

SEBASTIEN G. GREKI

Département des Sciences économiques
Faculté des Arts et des Sciences

**L'ESTIMATION DE MODÈLES DE SÉRIES CHRONOLOGIQUES
AVEC ERREURS SUR LES VARIABLES**

Rapport de recherche présenté à la
Faculté des études supérieures
en vue de l'obtention du
GRADE DE MAÎTRE ES SCIENCES (M. Sc)
(option économétrie)

Directeur de recherche:
M. Marcel Dagenais

UNIVERSITÉ DE MONTRÉAL
(Aout 95)

Dédicace

*<< Ce rapport de recherche est dédié à mon épouse, Ann ,
dont l'amour et les conseils ont été d'une valeur inestimable
durant la rédaction du présent document >>*

TABLE DES MATIÈRES

REMERCIEMENTS.....	i
SOMMAIRE.....	ii
INTRODUCTION.....	1
I- REVUE SÉLECTIVE DE LA LITTÉRATURE.....	3
A- Malinvaud, 3 ième édition, Méthode statistique de l'économétrie, PP.420-458...	3
B-Fomby and Hill, Advanced econometrics, PP.205-277.....	4
C- Dagenais, Journal of econometrics, 1994.....	11
D- Dagenais, Centre de recherche en développement économique (CRDE), 1995.	14
II- LA MÉTHODE.....	17
A- Le modèle.....	17
B- Choix de l'estimateur: la méthode des variables instrumentales.....	19
C- Estimateur de Fuller.....	25
III- SIMULATION DE MONTE CARLO.....	27
A- Description.....	27
B- Analyse des résultats.....	30
CONCLUSION.....	41
ANNEXES MATHÉMATIQUES.....	42
REFERENCES.....	51

REMERCIEMENTS

Nous remercions M. Marcel Dagenais, Professeur à l'université de Montréal et Directeur du présent rapport de recherche. Sa disponibilité, sa rigueur et ses précieux commentaires nous ont permis de mener à terme ce travail.

Nous remercions également M. René Garcia, Professeur à l'Université de Montréal qui a bien voulu servir de deuxième lecteur de ce rapport de recherche.

Nous avons aussi bénéficié de l'assistance de Cosmé Vodonou, étudiant en doctorat d'économétrie, dont l'amitié et la camaraderie sincères demeurent un souvenir agréable.

Nous pensons également à tous les étudiants gradués du département de sciences économiques de l'Université de Montréal, qui de près ou de loin nous ont encouragé sur cette voie.

Ce rapport a été rendu possible également grâce au soutien financier du CRDE de l'Université de Montréal où nous avons travaillé en tant qu'assistant de recherche durant 12 mois.

SOMMAIRE

Dans cette étude, notre objectif a été de comparer la performance de divers estimateurs du modèle:

$$y_t = bx_t + gy_{t-1} + u_t$$

lorsque les variables endogènes et exogènes peuvent contenir des erreurs de mesure avec des u_t qui peuvent être autocorrélés. Il s'agit, en fait, de faire une étude comparative de la vitesse de convergence des différents estimateurs dans les cas où ils sont recommandés. Nous avons retenu, sur la base d'une analyse théorique, quatre types d'estimateurs:

- Les moindres carrés ordinaires couramment recommandés lorsqu'il n'y a ni erreur de mesure sur les variables, ni autocorrélation des erreurs résiduelles.

- La méthode des variables instrumentales utilisant l'instrument $Z_1 = (x_t, x_{t-1})$ pour le cas où il y a de l'autocorrélation mais absence d'erreur sur les variables.

- L'estimateur de Fuller pour le cas où il y a coexistence des erreurs sur les variables et l'autocorrélation.

- Et enfin, la méthode de variable instrumentale classique associée à l'estimateur de Fuller utilisant d'abord l'instrument $Z = (x_{t+1}, x_{t+2})$ puis $Z_2 = (x_{t+1}, x_{t-1})$ comme solution alternative à la non convergence du premier estimateur.

Le critère que nous avons retenu pour tester la performance des différents estimateurs est la racine carrée de l'erreur quadratique moyenne (REQM).

Les résultats de notre recherche indiquent qu'aucune des méthodes testées n'est satisfaisante en raison de la taille beaucoup trop grande de l'échantillon requis pour la convergence des différents estimateurs.

INTRODUCTION

Lorsque l'on veut estimer les paramètres d'un modèle où il y a de l'autocorrélation des erreurs, la technique qui consiste à corriger d'abord cette autocorrélation est couramment utilisée. Or une telle méthode est inefficace lorsqu'il y a des erreurs de mesure sur les variables. L'utilisation des moindres carrés ordinaires (MCO), bien que déconseillée, se révèle préférable à une telle méthode en présence d'autocorrélation et d'erreurs de mesure sur les variables.

Selon une étude réalisée par simulation de Monte Carlo, lorsque les variables indépendantes sont mesurées avec une erreur de moyenne égale à zéro et de variance égale à 5% de la variance de la variable originale, la moyenne des estimations des MCO donne une valeur de 0.28 pour une vraie valeur du paramètre égale à 0.3. La technique couramment proposée quant à elle donne une valeur de 0.15 (voir Dagenais, CRDE, 1994). La même tendance se vérifie en terme d'erreur quadratique moyenne: MCO (0.065) alors que la technique couramment proposée donne (0.197).

Ces résultats s'expliquent par le fait que corriger l'autocorrélation revient à appliquer la technique d'estimation sur des données transformées. Cela peut conduire à des erreurs beaucoup plus grandes non seulement en terme d'erreur quadratique moyenne mais en terme d'erreur de type I (rejeter la vraie valeur du paramètre estimé par un test de student). On surestime le niveau de l'intervalle de confiance ce qui augmente les chances de commettre l'erreur de première espèce.

L'une des raisons pour laquelle la méthode qui corrige l'autocorrélation des erreurs est systématiquement proposée est due au fait que le problème des erreurs de mesure sur les variables n'a pas été suffisamment traité dans la littérature de sorte que les chercheurs y ont été très peu sensibilisés. Pourtant il s'agit d'un problème trop important pour être négligé. Pour s'en convaincre, il suffit d'analyser par exemple les données relatives aux pays en voie de développement. La plupart de ces données sont sujettes à des erreurs de mesure du fait même de l'absence d'instrument de collecte et de valorisation fiable et adéquat.

Il est aussi courant en économie qu'une variable soit approximée par une autre plus facilement observable. Cela peut être le cas lors de la mesure du PNB d'un pays. On peut l'approximer par le produit intérieur brut (PIB), faisant de ce fait abstraction des transferts avec l'étranger. Le PNB ainsi exprimé comporte nécessairement des erreurs de mesure.

Un autre exemple pourrait concerner la valorisation des revenus d'un ménage. On retient souvent comme approximation du revenu les dépenses totales, négligeant ainsi l'épargne qui en est une des composantes essentielles. Ces approximations contiennent un biais systématique et une erreur aléatoire de moyenne zéro. C'est cette erreur aléatoire qui nous intéresse ici car c'est la présence de celle-ci qui biaise les estimations des paramètres autres que la constante de la régression.

Nous allons faire une étude comparative de divers estimateurs en prenant comme norme le cas où il n'y a pas d'erreurs de mesure et les estimateurs utilisés sont les MCO ou l'estimateur des variables instrumentales où les instruments sont x_t et x_{t-1} , tel que recommandé dans les manuels. Les biais de ces estimateurs et la rapidité de convergence des MCO lorsque les u ne sont pas autocorrélés serviront de point de comparaison pour évaluer la performance des estimateurs des variables instrumentales et de l'estimateur de Fuller lorsque toutes les hypothèses restrictives sont relâchées.

I-REVUE SÉLECTIVE DE LA LITTÉRATURE

A - Malinvaud, 3^{ème} édition, Méthode statistique de l'économétrie, PP.420-458

Il est souvent admis que les variables sont observables sans erreur, ce qui est contraire à la réalité. On substitue souvent les vraies variables à d'autres variables en relation étroite mais non égales. Prenons l'exemple du revenu d'un ménage approximé par ses dépenses totales. Ceci conduit à une marge d'erreur positive entre la variable du modèle et la réalité qui contient un biais systématique et une erreur aléatoire.

Ce chapitre montre qu'en présence d'erreur de mesure sur les variables, les techniques d'estimation usuelles telles que les moindres carrés (MCO) par exemple ne peuvent être appliquées car certaines des hypothèses classiques ne sont plus vérifiées.

Pour rappel, les hypothèses du modèle linéaire sont :

$$x = Az + \varepsilon \quad (1)$$

où x est le vecteur des variables endogènes aléatoires

z est le vecteur des variables exogènes non aléatoires

ε est le terme d'erreur qui traduit l'effet des facteurs non observables du modèle

Les hypothèses couramment admises pour le traitement statistique du modèle sont:

hypothèse 1: indépendance des erreurs

Le vecteur des erreurs est indépendant de z et suit une loi indépendante du temps.

hypothèse 2: le vecteur ε est identiquement et indépendamment distribué.

La matrice des variances et covariances est une matrice non singulière.

hypothèse 3: Les erreurs suivent une loi normale.

hypothèse 4: Lorsque le nombre d'observations T croît indéfiniment, la matrice

$M_{zz} = \frac{1}{T} \sum zz'$ tend vers une matrice finie non singulière M_0 .

Si les variables du modèle (1) ne sont pas observées directement et que l'on dispose seulement d'une évaluation approchée, x^* et z^* telle que :

$$x^* = x + u \text{ et}$$

$$z^* = z + v$$

où u et v sont des erreurs de mesure indépendantes entre elles et entre observations, alors (1) devient:

$$\begin{aligned} x^* &= Az^* + (\varepsilon + u - Av) \\ &= Az^* + \eta \end{aligned}$$

La question qui se pose évidemment est de savoir si les hypothèses formulées plus haut et qui sont les conditions essentielles à l'application des MCO peuvent s'appliquer à $\eta = \varepsilon - u - Av$.

Il ressort à partir des affirmations faites sur les variables du modèle que η est indépendant de x , z et ε mais non de z^* , ce qui porte une entorse à l'hypothèse 1. Les MCO appliqués à ce modèle donnent donc des estimateurs non convergents.

Le recours à la méthode des variables instrumentales est la solution la plus appropriée car elle donne des estimateurs convergents et possiblement relativement efficaces. Les variables instrumentales doivent être choisies de telle sorte qu'elles ne doivent pas être orthogonales au vecteur des variables exogènes. Cependant, la question qui se pose est de connaître la vitesse de convergence de ces estimateurs car les moments dans les petits échantillons ne sont pas toujours définis.

B - Fomby and Hill, Advanced econometrics, PP.205-277

Fomby et Hill traitent essentiellement dans cette partie de l'autocorrélation, de ses causes et des méthodes d'estimation appropriées pour obtenir des estimateurs efficaces.

Nous nous intéresserons d'abord à l'autocorrélation d'ordre 1, et à la technique d'estimation appropriée, à partir d'un modèle simple. Dans la partie suivante, les effets de l'ajout d'une variable dépendante retardée au modèle sont analysés en terme d'estimation des paramètres. La dernière partie introduit des erreurs de mesure sur les variables.

-Le processus autorégressif d'ordre 1

Il est souvent admis que l'autocorrélation existe dans un modèle parce que les variables omises sont autocorrélées.

