellement des prix relatifs.

Bien sûr, elle ne permet pas une estimation globale. Les équations individuelles de flux commerciaux devront être soit calibrées (peut-être en utilisant des valeurs obtenues par des estimations globales), soit estimées à l'aide des techniques de panel.

À partir de maintenant, nous allons nous concentrer sur cette technique.

10.4.3 UNE METHODE COHERENTE: MACSIM

Nous présentons maintenant une méthode cohérente pour définir successivement les éléments ci-dessus. Il est utilisé par le système MacSim (2001, 2013 ?).

Nous allons commencer, non pas avec les importations, mais avec le système des prix, car les prix ne dépendent que des éléments locaux, tandis que les importations dépendent de la compétitivité-prix.

10.4.3.1 Le système de prix

Nous commençons par établir un système cohérent pour les prix du commerce. Pour cela, nous partirons des formulations des prix à l'exportation (formulations estimées ou calibrées). Pour plus de clarté, nous utiliserons les équations calibrées, mais la méthode s'applique aussi aux équations estimées.

Nous supposons que l'équation traditionnelle s'applique à chaque pays client. Mais cette fois, la destination des exportations sera identifiée.

$$\text{Log}(pex_{i,j}) = a \cdot \text{Log}(pp_i) + (1-a) \cdot (\text{Log}(pp_j ch_i / ch_j)) + e \cdot t + f$$

Où i est l'exportateur et j le client, et ch_i un indice représentant l'évolution (à partir de l'année de référence) de la valeur d'une monnaie commune (en général, le dollar US) par rapport à la monnaie du pays i . Si la monnaie se déprécie de i , l'indice croît.

Pour obtenir les prix à l'importation individuels, nous corrigeons par le taux de change:

$$pim_{j,i} = pex_{i,j} ch_j / ch_i$$

et nous obtenons le prix à l'importation global grâce une moyenne pondérée:

$$pim_j = \frac{\sum_i m_{j,i} pim_{j,i}}{\sum_i m_{j,i}}$$

(nous verrons plus tard comment on calcule les flux commerciaux à prix constants).

10.4.3.2 Les taux d'utilisation étrangers

Dans les modèles à un seul pays, nous avons vu le rôle fondamental du taux d'utilisation des capacités pour le pays local. Mais nous avons supposé les capacités étrangères infinies.

Maintenant, nous pouvons considérer, pour définir les importations d'un pays, le taux d'utilisation des pays qui le fournissent. Si le taux d'utilisation augmente en Allemagne, elle perdra des parts de marché en France, par rapport à d'autres exportateurs, mais aussi aux producteurs français.

Pour obtenir la moyenne, nous allons utiliser la même méthode que ci-dessus (sans corrections de change).

$$urx_j = \sum_i m_{j,i} ur_i / \sum_i m_{j,i}$$

10.4.3.3 Les importations globales

Nous pouvons maintenant déterminer les importations globales du pays i , en modifiant légèrement l'équation de notre modèle à un seul pays, pour tenir compte de la capacité des exportateurs:

$$\text{Log}(m_i / td_i) = b \cdot [\text{Log}(ur_i) - c \cdot \text{Log}(urx_i)] + d \cdot \text{Log}(pm_i / pp_i) + e$$

Pour simplifier l'explication, nous avons présenté uniquement la partie long terme. L'équation dynamique devra subir la même modification.

Cela signifie que, comme dans les modèles à un seul pays, une diminution générale de la capacité disponible des exportateurs va réduire leurs exportations, grâce à un effet de substitution. Le coefficient c , inférieur à l'unité, tient compte de la plus grande taille de cet ensemble de pays.

10.4.3.4 Les flux bilatéraux

Enfin, nous devons séparer les importations en exportations individuelles. Encore une fois, nous prendrons en compte la compétitivité relative, et les fluctuations des capacités disponibles, par rapport à la moyenne utilisée ci-dessus. En fait, plutôt que d'estimer une équation (nous ne disposons pas des données associées) nous devons utiliser:

$$m_{i,j} = m_i \cdot b_{i,j} [1 - \alpha_i (pex_j ch_i / ch_j - pim_i) - \beta_i (ut_j - utx_i)]$$

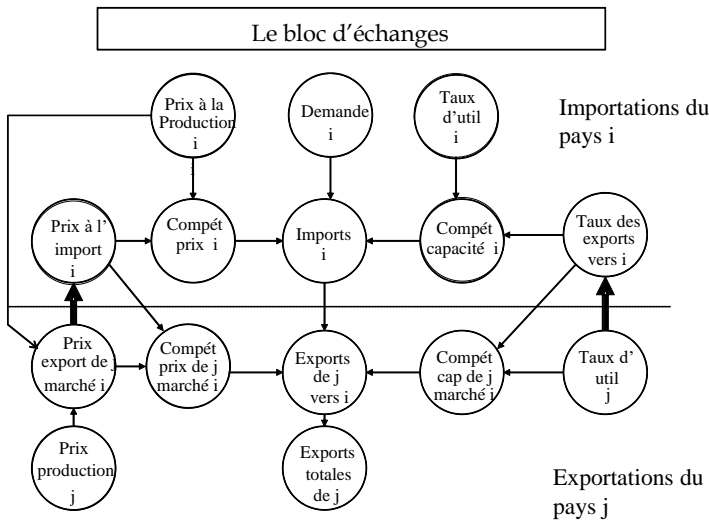
ce qui signifie que (comme pour les modèles à un seul pays) les exportateurs vers un pays augmenteront une part «naturelle» s'ils gagnent de la compétitivité et des capacités disponibles, cette fois par rapport à leurs concurrents. Plus précisément, les matrices A et B (l'une se déduisant de l'autre) représentent le

partage du commerce bilatéral associé à des taux d'utilisation et une compétitivité identiques dans tous les pays (pas nécessairement l'hypothèse la plus naturelle, ni la situation actuelle). α et β sont des scalaires fixes, et on leur attribuera une valeur plus élevée que pour l'équation d'importations (entre 0,5 et 1,0). Cela devrait représenter la plus grande polyvalence du choix de ses fournisseurs par un pays, une fois que la décision a été prise de ne pas acheter des produits locaux.

On observera que cette technique garantit l'identité de la somme des exportations individuelles avec sa valeur globale, sans aucune correction.

Bien sûr, les coefficients peuvent être différents d'un marché à l'autre, mais pas sur un même marché.

Le système peut être résumé par le graphique suivant:



En plus de ce qui précède, nous pouvons maintenant introduire des équations comptables:

Les exportations du pays i vers j :

$$x_{i,j} = m_{j,i} \cdot \text{usd}0_j / \text{usd}0_i$$

Où $\text{usd}0_i$ représente la valeur de la monnaie du pays i à l'année de base, en dollars américains.

Les exportations totales sont calculées comme la somme

$$x_i = \sum x_{i,j}$$

Le prix moyen à l'exportation du pays i :

$$pex_i = \sum pex_{i,j} \cdot x_{i,j} / x_i$$

10.4.3.5 Troisième sujet: les comportements communs

Mais nous devons aussi tenir compte des comportements communs à un ensemble de pays, qui prennent en compte des variables globales. Le cas le plus évident est la présence d'un accord global formel, comme l'utilisation d'une monnaie commune par les pays de l'Union Monétaire Européenne.

Cette option est très facile à mettre en œuvre: il suffit de créer un nouveau «pays» à l'aide des concepts globaux, en ne considérant que les variables pertinentes.

10.4.3.6 Le reste du monde

Si le champ des pays décrits est assez grand, on ne peut pas s'attendre à ce que les chocs appliqués à un grand sous-ensemble ou au modèle entier n'ait aucun effet sur l'économie du reste du monde. Cela est vrai dans notre système MacSim, qui modélise la plus grande partie de l'Europe (la zone euro + le Royaume-Uni) et aussi les États-Unis et le Japon. En outre, avec la technique que nous proposons (le calcul des parts de chaque pays dans les importations de chaque pays), endogénéiser le reste du monde permet à ses exportations de se conformer au cadre global, en évitant tout déséquilibre de comportement.

Mais bien sûr, nous n'allons pas estimer un modèle «Reste du Monde»: nous n'avons pas les données, et l'hétérogénéité y est trop élevée.

Ce que nous proposons est:

- de définir les éléments commerciaux comme d'habitude: les importations sont partagées selon le cadre habituel, et les exportations participent à la compétition comme un pays normal ;
- en outre, de créer un ensemble de multiplicateurs, définissant la rétroaction des exportations mondiales sur son PIB et ses importations, et en reliant ses importations à sa demande finale.

Ils doivent être chiffrés pour donner une réponse similaire aux modèles à un seul pays, à l'exception d'une capacité infinie (ou quasi infinie) et de l'inertie des prix.

Cela peut être fait assez simplement.

10.4.3.7 La séquence de production du modèle

Pour produire un modèle multi pays, il faut appliquer la séquence suivante.

10.4.3.7.1 Décider de l'architecture

Nous supposons que nous construisons notre modèle à partir de zéro, ce qui dans ce cas doit conduire à un ensemble de modèles de pays utilisant le même cadre économique, mais estimés séparément (ou semi séparément comme proposé plus haut), avec des originalités limitées pour les formules estimées.

Dans ce cas, il faut tenir compte:

- des données disponibles,
- du type de comportements que doit décrire le modèle,
- du degré de détail attendu.

Cela devrait permettre de décider:

- des variables utilisées par le modèle (elles doivent être utiles et disponibles),
- de la forme complète des identités,
- des comportements qui devraient être estimés, y compris leurs éléments explicatifs.

10.4.3.7.2 Produire un ensemble de modèles mono pays

Ensuite, il faut construire un ensemble de modèles mono pays, en utilisant la technique décrite plus haut. Bien sûr, un ensemble global de données doit être produit, contenant des séries pour chaque modèle, avec le même nom, mis à part un identifiant. Par exemple les exportations françaises peuvent être appelées FR_X.

Comme d'habitude, en ce qui concerne les équations de comportement, il convient de commencer par une déclaration d'intention avant de passer à des estimations.

Plusieurs séquences peuvent être prises en considération. On peut construire les modèles un par un, ou appliquer simultanément chacune des phases (production de données, le cadre, l'estimation, tests) à l'ensemble. La meilleure solution est probablement de commencer avec un pays, de dérouler la totalité du processus, et de l'appliquer à l'ensemble des autres pays, cette fois phase par phase en même temps.

Une telle stratégie permet de mesurer les difficultés, en résolvant le problème général sur un cas unique, et de définir à un coût limité la procédure complète, qui sera ensuite appliquée au cas global. En particulier, si des problèmes doivent être rencontrés une fois le projet global bien avancé, il faudra sans doute reconsidérer les premières étapes. Avoir à modifier l'ensemble des programmes peut alors devenir très coûteux.

Le choix n'est pas si simple, cependant, comme:

- L'estimation des mêmes équations sur un ensemble de pays peut être organisée assez facilement en utilisant les boucles EViews.
- Mais les processus d'estimation doivent fonctionner pour l'ensemble, car les équations ne sont pas autorisées à différer trop d'un pays à l'autre. Une direction intéressante trouvée pour un pays donné, sans être associée à ses caractéristiques spécifiques, ne peut être appliquée que si elle fonctionne aussi pour d'autres pays.

