

Article

« Quelques développements récents des méthodes macroéconométriques »

Alain Monfort

L'Actualité économique, vol. 68, n° 1-2, 1992, p. 305-324.

Pour citer cet article, utiliser l'information suivante :

URI: <http://id.erudit.org/iderudit/602069ar>

DOI: 10.7202/602069ar

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <http://www.erudit.org/apropos/utilisation.html>

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : erudit@umontreal.ca

QUELQUES DÉVELOPPEMENTS RÉCENTS DES MÉTHODES MACROÉCONOMÉTRIQUES*

Alain MONFORT

Département de la Recherche de l'INSEE

RÉSUMÉ — Au cours des dix ou quinze dernières années, les méthodes macroéconomiques ont profondément évolué. Ces développements se sont effectués dans des directions très variées, si bien qu'à l'heure actuelle le paysage est beaucoup plus complexe qu'il ne l'était au milieu des années 70. Le but de cet article est de fournir aux non-spécialistes de l'économétrie une description rapide de ce nouveau paysage. De façon à obtenir une image aussi fidèle que possible il nous a semblé qu'il fallait aborder cinq types de modèles : les modèles VAR, les modèles non stationnaires, les modèles ARCH, les modèles de moments généralisés et les modèles de déséquilibre.

ABSTRACT — *Some Recent Developments in Macroeconometrics.* During the last ten or fifteen years, the macroeconometric methods have been deeply transformed. These developments have been made in various directions and the situation is now much more complicated than it was in the middle of the seventies. The aim of this article is to propose a brief description of this situation for people who are not specialists of econometric methodology. Among the various possible topics we found five of them particularly important: the VAR models, the nonstationary models, the ARCH models, the GMM models and the disequilibrium models.

INTRODUCTION

Les développements des méthodes statistiques appliquées aux données macroéconomiques ont été très variés et très importants au cours des dix ou quinze dernières années. Les causes de ces évolutions sont diverses et il nous semble que les principales sont au nombre de quatre.

Tout d'abord l'économétrie, comme la plupart des disciplines scientifiques, n'échappe pas à certaines oscillations entre des écoles prônant l'empirisme et des écoles mettant en avant la théorie. Or, sous l'influence des travaux fondateurs de la *Cowles Commission* puis de la *Cowles Foundation*, une vision théorique reposant sur la notion de modèle structurel et expliquant les comportements des agents a

* Une première version de ce texte a été publiée dans les Actes du colloque de l'Association Française de Sciences Économiques *La méthodologie théorique et appliquée aujourd'hui*, Nathan, Paris, 1990.

été très longtemps dominante. Il n'est donc pas surprenant que certains économètres aient tenté de faire revenir le balancier vers davantage d'empirisme; c'est le cas en particulier de «l'école des VAR» qui incontestablement a pris une place importante dans les développements récents de la macroéconométrie.

D'autres évolutions des méthodes économétriques se sont produites pour des raisons strictement opposées aux précédentes, à savoir l'influence de la science économique; il faut citer en particulier le rôle de percées aussi différentes que la théorie du déséquilibre ou la théorie des anticipations rationnelles. Ces théories ont posé, et posent encore, des problèmes économétriques nouveaux qui ont généré de nombreux travaux et des résultats importants.

Une troisième raison du développement rapide des méthodes macroéconométriques est certainement la prodigieuse évolution des matériels et des logiciels informatiques. La micro-informatique, en particulier, permet maintenant d'effectuer à un faible coût des calculs que l'on pouvait difficilement envisager sur gros système il y a une quinzaine d'années. Ces nouvelles possibilités ont stimulé la recherche dans de nombreux domaines de l'économétrie: économétrie non linéaire, économétrie semi-paramétrique ou non paramétrique, économétrie dynamique...

Enfin, il est certain que l'économétrie a bénéficié des développements de deux disciplines voisines: la théorie des probabilités et la statistique mathématique. Ainsi l'économétrie des modèles non stationnaires a utilisé des résultats récents de la théorie des processus aléatoires, les méthodes économétriques fondées sur des simulations se sont appuyées sur divers théorèmes de théorie asymptotique, l'économétrie non paramétrique a largement puisé dans le domaine statistique correspondant...

Ces causes variées ont entraîné des développements dans des directions très différentes, si bien que le paysage actuel des méthodes économétriques n'a plus l'unité qu'il pouvait avoir jusqu'au milieu des années soixante-dix. Jusqu'à cette époque la recherche s'articulait autour d'un modèle dominant, le modèle à équations simultanées linéaires, et son objectif central était la construction de méthodes simples à mettre en œuvre pratiquement. Aujourd'hui la situation est beaucoup plus complexe et il n'est pas question de la décrire de façon exhaustive; nous allons cependant essayer d'en donner une image assez fidèle en présentant, aussi simplement que possible, cinq types de modèles introduits récemment: les modèles VAR, les modèles non stationnaires, les modèles ARCH, les modèles de moments généralisés et les modèles de déséquilibre.

1. LES MODÈLES VAR

Dans les années trente, les membres de la *Cowles Commission* ont préconisé une approche théorique de l'économétrie fondée d'une part sur des modèles structurels expliquant les comportements des agents et d'autre part sur des méthodes statistiques, qualifiées alors de «modernes», qui sont devenues depuis les méthodes classiques de la statistique inductive probabiliste [voir Haavelmo, 1943 et 1944; Koopmans, 1947; Koopmans-Rubin-Leipnik, 1950; Koopmans-Hood, 1953].

Cette conception s'est opposée à celle du *National Bureau of Economic Research* qui reposait sur une description des phénomènes, sans recherche explicative, et qui récusait l'utilisation de la théorie des probabilités dans les méthodes statistiques. Ce dernier point n'est plus aujourd'hui un sujet de controverse parmi les économètres, la statistique inductive s'étant largement imposée, en revanche l'opposition entre explication et description est toujours d'actualité.

La manifestation la plus récente et la plus importante de cette opposition est l'introduction des modèles VAR par Christopher Sims (1980). Pour Sims, la méthode traditionnelle de construction des modèles économétriques, celle de la *Cowles Commission*, comporte trop d'hypothèses non testées, comme l'exclusion de variables de certaines équations pour atteindre l'identification, le choix des variables exogènes, la forme de la distribution des décalages temporels... Il propose donc de traiter toutes les variables de façon symétrique, sans condition d'exclusion ou d'exogénéité, et de faire intervenir autant de retards pour toutes les variables dans toutes les équations. Le modèle général qu'il propose est un modèle auto-régressif vectoriel (VAR) qui peut s'écrire sous la forme :

$$Y_{1t} = \phi_{11}^{(1)} Y_{1,t-1} + \dots + \phi_{11}^{(p)} Y_{1,t-p} + \dots + \phi_{1n}^{(1)} Y_{nt-1} + \dots + \phi_{1n}^{(p)} Y_{nt-p} + \varepsilon_{1t} \tag{1}$$

$$Y_{nt} = \phi_{n1}^{(1)} Y_{1,t-1} + \dots + \phi_{n1}^{(p)} Y_{1,t-p} + \dots + \phi_{nn}^{(1)} Y_{nt-1} + \dots + \phi_{nn}^{(p)} Y_{nt-p} + \varepsilon_{nt}$$

ou en utilisant l'opérateur retard L et des notations plus concentrées :

$$Y_t = \Phi^*(L)Y_t + \varepsilon_t \tag{2}$$

ou encore :

$$\Phi(L)Y_t = \varepsilon_t \text{ [avec } \Phi(L) = I - \Phi^*(L)\text{]} \tag{3}$$

$\Phi(L)$ étant une matrice ($n \times n$) de polynômes en L de degré p .