Soit le vrai modèle suivant:

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 z_t + e_t$$

Et le modèle utilisé pour la régression est :

$$y_t = \beta_1 + \beta_2 x_t + u_t$$

avec $u_t = \beta_3 z_t + e_t$

Les hypothèses posées sur z_t sont cruciales pour déterminer la technique d'estimation utilisée. Étant donné que e_t est identiquement et indépendamment distribué (iid), si z_t est autocorrélé, alors u_t le sera aussi. Mais cette autocorrélation peut prendre plusieurs formes selon que z_t est dépendant ou non de x_t et la technique d'estimation devra en tenir compte.

Si l'autocorrélation apparaît du fait de l'autocorrélation de la variable z_t qui se trouve elle-même liée à x_t , il faut chercher à inclure cette variable. Car les MCO et la technique qui corrige l'autocorrélation peuvent conduire à des estimateurs non convergents.

Si on suppose l'indépendance de z_t vis à vis des variables explicatives et que l'autocorrélation qui en résulte peut s'approximer par un processus autorégressif d'ordre 1, le modèle pourrait s'écrire simplement:

$$Y = X\beta + u$$

où $u_t = \rho u_{t-1} + \varepsilon_t, t = 1, 2, \dots, T$

et ε_t est iid

On retient ici le processus autorégressif (AR) d'ordre 1 mais il reste entendu que d'autres formes d'autocorrélation pourraient être envisagées telles que les AR d'ordre n ou les ARMA(p,q). Les

résultats que nous dégagerons de notre analyse restent valables quelle que soit la forme d'autocorrélation choisie.

Il convient par ailleurs de poser que u_t suit un processus stationnaire .

L'espérance mathématique du processus se calcule aisément en faisant les substitutions successives:

$$u_t = \rho u_{t-1} + \varepsilon_t$$

En remplaçant:

$$u_{t-1} = \rho u_{t-2} + \varepsilon_{t-1},$$

on a :

$$u_t = \rho^2 u_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t$$

Par itération successive, il ressort que :

$$u_t = \rho^T u_{t-T} + \sum \rho^i \varepsilon_{t-i}$$

quand T tend vers l'infini, $u_t = \sum \rho^i \varepsilon_{t-i}$

L'espérance mathématique :

$$E(u_t) = \sum \rho^i E(\varepsilon_{t-i}) = 0$$

La variance :

$$\begin{aligned} \text{Var}(u_t) &= E(u_t)^2 \\ &= E(\sum \rho^i (\varepsilon_{t-i}))^2 = E(\sum \rho^{2i} \varepsilon_{t-i}^2) \\ &= E(\varepsilon_t)^2 + \rho^2 E(\varepsilon_{t-1})^2 + \rho^4 E(\varepsilon_{t-2})^2 + \dots \\ &= \sigma_\varepsilon^2 \left(\frac{1}{1-\rho^2} \right) \end{aligned}$$

De la même manière, on calcule la covariance :

$$\text{Cov}(u_t, u_{t-i}) = \rho^i \text{var}(u)$$

Il apparaît donc que la matrice des variances et covariances peut s'écrire sous la forme :

$$E(u_t u_t') = \sigma_\varepsilon^2 \Omega$$

où Ω est une matrice différente de la matrice identité.

La matrice de covariance est non sphérique donc les MCO, bien que sans biais ne sont pas efficaces. La méthode des moindres carrés généralisés est dans ce cas utilisée.

$$b = (X' \Omega^{-1} X)^{-1} (X' \Omega^{-1} y)$$

D'autres méthodes similaires peuvent être utilisées si la matrice est connue

(voir les méthodes de Prais-Winsten (1954), de Dhrymes (1971), de Cochrane-Orcutt et de Hildreth-Lu etc...)

-Variables dépendantes retardées et autocorrélation:

Lorsque l'on analyse l'effet des MCO dans un modèle simple contenant une variable dépendante retardée, l'hypothèse sur le terme d'erreur permet de choisir la technique d'estimation appropriée.

Soit le modèle simple :

$$y_t = \alpha y_{t-1} + e_t \quad (1)$$

e_t est le terme d'erreur de moyenne nulle, de variance finie σ_ε^2 et non autocorrélé.

Les MCO sont biaisés pour les petits échantillons à cause de la corrélation contemporaine entre y_{t-1} et e_{t-1} .

Selon une étude réalisée par White (1961), ce biais est approximativement égal à $-2\alpha/T$.

Mais ce biais disparaît dès lors qu'on augmente l'échantillon à l'infini. Les MCO deviennent donc des estimateurs efficaces asymptotiquement et s'apparentent aux estimateurs du maximum de vraisemblance.

Par contre, lorsqu'on considère le même modèle et qu'on admet que les erreurs sont autocorrélées; soit $e_t = \rho e_{t-1} + u_t$ et u_t est de moyenne zéro et de variance finie σ_u^2 , l'estimateur des MCO est non seulement biaisé mais aussi non convergent:

$$p \lim(\hat{\alpha} - \alpha) \neq 0.$$

Des estimateurs convergents ont été proposés par plusieurs auteurs. Nous nous contenterons de les énumérer sans entrer dans les détails des démonstrations.

Si rho est connu, la transformation de Prais-Winsten permet d'obtenir, par exemple, des estimateurs convergents. Considérons un modèle beaucoup plus général de la forme:

$$y_t = \beta x_t + \alpha y_{t-1} + e_t \quad ; \quad t = 2, 3, \dots, T$$

avec $e_t = \rho e_{t-1} + u_t$

et u_t est normal et i.i.d

En utilisant la transformation suivante

$$\sqrt{1 - \rho^2} y_t = \beta x_t \sqrt{1 - \rho^2} + \alpha y_{t-1} \sqrt{1 - \rho^2} + e_t \sqrt{1 - \rho^2}$$

avec $y_t - \rho y_{t-1} = \beta(x_t - \rho x_{t-1}) + \alpha(y_{t-1} - \rho y_{t-2}) + u_t \quad ;$
 $t = 3, 4, \dots, T$

ou plus simplement:

$$P'y = P'X\beta + P'e \quad (2) \quad \text{où } P \text{ est une matrice } (T-1) \times (T-1) \text{ de Prais-Winsten}$$

$$P' = \begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 \\ -\rho & 1 & 0 & \dots & 0 \\ 0 & -\rho & 1 & & : \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & -\rho & 1 \end{pmatrix}$$

En appliquant les MCO à l'équation (2), on obtient des estimateurs convergents dans le cas où ρ est connu.

Si par contre ρ n'est pas connu, la procédure itérative du maximum de vraisemblance permet d'estimer ρ et l'ensemble des paramètres.

Comme on peut le voir dans notre modèle, la difficulté essentielle réside dans le fait que y_{t-1} et e_t sont contemporanément corrélés. Le recours aux méthodes mentionnées plus haut ou surtout l'utilisation d'une méthode des variables instrumentales sont autant de solutions qui conduisent à des estimateurs convergents et possiblement assez efficaces.

La situation se complique davantage dès lors que des erreurs de mesure sont ajoutées aux variables. Aucune méthode satisfaisante pour les petits échantillons n'a encore été proposée dans ce sens et le but de notre étude est d'investiguer dans cette direction en utilisant une approche progressive.

La démonstration qui va suivre s'appuie sur un modèle simple mais il reste entendu que les résultats obtenus peuvent être généralisés à des modèles beaucoup plus complexes.

Soit le modèle simple:

$$y_t^* = \beta x_t^*$$

où les variables sont non observables et sont telles que les valeurs disponibles obéissent à la relation suivante:

$$y_t = y_t^* + e_t$$

$$x_t = x_t^* + u_t$$

$$Ee_t = 0, \quad Ee_t e_{t'} = 0, \quad Ee_t^2 = \sigma_e^2$$

$$Eu_t = 0, \quad Eu_t u_{t'} = 0, \quad Eu_t^2 = \sigma_u^2$$

$$Ee_t u_t = Ey_t^* e_t = Ex_t^* u_t = Ey_t^* u_t = Ex_t^* e_t = 0$$

$$\Rightarrow y_t = \beta x_t + (e_t - \beta u_t)$$

$$\hat{\beta} = \frac{\sum xy}{\sum x^2} = \frac{\sum x(\beta x + v_t)}{\sum x^2} = \beta + \frac{\sum xv/T}{\sum x^2/T}$$

avec $v_t = e_t - \beta u_t$

En posant $p \lim \sum x^2/T = \sigma_x^2$

$$\Rightarrow p \lim \hat{\beta} = \beta - \beta \sigma_u^2 / \sigma_x^2$$

Les MCO restent biaisés même asymptotiquement.

On pourrait également penser à la méthode du maximum de vraisemblance, mais cette approche est très peu utilisée du fait des hypothèses fortes qui sont requises sur les variables observables et sur les erreurs de mesure qui doivent être multi-normales. En outre, la connaissance à priori de la variance des erreurs de mesure est indispensable.

La méthode la plus utilisée dans les cas d'espèce est donc la méthode des variables instrumentales. Elle est de la forme:

$$\hat{\beta} = (z'x)^{-1} z'y$$

où z est la variable instrumentale choisie de telle sorte qu'elle soit indépendante des erreurs mais liée à x .

C - Dagenais , Journal of econometrics , 1994

Cet article traite de l'estimation des paramètres de la régression en cas d'erreurs de mesure sur les variables et d'autocorrélation des erreurs.

Comme mentionné dans notre introduction, lorsque autocorrélation et erreurs de mesure coexistent dans un modèle , toute technique d'estimation qui corrige cette autocorrélation en faisant fi des erreurs de mesure peut conduire à des estimateurs non convergents. Cela se vérifie surtout lorsque les erreurs de mesure sont faiblement corrélées , ce qui est le cas dans la plupart des données économiques.

Cet article qui étudie un modèle simple de régression où coexistent autocorrélation et erreurs sur les variables montre que l'utilisation de toute technique d'estimation qui tient compte de l'autocorrélation mais ne tient pas compte des erreurs de mesure (technique proposée) conduit à des estimateurs non convergents surtout lorsque les erreurs de mesure sont moins corrélées que les vraies valeurs.

Ces résultats ont été démontrés tant du point de vue asymptotique que de celui des échantillons finis (150 observations) en utilisant la simulation de Monte Carlo.

Les résultats se résument comme suit:

Variance erreur = 5% variance de x_t

Vraie valeur de Bêta = 0.3

	Bêta	EQM
MCO	0,28	0,07
Technique proposée ¹	0,15	0,19

¹ *technique qui corrige l'autocorrélation des erreurs; se retrouve dans la plupart des ouvrages d'économétrie.*

Le fait que x_t suive ou non un processus stationnaire n'influence pas significativement les résultats de l'analyse ci-dessus.

Cette simulation a été faite avec le modèle suivant:

$$y_t = \beta x_t + u_t, \quad t = 1 \dots n$$

x_t est la variable exogène et stochastique de moyenne zéro,
 β_t est le coefficient fixe inconnu,
 u_t est l'erreur normalement distribuée et indépendante de x_t

On assume que x_t est non observable et que seul $x_t^* = x_t + v_t$ l'est.

Notons que v_t est une erreur de mesure aléatoire normalement distribuée et indépendante de x_t .