Cela signifie qu'un plus grand nombre d'itérations sera nécessaire que pour un modèle mono pays. Le seul problème est de limiter leur nombre.

10.4.3.7.3 Assembler les modèles

Une fois des versions satisfaisantes obtenues pour les modèles mono pays, les assembler est essentiellement technique.

10.4.3.7.3.1 CALCULER LES FLUX BILATERAUX

Si la formalisation du commerce entre les pays individualise des flux explicites (notre option préférée) on a besoin d'une matrice de transfert. Elle peut représenter la part des exportateurs dans les importations de chaque pays, ou la part des clients dans les exportations de chaque pays. Les deux matrices sont nécessaires, mais on peut construire l'une à partir de l'autre de manière exacte. Les éléments de la matrice doivent être considérés comme des séries. Si elles ne sont disponibles que pour certaines périodes, leurs valeurs devront être interpolées (ou extrapolées), mais dans tous les cas le modélisateur doit être autorisé à les faire varier avec le temps sur le futur. Cela est particulièrement essentiel pour les économies en transition: la part de la Russie dans les exportations polonaises a diminué au cours des dernières décennies, mais elle n'est pas stabilisée et pourrait augmenter à nouveau à l'avenir.

À partir de ces matrices, des flux d'échanges individuels peuvent être calculés. Si la matrice n'est connue qu'à prix courants (ou constants), une hypothèse globale doit être faite, la plus simple étant que le même prix à l'exportation (ou à l'importation) s'applique à un pays quelle que soit sa destination (ou son origine).

Bien sûr, réagréger le résultat dans l'autre dimension ne donne pas la valeur réelle. Par exemple, si nous partons des importations et les partageons entre les exportateurs, la somme des exportations individuelles d'un pays présentera une différence avec les comptes nationaux de ce pays. La question est la taille de cette différence: elle ne doit pas être supérieure à la valeur du résidu de l'équation estimée, sinon une correction pourrait être introduite (mais cela réintroduira une incohérence au niveau mondial).

Bien sûr, en prévision les résultats seront cohérents, par construction du système modélisé. Et les conséquences de chocs aussi, comme la différence entre les deux systèmes cohérents. Nous considérons

l'erreur observée sur le passé comme une indication de la précision globale du modèle sur l'avenir, et non pas de sa cohérence interne.

10.4.3.7.3.2 FORMALISER LE COMMERCE À PRIX CONSTANTS

Nous pouvons le faire en utilisant le cadre défini plus haut. Nous avons seulement à examiner les questions techniques suivantes:

- Les équations d'exportation des modèles simples disparaissent.
- Les flux commerciaux ont deux dimensions, et appellent à une boucle supplémentaire.

Si les flux commerciaux ne sont pas disponibles sous forme de séries, les coefficients devront être calibrés. Ils doivent être identiques pour un marché donné, et nous ne pouvons pas attendre que le choix entre les origines des produits importés utilise la même élasticité-prix que le choix entre les importations et les produits locaux.

Un des avantages de notre proposition est que la somme des flux commerciaux va correspondre aux importations, sans aucune correction, car les différences pondérées à la moyenne pondérée s'annulent naturellement.

10.4.3.7.3.3 FORMALISER LES DEFLATEURS DES ECHANGES

Si les importations sont décidées par les importateurs, le prix de la transaction est décidé par l'exportateur.

Cette fois, nous pouvons estimer le prix global à l'exportation de chaque pays (en fonction des coûts locaux et du prix pratiqué par ses concurrents). Ensuite, nous pouvons appliquer les résultats à chaque marché à l'exportation, en utilisant comme prix des concurrents le prix spécifique sur ce marché.

En dehors de cela, nous devons supposer que pour chaque exportateur la même formule s'applique à tous ses clients.

10.4.3.7.3.4 INTRODUIRE LES IDENTITES

Enfin, nous devons calculer un certain nombre d'identités, comme les importations et les exportations à prix courants, globales et individuelles.

10.4.3.7.3.5 PRODUIRE LES ELEMENTS COMMUNS

Un autre problème est la présence d'accords entre les pays, nécessitant la définition de variables communes, fondée sur un ou plusieurs éléments communs.

Par exemple, sous l'hypothèse de parité de pouvoir d'achat, le taux de change de l'euro applicable à tous les pays de la zone dépendra de l'inflation de la zone.

Des équations devront être établies:

- pour résumer les indicateurs explicatifs,
- pour le calcul de l'élément commun, donnant généralement un choix entre plusieurs options (taux d'intérêt réel ou règle de Taylor, taux de change fixe réel ou nominal).

La formule doit permettre de choisir la liste des pays concernés. Ceci est très facile:

- pour le calcul des indicateurs, l'élément de chaque pays sera multiplié par un indicateur (1 = oui, 0 = non) pour indiquer s'il participe à la somme pondérée ;
- pour l'utilisation du résultat commun, une option doit être présente dans l'équation de chaque pays, en utilisant le même type d'indicateur (si 1, utiliser la valeur commune, si 0, utiliser une formule spécifique au pays).

Même en l'absence d'accord, il peut être utile de calculer des indicateurs communs (par exemple, le PIB de l'Union Européenne).

10.4.3.7.3.6 LE RESTE DU MONDE

Le «modèle» Reste du Monde aura besoin des équations habituelles décrivant les flux commerciaux, et en plus:

- d'une équation reliant le PIB et la demande locale aux exportations ;
- d'une équation d'importations globales liée à la demande du reste du monde et à sa compétitivité prix (les prix peuvent être affectés par l'inflation des autres pays). L'introduction d'une influence du taux d'utilisation devrait dépendre de la taille de la zone.
- la répartition des importations entre les différents exportateurs se fera de la manière habituelle.

10.4.3.7.3.7 RESOUDRE LE MODELE

La procédure pour résoudre le modèle est la même. Le nombre d'équations est simplement plus élevé, et les calculs plus complexes.

Un autre problème concerne la convergence: avec moins de concepts exogènes, la solution est moins ancrée sur des hypothèses, et la probabilité de divergence risque d'augmenter. Mais on peut aussi considérer qu'un problème individuel peut se diluer dans le processus global d'équilibrage. Rien n'est clair sur cet aspect des choses.

La question la plus importante est peut-être le système de prix: avec moins d'ancrages, même si l'inflation de chaque pays converge sur le long terme vers une valeur commune, la stabilisation peut se produire à des niveaux très hétérogènes. Si les indices de prix des pays convergent vers des valeurs éloignées, l'inflation créée normalement par un choc de demande peut être compensée par l'augmentation de la part d'importations valorisées à un niveau inférieur. L'effet global peut être déflationniste, en introduisant des propriétés qui doivent être considérées comme anormales, d'autant plus qu'elles sont dues à un élément artificiel: la distance par rapport à l'année de base.

Ceci peut être résolu par la gestion de résidus, qui aura pour but d'assurer que les prix à long terme restent assez proches les uns des autres. La technique de ciblage que nous présenterons bientôt peut être utilisée ici¹.

10.4.3.7.3.8 UTILISER LE MODELE

Bien sûr, si ce modèle est plus cher à produire, il permet également un plus large éventail d'études, et les rend plus fiables par rapport au cas d'un seul pays. Comme nous l'avons dit plus tôt, le multiplicateur de la demande peut être de 20 % plus élevé si nous intégrons un modèle français dans un contexte multi pays, comparativement à une utilisation autonome du même modèle.

En plus d'une meilleure étude des changements associés aux hypothèses locales d'un pays, pour ce pays et l'étranger, ce type de modèle permet en particulier d'étudier les conséquences:

- d'événements mondiaux, tels que le ralentissement de la croissance mondiale, ou un choc sur le prix du pétrole, en tenant compte automatiquement de ses conséquences pour tous les pays. Nous n'avons plus besoin de faire des hypothèses cohérentes pour le reste du monde ;
- d'accords commerciaux mondiaux ;
- de politiques internationales (par exemple une réduction de la TVA dans l'Union Européenne).

Ces éléments peuvent être évalués dans leur ensemble, et en comparant leurs conséquences pour chaque pays.

¹ Nous l'utilisons souvent dans la pratique.

Tout cela peut être fait en tenant compte du rôle des règles communes, telles que la création de la zone euro, qui peuvent être modifiées à volonté.

10.4.4 LE PROGRAMME EViews

La séquence des programmes EViews est trop compliquée et contient trop de problèmes spécifiques pour être présentée ici. Cependant, le modèle fonctionne et peut être utilisé dans les études opérationnelles, avec des propriétés raisonnables. Il est le cœur du programme MacSim utilisé pour enseigner les mécanismes macroéconomiques internationaux. Ses éléments sont disponibles et peuvent être fournis grâce à un accord spécial.

10.5 UN MODELE REGIONAL

La construction d'un modèle régional présente les mêmes caractéristiques que le cas multi pays, avec quelques différences:

- Les régions utilisent une monnaie commune, ainsi que la même règle de taux d'intérêt. Mais des éléments spécifiques des régions peuvent faire varier le niveau de ce dernier.
- Des variables telles que les divers déflateurs sont assez voisines, mais des différences de niveau peuvent apparaître d'une région à l'autre.
- Les importations en provenance d'autres régions sont facilitées. Cela signifie à notre sens qu'au lieu d'une séquence: définition des importations puis séparation entre fournisseurs, toutes les formes de la demande d'une région doivent être séparées immédiatement entre prestataires, la région elle-même étant traitée comme les autres. Mais bien sûr, cette région doit recevoir une part plus élevée, pour un même niveau de compétitivité. Certains produits ne peuvent provenir que de la région, en particulier certains services (mais pas seulement). Un coût de transport peut également être introduit.

Mais ces informations ont moins de chances d'être disponibles.

Les migrations peuvent souvent être obtenues, et doivent donc être prises en compte, à la fois par leurs déterminants et leurs effets. Le transfert des revenus des ménages doit alors être formalisé.

On peut également tenir compte de la difficulté technique à transporter des marchandises d'une région à l'autre.

10.6 UN MODÈLE MULTI PAYS, MULTI PRODUITS

Fondamentalement, les problèmes viennent de la combinaison de deux dimensions: les pays et les produits.

En plus de combiner les fonctions décrites plus haut, la principale difficulté réside dans la complexité beaucoup plus élevée du problème. Même si les programmes EViews peuvent être rendus compacts grâce à l'utilisation de boucles, le nombre réel de variables est beaucoup plus élevé, ce qui signifie que le contrôle des propriétés du modèle sera plus complexe. En pratique, il devra être réalisé en premier lieu au niveau global, puis au niveau détaillé.

En outre, les hypothèses retenues pour les prévisions et la réponse aux chocs seront plus difficiles à établir, surtout si l'on veut profiter du détail supplémentaire (ce qui paraît logique).

Les exigences en matière de données sont bien sûr plus importantes.

10.6.1 LE PROGRAMME EIEWS

Même si un tel programme a été effectivement créé, avec succès, sa taille est trop importante pour qu'il soit présenté ici.