Dans cette modélisation le processus $\{\varepsilon_t\}$ est un bruit blanc, de matrice de variance-covariance Ω , s'interprétant comme l'innovation du processus $\{Y_t\}$ et chaque équation est l'équation de régression d'une composante de Y_t sur son passé et sur le passé des autres composantes. Ces équations n'ont donc pas pour but de représenter des comportements économiques mais de mettre en évidence des régularités statistiques, en particulier de capter aussi finement que possible les interactions dynamiques entre ces variables.

L'estimation du modèle (3) est particulièrement simple puisqu'elle peut être effectuée par la méthode des moindres carrés ordinaires (m.c.o.) appliquée à chaque équation; les propriétés asymptotiques des estimateurs obtenus sont les propriétés habituelles de convergence et de normalité si le processus $\{Y_t\}$ est stationnaire, c'est-à-dire si les racines du polynôme en L $\det\Phi(L)$ sont à l'extérieur du cercle unité.

L'inversion de la matrice $\Phi(L)$ permet d'exprimer Y_t en fonction des innovations présentes et passées, on obtient la forme moyenne mobile :

$$Y_t = \Phi^{-1}(L)\varepsilon_t \tag{4}$$

$$Y_t = \Theta(L)\varepsilon_t \text{ en posant } \Phi^{-1}(L) = \Theta(L).$$

d'un modèle structurel en variante; ces phénomènes étaient déjà connus, mais quelquefois oubliés et rarement testés.

2. LES MODÈLES NON STATIONNAIRES

Les économètres, qu'ils soient classés du côté «empirique» ou du côté «théorique», se sont longtemps limités aux processus stationnaires, c'est-à-dire aux processus possédant des propriétés invariantes par translation dans le temps. Pour se ramener au cas stationnaire, diverses procédures ont été utilisées, en particulier l'élimination d'une tendance déterministe ou, selon les recommandations de Box et Jenkins (1970), la transformation par différences premières, ou éventuellement différences secondes, des séries initiales.

Cette méfiance à l'égard du non stationnaire reposait sur le fait que les procédures économétriques classiques, construites pour un monde stationnaire, peuvent avoir des propriétés curieuses quand elles sont appliquées à des séries non stationnaires. Prenons quelques exemples.

Considérons deux marches aléatoires indépendantes $\{Y_t\}$ et $\{X_t\}$, c'est-à-dire deux processus définis par:

$$\begin{aligned} Y_t &= Y_{t-1} + \varepsilon_t \\ X_t &= X_{t-1} + \eta_t \end{aligned} \tag{6}$$

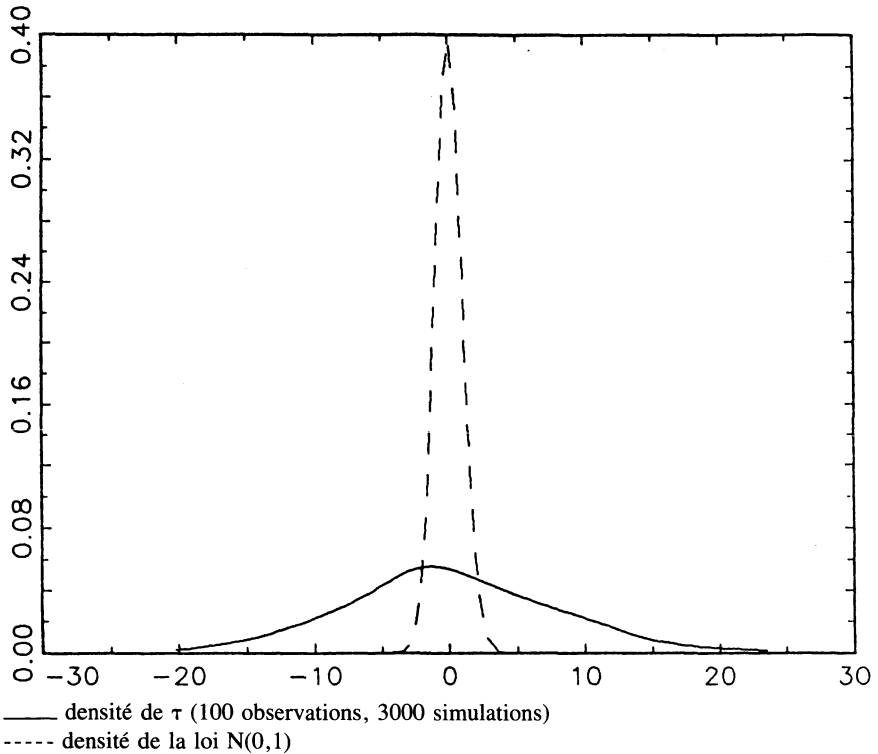
où $\{\varepsilon_t\}$ et $\{\eta_t\}$ sont des bruits blancs indépendants et de même variance. Le processus $\{X_t\}$ étant indépendant du processus $\{Y_t\}$ on pourrait penser que dans la régression de Y_t sur X_t :

$$Y_t = \hat{\alpha} + \hat{\beta}X_t + \hat{w}_t, 1, \dots, T \tag{7}$$

l'estimateur des moindres carrés ordinaires $\hat{\beta}$ tend vers zéro, lorsque T tend vers l'infini et que la probabilité de la région critique $W = \{|\tau| \geq 1.96\}$, où τ est le rapport de Student associé au coefficient de X_t , tend vers 0.05. Or on montre [voir Phillips, 1986] que $\hat{\beta}$ a une loi limite non dégénérée et que $T^{-1/2} \tau$ a également une loi limite non dégénérée ce qui implique que la probabilité limite de W est 1. Ces résultats montrent, par exemple, qu'en appliquant mécaniquement les procédures de tests construites pour le cas stationnaire on sera souvent amené à refuser l'hypothèse selon laquelle le coefficient de X_t est nul et donc à accepter l'hypothèse que le processus $\{X_t\}$ influence le processus $\{Y_t\}$, alors qu'ils sont indépendants. Pour mesurer l'ampleur de ce phénomène, nous avons déterminé par une méthode de simulation la loi de τ lorsque $T = 100$ et nous l'avons comparé à la loi normale $N(0,1)$ sur laquelle est fondé le test habituel [voir aussi Granger-Newbold, 1974]. La figure 1 montre que ces deux lois sont très différentes; en particulier la loi de τ est beaucoup plus dispersée que la loi $N(0,1)$ et donne à la région W une probabilité égale à 0.76. Dans cet exemple on accepterait donc trois fois sur quatre que $\{X_t\}$ influence $\{Y_t\}$.