On a également les relations suivantes:

$$u_t = \rho_u u_{t-1} + \varepsilon_t \quad |\rho_u| < 1$$

$$x_t = \rho_x x_{t-1} + \xi_t \quad |\rho_x| < 1$$

$$v_t = \rho_v v_{t-1} + w_t \quad |\rho_v| < 1$$

Certaines conclusions ont été obtenues pour différentes valeurs du paramètre λ défini comme le rapport des variances de v_t et x_t .

On suppose, bien entendu, que les v_t sont indépendantes entre observations, autrement dit leurs covariances sont nulles.

Plus λ est grand, plus les paramètres estimés s'écartent de leur vraie valeur

Les MCO restent préférables dans tous les cas de figure.

Lorsque λ augmente, l'estimateur du maximum de vraisemblance qui ignore la présence des erreurs de mesure se détériore encore plus.

Les résultats obtenus sont similaires quand bien même x_t suit une marche aléatoire et donc est non stationnaire.

L'explication de ces résultats est que la procédure du maximum de vraisemblance consiste à estimer ρ et à appliquer la méthode sur des données transformées ;

$$\text{soit, } x_t - \hat{\rho}_u x_{t-1}$$

Le biais peut devenir plus élevé avec les données transformées.

Résultats analytiques

MCO: $p \lim \hat{\beta} = \beta / (1 + \lambda)$

où $\hat{\beta} = \sum \tilde{x}_t y_t / \sum \tilde{x}_t^2$

MV: en ignorant la présence des erreurs de mesure sur les variables:

$p \lim \hat{\beta} = \beta \sigma_x^2 (1 + \rho_u^{*2} - 2\rho_x \rho_u^*) / D_3$
avec $D_3 = \sigma_x^2 (1 + \rho_u^{*2} - 2\rho_x \rho_u^*) + \sigma_v^2 (1 + \rho_u^{*2} - 2\rho_x \rho_u^*)$

et $\rho_u^* = \frac{\rho_u \sigma_u^2 + \beta^2 \psi^2 \rho_x \sigma_x^2 + \beta^2 (1 - \psi)^2 \rho_v \sigma_v^2}{\sigma_u^2 + \beta^2 \psi^2 \sigma_x^2 + \beta^2 (1 - \psi)^2 \sigma_v^2}$

où $\psi = \sigma_v^2 / (\sigma_x^2 + \sigma_v^2) = \lambda / (1 + \lambda)$

N.B.: variance est $O(N^{-1})$

biais est $O(1)$

Si bien que lorsque N augmente, la variance devient négligeable

Dans le cas où $\rho_v = 0.75$ et $\rho_x = 0.7$ le maximum de vraisemblance donne des résultats meilleurs que les MCO.

Car lorsque $\rho_v = \rho_x$, $\Rightarrow p \lim \tilde{\beta} = p \lim \hat{\beta}$.

Ainsi, pour les grands échantillons, si $\rho_x = \rho_v$, alors le biais de $\tilde{\beta}$ est similaire à celui de $\hat{\beta}$.

$\rho_v \ll \rho_x \Rightarrow EQM \quad MV \gg MCO$

$\rho_v \gg \rho_x \Rightarrow EQM \quad MV \ll MCO$

Étant donné que les erreurs de mesure sur les données économiques sont en général moins corrélés que les vraies valeurs, on ne devrait pas songer à corriger l'autocorrélation des erreurs qui conduisent à des estimateurs non convergents.

Dans le cas de variables strictement positives, une mesure de l'importance des erreurs sur les variables plus adéquate que λ peut être RI où :

$$RI = E|v_v| / \bar{x}, \quad \text{où } \bar{x} = \sum x_t / N$$

Le maximum de vraisemblance peut conduire à une erreur de type 1 beaucoup plus élevée que dans le cas des MCO.

D - Consistent estimators for regression models with errors in the variables and autocorrelated disturbances.

Cet article présente la méthode d'estimation proposée par Fuller en ce qui concerne le modèle suivant:

$$\tilde{Y}_t = \alpha + \tilde{X}_t \beta + \delta \tilde{Y}_{t-1} + u_t$$

$$u_t = \rho u_{t-1} + \varepsilon_t$$

\tilde{X}_t est un vecteur (1 x k) de variables exogènes

ε_t sont identiquement et indépendamment distribués de variance σ^2

β est de dimension (k x 1)

α, δ, σ^2 sont des paramètres à estimer.

Les vraies valeurs de \tilde{Y} et de \tilde{X} ne sont pas observées. On observe seulement X_t et Y_t avec des erreurs de mesure.

$$\begin{aligned} X_t &= \tilde{X}_t + v_t \\ Y_t &= \tilde{Y}_t + s_t \end{aligned}$$

où v_i et s_i , de dimension $(1 \times k)$, sont des erreurs de mesure indépendamment distribuées entre observations et $E(v_i, s_i) = 0$

Sous forme matricielle, et en considérant qu'on a T observations:

$$Y = \alpha L + X\beta + \delta Y_{-1} + \eta$$

$$\eta = u - v\beta + S - \delta S_{-1}$$

où l est un vecteur unité $(T \times 1)$

Y, Y_{-1}, η, u, S et S_{-1} sont des vecteurs $(T \times 1)$
et X et v sont des matrices de dimension $(T \times k)$

L'estimateur

L'instrument proposé est:

$$Z_{+2} = (L, X_{+1}, X_{+2})$$

La procédure de calcul est la suivante:

$$i) f = [(Y, L, X)'(Y, L, X) - (Y, L, X)'Z_{+2}(Z_{+2}'Z_{+2})^{-1}Z_{+2}(Y, L, X)]/(N - q)$$

où N est le nombre d'observations

et q est le nombre de colonnes de Z_{+2} soit $2k+1$.

X est le vecteur des variables $X = (x_{y_1})$

L est le vecteur unitaire pour tenir compte de la constante

$$ii) (\hat{Y}, \hat{X}) = Z_{+2}(Z_{+2}'Z_{+2})^{-1}Z_{+2}'(Y, X)$$

La troisième étape consiste à calculer la plus petite valeur propre de

$$iii) \left| (\hat{Y}, L, \hat{X})'(\hat{Y}, L, \hat{X}) - \nu f \right| = 0$$

$$\text{iv) } \hat{\theta} = \left[(L, \hat{X})' (L, \hat{X}) - (\tilde{V} - \sigma) f_{22} \right]^{-1} \left[(L, \hat{X})' \hat{Y} - (\tilde{V} - \sigma) f_{21} \right]$$

$$\text{où } \hat{\theta} = (\tilde{\alpha}, \tilde{\beta}, \tilde{\delta})'$$

et f_{21} est une sous-matrice de f contenant les lignes 2 à k et la colonne 1 de f

f_{22} est une matrice carré $[2 \text{ à } (k+2)] \times [2 \text{ à } (k+2)]$

C'est cette méthode que nous adapterons à notre modèle. La seule différence réside dans le fait qu'il n'y a pas de constante dans notre modèle.

II- LA MÉTHODE

A- LE MODÈLE

Dans le cadre de ce rapport de recherche , le modèle de série chronologique que nous étudierons est le suivant:

$$\tilde{y}_t = \alpha + \beta \tilde{x}_t + \gamma \tilde{y}_{t-1} + u_t$$

$$\text{Avec } u_t = \rho u_{t-1} + \varepsilon_t$$

$$\text{et } \alpha = 0$$

Il s'agit d'un modèle qui comporte une variable exogène et une variable indépendante retardée. On fait également l'hypothèse que le terme d'erreur suit un processus autorégressif d'ordre 1.

Dans la littérature, ce modèle est appelé modèle d'ajustement dynamique (Hatanaka, 1974).

Ce peut être un modèle qui explique le comportement du prix d'un bien donné. Le prix étant expliqué par exemple par la quantité de ce bien disponible sur le marché et par sa valeur passée.

Le terme d'ajustement dynamique est utilisé pour dire que la variable expliquée s'ajuste par rapport à sa valeur passée.

On propose également que le modèle suive un processus stationnaire d'où les conditions suivantes à poser sur les paramètres:

Rappelons que pour un processus stationnaire, la variance et l'espérance mathématique sont finies et indépendantes du temps et la covariance ne dépend que du retard.

Soit:

$$\tilde{y}_t = \beta \tilde{x}_t + \gamma \tilde{y}_{t-1} + u_t$$

$$\Rightarrow \tilde{y}_{t-1} = \beta \tilde{x}_{t-1} + \gamma \tilde{y}_{t-2} + u_{t-1}$$

$$\Rightarrow \tilde{y}_t = \beta \tilde{x}_t + \gamma \beta \tilde{x}_{t-1} + \gamma^2 \tilde{y}_{t-2} + u_t + \gamma u_{t-1}$$

De la même manière, si on remplace \tilde{y}_{t-2} par sa valeur, on obtient:

$$\tilde{y}_t = \beta \tilde{x}_t + \gamma \beta \tilde{x}_{t-1} + \gamma^2 \beta \tilde{x}_{t-2} + \gamma^3 \tilde{y}_{t-3} + u_t + \gamma u_{t-1} + \gamma^2 u_{t-2}$$

Par itérations successives, on obtient :

$$\tilde{y}_t = \gamma^T \tilde{y}_0 + \beta \sum_{i=0}^T \gamma^i \tilde{x}_{t-i} + \sum_{i=0}^T \gamma^i u_{t-i}$$

$$\text{Avec } \tilde{x}_t = \xi \tilde{x}_{t-1} + \omega_t$$

$$\Rightarrow \tilde{x}_{t-i-1} = \xi \tilde{x}_{t-i-2} + \omega_{t-i-1}$$

En faisant les substitutions successives comme précédemment, on obtient:

$$\tilde{x}_{t-i} = \xi^T \tilde{x}_0 + \sum_{j=0}^T \xi^j \omega_{t-i-j}$$

$$\text{Avec } u_t = \rho u_{t-1} + \varepsilon_t$$

En faisant un raisonnement analogue à ce qui est ci-dessus, on déduit:

$$u_{t-i} = \rho^T u_0 + \sum_{k=0}^T \rho^k \varepsilon_{t-i-k}$$

La relation finale s'écrit:

$$\tilde{y}_t = \gamma^T \tilde{y}_0 + \beta \sum_{i=0}^T \gamma^i (\xi^T \tilde{x}_0 + \sum_{j=0}^T \xi^j \omega_{t-i-j}) + \sum_{i=0}^T \gamma^i (\rho^T u_0 + \sum_{k=0}^T \rho^k \varepsilon_{t-i-k})$$

$$\text{ou } \tilde{y}_t = \gamma^T \tilde{y}_0 + \beta \xi^T \tilde{x}_0 \sum_{i=0}^T \gamma^i + \beta \sum_{i,j=0}^T \gamma^i \xi^j \omega_{t-i-j} + \rho^T u_0 \sum_{i=0}^T \gamma^i + \sum_{i,k=0}^T \gamma^i \rho^k \varepsilon_{t-i-k}$$

$$E(\tilde{y}_t) = \gamma^T E(\tilde{y}_0) + \beta \xi^T E(\tilde{x}_0) \frac{1-\gamma^T}{1-\gamma} + \rho^T \frac{1-\gamma^T}{1-\gamma} E(u_0)$$

Pour que cette espérance soit finie ; condition pour avoir un processus stationnaire, il suffit que :

$$|\gamma| < 1, \quad |\xi| < 1 \text{ et } |\rho| < 1$$

Et dans ce cas , l'espérance du processus est égal à zéro. De la même manière, on calcule la variance du processus et il apparaît aussi que pour qu'elle soit finie, il faut que les conditions ci-dessus soient respectées.