La technologie n'est pas trop complexe, car elle combine simplement les éléments de ses deux dimensions.

CHAPITRE 11 : NOUVEAUX ELEMENTS D'EVIEWS 8 POUR LA MODELISATION

Nous allons maintenant résumer les nouvelles options et fonctions proposées par la version 8 d'EViews, qui s'appliquent à la tâche de modélisation. Chacune d'entre elles a déjà été abordée dans la partie correspondante de l'ouvrage, mais nous avons jugé intéressant de donner aux utilisateurs une présentation synthétique.

Certaines de ces fonctions ne sont pas directement liées à la modélisation, mais permettent d'améliorer le processus.

11.1 LA GESTION DES SERIES

EViews 8 permet à l'utilisateur d'introduire ses propres labels. Ceci peut être utilisé pour caractériser les éléments du modèle, par exemple le pays pour un modèle multinational, l'agent en comptabilité nationale, le type (endogène / exogène), ou tout simplement le fait que la série est utilisée par un modèle particulier. Vous avez juste à appliquer déclarations «label» appliqué à une nouvelle catégorie. Aucune déclaration préalable n'est nécessaire.

Par exemple, vous pouvez utiliser:

```
HI.label (agent) Ménages
MARG.label (agent) Entreprises
```

Une fois qu'un type a été créé, il doit être déclaré dans son intégralité (vous ne pouvez pas utiliser «a» pour «agent »).

Si l'écran de la fenêtre workfile est en mode «Display +», vous pouvez trier les éléments en fonction de leurs caractéristiques. Outre le nom, le type et la date de dernière modification (ou création) vous avez accès à la description.

En définissant les éléments selon un certain code, vous pouvez obtenir un affichage dans lequel les éléments liés apparaîtront ensemble, dans l'ordre alphabétique.

Par exemple en utilisant:

```
HI.label (d) Modèle Fra / Ménages: revenu disponible
MARG.label (d) Modèle fra / Entreprises: Marges
```

Et lors d'un tri sur la description, les variables du modèle seront identifiées et réparties en fonction de l'agent.

Par défaut, l'écran présentant les éléments ligne à ligne (obtenu par appel à «details+» dans la barre supérieure principale) se limite aux seules descriptions. Pour accéder aux autres éléments (et donc au tri) sur les autres éléments, il convient de cliquer avec le bouton droit sur la barre supérieure (par exemple sur «Description»). Ceci fait apparaître un menu contextuel où l'on sélectionnera l'élément «Edit Columns...» (le dernier). On peut alors choisir les éléments affichés sur la ligne (y compris ceux créés par l'utilisateur).

Vous pouvez comparer les éléments entre fichiers de travail et pages à l'intérieur du même fichier de travail. EViews affiche une ligne par élément, dans lequel sera indiqué son statut, entre: inchangé, modifié (numériquement), ajouté, supprimé, remplacé (logiquement, ce dernier cas s'applique par exemple à une variable liée dont le lien a été modifié). Un filtre peut être appliqué.

Pour les séries, un niveau de tolérance peut être défini, en deçà duquel la série ne sont pas considérée comme modifiée. L'affichage indiquera le nombre de périodes montrant une différence supérieure.

Par défaut, tous les éléments seront affichés, mais on peut limiter le cas (par exemple, pour toutes les variables présentes dans les deux pages avec une différence supérieure au critère).

Les équations et les modèles ne sont pas comparés mais apparaissent dans la liste.

La commande associée est `wfcompare`.

```
wfcompare(tol = critère, list = type-de-comparaison) liste-des-series-de-référence liste-des-séries-comparées
```

Pour plus de détails veuillez vous reporter à l'Aide EViews.

Par exemple, si vous voulez comparer toutes les séries françaises (à commencer par «FRA_») entre les pages «base» et «updated», pour un niveau de tolérance de 0,00001, vous indiquerez:

```
wfcompare(tol = 1E-5, list = m) updated\fra_* base\ fra_*
```

Cette commande peut être particulièrement utile

- pour contrôler l'évolution des valeurs historiques pour un ensemble, montrant les équations qui devront être estimée à nouveau avec les données du modèle ;

- pour résumer les résultats d'une vérification résiduelle, montrant pour quelles équations le terme de droite (en utilisant les valeurs historiques de la variable expliquée) est différent du terme de gauche (le résultat du calcul). En fixant un seuil de tolérance légèrement supérieur à zéro (par exemple 0,0001) on peut limiter l'affichage aux erreurs jugées significatives.

11.2 LES PROGRAMMES

EViews 8 améliore la façon dont les programmes peuvent être exécutés.

- Vous pouvez exécuter une partie d'un programme, en sélectionnant avec la souris (selon la méthode habituelle), en cliquant sur le bouton droit et en choisissant «Run selected».

Cela est généralement plus efficace que la méthode précédente consistant à copier la partie sélectionnée dans un programme vide, et en l'exécutant. Toutefois, cette nouvelle méthode ne permet pas la modification.

- Symétriquement, on peut exclure temporairement une partie de l'exécution d'un programme, en la transformant en commentaire. Pour ce faire, il faut sélectionner la partie concernée, cliquer sur le bouton droit et choisir «Comment Selection». Pour réactiver les instructions, il faut les sélectionner à nouveau et utiliser «Uncomment Selection».

Cela peut se révéler un peu dangereux, surtout si vous (comme moi) avez le réflexe de sauvegarder le programme avant chaque exécution. Pour éviter de détruire l'original, on peut sauvegarder d'abord le programme modifié sous un autre nom.

Enfin, on peut demander d'afficher une colonne de chiffres à gauche des lignes de programme. Ceci est particulièrement efficace si vous utilisez la fonction «Go To Line».

11.3 LA GESTION DES MODELES

EViews 8 améliore fortement la gestion des équations du modèle, en particulier des identités.

Jusqu'à EViews 7:

- La suppression d'une identité n'était pas possible.
- L'ajout d'une nouvelle version dupliquait de fait sa définition, ce qui donnait plus d'équations que d'endogènes.
- Supprimer une équation estimée ne pouvait se faire qu'à travers l'ordre EXCLUDE, difficile à gérer, et pour conserver l'exclusion, il fallait préciser à nouveau celle-ci dans toutes les déclarations d'exclusion suivantes.

- Pour remplacer une équation estimée vous devez utiliser le même nom, sinon vous avez été confronté au même problème de duplication que ci-dessus.

Maintenant vous pouvez (si vous choisissez cette option):

- Supprimer n'importe quel type d'équation, en utilisant:

```
nom-du-modèle.drop nom-de-variable (pour les identités)  
nom-du-modèle.droplink nom-d'équation (pour les équations estimées)
```

Par exemple, vous pouvez spécifier:

```
PIB_fra_1.drop ' PIB devient exogène  
_fra_1.droplink eq_i ' l'équation eq_i qui donne l'investissement est abandonnée, et I devient  
exogène
```

- Remplacer la formule à l'aide de:

```
modél_name.replace nom-de-variable (pour les identités)  
modél_name.replacelink ancien-nom-équation nouveau-nom-équation (pour les équations  
estimées)
```

```
_fra_1.replace PIB= .....  
_fra_1.replacelink eq_i eq_i_new
```

- En outre, vous pouvez également remplacer toutes les occurrences d'une variable du modèle en utilisant un nouveau nom, grâce à:

```
nom-du-modèle.replacevar ancien-nom-de-variable nouveau-nom-de-variable
```

Par exemple:

```
_fra_1.replacevar PIB Q
```

remplace PIB par Q dans l'ensemble du modèle (mais pas dans les relations estimées)

11.4 LA GESTION DE LA RESOLUTION DU MODELE

Dans le chapitre sur la résolution du modèle, nous avons proposé de remplacer (override) une liste des séries en créant une nouvelle liste à partir de zéro, et en particulier de créer une liste vide pour exécuter une solution sans aucun remplacement.

EViews 8 offre une autre possibilité: vous pouvez éliminer individuellement les remplacements par la commande REVERT appliquée à une liste de séries devant revenir à la version de base:

```
nom-du-modèle.revert liste-des-séries
```

Mais l'ajout le plus important est la possibilité de spécifier directement les changements dans les hypothèses, en utilisant ADJUST.

La syntaxe est:

```
nom-du-modèle.adjust(init=série_initiale) nom-de-série ajustement
```

Par exemple, les instructions suivantes:

```
_mod_1.scenario(a=_2) "scenario 2"  
_mod_1.adjust(init="scenario 1") gdp =+10000
```

vont créer une série appelée gdp_2 avec une valeur supérieure de 10 000 à la valeur de «scénario 1» (peut-être appelé gdp_1). La série sera ajoutée à la liste «OVERRIDE», et à la liste «EXCLUDE» également si la variable est endogène.

Prenez soin d'introduire un espace avant le signe «=», sinon PIB prendra la valeur 10 000 !

L'introduction d'un changement comportant des éléments, tels que des paramètres ou des séries, est possible, mais plus difficile. On peut préférer utiliser des expressions réelles, comme nous l'avons fait dans nos exemples.

11.5 L'AFFICHAGE

Les fonctions graphiques de EViews ont été améliorées dans la version 8. Vous pouvez maintenant:

- décider de la portée du graphique à l'aide d'une règle, disponible en bas de celui-ci ;
- ajouter des flèches personnalisées ;
- exporter le graphique dans un fichier PDF.
- Mais l'ajout le plus intéressant est la commande MAKEGRAPH, spécialement adaptée aux solutions d'un modèle. La syntaxe en est:

```
nom-du-modèle.makegraph (options) nom-des-séries nom-du-graphe
```

Si aucune option n'est spécifiée la solution de scénario courant est affichée.

Les principales options sont les suivantes:

```
a: ajouter les valeurs observées  
c: comparer le scénario actif à la référence («baseline»)  
d: ajouter les écarts par rapport à la référence (comme un graphique supplémentaire dans le même cadre)  
n: ne pas inclure le scénario actif.
```

Les résultats peuvent être présentés bruts ou transformés, et pour les simulations stochastiques un intervalle de confiance peut être affiché.

Une fois que le graphique est créé, les modificateurs habituels peuvent être appliqués (légende, type de ligne, couleurs ...).

Il faut noter que le graphique ne doit pas préexister. Si c'est le cas il faut d'abord le supprimer (en utilisant le modificateur «noerr» pour éviter un message d'erreur possible).

```
delete(noerr) nom-du-graphe
```

ANNEXE : UNE METHODE SIMPLE POUR INVERSER LE STATUT ENDOGENE – EXOGENE DES VARIABLES DU MODELE

Nous présenterons maintenant une méthode permettant de contraindre un modèle à atteindre des valeurs données pour un ensemble de variables endogènes, en libérant les valeurs d'un ensemble équivalent d'éléments exogènes.

Cette méthode permet de traiter des simulations à périodes uniques ou multiples, les anticipations rationnelles pouvant être traitées sans augmenter fortement sa complexité (mais beaucoup le temps de calcul)

Dans un premier temps, nous allons considérer la résolution pour une seule période d'un modèle tourné vers le passé.