FIGURE 1

RÉGRESSION D'UNE MARCHÉ ALÉATOIRE SUR UNE AUTRE MARCHÉ ALÉATOIRE
INDÉPENDANTE (ET DE MÊME VARIANCE)



Comme deuxième exemple considérons l'estimation par les moindres carrés ordinaires du coefficient ρ dans le modèle autorégressif :

$$Y_t = \rho Y_{t-1} + \varepsilon_t \quad (8)$$

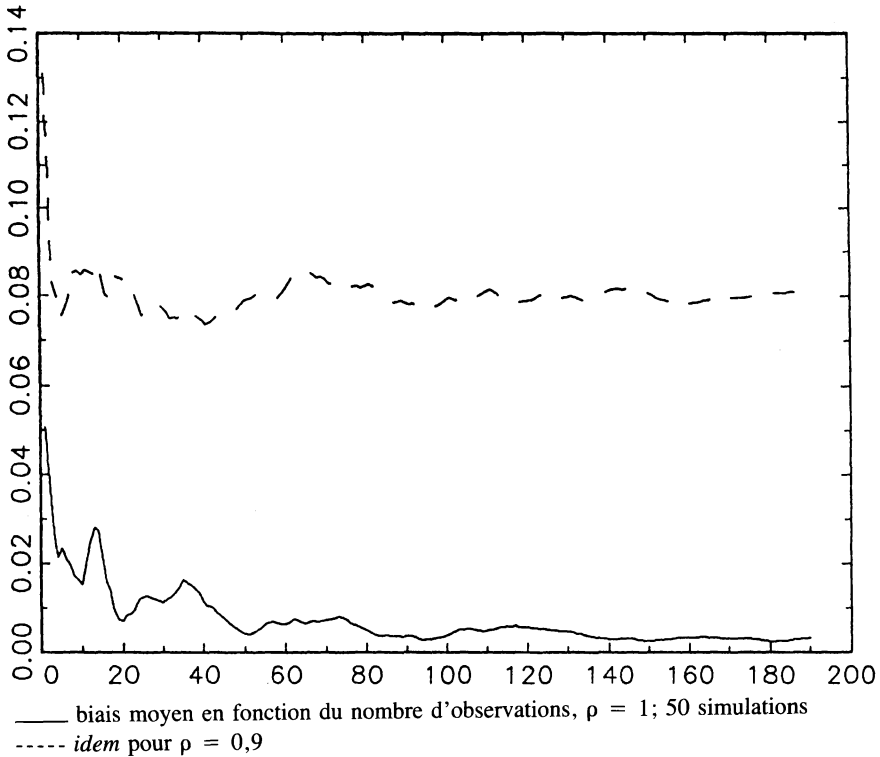
où $\{\varepsilon_t\}$ est un bruit blanc.

Si on suppose que la vraie valeur de ρ est 1, c'est-à-dire que $\{Y_t\}$ est une marche aléatoire, on peut montrer que l'estimateur des m.c.o. $\hat{\rho}$ converge vers la vraie valeur 1. On peut même montrer que $T(\hat{\rho}-1)$ converge en loi vers une loi non dégénérée et donc que la vitesse de convergence de $\hat{\rho}$ vers 1 est en $\frac{1}{T}$, et non en $\frac{1}{T^{1/2}}$ comme c'est le cas dans le domaine stationnaire; on dit que, si ρ vaut 1, $\hat{\rho}$ est «superconvergent». Cette superconvergence, dans le cas $\rho = 1$, entraîne en particulier que $\hat{\rho}$ reste convergent même si $\{\varepsilon_t\}$ est un processus stationnaire autocorrélé, ce qui n'est plus vrai si $|\rho| < 1$. La figure 2 montre l'évolution en fonction du nombre d'observations, de la moyenne du biais de l'estimateur des m.c.o. dans la régression de Y_t sur Y_{t-1} , le vrai modèle étant soit $Y_t = Y_{t-1} + \varepsilon_t$ avec

$\varepsilon_t = 0,7\varepsilon_{t-1} + u_t$, soit $Y_t = 0,9 Y_{t-1} + \varepsilon_t$ avec $\varepsilon_t = 0,7\varepsilon_{t-1} + u_t$, u_t étant dans les deux cas un bruit blanc gaussien. Dans le premier cas (non stationnaire) ce biais tend rapidement vers zéro alors que dans le deuxième cas (stationnaire) ce biais tend vers 0,082.

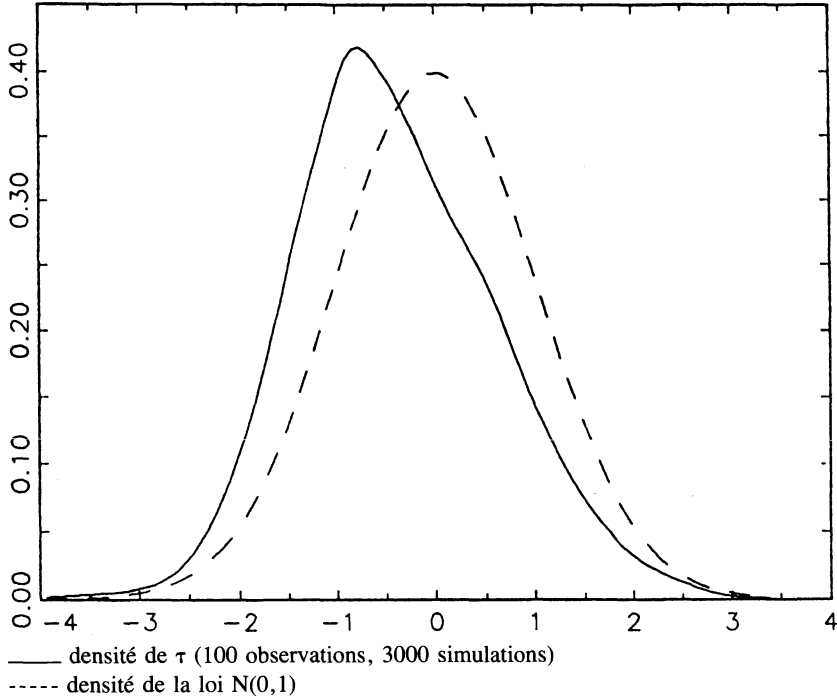
FIGURE 2

RÉGRESSION DE Y_t SUR Y_{t-1} AVEC ERREUR AUTOCORRÉLÉE



Cette propriété non classique, mais satisfaisante, des modèles non stationnaires coexiste avec d'autres propriétés également non classiques, mais plus négatives. Ainsi la statistique de test de Student τ associée au test de $\rho = 1$ n'a pas pour loi asymptotique la loi $N(0,1)$ et il faut donc modifier les procédures de test habituelles [voir Dickey-Fuller, 1979; Fuller, 1976]. La figure 3 présente la densité de τ pour $T = 100$ et $\rho = 1$, calculée par simulation (lorsque le bruit blanc est gaussien). On constate que, par rapport à la densité de la loi $N(0,1)$, cette densité est nettement décalée vers la gauche. En particulier si on utilise la région critique classique $\{\tau < -1,65\}$ pour tester $\tau = 1$ au niveau 5%, on refuse deux fois trop souvent l'hypothèse $\rho = 1$ lorsqu'elle est vraie, puisque la probabilité de cette région vaut environ 0,1 lorsque ρ vaut 1.