Ce sont donc ces hypothèses que nous retiendrons pour le traitement statistique de notre modèle.

En reprenant le modèle, on sait que les vraies valeurs \tilde{x}_t et \tilde{y}_t ne sont pas observables. Seuls x_t et y_t sont observables avec une erreur de mesure, c'est à dire qu'il s'agit d'une évaluation approchée .

$$x_t = \tilde{x}_t + v_t \quad \text{et} \quad y_t = \tilde{y}_t + s_t$$

v_t et s_t ne sont ni autocorrélées ni corrélées avec les autres variables.

Le modèle peut se réécrire:

$$y_t - s_t = \beta(x_t - v_t) + \gamma(y_{t-1} - s_{t-1}) + u_t$$

$$y_t = \beta x_t + \gamma y_{t-1} + \eta_t$$

$$\text{où } \eta_t = u_t + s_t - \gamma s_{t-1} - \beta v_t$$

$$y = W\theta + \eta \quad \text{avec } W = (x \ y_{-1}) \quad \text{et } \theta' = (\beta \ \gamma)$$

C'est sur cette décomposition du modèle que va s'appuyer toute la suite de l'analyse.

B - Choix de l'estimateur: la méthode des variables instrumentales

L'indépendance de η_t vis à vis de y_{t-1} et de x_t n'est pas vraie de sorte qu'il n'est pas indiqué d'appliquer les MCO: $P \lim \frac{Y_t - 1\eta_t}{N} \neq 0$

L'estimateur des MCO s'écrit comme suit:

$$\hat{\theta} = (W'W)^{-1}(W'Y)$$

$$= \theta + (W'W)^{-1}W'\eta$$

$$p \lim \hat{\theta} = \theta + (p \lim \frac{W'W}{T})^{-1} p \lim (\frac{W'\eta}{T})$$

$$p \lim \hat{\theta} \neq \theta$$

$$\text{car } p \lim (\frac{W'\eta}{T}) \neq 0$$

et $p \lim (\frac{W'W}{T}) \rightarrow Q_{ww}$. (matrice...finie) MCO non convergent

Il est possible de calculer le biais d'un tel estimateur avec ce modèle:

$$p \lim (\frac{W'W}{T}) = p \lim \left(\begin{array}{cc} \frac{X'X}{T} & \frac{X'Y_{-1}}{T} \\ \frac{Y'_{-1}X}{T} & \frac{Y'_{-1}Y_{-1}}{T} \end{array} \right)$$

$$P \lim (\frac{X'X}{T}) = \sigma_{\tilde{X}}^2 + \sigma_v^2$$

$$P \lim (\frac{Y'_{-1}Y_{-1}}{T}) = \sigma_Y^2 + \sigma_s^2$$

$$P \lim (\frac{X'Y_{-1}}{T}) = P \lim (\frac{Y'_{-1}X}{T}) = \frac{\beta\xi}{1-\gamma\xi} \sigma_{\tilde{X}}^2$$

D'où : $P \lim (\frac{W'W}{T}) = \begin{pmatrix} \sigma_{\tilde{X}}^2 + \sigma_v^2 & \frac{\beta\xi}{1-\gamma\xi} \sigma_{\tilde{X}}^2 \\ \frac{\beta\xi}{1-\gamma\xi} \sigma_{\tilde{X}}^2 & \sigma_Y^2 + \sigma_s^2 \end{pmatrix}$ est une matrice finie non singulière différente de

zéro.

De la même manière, on calcule:

$$P \lim (\frac{W'\eta}{T}) = p \lim \left(\begin{array}{c} \frac{X'\eta}{T} \\ \frac{Y'_{-1}\eta}{T} \end{array} \right)$$

$$P \lim (\frac{X'\eta}{T}) = P \lim \frac{1}{T} ((\tilde{X}' + v')u + (\tilde{X}' + v')s + (\tilde{X}' + v')(-\gamma s) + (\tilde{X}' + v')(-\beta v))$$

$$\Leftrightarrow P \lim\left(\frac{X'\eta}{T}\right) = -\beta\sigma_v^2$$

$$P \lim\left(\frac{Y'_{-1}\eta}{T}\right) = \frac{\rho}{1-\gamma\rho}\sigma_u^2$$

Pour résumer, le biais est égal à :

$$\text{Biais} = \text{inv}\left(\begin{array}{cc} \sigma_{\bar{X}}^2 + \sigma_v^2 & \frac{\beta\xi}{1-\gamma\xi}\sigma_{\bar{X}}^2 \\ \frac{\beta\xi}{1-\gamma\xi}\sigma_{\bar{X}}^2 & \sigma_{\bar{Y}}^2 + \sigma_s^2 \end{array} \right) * \left(\frac{-\beta\sigma_v^2}{\frac{\rho}{1-\gamma\rho}\sigma_u^2} \right)$$

Comme on peut le voir, les MCO sont biaisés même asymptotiquement. Ce résultat nous permet faire deux remarques fondamentales:

- i) En l'absence d'autocorrélation du terme d'erreur u_t et d'erreur de mesure sur la variable x_t , les MCO deviennent des estimateurs convergents.
- ii) La seconde remarque est qu'il y ait ou non une erreur de mesure sur la variable endogène, y_t n'a pas d'incidence sur le biais des MCO si s_t est non corrélée avec l'erreur de mesure sur x_t .

L'alternative aux MCO en présence d'autocorrélation et d'erreur de mesure est la méthode des variables instrumentales et le choix de celles-ci doit répondre à deux exigences essentielles:

d'une part les variables instrumentales doivent être reliées avec les variables explicatives mais non avec le terme d'erreur. D'autre part, et d'un point de vue pratique, les variables instrumentales doivent être faciles à générer, c'est à dire que leur choix ne doit pas faire appel à la collecte d'autres données que celles déjà disponibles afin de minimiser les coûts en temps et en argent.

Étant donné que les x_t sont autocorrélés, il devient utile de prendre comme variables instrumentales les x retardés ou avancés. On passe outre la deuxième difficulté car les x_t sont disponibles et aucune autre variable supplémentaire n'est nécessaire.

La convergence sera testée à partir des développements ci-dessous.

Dans un premier temps, comme solution intermédiaire, nous testerons l'estimateur des variables instrumentales utilisant comme instruments $Z_1 = (x_t, x_{t-1})$. C'est la méthode recommandée par la plupart des ouvrages économétriques dès lors qu'il y a autocorrélation des erreurs.

L'estimateur s'écrit :

$$\hat{\theta} = (Z_1' W)^{-1} (Z_1' y) \quad \text{avec} \quad Z_1 = (x_t, x_{t-1})$$

$$\hat{\theta} = \theta + (Z_1' W)^{-1} Z_1' \eta$$

$$p \lim(\hat{\theta} - \theta) = p \lim \left[(Z_1' W)^{-1} (Z_1' \eta) \right]$$

$$p \lim(\hat{\theta} - \theta) = p \lim \left(\frac{Z_1' W}{N} \right)^{-1} p \lim \left(\frac{Z_1' \eta}{N} \right)$$

$$p \lim \left(\frac{Z_1' W}{N} \right) = \begin{pmatrix} p \lim \left(\frac{X' X}{N} \right) & p \lim \left(\frac{X' Y_{-1}}{N} \right) \\ p \lim \left(\frac{X'_{-1} X}{N} \right) & p \lim \left(\frac{X'_{-1} Y_{-1}}{N} \right) \end{pmatrix}$$

avec $p \lim \left(\frac{X' X}{N} \right) = \sigma_v^2 + \sigma_X^2$

$$p \lim \left(\frac{X' Y_{-1}}{N} \right) = \frac{\beta \xi}{1 - \gamma \xi} \sigma_X^2$$

$$p \lim \left(\frac{X'_{-1} X}{N} \right) = \xi \sigma_X^2$$

$$p \lim \left(\frac{X'_{-1} Y_{-1}}{N} \right) = \frac{\beta}{1 - \gamma \xi} \sigma_X^2$$

D'où $p \lim \left(\frac{Z_1' W}{N} \right)$ tend vers une matrice finie.

D'un autre côté, on montre aisément que :

$$p \lim \left(\frac{Z_1' \eta}{N} \right) = \begin{pmatrix} p \lim \left(\frac{X' \eta}{N} \right) \\ p \lim \left(\frac{X'_{-1} \eta}{N} \right) \end{pmatrix} = \begin{pmatrix} -\beta \sigma_v^2 \\ 0 \end{pmatrix}$$

Avec les hypothèses générales de notre modèle, à savoir l'existence d'autocorrélation et d'erreurs sur les variables, l'estimateur ci-dessus n'est pas convergent même asymptotiquement car :

$$p \lim(\hat{\theta} - \theta) \neq 0$$

Par contre, il apparaît que pour avoir un estimateur convergent, il suffit de supposer l'absence d'erreurs sur la variable x_t .

La seconde constatation est identique à la première: le fait qu'il y ait ou non une erreur sur y_t n'a aucune incidence sur les résultats obtenus, ceteris paribus.

Le dernier estimateur que nous proposerons devant la non convergence des deux méthodes précédentes est aussi un estimateur des variables instrumentales avec $Z = (x_{t+1}, x_{t+2})$.

En reprenant les mêmes notations:

$$\tilde{x}_t = \xi \tilde{x}_{t-1} + \omega_t$$

L'estimateur des variables instrumentales s'obtient par :

$$\hat{\theta} = (Z'W)^{-1} Z' Y \text{ avec } Z = (x_{t+1}, x_{t+2}) \text{ en utilisant une écriture matricielle.}$$