LE PROBLEME

Considérons le modèle

$$y = f(y, x)$$

avec:

y vecteur de dimension endogène m

x vecteur de dimension n endogène

Nous avons supprimé l'indice de temps, car les valeurs passées (et futures) ne changeront pas dans le processus de résolution.

Supposons que nous voulons obtenir:

$$y^s = y^{s*}$$

Pour un sous-ensemble de y de dimension p

par l'endogénéisation d'un sous-ensemble de x avec la même dimension, appelé x^s .

LE CAS LINEAIRE

Si f est une fonction linéaire:

$$y = A \cdot y + B \cdot x + b$$

nous aurons

$$y = (I - A)^{-1}(B \cdot x + b)$$

Appelons C la sous matrice $(I - A)^{-1} \cdot B$ (dimension $(p \times n)$ réduite aux lignes y^s et C^s la sous matrice $(p \times p)$ de C associée aux éléments de x^s .

Nous avons :

$$y^{s*} = C^s \cdot (x^{s*} - x^{s0}) + C \cdot x^0 + c$$

Les valeurs permettant d'atteindre y^{s*} sont obtenues par:

$$x^{s*} = x^{s0} + (C^s)^{-1}(y^{s*} - C \cdot x^0 - c)$$

$$x^{s*} = x^{s^0} + (C^s)^{-1}(y^{s*} - y^{s^0})$$

La matrice C peut être facilement obtenue par une transformation, une fois que A, B et b sont connus.

LE CAS NON-LINEAIRE

Considérons maintenant le cas non-linéaire. Nous avons encore:

$$y = f(y, x)$$

et nous voulons toujours obtenir:

$$y^s = y^{s*}$$

Nous commencerons par supposer que le modèle est quasi-linéaire, et appliquer la méthode ci-dessus. La matrice C devient:

$$\partial g(y) / \partial g(x)$$

Où g représente la fonction $y = g(x)$ où le vecteur y a été éliminé du terme de droite.

Il est malheureusement clair que l'identification de la fonction g est difficile, voire impossible dans la pratique. Par contre, les valeurs de la matrice jacobienne sont faciles à obtenir, ou son approximation par différences finies.

Il suffit pour cela de résoudre le modèle pour des valeurs données de x, puis d'appliquer de petits chocs successifs à chacune des exogènes, et calculer la solution associée. Nous obtenons:

$$\partial g(y_i) / \partial g(x_j) \approx (g_i(x + \Delta x_j) - g_i(x)) / \delta x_j$$

où Δx_j est un vecteur de dimension n aux valeurs nulles sauf pour x_j , modifié par δx_j .

Pour obtenir l'équivalent de la matrice C , il suffit de sélectionner dans la matrice jacobienne les colonnes et les lignes associées à x^s et y^s .

La solution utilise la même formule que le cas linéaire:

$$x^{s*} = x^{s0} + (C^s)^{-1}(y^{s*} - y^{s0})$$

On remarquera que la résolution du modèle et le calcul des modifications sont nécessaires uniquement pour les exogènes concernées, et que l'inversion s'applique à une matrice ($p \times p$) très réduite.

La seule différence avec le cas linéaire, c'est que la solution n'est pas exacte, en raison de l'approximation linéaire. Mais il suffit de poursuivre le processus, à partir des nouvelles valeurs obtenues pour x (en fait limité à x^s). Le processus reste exactement le même, et peut être répété autant de fois que nécessaire. Et comme l'approximation linéaire devient de plus en plus précise (en termes de propriétés) au fur et à mesure que nous nous approchons de la solution (surtout si le modèle est localement convexe), la part de l'erreur restante éliminée à chaque fois devient de plus en plus élevée, et la convergence devrait être atteinte après quelques itérations.

Nous rencontrons ici une méthode de type Newton, avec une propriété de convergence supra linéaire. Une caractéristique intéressante est que nous connaissons la précision réelle de la solution, contrairement à d'autres algorithmes qui ne considèrent que la vitesse de convergence.

QUELQUES INTERROGATIONS

Cependant, plusieurs questions apparaissent:

- La méthode que nous suggérons semble assez simple pour un logiciel qui ne propose pas un processus automatique (comme Troll), et représente une alternative efficace à des outils plus complexes (et sans doute moins efficaces). On peut se demander pourquoi il n'est pas proposé d'emblée par les logiciels, et appliqué plus souvent par les modélisateurs.

Plusieurs réponses peuvent être envisagées:

- Il ne fonctionne pas dans la pratique, ou seulement dans des cas très simples (non opérationnels). Nous allons prouver qu'il fonctionne de manière générale. Il n'y a aucune raison (mais cela reste à prouver) qu'il ne fonctionne pas aussi pour la majorité des cas raisonnables.
- Il ne fonctionne pas toujours. Mais nous pouvons au moins essayer, et peut-être l'améliorer dans les cas «difficiles». On peut s'attendre à deux cas principaux d'échec:
 - Le modèle est fortement non linéaire. C'est le principal problème avec la méthode de Newton, qui peut entraîner la solution vers des valeurs absurdes si la matrice jacobienne calculée est très éloignée de sa valeur «exacte» (à la solution). Il est essentiel de partir des valeurs assez proches de la solution, ce qui devrait être le cas dans les situations opérationnelles, si la qualité de la prévision du modèle est suffisamment élevée pour fournir spontanément des solutions raisonnables, sinon acceptables.
 - Le Jacobien est très colinéaire, et ne peut pas être inversé, ou suggère une fois inversé des changements très élevés des exogènes (au moins de deux d'entre elles). Ce cas est moins problématique, car une autre méthode («à la main» incluse) sera confrontée aux mêmes difficultés, et ici au moins les problèmes seront observés immédiatement, et peut-être interprétés à travers l'observation de la matrice. Le temps gagné peut être utilisé pour chercher dans d'autres directions.
- Elle ne s'applique pas au problème traité. En effet, le procédé est limité à la recherche d'un ensemble donné de n valeurs endogènes en libérant n éléments exogènes. Un modélisateur professionnel voudra peut-être répartir le poids sur plusieurs éléments exogènes. Mais si la pondération de la répartition entre instruments est connue, la méthode peut encore être appliquée: plusieurs exogènes seront affectées, mais l'algorithme ne recherchera qu'une seule valeur. Par exemple, si les taux de cotisations sociales des employeurs et des travailleurs sont utilisés simultanément avec la même variation, cette variation devient l'unique élément endogénéisé.

Plus problématique est le cas des limites posées sur les changements¹: une donnée exogène est utilisée jusqu'à ce qu'il atteigne une limite, et dans ce cas le calcul bascule vers un nouvel élément. Ce cas peut également être traité automatiquement, mais le programme devient plus complexe, et la solution n'est pas garantie dans les limites acceptables (ni la méthode "à la main" de toute façon).

- Enfin, si la présentation de la méthode a été faite sur un processus rétrospectif, elle peut également être appliquée aux modèles à anticipations rationnelles. Deux techniques peuvent être envisagées, déjà présentées pour le cas simple:
 - Fair Taylor: nous itérerons sur la solution période par période, initialisant les valeurs futures par celles de la dernière itération (au début, les valeurs historiques, théoriques ou calculées précédemment).
 - Lafargue: nous résolvons un modèle global où les équations sont répétées autant de fois que la dimension temporelle.

Mais maintenant, nous devons tenir compte de deux boucles imbriquées, et le processus peut devenir d'une lenteur inacceptable. Des améliorations spécifiques n'ont pu encore être trouvées.

UN TEST SUR UN TRES PETIT MODELE.

Nous présenterons ici les conditions et les résultats d'un essai effectué sur un modèle très petit, très voisin de l'exemple utilisé au début (17 équations).

Nous avons introduit un écart sur le PIB (appelé Q) et les importations à prix constants (M) respectivement de +5% et -5% par rapport aux valeurs simulées pour la même période. Le changement de la valeur va bien au-delà des cas habituels, et leur sens est plutôt opposé: une augmentation du PIB génère des importations, même si toutes choses étant égales par ailleurs une augmentation des importations réduit le PIB.

Nous allons utiliser comme instruments le taux de salaire réel WR et la demande mondiale WD. Ces éléments sont plutôt indépendants, d'abord les uns des autres (cela devrait accroître leur efficacité et leur précision, comme pour une équipe aux compétences différentes), et aussi des éléments ciblés.

¹ Essentiellement dûr des raisons sociologiques : on ne peut supposer qu'une augmentation de 20% de l'impôt sur le revenu soit acceptée par les ménages.

Nous allons donner:

- Le programme EViews, assez court en instructions (mais pas en commentaires). Il peut s'appliquer à n'importe quel modèle, sous forme de sous-programme, les seules informations demandées (en plus du nom du modèle lui-même) étant la période concernée, les listes de variables ciblées (avec leurs valeurs) et d'éléments de ciblage.
- Les résultats des itérations. Dans cette liste, le suffixe "_cur" est associé aux valeurs actuelles (endogènes et exogènes) et le suffixe "_star"aux cibles endogènes.

Les valeurs «delta» représentent la différence entre la cible et la solution courante, pour chaque endogène ciblée. Nous constatons que nous avons besoin dans la pratique de 4, souvent 3 itérations (la première consistant à simuler le modèle avec les valeurs initiales). Le gain relatif le plus élevé est obtenu lors de la deuxième (la troisième doit faire face à l'imprécision d'EViews).

A titre indicatif, l'exécution du programme prend 0,003 secondes sur un ordinateur portable de puissance moyenne.

Nous allons ensuite reproduire ce test sur un modèle de plus grande taille.