FIGURE 3

RÉGRESSION DE Y_t SUR Y_{t-1} (MARCHE ALÉATOIRE)

En troisième lieu considérons un problème analogue au précédent, mais dans le cas bivarié, le modèle est :

$$Y_t = Y_{t-1} + \varepsilon_t \text{ avec } Y_t = \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} \quad (9)$$

où $\varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$ est une moyenne mobile définie par :

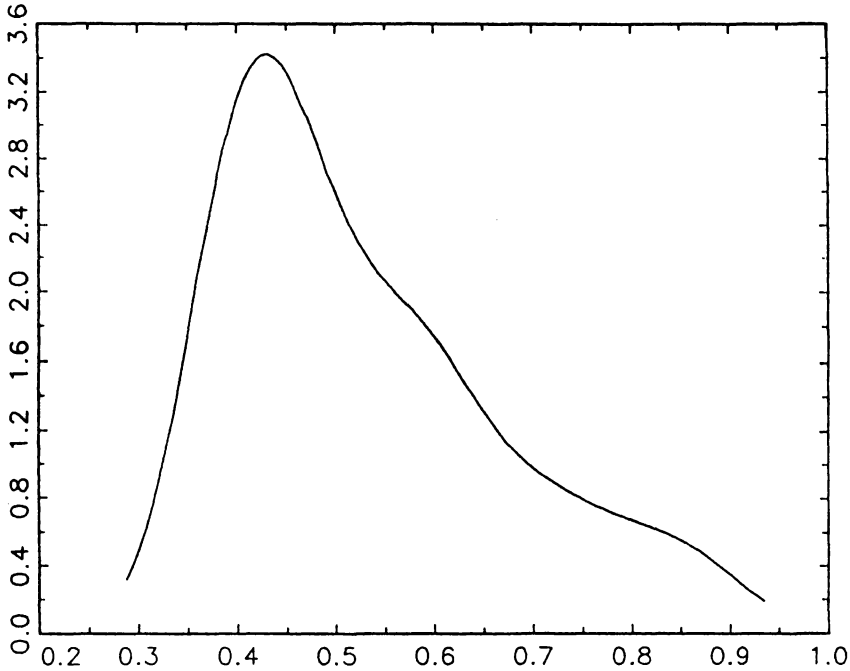
$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} + \begin{bmatrix} -0,8 & 0,4 \\ 0,2 & -0,6 \end{bmatrix} \begin{pmatrix} u_{1,t-1} \\ u_{2,t-1} \end{pmatrix} \quad (10)$$

où $u_t = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$ est un bruit blanc, que l'on supposera gaussien centré réduit.

Le processus $\{Y_t\}$ est non stationnaire et on pourrait donc penser, en extrapolant les résultats précédents, que la matrice (2x2) des coefficients des estimateurs des m.c.o. dans la régression de Y_t sur Y_{t-1} , va converger vers la matrice identité, or ce n'est pas le cas. Par exemple la figure 4 fournit, dans une régression sur 500 observations, la densité de l'estimateur du coefficient de $Y_{1,t-1}$ dans la première équation. On voit que cet estimateur est en général très inférieur à la vraie valeur, à savoir 1, en particulier le mode de cette distribution vaut approximativement 0,42.

FIGURE 4

RÉGRESSION DE Y_t , VECTEUR DE DIMENSION 2 INTÉGRÉ D'ORDRE 1, SUR Y_{t-1} .



Densité de l'estimateur du coefficient de $Y_{1,t-1}$ dans la régression de Y_{1t} sur $Y_{1,t-1}$ et $Y_{2,t-1}$. Cas cointégré (500 observations, 500 simulations).

Si on remplace $-0,8$ par $-0,4$ dans (10) cette densité change complètement (voir figure 5) et en particulier son mode est, cette fois, proche de la vraie valeur.

La raison de cette propriété inattendue est la suivante. Bien que dans les deux modèles, le processus $\{Y_t\}$ soit non stationnaire, il existe, dans le premier modèle, une combinaison linéaire des processus composants $\{Y_{1t}\}$ $\{Y_{2t}\}$ qui est stationnaire, on dit que ces deux processus sont cointégrés. En effet, les relations de définitions (9) et (10) s'écrivent :

$$(1-L)Y_t = (I + \Theta L)u_t \text{ où } \Theta = \begin{bmatrix} -0,8 & 0,4 \\ 0,2 & -0,6 \end{bmatrix}$$

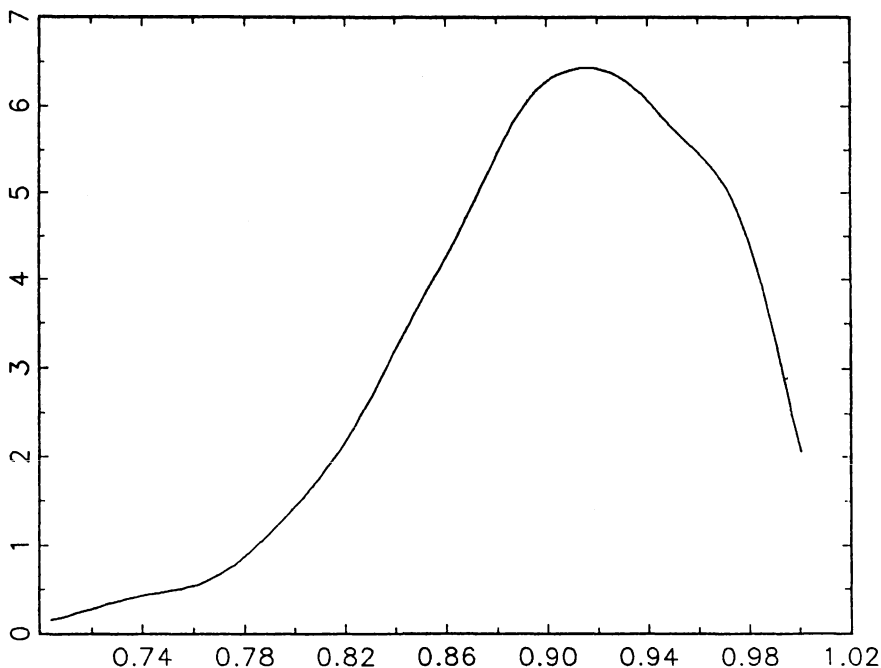
En multipliant à gauche par le vecteur $(1,-1)$ et en posant $Z_t = Y_{1t} - Y_{2t}$, on obtient :

$$(1-L)Z_t = [1-L, -1+L]u_t$$

$$Z_t = u_{1t} - u_{2t}$$

Le processus $\{Z_t\}$ est stationnaire et on retrouve le problème de non convergence dans le cas stationnaire et erreurs autocorrélées.

FIGURE 5

RÉGRESSION DE Y_t , VECTEUR DE DIMENSION 2 INTÉGRÉ D'ORDRE 1, SUR Y_{t-1} .

Densité de l'estimateur du coefficient de $Y_{1,t-1}$ dans la régression de Y_{1t} sur $Y_{1,t-1}$ et $Y_{2,t-1}$.
Cas non cointégré (500 observations, 500 simulations).

Cette liste des propriétés inattendues n'est évidemment pas exhaustive, on pourrait par exemple ajouter le fait que si une constante est présente dans le modèle (8) et dans la régression associée alors la statistique de Student redevient asymptotiquement normale même si $\rho = 1$ [voir Gouriéroux-Monfort, 1990-a, chapitre 13], ou encore le fait que si on ajuste le modèle (8), avec $\rho = 1$, sur une droite pour éliminer la tendance, les résidus peuvent présenter des propriétés cycliques artificielles [voir Nelson-Kang, 1981]...

On comprendra donc que, jusqu'à une date récente, les économètres ont eu tendance à se ramener systématiquement au cas stationnaire même au prix de certains dangers comme l'indique la dernière remarque sur l'élimination d'une tendance déterministe. En fait la procédure la plus populaire a été le passage aux différences premières et nous allons discuter ce point dans le cadre du modèle VAR général :

$$\Phi(L)Y_t = \varepsilon_t \quad \text{avec } \Phi(0) = I \quad (11)$$

où $\{Y_t\}$ est un vecteur de taille n supposé non stationnaire, ce qui implique que $\det\Phi(L)$ admet 1 comme racine, mais où $\Delta Y_t = (1-L)Y_t = Y_t - Y_{t-1}$ est supposé stationnaire.