En effet on a :

$$\hat{\theta} = (Z'W)^{-1} Z' (W\theta + \eta)$$

$$= \theta + (Z'W)^{-1} Z' \eta$$

$$p \lim \hat{\theta} = \theta + (p \lim \frac{Z'W}{N})^{-1} p \lim \frac{Z' \eta}{N}$$

$$p \lim \left(\frac{Z'W}{N} \right) = \begin{pmatrix} p \lim \frac{X'_{t+1} X}{N} & p \lim \frac{X'_{t+1} Y_{-1}}{N} \\ p \lim \frac{X'_{t+2} X}{N} & p \lim \frac{X'_{t+2} Y_{-1}}{N} \end{pmatrix}$$

$$p \lim \left(\frac{X'_{t+1} X}{N} \right) = p \lim \left(\frac{1}{N} (\tilde{X}'_{t+1} + v'_{t+1}) (\tilde{X} + v) \right)$$

$$= p \lim \left(\frac{1}{N} \tilde{X}'_{t+1} \tilde{X} \right)$$

$$= p \lim \left(\frac{\xi \tilde{X} + \omega}{N} \right)' \tilde{X} = \xi p \lim \left(\frac{\tilde{X}' \tilde{X}}{N} \right) = \xi \sigma_{\tilde{X}}^2$$

$$p \lim \left(\frac{X'_{t+1} Y_{-1}}{N} \right) = p \lim \left(\frac{1}{N} (\tilde{X}'_{t+1} + v'_{t+1}) (\tilde{Y}_{-1} + s_{-1}) \right)$$

$$\begin{aligned}
&= p \lim \left(\frac{1}{N} (\tilde{X}'_{+1} \tilde{Y}_{-1}) \right) \\
&= p \lim \left(\frac{1}{N} (\tilde{X}'_{+1} (\beta \tilde{X}_{-1} + \gamma \tilde{Y}_{-2} + u_{-1})) \right) \\
&= \beta \cdot p \lim \left(\frac{1}{N} (\tilde{X}'_{+1} \tilde{X}_{-1}) \right) + \gamma \cdot p \lim \left(\frac{1}{N} (\tilde{X}'_{+1} \tilde{Y}_{-2}) \right) \\
&= \beta \xi \cdot p \lim \left(\frac{1}{N} (\tilde{X}' \tilde{X}_{-1}) \right) + \gamma \xi \cdot p \lim \left(\frac{1}{N} (\tilde{X}' \tilde{Y}_{-2}) \right) \\
&= \beta \xi^2 p \lim \left(\frac{1}{N} (\tilde{X}'_{-1} \tilde{X}'_{-1}) \right) + \gamma \xi \cdot p \lim \left(\frac{1}{N} (X' Y_{-2}) \right) \\
&= \beta \xi^2 \sigma_{\tilde{X}}^2 + \gamma \xi \cdot p \lim \left(\frac{1}{N} (X'_{+1} Y_{-1}) \right)
\end{aligned}$$

$$\text{d'où } p \lim \left(\frac{X'_{+1} Y_{-1}}{N} \right) = \frac{\beta \xi^2 \sigma_{\tilde{X}}^2}{1 - \gamma \xi}$$

De la même manière, on calcule :

$$p \lim \left(\frac{X'_{+2} X}{N} \right) = \xi \sigma_{\tilde{X}}^2 \quad \text{et} \quad p \lim \left(\frac{X'_{+2} Y_{-1}}{N} \right) = \frac{\beta \xi^3 \sigma_{\tilde{X}}^2}{1 - \gamma \xi}$$

Il s'en suit d'après les calculs faits ci-dessus que la matrice $(Z'W)$ est finie.

Il reste à dériver $p \lim \left(\frac{1}{N} (Z' \eta) \right)$;

$$p \lim \left(\frac{1}{N} (Z' \eta) \right) = \begin{pmatrix} p \lim \frac{X'_{+1} \eta}{N} \\ p \lim \frac{X'_{+2} \eta}{N} \end{pmatrix}$$

$$\begin{aligned}
p \lim \frac{X'_{+1} \eta}{N} &= p \lim \left(\frac{1}{N} (\tilde{X}'_{+1} + v'_{+1}) \eta \right) \\
&= p \lim \left(\frac{1}{N} (\tilde{X}'_{+1} \eta) \right) = 0
\end{aligned}$$

$$p \lim \frac{X'_{+2} \eta}{N} = p \lim \left(\frac{1}{N} (\tilde{X}'_{+2} + v'_{+2}) \eta \right) = p \lim \left(\frac{1}{N} (\tilde{X}'_{+2} \eta) \right) = 0$$

Cette méthode des variables instrumentales¹ donne des estimateurs convergents et est donc préférable aux autres estimateurs. Ces résultats sont surtout valables asymptotiquement et lorsque l'on fait une simulation de Monte Carlo avec des échantillons finis, cet estimateur peut tout de même conduire à des résultats inacceptables.

Il peut être démontré qu'on obtient les mêmes résultats en utilisant l'instrument $Z_2 = (x_{t+1} \ x_{t-1})$ et l'estimateur des variables instrumentales résultant pourrait être utilisé à l'occasion comme solution de rechange.

Füller (1977) a proposé un estimateur qui est censé conduire à des résultats meilleurs que la méthode originale des variables instrumentales même pour les échantillons finis. C'est surtout cet estimateur que nous allons tester dans le cadre de ce travail. Les autres estimateurs seront utilisés comme points de comparaison.

C- Estimateur de Füller

La méthode utilisée est une adaptation de celle de Füller développée dans la première partie. La différence par rapport à la méthode de Füller réside dans le fait qu'il n'y a pas de constante dans notre modèle. Elle se présente comme suit avec les mêmes notations:

$$(1) \quad y_t = \beta x_t + \gamma y_{t-1} + \eta \quad , \text{ les } y_t \text{ et } x_t \text{ sont des vecteurs de dimension égale à un.}$$

En posant $\theta = (\beta, \gamma)'$

et $X = (x \ y_{-1})$

(1) peut s'écrire alors sous forme matricielle:

$$Y = X\theta + \eta$$

$$(2) \quad \hat{\theta} = \left[(\hat{X}'\hat{X}) - (\bar{v} - \gamma)f_{22} \right]^{-1} \left[(\hat{X}'\hat{Y}) - (\bar{v} - \gamma)f_{21} \right],$$

¹ Comme la plupart des estimateurs de variables instrumentales, les moments dans les petits échantillons ne sont pas toujours définis; cet estimateur peut donc conduire à des résultats très erronés lors d'une simulation de Monte Carlo pour un échantillon de taille avoisinant 150 observations.

où γ est une constante supposée égale à 1 et les autres éléments sont définis comme suit:

$$(3) \quad (\hat{Y}, \hat{X}) = Z_{+2}(Z'_{+2}Z_{+2})^{-1}Z'_{+2}(Y, X)$$

où $Z_{+2} = (x_{t+1}, x_{t+2})$ et x_{t+1} et x_{t+2} sont des vecteurs colonnes commençant successivement par : $t = 2$ à $T+1$ et $t = 3$ à $T+2$ en supposant ici que ces observations sont disponibles.

\bar{v} est la plus petite valeur propre telle que:

$$(4) \quad |(\hat{Y}, \hat{X})'(\hat{Y}, \hat{X}) - v f| = 0$$

où f est une matrice (2 x 2) définie comme suit:

$$(5) \quad f = [(Y, X)'(Y, X) - (Y, X)'Z_{+2}(Z'_{+2}Z_{+2})^{-1}Z'_{+2}(Y, X)] / (N - q)$$

où q correspond au nombre de colonnes de Z_{+2} , soit 2.

L'intérêt d'utiliser la méthode de Fuller en lieu et place de la méthode des variables instrumentales classique s'explique par le fait que cette dernière n'a pas toujours des moments d'ordre fini. La simulation de Monte Carlo avec la méthode des variables instrumentales classique peut donc conduire à des résultats erronés.

L'avantage d'utiliser la méthode de Fuller est d'annuler cet effet avec l'utilisation de l'équation (2). En outre, la méthode de Fuller étant avant tout un estimateur des variables instrumentales utilisant l'instrument $Z = (x_{t+1}, x_{t+2})$, elle produit aussi un estimateur convergent.

III- SIMULATION DE MONTE CARLO

A- Description

La simulation de Monte Carlo a pour objectif de tester la performance de l'estimateur proposé en ce qui concerne les effets dans les échantillons de taille moyenne.

Le principe consiste à fixer d'avance les valeurs des paramètres à estimer et cela dans le respect des hypothèses du modèle. Ceci étant, les données sont générées pour être utilisées afin de calculer la statistique demandée. Cette dernière est considérée acceptable lorsque les valeurs calculées des paramètres convergent vers leurs vraies valeurs à mesure que la taille de l'échantillon grandit.

En ce qui nous concerne, les données utilisées ont été générées de manière aléatoire par ordinateur grâce au logiciel MATLAB. Cette étude reste donc théorique dans un certain sens mais il va sans dire que les conclusions obtenues peuvent aussi bien s'appliquer à des données réelles. Le fait d'utiliser des données théoriques n'enlève rien au caractère pratique de notre expérience. L'hypothèse de la loi normale utilisée lors de la simulation reste tout de même une hypothèse réaliste. Il est couramment admis, par application du théorème limite central, que bon nombre de données économiques suivent une loi normale conditionnelle. Ou encore, le fait de poser que la variable explicative X est autocorrélée mais stationnaire et de la générer telle quelle, ou de faire l'hypothèse que les erreurs de mesure ne sont pas autocorrélées ni corrélées avec d'autres erreurs de mesure, sont autant de situations qui se rencontrent en pratique. Pour n'expliciter que le dernier exemple, il n'y a souvent pas de raison que les erreurs de mesure soient corrélées entre elles ou avec d'autres erreurs de mesure. En effet, lors de la valorisation d'une variable économique, l'erreur de mesure peut provenir entre autres du degré de précision de l'appareil utilisé, de l'habileté de l'utilisateur, de son expérience, de son appréciation de la lecture du résultat et bien d'autres facteurs subjectifs etc... Non seulement le résultat peut varier d'une personne à l'autre mais il peut également varier d'un pays à l'autre selon son avancée technologique. Ce sont souvent des facteurs exogènes qui n'ont en général aucun lien entre eux qui sont responsables des erreurs sur les variables.

Les étapes de la simulation consistent à :

- fixer d'avance les vraies valeurs des paramètres compatibles avec les hypothèses du modèle.

- fixer un niveau de R^2 soit par exemple $R^2 = 0.75$ où R^2 est le coefficient de détermination qui traduit la part de la régression qui est expliquée par les variables indépendantes. Un $R^2 = 0.75$ est une valeur très représentative. Elle est couramment rencontrée dans les modèles classiques de régression.

$$R^2 = 1 - \frac{P \lim(\epsilon' \epsilon) / N}{P \lim(\check{Y}' \check{Y})} \quad \text{avec } \check{Y} = \tilde{Y}_t - \rho \tilde{Y}_{t-1}$$

Le calcul de la variance de ϵ se déduit de la formule ci-dessus en notant bien que :

$$\text{var}(\epsilon) = p \lim(\epsilon' \epsilon) / N \quad (\text{voir annexe 2})$$

La seconde étape consiste à calculer la variance de u_t à partir de celle de ϵ d'après la relation :

$$\text{var}(u_t) = (1 / (1 - \rho^2)) \cdot \text{var}(\epsilon)$$

La variance de y_t quant à elle se calcule aisément dès lors qu'il a été posé que celle de x_t est connue et égale à un (voir annexe 1).

Dès lors que les variances ont été calculées, on tire toutes les conditions initiales des variables à savoir :

- Tirer ϵ_0 à partir d'une normale de moyenne zéro et ayant une variance tirée de la relation ci-dessus. u_0 , y_0 et x_0 seront tirés de manière analogue.

Les autres observations des variables sont obtenues par itérations successives en tenant compte des relations qui existent entre leurs valeurs passées et présentes.

Les x_t étant corrélés entre eux, il est par exemple possible de générer les autres observations dès lors que la valeur initiale est connue et cela en faisant une boucle.

- Des erreurs de mesure de variance égale à 20% des variances des vraies variables sont ensuite ajoutées aux valeurs de x_t et de y_t pour former les variables observables sur lesquelles sont appliquées les formules d'estimateurs.

Quatre méthodes d'estimation retiendront notre attention:

- 1) Les moindres carrés ordinaires (MCO).
- 2) La méthode de Fuller utilisant comme variables instrumentales $Z = (x_{t+1}, x_{t+2})$
- 3) La méthode des variables instrumentales classique associée à Z
- 4) La méthode des variables instrumentales utilisant comme instruments $Z_1 = (x_t, x_{t-1})$

Il est également envisagé de développer une méthode de Fuller utilisant comme instrument $Z_2 = (x_{t+1}, x_{t-1})$. Nous l'appellerons méthode de Fuller modifiée et cela en raison de la convergence très lente des autres estimateurs ci - dessus.