Le programme

```
'=====
```

```
' Ce programme calcule pour un ensemble de variables exogènes
' les valeurs qui permettent à un ensemble de facteurs endogènes
' de trouver des solutions données
```

```
' Les deux ensembles doivent avoir la même dimension
```

```
' Le modèle pic_b est utilisé comme un exemple
' Mais le programme est assez général
```

```
' Le présent programme s'applique à une seule période
' mais il peut être généralisé assez facilement
```

```
'=====
```

```
' Nous décidons du répertoire
```

```
cd "c:\program files\reviews5\pic\"
```

```
' Les résultats seront transférés sous forme de texte dans un fichier appelé jacob.txt
' remplaçant toute version précédente
```

```
output(r,o) jacob
```

```
'
' Nous créons un fichier de travail spécifique appelé pic_star
' à partir du pic_b original

' Nous veillons à ce qu'aucun fichier de ce nom ne soit ouvert à l'heure actuelle
```

```
close pic_b.wf1
close pic_star
```

```
wfopen pic_b
wfsave pic_star
```

```
' Nous créons un nouveau modèle
```

```
if @isobject("_pic_b")=1 then
delete _pic_b
endif
```

```
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

```

' Nous commençons le programme lui-même

' Nous utiliserons la période 2001S1

```

smpl 2001S1 2001S1

```

' Nous définissons les listes d'éléments qui changent de statut

' d'exogène à endogène

' d'endogène à exogène

' Les nombres doivent être les mêmes

' Sinon, le programme s'arrête avec un 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 est le nombre commun

```

scalar nstar=g_vexog1.@count

```

' Nous supprimons Scenario 1 (il existe)

```

_pic_b.scenario(d) "Scenario 1"

```

' Nous construisons le scénario 0 comme base
 ' Il contiendra la solution courante
 ' Le suffixe associé sera "_0"
 ' Nous déclarons toutes les exogènes modifiées
 ' (Avec la meilleure solution courante)

' (Note: on commence avec une liste vide que l'on remplit élément par élément)

```

_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

```

' Pour le calcul du Jacobien
 ' Les scénarios sont numérotés de 1 à nstar (ici 2)
 ' Le suffixe est "_n"
 ' Les exogènes sont les mêmes que pour le scénario 0

- ' Remarque: tous les exogènes vont être remplacées
- ' Mais une seule va changer à partir de la valeur _0
- ' Pour un scénario donné

```
for !i=1 to nstar
_pic_b.scenario(n,a={!i},i="Scenario 0") "Scenario {!i}"
_pic_b.append assign @all _{!i}
Next
```

'

Maintenant nous donnons les valeurs des cibles

```
smpl 2000S1 2001S1
genr q_star=q*(1+0.05*(t=2001))
genr m_star=m*(1-0.05*(t=2001))
```

- ' Nous définissons le Jacobien

```
matrix(2,2) jacob
```

- ' Maintenant, nous initialisons les instruments exogènes
- ' utilisés pour atteindre l'objectif

```
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
```

' et les variables endogènes cibles

```
for !i=1 to g_vendo1.@count
%1=g_vendo1.@seriesname(!i)
genr {%1}_0={%1}
genr {%1}_cur={%1}_0
next
```

' Nous fixons le nombre maximum d'itérations
' et le critère de convergence

```
scalar nitmax = 20
scalar sconv = 1e-6
```

' Nous initialisons les variables de contrôle
' Le nombre d'itérations
' et la variable contrôlant la convergence

```
!niter = 1
!iconv = 0
```

' Nous doublons les éléments afin qu'ils puissent être affichés
' (ce n'est pas possible pour les paramètres « ! »)

```
scalar niter=0
scalar iconv=0
```

' nous affichons certains éléments pour le contrôle du processus

```
print g_vexog1 g_vendo1
```



```
print *_cur *_star
```

```
'-----  
'   Nous commençons la boucle  
'-----
```

```
smpl 2001S1 2001S1
```

```
' Elle sera répétée tant que la convergence n'est pas atteinte  
' et le nombre maximum de boucles non plus
```

```
while !niter<=nitmax and !iconv=0
```

```
'   Nous testons la convergence  
'   Nous supposons d'abord qu'elle est atteinte (!Iconv = 1)  
'   Ensuite, nous observons la différence relative  
'   entre la solution et la cible présente  
'   pour chaque endogène ciblée  
  
'   Si l'écart relatif est supérieur au critère  
'   pour n'importe quelle cible  
'   la convergence n'est pas encore atteinte(! Iconv = 0)  
'   et nous nous arrêtons
```

```

!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

```

' Nous incrémentons le compteur d'itérations

```

!niter=!niter+1
scalar niter=niter+1

```

' Nous affichons les différences

```

print delta_*

```

' Nous calculons la solution de base
' en tant que Scenario 0

```

_pic_b.scenario "Scenario 0"

```

```
solve _pic_b
```

' Les valeurs de la solution sont définies comme courantes

```
for !i=1 to g_vendo1.@count
%1=g_vendo1.@seriesname(!i)
genr {%1}_cur={%1}_0
next
```

' Nous affichons la solution courante (exogènes + endogènes) et les cibles

```
print niter
print *_cur *_star
```

' Nous calculons le Jacobien
' Nous commençons une boucle sur les éléments choqués

```
for !i=1 to g\_vexog1.@count
```

' Pour chaque exogène, nous calculons l'ensemble des instruments sélectionnés
' en choquant la valeur courante de 0,1%
' Tandis que les autres conservent leur valeur de base
' Cela est nécessaire pour mettre en œuvre les changements en cours

```
for !j=1 to g_vexog1.@count
%2=g_vexog1.@seriesname(!j)
genr {%2}_(!i)={%2}_cur*(1+.001*(!i={!j}))
next
```

' Nous résolvons le modèle sous le scénario !i

```
_pic_b.scenario "Scenario {!i}"
_pic_b.solve
```

- ' Maintenant, nous calculons la variation relative de la cible endogène
- ' Cela vous donnera une colonne de la matrice jacobienne

```
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
```

- ' Ceci termine le calcul du Jacobien
- ' Nous calculons son inverse

```
matrix jacob_inv=@inverse(jacob)
```

- ' Nous appliquons à chaque exogène un changement égal à l'erreur résiduelle.
- ' multipliée par l'inverse de la matrice jacobienne
- ' Cela donne une nouvelle solution pour l'exogène
- ' Nous recommençons le processus jusqu'à la convergence (espérons-le)

```
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

```

LES RÉSULTATS

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  la solution du modèle donne des erreurs importantes.
DELTA_2 -0.095208   Le ciblage n'a pas encore commencé
=====

```

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

```

UN TEST SUR DE PLUS GRANDS MODELES

Nous avons reproduit le test sur des cas plus opérationnels.

Remarque: l'application du programme à un nouveau modèle est immédiate, à condition qu'il soit disponible dans un fichier de travail EViews avec les données requises. Il suffit de mettre à jour les deux listes de cibles et d'instruments, et définir des valeurs pour les cibles. En tout, cela devrait prendre moins de dix minutes.

Premier cas: le modèle trimestriel de l'économie française, que nous venons d'utiliser pour illustrer la modélisation sous EViews. Il comprend 88 équations, dont 11 sont des comportements. Ses

caractéristiques sont tout à fait traditionnelles, et toutes les estimations sont validées par l'économétrie. Il contient une boucle keynésienne et une boucle prix-salaires.

Nous avons également choisi un cas plus opérationnel. Nous demandons au modèle de respecter :

- Un taux de chômage de 6%
- Un taux d'inflation de 2%.
- Une balance commerciale nulle (= exportations importations à prix courants).
- Un déficit budgétaire de 1% du PIB.

Pour cela nous allons utiliser les instruments suivants:

- Les taux de cotisations sociales payées par les entreprises et les travailleurs.
- Les investissements publics à prix constants.
- Le nombre de fonctionnaires.

L'algorithme converge à nouveau, cette fois en 5 itérations. Le temps de calcul est 0,231 secondes.

Deuxième cas: un modèle multi-pays annuel pour la Communauté Andine (5 pays dont le Venezuela et un reste du monde simplifié). Il comprend 770 équations, dont 11 x 5 de comportement. Ses spécifications sont encore assez traditionnelles, et toutes équations de comportement sont calibrées pour le moment. Il formalise complètement les flux commerciaux. Sa logique suit globalement celle que nous avons décrite plus haut (modèle MacSim).

Cette fois, nous avons pris pour cibles le déficit public de chaque pays (à 3% du PIB) et les instruments successifs:

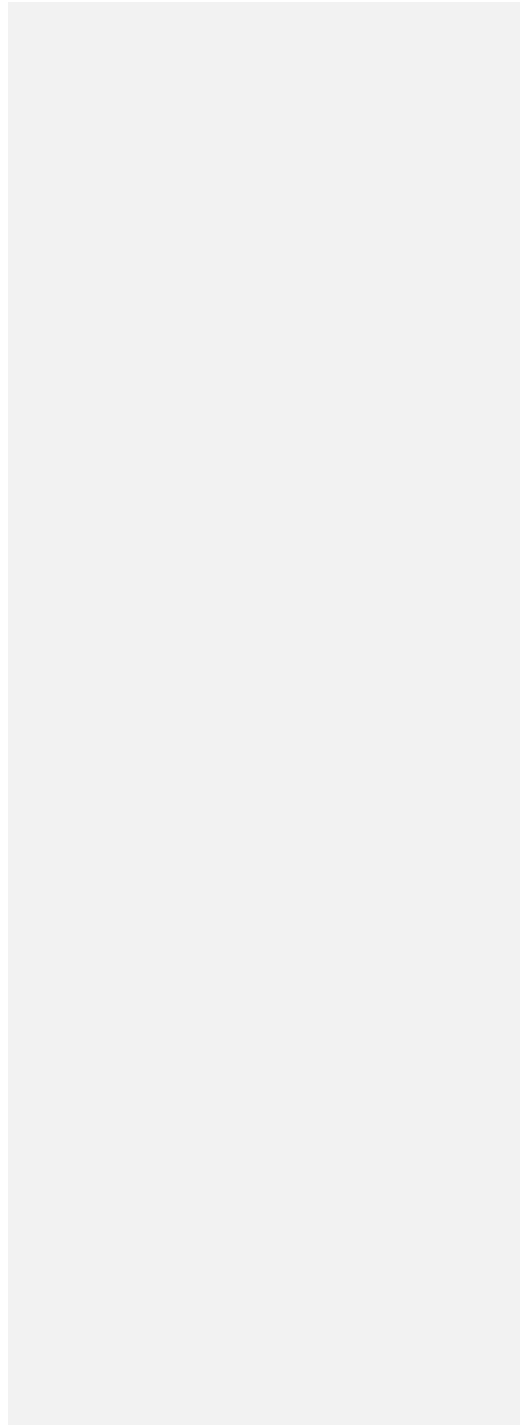
- Les investissements publics
- Le taux de TVA
- Le taux des cotisations de sécurité sociale payées par les travailleurs
- Le taux des cotisations sociales payées par les entreprises.
- Le taux de change
- Le taux d'intérêt à long terme.

Le système converge dans tous les cas, en un temps compris entre 3,7 et 5,2 secondes (5 itérations la plupart du temps).

et enfin avec deux éléments:

- objectifs: Le déficit public de chaque pays (à 3% du PIB) et l'équilibre des échanges (exportations nettes = 0)
- instruments: l'investissement du gouvernement et des subventions aux entreprises.

converge en 7 .2 secondes.



BIBLIOGRAPHIE

Nous avons limité notre exposé aux notions élémentaires d'économétrie, et n'avons utilisé la macro-économie que pour illustrer certaines méthodes. Nous vous donnons maintenant une liste d'ouvrages de référence moins basiques ou plus spécialisés.

ECONOMETRIE ET STATISTIQUE

BALTAGI B.H. (2008) *Econometrics*, Springer

BALTAGI, B.H., (2007) *Econometric Analysis of Panel Data*, Wiley, 4th Ed.

BANERJEE A., GALBRAITH J. W., DOLADO J., HENDRY D. F., (1993) *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data (Advanced Texts in Econometrics)*, Oxford University Press

CARTER HILL R., GRIFFITHS W.E. (2011), *Principles of Econometrics*, 4th edition, John Wiley & Sons

CARTER HILL R., LIM G.C. (2011), *Using Eviews for Principles of Econometrics*, 4th edition, John Wiley & Sons

CHAREMZA W., DEADMAN D.,(1997) *New Directions in Econometric Practice: General to Specific Modelling, Cointegration, and Vector Auto regression*, Edward Edgar Pub

CLEMENTS P., HENDRY D. F. , (1998), *Forecasting Economic Time Series*, Cambridge University Press.