Le modèle (11) peut être réécrit, sous forme plus détaillée :

$$Y_t = \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + \varepsilon_t \tag{12}$$

ou encore, sous la forme équivalente :

$$\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \dots + \Gamma_{p-1} \Delta Y_{t-p+1} - \Pi Y_{t-p} + \varepsilon_t \tag{13}$$

Les matrices Γ_i et Π s'expriment facilement en fonction des matrices Φ_j , en particulier on a $\Pi = \Phi(1) = I - \sum_{i=1}^p \Phi_i$. Comme $\{Y_t\}$ est non stationnaire on a vu

que $\det \Phi(1) = 0$ et donc Π est de rang $r < n$. On peut alors écrire $\Pi = D \times F$, où D et F sont des matrices ($n \times r$) et ($r \times n$) de rang r , et, en posant $Z_t = F Y_t$, le modèle (13) devient :

$$\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \dots + \Gamma_{p-1} \Delta Y_{t-p+1} - D Z_{t-p} + \varepsilon_t \tag{14}$$

Comme le processus $\{\Delta Y_t\}$ est stationnaire, la forme (14) implique que le processus $\{Z_t\}$, de taille r , est également stationnaire; on a donc mis en évidence r combinaisons linéaires des composantes de Y_t , à savoir $F Y_t$, qui sont stationnaires. Ces relations $Z_t = F Y_t$ sont appelées relations de cointégration [voir Granger, 1981, Engle-Granger, 1987] et sont interprétées comme des relations de long terme entre les processus non stationnaires $\{Y_{1t}, \dots, Y_{nt}\}$, les Z_t étant les «erreurs» de ces relations de long terme. La forme (14) est, pour cette raison, appelée forme à correction d'erreurs, car les «erreurs» passées Z_{t-p} de la relation de long terme interviennent explicitement dans la spécification [Engle-Granger, 1987; Hendry-Von Ungern Sternberg, 1981]. Grâce aux travaux de Engle-Granger (1987), Johansen (1988), Phillips (1990), on connaît maintenant des méthodes statistiques adaptées à ce type de modèle. En particulier, il existe des tests (non classiques) sur la valeur du rang r de la matrice Π . Une fois ce rang r fixé, la forme à correction d'erreurs (14) et certaines conditions d'identification sur D ou F fournissent le paramétrage le plus «parcimonieux»; la méthode du maximum de vraisemblance appliquée à ce modèle et les méthodes de tests asymptotiques possèdent les propriétés classiques. Engle-Granger ont également proposé une procédure en deux étapes plus simple : estimation de F par les m.c.o., soit \hat{F} , remplacement de Z_{t-p} (inobservable) dans (14) par $\hat{Z}_{t-p} = \hat{F} Y_{t-p}$ et estimation des paramètres Γ_i et D par les m.c.o. Les estimateurs de Γ_i et D ainsi obtenus sont asymptotiquement efficaces, l'estimateur \hat{F} est convergent (même superconvergent) mais les tests sur les coefficients F à partir de \hat{F} sont difficiles à mettre en œuvre.

Cette nouvelle théorie statistique éclaire bien le vieux dilemme : faut-il travailler sur les niveaux ou sur les différences premières? Si on admet que les variables en niveau Y_t suivent un processus non stationnaire de type autoregressif (11) et si de modèle est estimé directement par la méthode du (pseudo) maximum de vraisemblance on obtient des estimateurs qui, en général n'ont pas les propriétés classiques [voir cependant Sims-Stock-Watson, 1990 pour des exceptions]; donc, dans le cas non stationnaire, un surparamétrage n'a pas seulement des conséquences sur l'efficacité asymptotique, comme dans le cas stationnaire, mais aussi sur la

forme des lois asymptotiques ce qui complique fortement la théorie des tests et des régions de confiance. Si au contraire on travaille sur les différences premières, cela revient à ignorer le terme ΠY_{t-p} dans (13), c'est-à-dire à supposer que le rang r de Π est nul, ce qui n'est pas le cas en général, et on commet donc une erreur de spécification. La bonne solution passe donc bien par une étude précise des propriétés de cointégration du modèle; ceci explique le succès de ces méthodes et la très abondante littérature sur le sujet [voir Engle-Yoo, 1989 pour une présentation plus détaillée des résultats].

3. LES MODÈLES ARCH

Une autre caractéristique commune de l'école «théorique» et de l'école «empirique» a longtemps été l'absence de flexibilité dans la spécification de la variance conditionnelle de la variable endogène (le conditionnement s'entendant par rapport au passé).

Ainsi le modèle à équations simultanées dynamiques, qui est le modèle explicatif central, s'écrit :

$$B(L)Y_t + C(L)X_t = \varepsilon_t \quad (15)$$

où $\{\varepsilon_t\}$ est un bruit blanc gaussien de variance Ω et $\{X_t\}$ un vecteur de variables exogènes. La matrice de variance-covariance de Y_t conditionnelle au passé, à savoir $B^{-1}(0)\Omega B^{-1}(0)'$, est donc supposée fixe.

De même, dans le modèle VAR $\Phi(L)Y_t = \varepsilon_t$, avec $\Phi(0) = I$, où $\{\varepsilon_t\}$ est un bruit blanc gaussien de matrice de variance-covariance Ω , la matrice de variance-covariance de Y_t conditionnellement au passé est égale à Ω , et donc fixe.

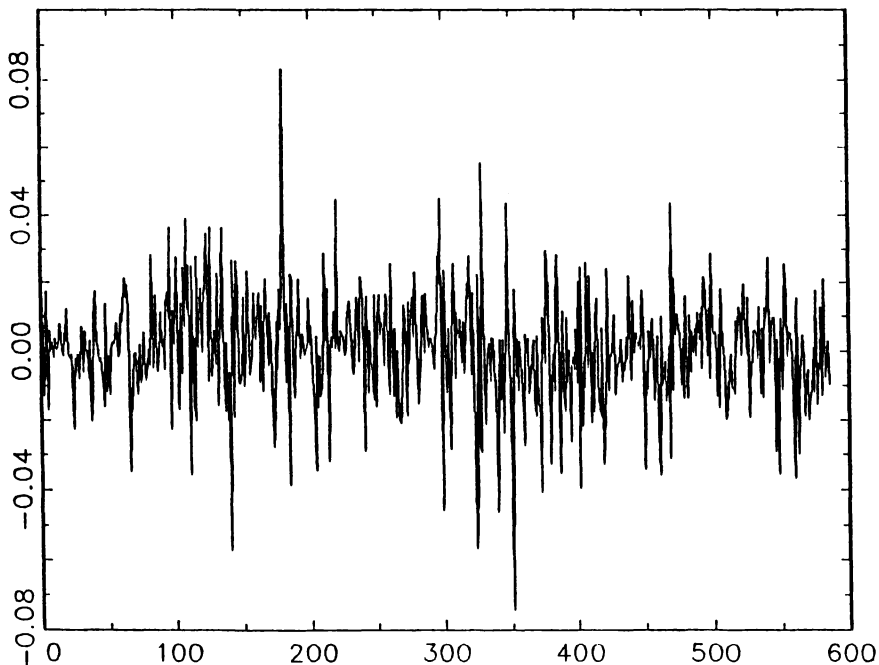
Cette hypothèse de stabilité dans le temps de la variance conditionnelle peut être très contraignante. Considérons par exemple, voir la figure 6, l'évolution entre janvier 79 et avril 90 de la variation relative hebdomadaire du cours du dollar en francs (en fait la différence des logarithmes).