Sauf mention expresse, toutes les expériences de Monte Carlo de cette étude reposent sur les hypothèses suivantes:

- Psi = 0.9
- Rho = 0.7
- Bêta = 1
- Gamma = 0.8

A l'occasion, la taille de l'échantillon sera progressivement accrue pouvant passer de 150 à plus de 1000 observations jusqu'à obtenir une convergence satisfaisante de l'estimateur et cela dans les limites de la capacité de l'ordinateur utilisé.

L'analyse finale retiendra comme résultat:

- La moyenne des estimateurs basée sur le nombre de répétitions,
- le biais des estimateurs et leurs variances et enfin ,
- la racine carrée de l'erreur quadratique moyenne (REQM).

On retient souvent cette dernière mesure comme critère fondamental pour apprécier un estimateur. C'est donc elle qui nous permettra de dire si un estimateur est préférable à un autre. Par contre, lorsque les situations l'exigeront, nous analyserons également la variance et le biais de manière distincte afin de relativiser certaines conclusions.

B- Analyse des résultats

L'approche que nous utiliserons consiste à tester les estimateurs concernés dans des cas où les hypothèses retenues permettent la convergence de l'un ou l'autre de ceux-ci. Ainsi, on s'intéressera beaucoup plus aux MCO dès lors que nous supposerons l'absence d'erreur sur les variables et d'autocorrelation. L'objectif de cette approche est d'avoir un point de repère pour évaluer la performance des autres estimateurs dans les mêmes conditions. Nous pourrions alors apprécier la taille de l'échantillon qu'il faut pour réaliser la convergence. De cette façon, on peut voir s'il est réaliste de disposer d'un nombre d'observations aussi élevé.

Cette approche nous donnera aussi l'occasion de faire une étude comparative de la vitesse de convergence des différents estimateurs.

Dans les différents tableaux qui seront soumis à notre appréciation, les notations suivantes seront retenues:

- 1) **mcomoy** : moyenne des estimations des moindres carrés ordinaires.
- 2) **estmoy** : moyenne des estimations de Fuller utilisant comme variables instrumentales $Z=(x_{t+1}, x_{t+2})$.
- 3) **estZmoy** : moyenne des estimations avec la méthode classique de variables instrumentales utilisant Z comme instruments.
- 4) **estZ1moy** : moyenne des estimations avec la méthode des variables instrumentales avec $Z1=(x_t, x_{t-1})$.

Les données entre parenthèses et placées à droite des moyennes des estimations représentent la racine carrée de l'erreur quadratique moyenne (REQM).

Rappelons à nouveau notre modèle

$$\bar{y}_t = \beta \bar{x}_t + \gamma \bar{y}_{t-1} + u_t$$

$$\text{avec } u_t = \rho u_{t-1} + \varepsilon_t$$

Cas 1: $\rho = 0$, $\lambda^1 = 0$, $\psi = 0.9$ et $R^2 = 0.75$

Comme démontré dans les résultats analytiques plus haut, les MCO sont recommandés dans ce cas car ils sont alors convergents. C'est donc les MCO qui serviront de point de comparaison, surtout en terme de vitesse de convergence. L'expérience dont nous relaterons certains résultats a concerné successivement un échantillon de 150, 250, 350, 500 puis 1000 observations.

Échantillon de 150 observations répétées 500 fois

	bêta = 1	gamma = 0.8
mcomoy*	1.0740 (0.3706)	0.7832 (0.0530)
estmoy	1.0557 (0.9236)	0.7833 (0.2436)
estZmoy	0.6876 (42.9681)	0.2654 (26.2530)
estZImoy	1.0158 (1.0824)	0.8186 (0.3687)

A 150 observations, les MCO donnent des résultats nettement meilleurs que les autres estimateurs en terme de REQ. Cela corrobore les résultats analytiques qui conseillent d'utiliser les MCO dans ce cas. On peut quand même s'étonner des résultats de estZmoy qui devraient être presque aussi bons. Ce mauvais résultat s'expliquerait par l'inexistence des moments. En effet, plus on utilise des variables instrumentales x avancés ou retardés significativement telles que (x_{t+1}, x_{t+2}) , plus le nécessaire lien qui doit exister entre celles-ci et les variables explicatives n'est plus évident. $p \lim(\sum ZX/T)$ n'est pas définie et cela d'autant plus que T tend vers l'infini puisque le numérateur va tendre vers zéro.

NB: Ici les X représentent toutes les variables explicatives et Z les variables instrumentales.

une augmentation de l'échantillon à 500 observations donnent les résultats suivants:

¹ mesure le ratio de la variance des erreurs de mesure par rapport à la variance de la vraie variable concernée; $\lambda = 0$ par exemple, traduit l'absence d'erreur de mesure sur la variable

Échantillon de 500 observations répétées 500 fois

bêta = 1

gamma = 0.8

mcomoy*	1.0171 (0.1801)	0.7951 (0.0256)
estmoy	0.9545 (0.7907)	0.8122 (0.2266)
estZmoy	-0.6241 (16.051)	1.2939 (4.4928)
estZ1moy	1.0134 (0.4003)	0.7968 (0.1203)

Le REQM des MCO s'est considérablement amélioré pour la moyenne des estimations. Bien que le biais soit différent de zéro, on peut parler de convergence satisfaisante avec 500 observations. EstZ1moy converge aussi rapidement même s'il reste moins performant que les MCO (son REQM supérieur à celui des MCO).

Échantillon de 1000 observations 100 fois

bêta = 1

gamma = 0.8

mcomoy*	1.0020 (0.1328)	0.7960 (0.0186)
estmoy	1.0879 (0.8302)	0.7814 (0.2539)
estZmoy	4.0404 (26.9874)	0.0257 (7.6455)
estZ1moy	0.9975 (0.2812)	0.7980 (0.0771)

Ces résultats nous permettent de conclure que les MCO sont l'estimateur à utiliser en présence d'un modèle de séries chronologiques comme le nôtre et respectant les hypothèses énoncées dans ce cas. Il s'agit d'un estimateur convergent comme le confirment les résultats analytiques et la simulation de Monte Carlo.

Cette conclusion nous amène au cas suivant dont les hypothèses se résument comme suit:

Cas 2: $\rho = 0.7$, $\lambda = 0$, $\psi = 0.9$ et $R^2 = 0.75$

Il a été démontré dans les résultats analytiques que $estZ1moy$ est l'estimateur recommandé dans ce cas. Comme précédemment, nous avons analysé les résultats des expériences pour 150, 250, 350, 500 puis 1000 observations.

Certains résultats ont été rapportés ci-dessous avec les mêmes notations:

Échantillon de 150 observations répétées 500 fois

	bêta = 1	gamma = 0.8
mcomoy	0.6250 (0.4831)	0.9371 (0.1391)
estmoy	0.6078 (0.7246)	0.9422 (0.1872)
estZmoy	3.0409 (46.9993)	-0.0870 (20.791)
estZ1moy*	1.6521 (12.4579)	0.6105 (4.2542)

A ce stade, les MCO restent préférables à l'estimateur qui est recommandé à savoir $estZ1moy$ car les MCO ont des REQM plus bas. Un constat général montre d'ailleurs que les deux estimateurs de variables instrumentales classiques donnent des résultats inacceptables. On expérimente encore probablement l'anomalie signalée plus haut en ce qui concerne l'inexistence des moments.

Pour revenir à notre démarche énoncée en début de chapitre, nous avons progressivement agrandi la taille de notre échantillon pour connaître à quel moment $estZ1moy$ pourrait converger.

A 350 observations, $estZ1moy$ s'améliore significativement comme en témoigne le tableau ci-dessous:

Échantillon de 350 observations répétées 500 fois.

	bêta = 1	gamma = 0.8
mcomoy	0.5636 (0.4792)	0.9437 (0.1444)
estmoy	0.5451 (0.6422)	0.9476 (0.1833)
estZmoy	-1.027 (32.0741)	1.4239 (8.1860)
estZ1moy*	1.0301 (0.3551)	0.7919 (0.2337)

L'estimation de gamma a un "REQM" beaucoup plus élevé que celui des MCO; la convergence peut être qualifiée ici de relative.

Par contre, lorsque l'échantillon passe à 500, la convergence de estZ1moy s'améliore nettement comme le montre le tableau suivant:

Échantillon de 500 observations répété 500 fois.

	bêta = 1	gamma = 0.8
mcomoy	0.5517 (0.4735)	0.9434 (0.1438)
estmoy	0.5338 (0.6232)	0.9502 (0.1841)
estZmoy	8.6523 (159.695)	-1.1595(40.1136)
estZ1moy*	1.0340 (0.2712)	0.7880 (0.0895)

A ce stade, il apparaît clairement que estZ1moy est de loin préférable aux autres estimateurs. Il a le plus bas REQM. Mais ce résultat est encore perfectible car lorsque la taille de l'échantillon passe de 500 à 1000 observations, le REQM passe de 0.2712 à 0.1879 pour bêta et de 0.0895 à 0.054 pour gamma. Les estimations moyennes quant à elles restent sensiblement identiques ce qui montre qu'on tend vers une variance plus petite quand la taille de l'échantillon augmente, améliorant ainsi l'estimateur estZ1moy. Le REQM des autres estimateurs reste très proche de leur valeur à 500 observations même si la taille de l'échantillon passe à 1000.

Ces résultats suggèrent que la notion de convergence demeure une notion assez relative car ce qui peut être considéré asymptotique pour telle méthode peut être différent pour telle autre. Il serait important que la vitesse de convergence soit mentionnée explicitement dans la littérature. Cette notion nous semble beaucoup plus importante, car même si un estimateur est convergent, s'il requiert un nombre d'observations très élevé, il peut ne pas être applicable à des situations économiques réelles, la taille des échantillons disponibles étant trop restreinte.

D'ailleurs, pour finir cette étude, nous nous proposons de faire une analyse comparative de la vitesse de convergence des estimateurs mais avant, abordons le cas suivant:

Cas 3: $\rho = 0.7$, $\lambda = 0.2$, $\psi = 0.9$ et $R^2 = 0.75$

En se basant sur les résultats asymptotiques, c'est $estZ_{moy}$ qui pourrait être recommandé dans ce cas. Mais nous nous intéresserons plus à l'estimateur de Fuller pour pallier à l'effet néfaste de la méthode classique de variables instrumentales pour les échantillons limités.

Échantillon de 150 observations 500 fois.

	bêta = 1	gamma = 0.8
mcomoy	0.9612 (0.5103)	0.7444 (0.0911)
estmoy*	1.3686 (1.2694)	0.6778 (0.2779)
estZmoy	0.4542 (59.6816)	-0.0269(22.0314)
estZ1moy	0.3006 (1.8829)	1.0391 (1.2651)

En terme de REQ, les MCO restent les estimateurs à choisir. L'estimateur de Fuller bien qu'ayant un niveau de REQ plus bas que celui des autres estimateurs de variables instrumentales est moins performant que les MCO.

Même à 500 observations, la convergence de l'estimateur de Fuller n'est pas apparente, bien que l'on note une légère amélioration.