CORBAE D., STINCHCOMBE M. B., ZEMAN J. (2009) *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*, Princeton University Press

DIEBOLD, F.X., (2000), *Elements of Forecasting*, South Western, 2ème Ed.

ENDERS, W., (2004), *Applied Econometric Time Series*, Wiley Series in Probability and Statistics, Wiley, 2nd Ed.

GOMEZ V., MARAVALL A.(1996). *Programs TRAMO (Time series Regression with ARIMA noise, Missing observations, and Outliers) and SEATS (Signal Extraction in ARIMA Time Series). Instructions for the User. Working Paper 9628, Banco de España.*

GOURIEROUX C., MONFORT A., GALLO G. (translator) (1997) *Time Series and Dynamic Models (Themes in Modern Econometrics)* Cambridge University Press.

- GOURIEROUX C., MONFORT A., VUONG Q. (translator) (1995) *Statistics and Econometric Models (Themes in modern econometrics)*, Cambridge University Press.
- GREENE W.H., (2011) *Econometric Analysis*, Prentice Hall, 7th ed.
- GREENE W.H., (2011), *Econometric analysis*, Pearson , 7th Ed.
- GUKARATI D. N. (2010) *Basic Econometrics*, 5th edition , McGraw-Hill Companies (first published 1995)
- HAMILTON J D., (1994), *Times Series Analysis*, Princeton University Press.
- HARVEY G. C. (1990) *The Econometric Analysis Of Time Series* MIT Press (MA)
- HAYASHI F. (2000) *Econometrics* , Princeton University Press
- HEIJ C., FRANSES P. H. (2004), *Econometric Methods with Applications in Business and Economics* Oxford University Press, USA
- HENDRY D. F., (1995) *Dynamic Econometrics*, Oxford University Press
- JOHNSTON J. (1990), *Econometric Methods* Irwin/McGraw-Hill
- JUSELIUS K., (2006), *The Cointegrated VAR Model: Methodology and Applications*, Oxford University Press.
- KENNEDY P. (2008) *A Guide To Econometrics* , , 6th edition, Blackwell Publishers
- LADIRAY D, QUENNEVILLE B, (2001), *Seasonal adjustment with the X-11 Method*, Springer.
- LIM G.C. (2011), *Studyguide for Using Eviews for Principles of Econometrics* , Academic Internet Publishers
- LÜTKEPOHL H., KRÄTZIG, (2004), *Applied Time Series Econometrics*, Cambridge.
- MACKINNON J. G. (2003), *Econometric Theory and Methods*, Oxford University Press, USA
- MADDALA GS, (1994), *Econometric Methods and Applications (Economists of the Twentieth Century)*, Edward Elgar Pub.
- NGURAH AGUNG I. G. (2009), *Time Series Data Analysis Using Eviews*, John Wiley & Sons
- PINDYCK R.S., RUBINFELD D. L. (2000) *Econometric Models And Economic Forecasts*
- RAO B, (2007), *Cointegration: For the Applied Economist*, St Martins Press, 2nd Edition.

ROBERT E. LUCAS, Jr. (1976), 'Econometric policy evaluation: a critique.' Carnegie-Rochester Conference Series on Public Policy' 1, pp. 19–46.

STOCK J. H., WATSON M. W. (2010), Introduction to Econometrics, (Addison-Wesley Series in Economics), 3rd edition, Wesley Publishing Company

THOMPSON J. R. (2011), Empirical Model Building: Data, Models, and Reality, John Wiley & Sons

WOOLRIDGE J. M. (2005) Introductory Econometrics: A Modern Approach, Thomson South-Western

X-12-ARIMA Reference Manual, (2011), Version 0.3, Time Series Research Staff, Statistical Research Division, U.S. Census Bureau, WWW: <http://www.census.gov/srd/www/x12a/>

MACROECONOMIE

AGENOR, P.-R., MONTIEL, P.J., (2008), Development Macroeconomics, Princeton University Press, 3rd Ed.

AGHION P, HOWITT P, (2008), The Economics of Growth, MIT Press.

ARNOLD R. A. (2013) Macroeconomics, 11th edition, South Western Educational Publishing (first published December 1995)

BARRO R. J. (1997) Macroeconomics MIT Press (MA) 5th edition

BARRO R.J. et X. SALA-I-MARTIN, (2003), Economic Growth, Mc Graw-Hill, New York (2nd edition).

BEHRMAN, J., & SRINIVASAN, T.N., (Eds.) 1995, Handbook of Development Economics, Vol III., Elsevier.

BLANCHARD O. (2012) Macroeconomics, 6th edition, Prentice Hall

BLANCHARD O. J. , FISHER S. (1989), Lectures on Macroeconomics MIT Press (MA)

BURDA M, WYPLOZ C (2009), Macroeconomics, An European text, 5th ED , Oxford University Press.

CHENERY, H., SRINIVASAN, 1989, Handbook of Development Economics, North Holland, Amsterdam.

GALI, J (2008), Monetary Policy, Inflation, and the Business Cycle. Princeton University Press

GROSSMAN, G., M., 1995, Handbook of International Economics, Vol.III, Elsevier, North Holland.

LEQUILLER F, BLADES D, (2006), Understanding National Accounts, OECD Publishing, downloadable at : <http://www.oecd.org/dataoecd/37/12/38451313.pdf>

MANKIW N. G. (2012) Macroeconomics, 8th edition, Wh Freeman (first published 1991)

MANKIW N. G., Principles of Macroeconomics, 6th Edition, Thomson South-Western

MUET P A, Théories et modèles de la macroéconomie, (1990) Economica.

OBSTFELD M., ROGOFF K. S. (1996), Foundations of International Macroeconomics, MIT Press (MA)

RAY, D., (1998), Development Economics, Princeton University Press,

ROMER D. (2011), Advanced Macroeconomics, 4th edition, Mcgraw Hill Higher Education

SACHS J. D., BONO (Foreword) (2006), The End of Poverty, Penguin Books

STACHURSKI J. (2009), Economic Dynamics: Theory and Computation, MIT Press.

TAYLOR, J.B., (1993) "Discretion versus Policy Rules in Practice," Carnegie-Rochester Conference Series on Public Policy, 39, pp.195-21

TAYLOR, J.B., and WOODFORD, M., (1999), Handbook of Macroeconomics, Elsevier, North Holland.

WOODFORD M (2003), Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press

MODELES ET MODELISATION

AMMAN H, KENDRICK D.A., RUST J., (eds.), (1996), Handbook of computational economics : H vol. 1. North-Holland, Amsterdam,

AMMAN H, RUSTEM B, WHINSTON A (1996) Computational Approaches to Economic Problems (Advances in Computational Economics, Vol. 6, Kluwer Academic Publishing.

ALLEN C., ed (1997) Macroeconomic Modelling In A Changing World: Towards A Common Approach John Wiley & Sons

ARTUS P., DELEAU M., MALGRANGE P. (1986) , Modélisation Macroéconomique, Economica.

BODKIN R., KLEIN L., MARWAH K., (1991), A History of Macroeconometric Model Building. Edward Elgar.

BRILLET JL (1994) Modélisation Econométrique: Principes et Techniques, Economica.

CANOVA F. (2007), Methods for Applied Macroeconomic Research, Princeton University Press.

CARNOT, N. , TISSOT, B (2002), La Prévision Economique, Economica.

DEJONG, D. N. with DAVE C. (2007), Structural Macroeconometrics. Princeton University Press

- DERVIS, K., DE MELO, J. et ROBINSON, S., 1982, General Equilibrium Models for Development Policy, World Bank, Washington D.C..
- FAIR R. C. (1994), Testing macroeconomic Models, Harvard University Press.
- FAIR R. C., TAYLOR, J.B. (1983), Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models, *Econometrica*, vol. 51(4), pages 1169-85, July.
- ISHER P. (1992), Rational Expectations in Macroeconomic Models (Advanced Studies in Theoretical & Applied Econometrics, Vol. 26), Kluwer.
- FOUQUET D. et alii (1978) DMS, Modèle Dynamique Multi Sectoriel, Collections de l'INSEE no 64-65.
- GANTMACHER, F. (1959), Theory of matrices, AMS Chelsea publishing
- GILLI M. (1996), Computational Economic Systems: Models, Methods & Econometrics (Advances in Computational Economics, Vol 5), Kluwer Academic Publishing.
- GILLI M., ROSSIER E. (1980), Understanding and Solving Complex Systems, *Automatica*, Vol. 17, pages 647-652.
- HALL S. G., HENRY S. C. B. (1989) Macroeconomic Modelling, North Holland
- INTRILIGATOR M., BODKIN G., HSIAO C.. (1995), *Econometric Models, Techniques and Applications*, Prentice Hall Press.
- JANSEN S. J., BARDESEN G., EITRHEIM O.
- , NYMOEN R. (2005), *The Econometrics of Macroeconomic Modelling*, Oxford University Press
- WALLIS K. F., ed. (1988) *Models of the U.K. Economy: A Fourth Review by the Esrc Macroeconomic Modelling Bureau* Oxford University Press, USA
- KLEIN L. (1991) *Comparative Performance of U.S. Econometric Models*, Oxford University press
- LAFFARGUE, J. P. (1990) Résolution d'un modèle économique avec anticipations rationnelles, *Annales d'Economie et de Statistiques* no 17.
- LAXTON D. (1998) *Multimod Mark III: The Core Dynamic and Steady-State Models*, FMI
- MALGRANGE P., MUET P. A. (1985), *Contemporary Macroeconomic Modelling*, Blackwell Publishers
- MCADAM P., HUGHES HALLETT A. J. (1999) *Advances in Computational Economics, Volume 12: Analyses in Macroeconomic Modelling*, Springer

NEPOMIASCHY P., RAVELLI A. (1978), Adapted Methods for Solving and Optimizing Quasi--Triangular Econometric Models , Annals of Economic and Social Measurement.

PINDYCK R., RUBINFELD D. (1997), Econometric Models and Economic Forecasts; Cambridge University Press.

PONTRAGIN L.S., BOLTIANSKII V.G., AMKRELIDZE R. V., MISCHENKO E. F. (1962) The Mathematical Theory of Optimal Processes (Russian), English translation: Interscience.