Il est clair qu'un choc important, a plutôt tendance à être suivi par un choc important, mais pas nécessairement dans le même sens. Par conséquent, il paraît difficile d'admettre que la variance conditionnelle est fixe comme on le fait en choisissant, par exemple, une modélisation par un processus stationnaire normal.

La modélisation par un processus stationnaire normal n'est donc pas acceptable en raison de cette hétéroscédasticité conditionnelle, en outre elle doit aussi être refusée pour une autre raison: le coefficient d'aplatissement empirique mesurant le rapport entre le moment centré d'ordre 4 et le carré de la variance a pour valeur 5,6, ce qui est nettement supérieur à sa valeur dans le cas gaussien, à savoir 3.

FIGURE 6

TAUX D'ACCROISSEMENT HEBDOMADAIRE DU COURS DU DOLLAR EN FRANCS
(JANVIER 79, AVRIL 90).



Pour prendre en compte ces phénomènes, Engle (1982) a introduit les modèles autorégressifs conditionnellement hétéroscédastiques (ARCH). Un processus ARCH (q) est défini de la manière suivante :

$$\varepsilon_t / \varepsilon_{t-1} \sim N \left(0, \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \right) \tag{16}$$

où les $\alpha_i, i = 0, \dots, q$ sont des réels positifs. Cette définition impose que, conditionnellement à son passé ε_{t-1} , la variable ε_t soit normale centrée et de variance $\alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$. Un choc passé important augmente donc la valeur de cette

variance conditionnelle. On peut montrer que si la somme $\sum_{i=1}^q \alpha_i$ est inférieure

à un, le processus $\{\varepsilon_t\}$ a une variance marginale finie ; par ailleurs toutes ses autocovariances sont nulles. Ce processus est donc un bruit blanc «au sens faible», c'est-à-dire au sens de la nullité de l'autocorrélation, mais ce n'est pas un bruit blanc «au sens fort», c'est-à-dire au sens de l'indépendance, puisque, d'après (16), les lois conditionnelles ne sont pas fixes. Il est aussi important de noter que, malgré la normalité conditionnelle, ce processus n'est pas normal puisque les variances

conditionnelles ne sont pas fixes. Un calcul simple montre d'ailleurs que le moment marginal d'ordre 4 de ε_t est supérieur à 3 fois le carré de la variance marginale, ce qui est cohérent avec l'observation empirique mentionnée ci-dessus.

Comme les modèles ARCH sont des bruits blancs faibles, ils peuvent servir de base à la construction de modèles plus complexes; par exemple, on peut définir un processus autorégressif-ARCH, noté $AR(p)$ -ARCH(q), par

$$\Phi(L) Y_t = \varepsilon_t \quad (17)$$

où $\phi(L)$ est un polynôme en L de degré p et ε_t un processus ARCH(q).

Un ensemble de travaux en expansion rapide a étudié les propriétés de ces modèles et de leurs généralisations sur le plan de l'estimation, des tests et de la prévision [voir Bollerslev, 1986; Weiss, 1986; Engle-Lillien-Robbins, 1987 et Bollerslev-Chou-Jayaraman-Kroner, 1990]. Ces modèles peuvent être étendus au cas multivarié mais, sur le plan statistique, une difficulté importante apparaît: le très grand nombre de paramètres. Pour résoudre ce problème, on peut, comme pour les modèles VAR, penser à deux types de réductions du modèle: une réduction fondée sur des critères purement statistiques et une réduction fondée sur des considérations économiques. Il faut cependant remarquer que la théorie économique impose rarement des contraintes sur les moments du deuxième ordre, des exceptions notables étant les théories financières du CAPM et de l'APT.

Malgré ces difficultés inhérentes à toute découverte importante, la modélisation de type ARCH a largement prouvé son efficacité, en particulier dans la modélisation de variables comme les prix, les taux d'intérêt, les taux de change, les rendements des actifs financiers...

4. LES MODÈLES DE MOMENTS GÉNÉRALISÉS

Les développements que l'on vient de décrire, modèles VAR, modèles non stationnaires, modèles ARCH ont été plus influencés par la théorie statistique que par la théorie économique. Certains développements, comme les modèles de déséquilibre étudiés dans le paragraphe suivant, ont au contraire été totalement fondés sur la théorie économique. Les modèles de moments généralisés que l'on va maintenant présenter occupent une position intermédiaire. En effet les conditions de type moments généralisés qui constituent le modèle économétrique sont obtenues à partir d'un modèle économique mais, loin d'être équivalentes à ce modèle, elles en sont des conséquences partielles permettant d'identifier certains paramètres. Le cas le plus fréquent est celui où le modèle économique est une optimisation intertemporelle non linéaire; dans ce cas il est en général impossible d'explicitement la solution et donc de fournir une famille de lois de probabilité pour les variables endogènes, ce qui exclut la méthode du maximum de vraisemblance. En revanche, il peut être possible de fonder une méthode d'estimation sur les conditions d'Euler, c'est-à-dire sur les conditions du premier ordre du problème d'optimisation. Ces conditions s'écrivent souvent:

$$E_t f(W_{t+1}, \dots, W_{t+h}; \theta) = 0 \text{ ou } E_t f_t(\theta) = 0 \quad (18)$$

où W_{t+i} est un vecteur aléatoire observé à la date $t+i$, E_t est l'opérateur espérance conditionnelle à l'information disponible à la date t , f est une fonction à valeurs dans \mathbb{R}^m et θ un paramètre inconnu de \mathbb{R}^p .

Si l'on dispose d'un vecteur Z_t de taille n de variables observables à la date t , on a les relations suivantes :

$$E[Z_t \otimes f_t(\theta)] = 0 \tag{19}$$

La méthode des moments généralisés [Hansen, 1982; Hansen-Singleton, 1982], repose sur la minimisation en θ de :

$$\sum_{t=1}^T [Z'_t \otimes f'_t(\theta)] \Omega_T \sum_{t=1}^T [Z_t \otimes f_t(\theta)] \tag{20}$$

où Ω_T est une matrice (nm, nm) symétrique positive convergeant vers Ω lorsque T tend vers l'infini.

Sous certaines conditions de régularité, on peut montrer que l'estimateur $\hat{\theta}_T$ ainsi obtenu est convergent et asymptotiquement normal. En outre un choix optimal pour Ω_T consiste à prendre une suite de matrices convergeant vers $Q^{-1}(\theta_0)$ où $Q(\theta_0)$

est la matrice de variance-covariance asymptotique de $\frac{1}{\sqrt{T}} \sum_{t=1}^T Z_t \otimes f_t(\theta_0)$. Si

on note $\hat{\theta}_T$ l'estimateur correspondant, la matrice de variance-covariance asymptotique de $\sqrt{T}(\hat{\theta}_T - \theta_0)$ est alors :

$$\left\{ E \left[Z'_t \otimes \frac{\partial f'_t}{\partial \theta}(\theta_0) \right] Q_0^{-1}(\theta) E \left[Z_t \otimes \frac{\partial f_t}{\partial \theta'}(\theta_0) \right] \right\}^{-1}.$$

En général le nombre de contraintes (19), à savoir nm , est supérieur à la taille p de θ . Le modèle est alors suridentifié et on peut tester cette suridentification à l'aide de la statistique

$$\xi_T = \frac{1}{T} \sum_{t=1}^T [Z'_t \otimes f'_t(\hat{\theta}_T)] \hat{\Omega}_T \sum_{t=1}^T [Z_t \otimes f_t(\hat{\theta}_T)] \tag{21}$$

Sous l'hypothèse de validité des contraintes de suridentification, ξ_T converge en loi vers une loi du χ^2 à $nm-p$ degrés de liberté; on refusera donc cette hypothèse au niveau α si $\xi_T \geq \chi^2_{1-\alpha}(nm-p)$.