Échantillon de 500 observations 500 fois

	bêta = 1	gamma = 0.8
mcomoy	0.9090 (0.2689)	0.7683 (0.0478)
estmoy*	1.2819 (0.9712)	0.7144 (0.2662)
estZmoy	-2.7438(85.2475)	2.0103 (30.2507)
estZ1moy	0.4377 (0.8035)	0.9598 (0.3430)

Estmoy est un estimateur stable mais qui semble converger très lentement. On se propose, comme les précédents estimateurs, de porter la taille de l'échantillon à 1000 observations.

Échantillon de 1000 observations 100 fois

	bêta = 1	gamma = 0.8
mcomoy	0.9036(0.2076)	0.7706 (0.0383)
estmoy*	1.3117 (0.8964)	0.7051 (0.2568)
estZmoy	0.5065 (14.4855)	1.0263 (5.2444)
estZ1moy	0.5065 (0.6122)	0.9260 (0.1768)

Même une taille de l'échantillon de 1000 observations n'améliore pas le résultat de estmoy. Il s'agit d'un estimateur qui semble converger très lentement.

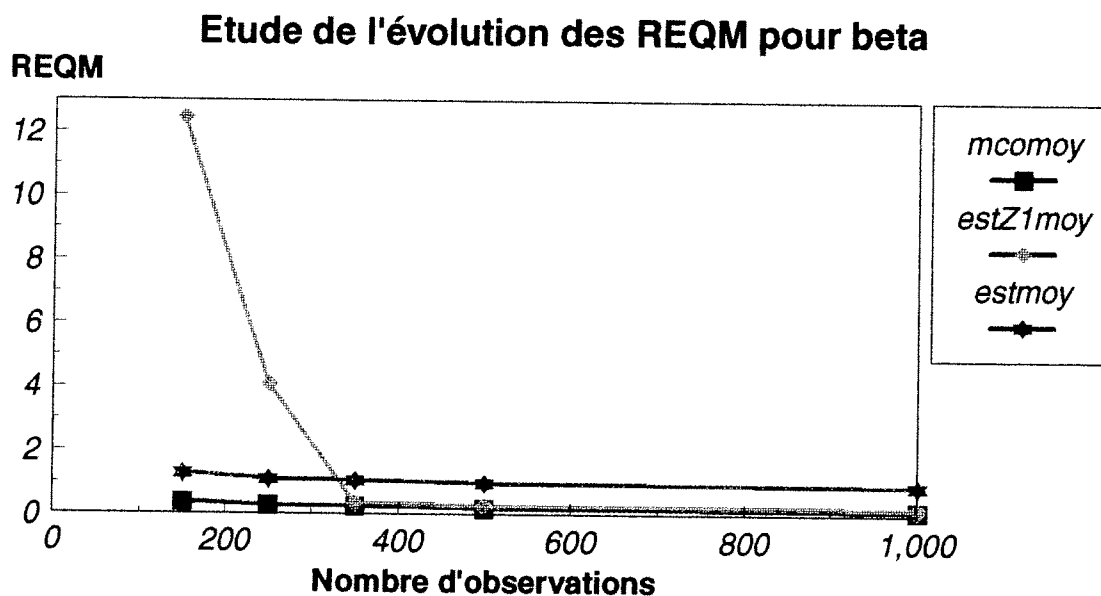
Nous avons décidé de faire une étude comparative de la vitesse de convergence des estimateurs selon les cas où ils sont recommandés. Le but de cette étude est de mieux apprécier la rapidité de convergence des différents estimateurs. On retient ici le REQM en lieu et place du biais, comme critère pour apprécier la convergence de l'estimateur. S'il en est ainsi, c'est parce que le REQM est un critère de choix plus global car il intègre et le biais et la variance.

Tableau comparatif¹ des REQM des estimations de bêta.

Nbre d'observ.	150	250	350	500	1 000
mcomoy(cas1)	0,3706	0,2788	0,248	0,1801	0,1328
estZ1moy(cas2)	12,4579	4,0699	0,3551	0,2712	0,1989
estmoy(cas3)	1,2694	1,0883	1,0487	0,9712	0,8964

Le graphique suivant illustre à partir de l'analyse de l'évolution des REQM, la vitesse de convergence des différents estimateurs de bêta qui ont été testés.

¹ Les résultats des mco dans le tableau croisé sont ceux du cas1, ceux de estZ1moy sont du cas2... Le but d'une telle démarche est d'apprécier la performance absolue des différents estimateurs dans les conditions requises pour leur application.



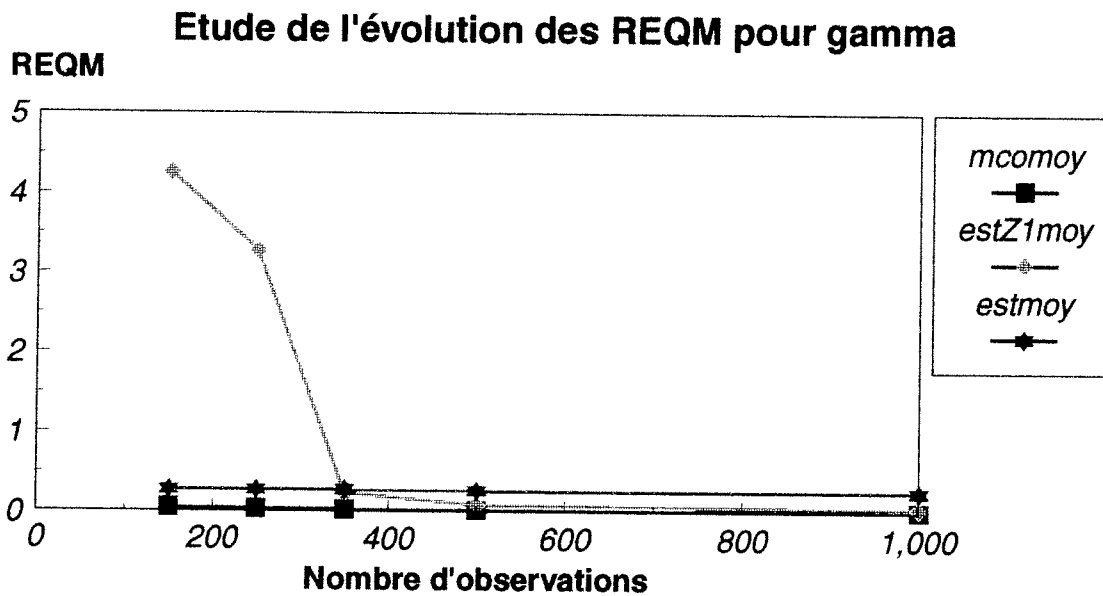
sources: extrait des résultats du programme matlab

A l'analyse du tracé de ces courbes, le REQM de estZ1moy est beaucoup plus élevé que celui de tout autre estimateur. Mais celui-ci s'améliore davantage lorsque la taille de l'échantillon atteint 350 observations, dépassant même celui de estmoy. Ce dernier a un REQM petit mais presque constant sur toute la ligne si bien qu'il ne pourrait converger qu'à une taille d'échantillon beaucoup trop élevée. Cela rend inapplicable un tel estimateur à la réalité économique étant donné que les données de séries chronologiques utilisées sont relativement de taille petite. Pour finir, on pourrait faire remarquer que les MCO sont les estimateurs qui ont tendance à converger plus rapidement lorsque les conditions qui assurent la convergence sont satisfaites.

Tableau comparatif des REQM des estimations de gamma.

Nbre d'observ.	150	250	350	500	1 000
mcomoy(cas1)	0,053	0,0391	0,0317	0,0256	0,0186
estZ1moy(cas2)	4,2542	3,275	0,2337	0,0895	0,0512
estmoy(cas3)	0,2779	0,2746	0,2793	0,2662	0,2568

De la même façon, voyons l'analyse de l'évolution des REQM pour les estimateurs de gamma:



L'étude de l'évolution des REQM de gamma ne fait que confirmer les conclusions que nous avons tirées pour bêta. On retient définitivement que l'estimateur de Fuller appliqué à notre modèle lorsqu'il y a des erreurs de mesure sur les variables, donne des résultats loin d'être convaincants. Comme solution alternative, nous proposons ce que nous sommes convenus d'appeler la méthode de Fuller alternative. C'est la technique de Fuller qui utilise comme variables instrumentales $Z_2 = (x_{t+1}, x_{t-1})$. Cette expérience a été réalisée successivement avec les mêmes tailles d'échantillon à savoir 150, 250, 350, 500 et 1000 observations.

Les mêmes notations que celles retenues lors des tableaux ci-dessus seront conservées à l'exception de estmoy qui devient estmoy2. Les hypothèses retenues sont ceux du cas 3 à savoir:

$\rho = 0.7$, $\lambda = 0.20$, $\psi = 0.9$ et $R^2 = 0.75$

La première expérience couvre un échantillon de 150 observations avec 500 répétitions.

	bêta = 1	gamma = 0.8
mcomoy	0.9612 (0.5103)	0.7444 (0.0911)
estmoy2*	1.2557 (1.1455)	0.7609 (0.2608)
estZ2moy	2.6253 (16.2697)	0.4382 (5.7754)
estZ1moy	0.3006 (1.8829)	1.0391 (1.2651)

Devant ce premier résultat, l'estimateur que nous recommanderons est encore celui des MCO. C'est bien celui qui a le niveau de REQM le plus bas.

En analysant le REQM de estZ2moy, il ressort qu'il est à peu près 10 fois plus élevé que celui de l'estimateur de Fuller alternatif. On peut déjà conclure que ce dernier est au moins meilleur que l'estimateur classique des variables instrumentales qui lui est associé.

Nous agrandissons la taille de notre échantillon à 500 puis à 1000 observations. Regardons les tableaux suivants :

Échantillon de 500 observations 500 fois

bêta = 1

gamma = 0.8

mcomoy	0.9090 (0.2689)	0.7683 (0.0478)
estmoy2*	1.1235 (0.8822)	0.7760 (0.2299)
estZ2moy	0.2042 (24.3404)	1.0450 (6.9825)
estZ1moy	0.4377 (0.8035)	0.9598 (0.3430)

A 500 observations, estmoy2 tend à s'améliorer pour converger vers sa vraie valeur. Tandis que les MCO s'écartent davantage de la vraie valeur du paramètre.