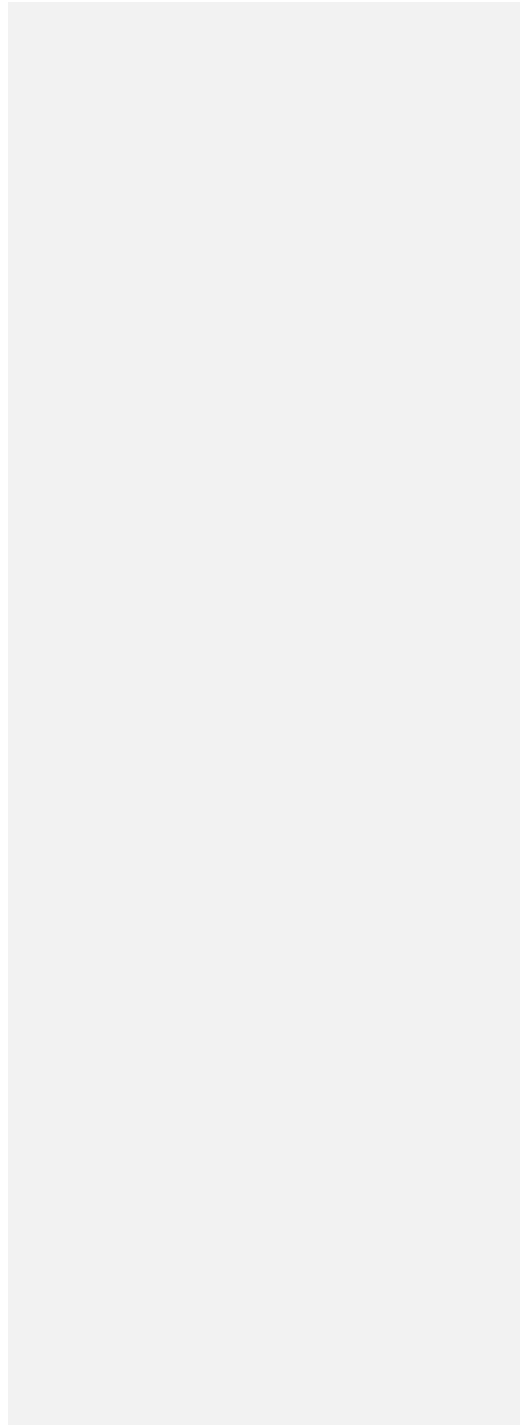
SADOLET E., DE JANVRY A., (1995), Quantitative Development Analysis, John Hopkins University Press.

SBODORNE A., TAMBALOTTI A., RAO K., WALSH K., (2010), 'Policy analysis using DSGE models: an introduction'. Federal Reserve Bank of New York Economic Policy Review 16 (2).

SHOVEN J., WHALLEY J. (1992) Applying General Equilibrium (Cambridge Surveys of Economic Literature) Cambridge University Press

TAYLOR, L. (Ed.) (1990), Socially Relevant Policy Analysis, Structuralist Computable General Equilibrium Models for the Developing World, MIT Press.

WHITNEY, J.D. (1994), A Course in macroeconomic modelling and forecasting, Harvester Wheatsheaf.



- @count, 118, 120, 122, 223, 284, 516
- @seriesname, 120
- Absence de solution, 113
- add, 119
- ADJUST, 283, 503
- agents économiques, 19, 20, 26, 29, 37, 61-62, 67, 83, 246, 287, 311, 344, 351-353, 363, 465
- agrégation temporelle, 81
- agriculture, 26, 465, 466, 469, 472, 474, 484
- algorithme, 41, 44, 71, 227, 232, 234, 235, 240, 244-247, 253-255, 271, 274, 296, 312, 337, 345, 435, 508, 520
- amortissement, 23, 92, 97, 176, 331, 356, 359, 405, 412
- analyse des modèles, 63
- année de base, 94, 114, 157, 367, 374, 443, 492, 497
- anticipations rationnelles, 37, 70, 287, 311, 505, 508, 526
- append, 105
- artisanal, 469
- artisanale (agriculture), 470, 472, 475, 484
- Census-X12-ARIMA, 34
- CES, 303, 355, 466, 487
- CGE, 19, 20, 297
- champ (du modèle), 61
- chocs, 57, 285, 339, 340
- chômage, 26, 27, 77, 157, 288, 296, 357, 358, 360, 364, 366, 402, 410, 415, 423, 448, 458, 459, 461, 465, 472, 474, 484, 519
- Chow breakpoint test, 182
- ciblage (des simulations), 338
- Clay-Clay, 356
- Cobb-Douglas, 28, 129, 255, 355, 395, 402, 412, 413, 426, 428, 434, 441, 447, 458, 466, 471, 472, 487, 488
- cointégration, 124, 150, 151, 156, 157, 158, 191, 197, 198, 201, 202, 203, 204, 205, 207, 213, 214, 216, 217, 239, 302, 316, 318, 322, 414, 416, 418, 433
- commerce extérieur, 209, 234, 247, 273, 286, 287, 292, 327, 339, 359, 366, 448, 452, 465, 466, 473, 475, 489
- compatibilité globale (spécifications), 45
- compétitivité, 157, 216, 341, 358, 359, 361, 368, 369, 425, 427, 448, 452, 453, 455, 457, 458, 465, 467, 468, 469, 472, 473, 489, 491, 496
- comportement, 19, 20, 21, 26, 27, 28, 30, 31, 32, 33, 35, 37, 39, 41, 47, 52, 54, 55, 60, 64, 65, 67, 72, 89, 91, 98, 112, 113, 117, 131, 144, 150, 161, 172, 179, 182, 183, 190, 191, 223, 225, 278, 279, 280, 285, 286, 287, 292, 298, 310, 311, 316, 317, 319, 326, 333, 337, 338, 341, 343, 344, 354, 356, 359, 360, 364, 365, 372, 374, 392, 395, 430, 448, 465, 520
- comptabilité nationale, 84, 351, 500
- comptes nationaux, 58, 64, 77, 158, 344, 353, 361, 365, 495

- consommation, 23, 27, 28, 34, 88, 89, 94, 95, 97, 190, 275, 286, 291, 326, 352, 365, 366, 367, 415, 423, 448, 452, 458, 465, 473, 479, 481, 482, 483
- consommation intermédiaire, 46, 93, 97, 207, 353, 359, 367, 368, 375, 465, 468, 470, 474, 482, 483
- continuité, 41
- contraintes de spécification, 54
- convergence, 48, 54, 71, 128, 149, 154, 178, 225, 228, 231, 232, 234, 235, 240, 244, 245, 246, 247, 248, 251, 252, 253, 254, 255, 256, 259, 261, 262, 263, 266, 270, 271, 274, 283, 292, 293, 295, 297, 298, 301, 303, 304, 305, 312, 316, 318, 322, 324, 325, 406, 433, 444, 463, 497, 507
- convergence (étude de la), 235
- convergence de la solution, 232
- convexité (du modèle), 45
- correction d'erreur, 124, 126, 150, 153, 155, 156, 157, 158, 162, 172, 178, 179, 270, 273, 280, 291, 299, 302, 304, 330, 332, 358, 395, 399, 412, 418, 419, 423, 424, 471, 473
- corrélation, 31, 126, 138, 151, 176, 217, 319, 322
- cotisations sociales, 49, 286, 355, 360, 361, 402, 405, 443, 483, 484, 520
- court terme, 19, 63, 64, 174, 203, 212, 213, 279, 298, 299, 303, 338, 340, 362, 366, 368, 369, 371, 423, 433, 434, 450, 460, 466, 469, 471, 480
- cycles, 180, 181, 265, 274, 292, 297, 298, 304, 305, 316, 435
- cyclique (processus), 273, 274, 291, 298, 323
- cycliques (processus), 232, 304
- déficit budgétaire, 91, 298, 337, 371, 519
- demande de l'Etat, 26
- demande finale, 23, 29, 88, 95, 207, 212, 358, 359, 368, 374, 447, 458, 460, 482, 483, 493
- demande locale, 23, 26, 88, 95, 96, 203, 207, 367, 448, 450, 465, 472, 481
- demande mondiale, 23, 26, 30, 88, 96, 97, 157, 192, 200, 203, 205, 287, 299, 306, 331, 332, 341, 353, 367, 368, 425, 472
- dérivées, 39, 40, 230, 254, 259, 272, 318
- désagregation temporelle, 82
- désaisonnalisation, 83
- déterministe, 261, 318, 320
- diagnostics opérationnels, 57
- Dickey-Fuller, 152, 184, 185, 214
- Dickey-Fuller (test), 151
- différences finies, 230, 253, 254, 255, 507
- différenciabilité, 42
- dimension temporelle, 33
- discrétisation, 33
- documentation des données, 85
- données (traitement), 77
- droits de douane, 59, 286, 360, 361, 453, 455, 461, 462, 465, 481, 482, 484
- drop, 116, 119, 171, 191
- droplink, 171
- DSGE, 19, 20, 527
- Durbin et Watson, 161

- Durbin et Watson (test), 137
- Durbin-Watson, 126, 138, 162, 165, 173, 185, 193, 196
- dynamique, 20, 21, 37, 144, 153, 154, 155, 156, 157, 178, 203, 204, 209, 261, 273, 277, 279, 280, 302, 318, 322, 323, 325, 327, 328, 330, 331, 332, 334, 358, 363, 365, 412, 418, 427, 433, 447, 459, 471, 491
- dynamiques (modèles), 35
- économétrie, 30
- emploi, 23, 28, 33, 77, 88, 91, 93, 97, 158, 174, 178, 179, 180, 183, 185, 187-190, 269, 288, 291, 355, 357, 358, 362, 366, 395, 397-402, 404, 410, 434, 450, 466, 469, 470, 481, 483, 484
- endogène, 76, 283, 303, 500, 505
- endogènes, 25, 26, 45, 49, 89, 97, 99, 105, 107, 124, 207, 222, 231, 261, 294, 301, 308, 311, 312, 318, 337, 338, 392, 432, 443, 448, 464, 502, 505
- endogènes (variables), 25
- énergie, 470
- entreprises, 23, 26, 28, 32, 35-37, 50, 61, 65, 72, 77, 83, 88-97, 159, 164, 166, 172, 175-179, 190, 195, 207, 212, 213, 248, 273, 286, 292, 316, 324, 333, 338, 341-362-369, 374,-381, 387, 391, 396, 402, 404, 405, 415, 441, 442, 448, 450, 455, 459, 465-467, 470, 473, 479-485, 488, 520
- épargne, 27, 28, 94, 95, 288, 352, 364-366, 423
- Équations, 25, 27
- équations de comportement, 27
- équilibre général calculable, 19
- équilibre offre-demande, 29, 37, 46, 47, 246, 290, 331, 347, 367, 447, 453, 472
- erreur (simulation), 114
- estimation, 67, 124
- estimation (méthodes), 124
- estimations, 69
- EViews 8, 85, 102, 119, 122, 171, 283, 373, 500, 501, 502, 503
- ex ante, 100, 290, 291, 333, 341, 365, 447, 448, 457
- ex post (simulations), 274, 278, 279, 281, 285, 290, 292, 319, 322, 333, 342, 365, 447, 448, 450
- Excel, 67, 69, 70, 71, 78, 79, 80, 81, 85, 89, 102, 110, 117, 265, 293, 295, 308, 348, 443
- EXCLUDE, 171, 265, 280, 283, 308, 502, 503
- existence d'une solution, 42
- exogène, 26, 50, 51, 55, 94, 127, 207, 247, 288, 361, 500, 505
- exogènes, 25, 30, 39, 45, 49, 76, 89, 97, 98, 107, 124, 270, 281, 302, 303, 306, 318, 323, 332, 338, 341, 362, 370, 377, 392, 431, 443, 464, 505
- exogènes(variables), 25
- exportation, 205, 360, 420, 448, 450, 458, 472, 479, 485, 495, 496
- exportations, 23, 26, 29, 30, 43, 88, 96, 157, 158, 191, 192, 195, 198, 200, 203-207, 213, 220, 234, 288, 292, 299, 352, 353, 358, 361, 367-369, 424, 452, 458, 460, 462, 465, 469, 470, 472, 479, 481, 482, 486, 488-493, 495, 519, 520
- ex-post (simulations), 280
- facteurs complémentaires, 28, 174, 178, 255, 355, 356, 395, 412, 434, 439, 447, 458, 466, 488
- facteurs de relaxation, 248, 249, 251, 266, 267, 272