5. LES MODÈLES DE DÉSÉQUILIBRE

Les modèles économétriques de déséquilibre sont largement fondés sur les travaux d'économie théorique comme ceux de Barro-Grossman (1971), Benassy (1975), Dreze (1975), Younès (1975), Malinvaud (1976).

Des synthèses sur les méthodes économétriques du déséquilibre ont déjà été effectuées [voir Laffont-Monfort, 1976; Gouriéroux-Laffont-Monfort, 1984; Quandt, 1982, 1988] et il n'est pas question de les reprendre ici. Signalons

seulement deux progrès récents concernant les problèmes importants de l'agrégation et de la dynamique.

Pour simplifier l'exposition considérons le cas d'un seul marché. Dans ce cas, le modèle canonique s'écrit :

$$\begin{aligned} D_t &= d(x_t, \theta) + u_{1t} \\ S_t &= s(x_t, \theta) + u_{2t} \\ Q_t &= \min(D_t, S_t) \end{aligned} \quad t = 1, \dots, T \quad (22)$$

où D_t représente la demande, S_t l'offre, Q_t la quantité échangée, x_t un vecteur de variables exogènes ou endogènes retardées contenant le prix et (u_{1t}, u_{2t}) un vecteur de perturbations que l'on supposera normales centrées de variances σ_1^2 et σ_2^2 .

Pour répondre à l'objection classique sur l'agrégation de micromarchés se trouvant dans des situations de déséquilibre différentes, on peut remplacer le modèle (18) par :

$$Q_t = E_{v_1, v_2} \{ \min [d(x_t, \theta) + u_{1t} + v_1, s(x_t, \theta) + u_{2t} + v_2] \} \quad (23)$$

où v_1 et v_2 sont des perturbations spatiales propres aux micromarchés que l'on supposera également normales.

La vraisemblance de modèles du type (23) est complexe ou même très complexe (par exemple dans le cas de plusieurs marchés avec des effets de report et des non linéarités), ce qui a longtemps rendu ces modèles inutilisables en pratique. Les techniques récentes du pseudo maximum de vraisemblance [Gouriéroux-Monfort-Trognon, 1984] et du pseudo maximum de vraisemblance simulé [Laroque-Salanié, 1989] fournissent maintenant des méthodes économétriques adaptées à ce type de modèles. Par exemple (23) peut être estimé en minimisant par rapport aux paramètres la pseudo vraisemblance :

$$\sum_{t=1}^T \left[\text{Log } V(Q_t) + \frac{(Q_t - EQ_t)^2}{V(Q_t)} \right] \quad (24)$$

les espérances EQ_t et les variances VQ_t , dont l'expression analytique est très complexe, étant remplacées par des moments empiriques fondés sur des simulations. Des études théoriques et pratiques montrent qu'il suffit d'un faible nombre de simulations pour obtenir des résultats très proches de ceux fournis par la méthode du pseudo maximum de vraisemblance exacte. La raison intuitive de ce résultat est qu'on n'exige pas une précision importante sur chaque EQ_t ou chaque VQ_t mais seulement sur la somme (24), et donc la loi des grands nombres joue. Ce type de propriété rend les méthodes économétriques fondées sur des simulations très prometteuses et elles ont déjà permis de résoudre des problèmes, considérés jusqu'ici comme hors de portée, dans des domaines très variés : agrégation temporelle, hétérogénéité individuelle, modèles qualitatifs dynamiques, modèles de choix discrets fondés sur des modèles structurels... [Voir Gouriéroux-Monfort, 1990-b, 1991].

Dans le domaine de la dynamique des modèles de déséquilibre divers progrès ont également été effectués. Ainsi on peut maintenant proposer des tests d'absence d'autocorrélation de u_{1t} ou u_{2t} dans (22), analogues au test classique de Durbin et Watson, dont le rôle est si important dans la modélisation traditionnelle. La statistique de test devient, par exemple pour u_{1t} :

$$DW = 2\left(1 - \frac{s}{T^{1/2}}\right) \tag{25}$$

avec
$$s = \sum_{t=2}^T \hat{u}_{1t} \hat{u}_{1,t-1} / \left(\sum_{t=2}^T \hat{u}_{1t}^2 \hat{u}_{1,t-1}^2 \right)^{1/2}$$

et
$$\hat{u}_{1t} = E(D_t / Q_t; \hat{\theta}, \hat{\sigma}^2) - d(x_t, \hat{\theta})$$

et la région critique est $2-c < DW < 2+c$, avec $c = 2\left(\frac{1}{T} \chi_{1-\alpha}^2(1)\right)^{1/2}$.

Les u_{1t} sont appelées résidus généralisés et jouent, dans la théorie des tests, un rôle tout à fait analogue à celui des résidus classiques du modèle linéaire [voir Gourieroux-Monfort-Trognon, 1985].

Les techniques du pseudo maximum de vraisemblance simulé sont également utiles pour l'estimation de modèles dans lesquels apparaissent des variables endogènes retardées inobservables [voir Salanié, 1988] et on peut penser, plus généralement, qu'elles lèvent beaucoup de difficultés techniques sur lesquelles la modélisation économétrique du déséquilibre avait buté.

CONCLUSION

Les quelques exemples précédents ne couvrent pas, loin s'en faut, tout le champ des développements récents des méthodes macroéconométriques; ainsi des domaines importants comme les méthodes semi-paramétriques, les tests non emboîtés, les méthodes de filtrage et de lissage de Kalman, les traitements des anticipations, n'ont pas été évoqués. On peut cependant espérer que ces exemples sont suffisamment variés pour donner une idée de l'extrême diversité de ces développements et pour montrer que l'on se trouve actuellement dans une phase de recherche à la fois très active et très hétérogène.

Cette situation n'est pas sans conséquence sur la pratique macroéconométrique. Tout d'abord, il faut admettre, même si ce n'est pas très confortable, l'absence d'une méthodologie dominante admise par tous et il serait vain de penser que le retour à une situation d'équilibre, analogue à celle des années cinquante ou soixante, est pour bientôt. Même si le bouillonnement actuel est suivi, comme c'est probable, d'une phase de décantation, la diversification des méthodes semble être un phénomène irréversible et il faut donc apprendre à rechercher la méthodologie la plus adaptée à chaque problème.

La seconde conséquence que l'on peut tirer est la nécessité d'un plus grand investissement sur le plan de la méthodologie statistique. En effet, la plupart des

innovations méthodologiques récentes sont relativement délicates à mettre en œuvre. Une mauvaise compréhension de leur principe peut entraîner de sérieuses erreurs mais, en même temps, on ne peut plus défendre la thèse selon laquelle il existe des bonnes méthodes simples, robustes, universelles que l'on peut toujours utiliser, les autres méthodes étant des raffinements de spécialistes. Chaque méthode a un objectif et un champ de validité précis; un usage aveugle de ces méthodes peut être très dangereux et il devient donc indispensable de connaître leurs règles d'utilisation.

Enfin ces deux conséquences, qui peuvent paraître négatives, ont une contrepartie positive. Ces divers coûts que l'on vient d'évoquer ne sont que le prix à payer pour une analyse plus fine, plus profonde, des phénomènes macroéconomiques, en particulier dans leurs aspects dynamiques. Il était d'ailleurs un peu illusoire de penser que la macroéconomie pouvait se développer indépendamment des méthodes statistiques et il est rassurant de constater que la macroéconomie moderne se confond de plus en plus avec la macroéconométrie.

BIBLIOGRAPHIE

- BARRO, R.J. et H.I. GROSSMAN (1971), «A General Disequilibrium Model of Income and Employment» *American Economic Review*, 61, pp. 82-93.
- BÉNASSY, J.P. (1975), «Neo-Keynesian Disequilibrium Theory on a Monetary Economy», *Review of Economic Studies*, XLII, pp. 503-24.
- BOSSERSLEV, T. (1986), «Generalized Autoregressive Conditional Heteroscedasticity», *Journal of Econometrics*, 31, pp. 307-327.
- BOLLERSLEV, T., R. CHOU, N. JAYARAMAN et K. KRONER (1990), «ARCH Modelling in Finance: A Selective Review of the Theory and Empirical Evidence, with Suggestions for Future Research», discussion paper.
- BLANCHARD, O. et D. QUAH (1989), «The Dynamic Effects of Aggregate Demand and Supply Disturbances», *The American Economic Review*, 79, n° 4, pp. 655-673.
- BLANCHARD, O. (1989), «A Traditional Interpretation of Macroeconomic Fluctuations», *The American Economic Review*, 79, n° 5, pp. 1146-1164.
- BOX, G.P. et G.M. JENKINS (1970), *Time Series Analysis, Forecasting and Control*, Holden-Day.
- DICKEY, D.A. et W.A. FULLER (1979), «Distribution of Estimates for Autoregressive Time Series with Unit Root», *Journal of the American Statistical Association*, 74, pp. 427-431.
- DOAN, T., R. LITTERMAN et C. SIMS (1984), «Forecasting and Conditional Projection Using Realistic Prior Distributions», *Econometric Review*, 3, pp. 1-144.
- DREZE, J.M. (1975), «Existence of an Exchange Equilibrium Under Price Rigidities», *International Economic Review*, 16, pp. 301-320.

- ENGLE, R. et C.W.J. GRANGER (1987), «Co-integration and Error Correction: Representation, Estimation and Testing», *Econometrica*, 55, n° 2, pp. 251-276.
- ENGLE, R. et B.S. YOO (1989), «Cointegrated Economic Time Series: A Survey with New Results», Discussion paper 87-26R, University of San Diego.
- FULLER, W. (1976), *Introduction to Statistical Time Series*, Wiley.
- GRANGER, C.W.J. (1981), «Some Properties of Time Series Data and Their Use in Econometric Model Specification», *Journal of Econometrics*, pp. 121-130.
- GOURIÉROUX, C., J.J. LAFFONT et A. MONFORT (1984), «Econométrie des modèles d'équilibre avec rationnement: une mise à jour», *Annales de l'INSEE*, 55/56, pp. 5-37.
- GOURIÉROUX, C., A. MONFORT et A. TROGNON (1984), «Pseudo Maximum Likelihood Methods: Theory», *Econometrica*, 52, pp. 681-700.
- GOURIÉROUX, C., A. MONFORT et A. TROGNON (1985), «A General Approach to Serial Correlation» *Econometric Theory*, pp. 315-340.
- GOURIÉROUX, C. et A. MONFORT (1990a), *Séries Temporelles et Modèles Dynamiques*, *Economica*, Paris.
- GOURIÉROUX, C. et A. MONFORT (1990-b), «Simulation Based Econometrics for Models with Heterogeneity», *Annales d'Économie et de Statistique*, n° 20/21.
- GOURIÉROUX, C. et A. MONFORT (1991), «Simulation Based Econometrics: A Survey», *Journal of Econometrics*, (à paraître).
- GRANGER, C.W.J. et P. NEWBOLD (1974), «Spurious Regressions in Econometrics», *Journal of Econometrics*, 2, pp. 111-120.
- HAAVELMO, T. (1943), «The Statistical Implications of a System of Simultaneous Equations», *Econometrica*, 11, pp. 1-12.
- HAAVELMO, T. (1984), «The Probability Approach in Econometrics», *Econometrica*, 12, Supplement.
- HANSEN, L.P. (1982), «Large Sample Properties of Generalized Method of Moments Estimators», *Econometrica*, 50, n° 4, pp. 1029-1084.
- HANSEN, L.P., et K.J. SINGLETON (1982), «Generalized Instrumental Variables Estimation Estimation of Nonlinear Rational Expectation Models», *Econometrica*, 50, n° 5, pp. 1269-1286.
- HENDRY, D.F. et T. Von UNGERN-STERNBERG (1981), «Liquidity and Inflation Effects on Consumer's Expenditure», in DEATON, A.S., *Essays in the Theory and Measurement of Consumer's Behavior*, Cambridge University Press.
- JOHANSEN, S. (1988), «Statistical Analysis of Cointegration Vector», *Journal of Economics Dynamics and Control*, 12, 2/3, pp. 231-254.
- KOOPMANS, T.C. (1947), «Measurement Without Theory», *Review of Economic Statistics*, 29, pp. 161-172.
- KOOPMANS, T.C., H. RUBIN et R.B. LEIPNIK (1950), «Measuring the Equations Systems of Dynamic Economics», in *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph n° 10, John Wiley.

- KOOPMANS, T.C. et W. HOOD (1953), «The Estimation of Simultaneous Linear Economic Relationships», in *Studies in Econometric Method*, Cowles Foundation Monograph n° 14, Yale University.
- LAFFONT, J.J. et A. MONFORT (1976), «Économétrie des modèles d'équilibre avec rationnement», *Annales de l'INSEE*, 24, pp. 4-39.
- LAROQUE, G. et B. SALANIE (1989), «Estimation of Multi-Market Fix-Price Models: An Application of Pseudo Maximum Likelihoods Methods», *Econometrica*, 57, n° 4, pp. 831-860.
- MALINVAUD, E. (1976), *The Theory of Unemployment Reconsidered*, Basil Blackwell.
- MONFORT, A. et R. RABEMANANJARA (1990), «From a VAR Model to a Structural Model, with an Application to the Wage Price Spiral», *The Journal of Applied Econometrics*, 5, pp. 203-227.
- NELSON, C.R. et H. KANG (1981), «Spurious Periodicity in Inappropriately Detrended Time Series», *Econometrica*, 49, 3, pp. 741-751.
- PHILLIPS, P.C.B. (1986), «Understanding Spurious Regressions in Econometrics», *Journal of Econometrics*, pp. 311-340.
- PHILLIPS, P.C.B. (1990), «Optimal Inference in Cointegrated Systems», Yale University.
- QUANDT, R. (1982), «Econometric Disequilibrium Models», *Econometric Reviews*, 1, pp. 1-63.
- QUANDT, R. (1988), *The Econometrics of Disequilibrium*, Basil Blackwell.
- SALANIE, B. (1988), «Wage and Price Adjustment in a Multimarket Disequilibrium Model», Document de travail ENSAE/INSEE n° 8812.
- SIMS, C. (1980), «Macroeconomics and Reality», *Econometrica*, 48, pp. 1-48.
- SIMS, C., J.H. STOCK et M.W. WATSON (1990), «Inference in Linear Time Series Models with Some Unit Roots», *Econometrica*, 58, N° 1, pp. 113-144.
- WEISS, A.A. (1986), «Asymptotic Theory for ARCH Models: Estimation and Testing», *Econometric Theory*, 2, pp. 107-130.
- YOUNES, Y. (1975), «On the Role of Money in the Process of Exchange and the Existence of a Non-Walrasian Equilibrium», *Review of Economic Studies*, pp. 489-501.