- Fair-Taylor, 312
- fiabilité (du modèle), 58, 68, 203, 207, 281, 285, 297, 318, 319
- fonction de production, 28, 129, 285, 354, 356, 357, 374, 395, 402, 412, 458, 471, 472, 487, 488
- format (des données), 78
- G.@COUNT, 105
- Gauss-Seidel, 52, 71, 227, 228, 234, 235, 240, 244, 245, 247, 248, 249, 253, 254, 255, 256, 261, 263, 264, 266-272, 274, 304, 312, 435, 444
- gestion des données, 69
- GROUP, 119
- groupes, 119
- Hodrick-Prescott (filtre), 357
- homogénéité, 51
- homogénéité dynamique, 330, 418
- homoscédasticité, 127, 158, 159
- horizon (des simulations), 63
- identification, 51
- identifié (système ou équation), 112, 228, 229, 232, 337, 363
- identités, 19, 27, 28, 55, 67, 72, 89, 98, 99, 111, 112, 116, 117, 124, 171, 226, 303, 370, 371, 372, 392, 430, 443, 474, 489, 494, 496, 502
- importations, 23, 34, 41, 43, 88, 95, 152, 157, 192, 206, 207, 209, 212-219, 234, 281, 288, 292, 352, 353, 358, 361, 367-369, 420, 427, 448, 452, 457, 458, 460, 462, 465, 469, 470, 472, 479, 482, 488-491, 493, 495, 496, 519
- impôt sur le revenu, 50, 286, 352, 363, 365, 483
- impôts indirects, 360, 361, 465, 481, 483
- inflation, 26, 27, 28, 30, 33, 39, 50, 127, 326, 341, 365, 366, 423, 448, 450, 452, 453, 458, 461, 497, 519
- information complète, 130
- information limitée, 130
- intervalle de confiance, 318, 319, 349, 504
- investissement, 23, 27, 37, 77, 88, 91, 94, 95, 97, 105, 158, 172, 173, 175, 203, 213, 291, 312, 314, 315, 316, 352, 353, 366, 370, 374, 395, 404, 405, 433, 448, 452, 455, 458, 465, 466, 470, 480, 482, 483, 484, 520
- investissements directs étrangers, 467, 484
- itératif (processus), 130
- jacobien, 230, 231, 240, 246, 250, 252, 253, 255, 270, 317, 507, 508
- Les sentiers de croissance équilibrée, 328
- linéarité (des formules), 38
- LINK, 103, 109
- linkto, 110
- lissage, 83
- logiciel, 21, 68-71, 78-81, 87, 225, 293, 294, 295, 345, 347, 508
- long terme, 19, 21, 26, 44, 51, 63, 151, 154-158, 168, 177, 180, 183, 187, 191, 197, 203, 206, 212, 213, 226, 279, 285, 287, 291, 297, 298-305, 308, 318, 323, 328, 330, 333, 340, 369, 370, 371, 374, 402, 412, 420, 425, 431, 433-435, 447, 450, 452, 453, 461, 463, 471, 473, 480, 488, 491, 497, 520

- macroéconomie, 60, 346, 350, 525
- MacSim, 49, 60, 63, 319, 367, 488, 489, 493, 497, 520
- MAKEGRAPH, 348
- makegroup, 115, 116
- matrice d'incidence, 235, 236, 239
- maximisation, 65, 70, 295, 296, 345, 362, 402
- ménages, 23, 27, 28, 29, 32, 33, 35, 36, 37, 42, 45, 46, 49, 52, 61, 62, 65, 72, 77, 83, 88, 89, 93, 94, 97, 99, 100, 247, 276, 285, 286, 288, 291, 324, 325, 327, 331, 340, 341, 351, 352, 353, 354, 357, 363, 364, 365, 366, 370, 376, 378, 389, 423, 448, 465, 467, 473, 474, 479, 480, 481, 482, 483, 484, 485, 486, 498
- mises à jour (des données), 84
- MODEL, 103
- modèle à correction d'erreur, 152
- modèle mondial, 26, 49, 75, 83, 256, 336, 340, 342, 367, 453
- modèle multi pays, 108, 353, 487, 494
- modèles opérationnels, 38, 62, 126, 157, 268, 350, 355, 369
- modèles théoriques, 60
- moindres carrés, 124
- moindres carrés (non linéaires), 128
- moindres carrés ordinaires, 125, 127, 128, 149, 160, 194, 203, 316
- monotone, 252, 316, 323, 463
- monotone (processus), 273
- Monte Carlo, 317
- moyen terme, 19, 20, 64, 279, 281, 292, 298, 302, 305, 339, 340, 447, 448, 452
- multiplicateur, 39, 212, 220, 226, 247, 248, 272, 273, 287, 288, 292, 325, 333, 367, 447, 448, 460, 497
- Newton, 228, 231, 232, 234, 240, 249, 250, 251, 252, 253, 254, 255, 256, 259, 261, 266, 267, 269, 270, 271, 272, 304, 312, 507, 508
- noerr, 104
- nomenclature, 83
- non récursif (bloc), 240
- non récursif (bloc), 263
- non récursif (bloc), 268
- non-linéarités, 299, 318
- normalisation, 51
- normalisé (modèle), 236
- organisation du travail, 87
- page, 109
- panel (données de), 131
- paramètres, 30
- pays en transition, 343, 344
- périodicité, 64
- politique économique, 21, 57, 68, 276, 285, 287, 336, 337, 340, 346
- politiques économiques, 226, 318
- POOL, 131, 132
- pools, 131

- préparation du modèle, 67
- présentation (des résultats), 70
- prévisions, 19-21, 41, 57, 59, 63, 68, 226, 235, 239, 261, 274, 276, 279-280, 287, 292, 293, 297, 299, 301, 304, 317, 336-338, 342-345, 361, 365, 435, 444, 488, 498, 508
- prévisions tendanciennes, 336, 338
- prix de la demande, 48, 359, 360, 420, 450, 458, 459, 474
- problèmes de convergence, 232
- production potentielle, 341, 354
- productivité, 23, 26, 91, 174, 175, 176, 178, 179, 180, 181, 182, 183, 185, 187, 190, 286, 287, 299, 302, 306, 307, 332, 341, 360, 397, 399, 407, 415, 434, 465, 469, 481, 488
- programmation, 119
- Putty-Clay, 356
- R2, 140-143, 147, 161, 162, 166, 168, 221, 410
- récuratif, 240, 256, 266, 440, 441, 443
- récuratif (bloc), 77, 108, 240, 264, 268, 440, 441
- réursive, 115, 240, 268
- réursive (structure), 76, 77
- règle de Taylor, 30, 326, 327, 328, 343, 371, 496
- replacelink, 171
- replacevar, 172
- RESID, 101, 102, 163, 184
- résidu d'estimation, 31
- résidus, 32, 68, 101, 113, 116, 117, 125, 138, 139, 142, 143, 149, 151, 162, 163, 166, 170, 181, 183, 193, 222, 223, 225, 239, 274, 277, 279, 280, 281, 283, 285, 288, 304, 309, 317, 319, 338, 341, 430, 443, 474, 497
- résolution, 69
- résolution (du modèle), 67
- retards échelonnés, 127
- REVERT, 283, 503
- Ritz-Jordan, 228, 271
- ruptures, 143, 145, 181, 182, 399
- saisonnalité, 34
- SCENARIO, 282, 294
- scénario, 57, 261, 266, 282, 283, 290, 305, 308, 336, 337, 348, 503, 504
- scénarios, 30, 57, 63, 71, 261, 281, 295, 336, 340, 341
- sentier stationnaire, 433
- simulation de base, 57, 178, 227, 299, 308, 314, 338, 339
- simulations, 274
- simultané (bloc), 54, 108, 441
- simultané (blocs), 440, 441
- simultanée (estimation), 129
- simultanée (structure), 58, 76
- simultanés (blocs), 240, 255, 440
- SOLVE, 115, 261, 267, 269, 301, 308
- spécification (du modèle), 54

- stationnaire (processus), 152, 156, 197, 198, 399
- stationnarité, 150
- statique, 261
- statiques (modèles), 35
- statistique de Student, 126, 128
- Statistique de Student, 147
- stochastique, 235, 261, 320, 322
- structure par blocs, 76
- structurel, 20, 26, 58, 287, 297, 341, 345, 350, 369, 464
- structurels (modèles), 19
- suppression (des données), 84
- systèmes, 133
- taille des modèles, 61
- taux apparent (impôt), 363, 370
- taux d'utilisation, 213, 314, 315, 444, 490
- taux de salaire, 28, 29, 45, 50, 94, 97, 100, 127, 277, 306, 345, 355, 358, 359, 364, 376, 387, 405, 414, 458, 465, 484
- taux d'intérêt, 26, 30, 50, 288, 326, 328, 356, 358, 364, 366, 370, 374, 405, 488, 496, 520
- taux d'utilisation, 157, 174-178, 195, 198, 203, 205, 207, 212-216, 219, 221, 291, 316, 331, 356, 357, 368, 369, 374, 395-397, 404, 408, 412, 425, 427, 434, 448, 450, 466, 474, 490, 491, 496
- tendance, 50, 140, 142, 143, 145, 148, 150, 159, 162, 163, 166-168, 179, 180-183, 187, 200, 209, 211, 277, 299, 300, 303, 370, 396, 399, 408, 420, 423, 425, 433, 434
- terme (court, moyen, long), 64
- terme aléatoire, 31
- terme constant, 149
- test de rupture de Chow, 144
- tests (estimation), 136
- Tramo-Seats, 34, 83
- Transferts, 351
- transformations formelles, 83
- transformations temporelles, 81
- TVA, 360, 361, 368, 370, 374, 384, 437, 465, 468, 481, 483, 484, 497, 520
- un modèle très simple, 23
- unicité de la solution, 44
- UPDATE, 117, 266
- valeurs propres, 246, 253, 255, 270, 271, 274, 304, 318, 322, 323, 326, 328, 332
- valeurs propres (du Jacobien), 323
- validation (du modèle), 274
- VAR, 19, 20, 201, 203-205, 220, 297, 299, 433, 524
- variable muette, 189, 190, 273, 400
- variables de bouclage, 108, 239, 240, 241, 244, 246, 252, 253, 254, 255, 267
- variables muettes, 34, 278, 400
- variantes, 57, 228
- variations de stocks, 159

vérification résiduelle, 87, 111, 114, 115, 222, 223,
262, 271, 302, 372, 501

wfcompare, 86, 87, 501

Windows, 88, 293, 294, 296

X-13ARIMA, 83