Échantillon de 1000 observations pour 500 répétitions

bêta = 1

gamma = 0.8

mcomoy	0.9116 (0.2097)	0.7719 (0.0377)
estmoy2*	1.0288 (0.7146)	0.7970 (0.1791)
estZ2moy	0.9863 (1.0095)	0.8115 (0.2722)
estZ1moy	0.5082 (0.6250)	0.9249 (0.1755)

Il ressort que estmoy2 converge de façon satisfaisante pour un échantillon de 1000 observations. Mais si on retient comme seul critère le "REQM", les MCO restent encore préférables. Par contre si l'analyse tient compte seulement du biais, estmoy2 devient meilleur. Il est connu que les MCO

sont des estimateurs qui ont une petite variance même lorsqu'ils sont biaisés , si bien que tout estimateur même nettement moins biaisé aura difficilement un REQM inférieur à celui des MCO. Nous avons accru l'échantillon à la taille maximale permise par la capacité de l'ordinateur, soit 4000 observations 500 fois. Les résultats sont les suivants :

Échantillon de 4000 observations pour 500 fois

	bêta = 1	gamma = 1
mcomoy	0.8971 (0.1397)	0.7755 (0.0272)
estmoy2*	1.0132 (0.3999)	0.7979 (0.1062)
estZ2moy	1.0088 (0.4187)	0.7992 (0.1116)
estZ1moy	0.4797 (0.5599)	0.9311 (0.1467)

Bien que les REQM de l'estimateur proposé (estmoytd) se soient largement améliorés, ils restent encore inférieurs à ceux des MCO. Mais la théorie nous enseigne qu'il arrivera, si la taille de l'échantillon peut s'accroître davantage ,que les REQM des estmoytd deviendront meilleurs que ceux des mcomoy; puisqu'il a été démontré dans les chapitres précédents que les MCO appliqués à ce modèle restent biaisés même asymptotiquement. Autrement dit, le REQM des MCO sera toujours différent de zéro aussi grande que soit la taille de l'échantillon. Par contre, d'après la théorie développée plus tôt, estmoytd est non biaisé asymptotiquement. La question qui se pose est évidemment de savoir à quelle taille d'échantillon peut on parler de phénomène asymptotique. Et quel est son intérêt si l'asymptoticité est réalisée à une taille d'échantillon très grande pouvant par exemple avoisiner 50000 observations.

Un tel estimateur présente peu d'intérêt pratique car dans des situations réelles, il est rare que l'on rencontre des échantillons atteignant plus de 5000 observations, et ce, même dans des séries chronologiques utilisant des données journalières. Est ce suffisant pour recommander les MCO lorsqu'il y a des erreurs de mesure quand on sait que l'estimateur n'est pas convergent.

Si on prend comme critère les REQM, aucun des estimateurs étudiés ne semble préférable. Cette étude, loin d'épuiser le sujet sur l'estimation des modèles à erreurs sur les variables, pourrait servir de base à d'autres investigations futures.

CONCLUSION

Nous avons passé en revue différents estimateurs des paramètres d'un modèle autorégressif simple selon qu'il y ait ou non autocorrelation du terme d'erreur et erreurs de mesure sur les variables. Les conclusions de cette étude recommandent les MCO en l'absence d'autocorrélation et d'erreur de mesure, l'estimateur de variables instrumentales $Z_1 = (x_t, x_{t-1})$ lorsqu'il y a seulement autocorrélation. En ce qui concerne la convergence des estimateurs dans les cas où chacun d'eux est indiqué lorsqu'on s'appuie sur la théorie asymptotique, les MCO restent l'estimateur qui converge le plus rapidement. La convergence de l'estimateur de Füller et de la méthode classique des variables instrumentales utilisant comme instrument Z_1 est très mitigée. L'estimateur de Füller est beaucoup plus stable mais converge beaucoup plus lentement. L'inexistence des moments lorsqu'on utilise comme instruments les x retardés ou avancés significativement rend inacceptables certaines méthodes classiques de variables instrumentales .

ANNEXE 1**Variance de Y**

$$\begin{aligned}\tilde{Y}_t &= \beta\tilde{X}_t + \gamma\tilde{Y}_{t-1} + u_t \\ \tilde{X}_t &= \xi\tilde{X}_{t-1} + \omega_t \quad ; \quad u_t = \rho u_{t-1} + \varepsilon_t\end{aligned}$$

$$(1) \quad E(\tilde{Y}_t) = \beta E(\tilde{X}_t) + \gamma E(\tilde{Y}_{t-1}) + E(u_t)$$

$$\begin{aligned}E(\tilde{X}_t) &= 0 \quad ; \quad E(u_t) = 0 \\ E(\tilde{Y}_t) &= E(\tilde{Y}_{t-1})\end{aligned}$$

$$\Rightarrow E(\tilde{Y}_t)(1 - \gamma) = 0 \quad \Leftrightarrow \quad E(\tilde{Y}_t) = 0$$

$$\begin{aligned}(2) \quad \text{Var}(\tilde{Y}_t) &= E(\tilde{Y}_t)^2 \\ &= E(\beta\tilde{X}_t + \gamma\tilde{Y}_{t-1} + u_t)^2 \\ &= E(\beta^2\tilde{X}_t^2 + \gamma^2\tilde{Y}_{t-1}^2 + u_t^2 + 2\beta\gamma\tilde{X}_t\tilde{Y}_{t-1} + 2\beta\tilde{X}_tu_t + 2\gamma\tilde{Y}_{t-1}u_t)\end{aligned}$$

Sachant que:

$$E(\tilde{X}_t\tilde{Y}_t) = \beta V(\tilde{X}) / (1 - \gamma\xi)$$

$$\text{et } E(\tilde{X}_tu_t) = 0$$

$$\Rightarrow \text{Var}(\tilde{Y}_t) = \beta^2 V(\tilde{X}) + \gamma^2 V(\tilde{Y}) + V(u_t) + 2\beta^2\gamma\xi V(\tilde{X}) / (1 - \gamma\xi) + 2\gamma\rho E(\tilde{Y}_t u_t) \quad (a)$$

le terme inconnu est $E(\tilde{Y}_t u_t)$

$$E(\tilde{Y}_t u_t) = E(\beta\tilde{X}_t u_t + \gamma\tilde{Y}_{t-1} u_t + u_t^2)$$

$$\Rightarrow E(\tilde{Y}_t u_t) = \gamma\rho E(\tilde{Y}_t u_t) + V(u_t)$$

$$\text{d'où } E(\tilde{Y}_t u_t) = V(u_t) / (1 - \gamma\rho)$$

Finalement :

$$V(\tilde{Y}_t) = (\beta^2 V(\tilde{X}) + V(u) + \frac{2\beta^2\gamma\xi}{1-\gamma\xi} V(\tilde{X}) + \frac{2\gamma\rho}{1-\gamma\rho} V(u)) / (1 - \gamma^2)$$

En simplifiant, on obtient:

$$V(\tilde{Y}_t) = \beta^2 \frac{(1+\gamma\xi)}{(1-\gamma\xi)(1-\gamma^2)} V(\tilde{X}_t) + \frac{(1+\gamma\rho)}{(1-\gamma^2)(1-\gamma\rho)} V(u_t)$$

ANNEXE 2

Variance de epsilon:

$$R_{\text{carré}} = 0.75 ; \quad \text{Bêta} = 1 ; \quad \text{gamma} = 0.8 ; \quad \text{psi} = 0.9 ; \quad \text{var}(x_t) = 1$$

$$\tilde{y}_t = \beta \tilde{x}_t + \gamma \tilde{y}_{t-1} + u_t$$

$$\tilde{y}_t - \rho \tilde{y}_{t-1} = \beta(\tilde{x}_t - \rho \tilde{x}_{t-1}) + \gamma(\tilde{y}_{t-1} - \rho \tilde{y}_{t-2}) + \varepsilon_t$$

$$\Leftrightarrow \tilde{y}_t^* = \beta \tilde{x}_t^* + \gamma \tilde{y}_{t-1}^* + \varepsilon_t$$

$$(1) \quad R^2 = 1 - \frac{p \lim(\varepsilon' \varepsilon)/N}{p \lim(\tilde{y}^{*'} \tilde{y}^*)/N}$$

Tout en remarquant que $p \lim(\varepsilon' \varepsilon)/N = \text{var}(\varepsilon)$

Calculons:

$$p \lim \left(\frac{\tilde{y}^{*'} \tilde{y}^*}{N} \right) = \beta^2 p \lim(\tilde{x}_t^{*'} \tilde{x}_t^*/N) + \gamma^2 p \lim(\tilde{y}_{t-1}^{*'} \tilde{y}_{t-1}^*/N) + 2\beta\gamma p \lim(\tilde{x}_t^{*'} \tilde{y}_{t-1}^*/N) + V(\varepsilon)$$

$$\Rightarrow p \lim \left(\frac{\tilde{y}^{*'} \tilde{y}^*}{N} \right) (1 - \gamma^2) = \beta^2 p \lim(\tilde{x}_t^{*'} \tilde{x}_t^*/N) + 2\beta\gamma p \lim(\tilde{x}_t^{*'} \tilde{y}_{t-1}^*/N) + V(\varepsilon)$$

$$\text{or } p \lim(\tilde{x}_t^{*'} \tilde{x}_t^*/N) = E\left(\frac{1}{N} \sum (\tilde{x}_t - \rho \tilde{x}_{t-1})^2\right)$$

$$= E\left(\frac{1}{N} \sum (\tilde{x}_t^2 + \rho^2 \tilde{x}_{t-1}^2 - 2\rho \tilde{x}_t \tilde{x}_{t-1})\right)$$

$$= (1 - 2\rho\xi + \rho^2) \text{Var}(\tilde{x}_t)$$

$$p \lim(\tilde{x}_t^{*'} \tilde{y}_{t-1}^*/N) = \beta p \lim(\tilde{x}_t^{*'} \tilde{x}_{t-1}^*/N) + \gamma p \lim(\tilde{x}_t^{*'} \tilde{y}_{t-2}^*/N) p \lim(\tilde{x}_t^{*'} \varepsilon_{t-1}/N)$$

$$= \beta\xi p \lim(\tilde{x}_{t-1}^{*'} \tilde{x}_{t-1}^*/N) + \gamma\xi p \lim(\tilde{x}_{t-1}^{*'} \tilde{y}_{t-2}^*/N)$$

$$= \beta\xi(1 - 2\rho\xi + \rho^2) \text{Var}(\tilde{x}_t) + \gamma\xi p \lim(\tilde{x}_t^{*'} \tilde{y}_{t-1}^*/N)$$

$$(2) \quad \Rightarrow p \lim(\tilde{x}_t^{*'} \tilde{y}_{t-1}^*/N) = [\beta\xi(1 - 2\rho\xi + \rho^2)/(1 - \gamma\xi)] \text{Var}(\tilde{x}_t)$$

$$\Rightarrow p \lim \left(\frac{\tilde{y}^{*'} \tilde{y}^*}{N} \right) = \left(\frac{1}{1 - \gamma^2} \right) [(1 - 2\rho\xi + \rho^2)(\beta^2 + (2\beta^2\gamma\xi)/(1 - \gamma\xi)) \text{Var}(\tilde{x}_t) + \text{Var}(\varepsilon)]$$

En simplifiant l'expression ci-dessus, on a:

$$p \lim \left(\frac{\tilde{y}^{*'} \tilde{y}^*}{N} \right) = \left(\frac{1}{1 - \gamma^2} \right) [\beta^2(1 - 2\rho\xi + \rho^2)(1 + \gamma\xi)/(1 - \gamma\xi) \text{Var}(\tilde{x}_t) + \text{Var}(\varepsilon)]$$

Et d'après (1), on en déduit la valeur de $\text{Var}(\varepsilon)$:

$$\text{Var}(\varepsilon) = \beta^2 ((1 + \gamma\xi)/(1 - \gamma\xi)) ((1 - R^2)/(R^2 - \gamma^2)) (1 - 2\rho\xi + \rho^2) \text{Var}(\tilde{x}_t)$$

ANNEXE 3:**Algorithme du programme informatique pour un échantillon de T observations m fois.**

Entrée de valeur pour les paramètres :

- $R_{\text{carré}} = 0.75$
- $\rho = 0.7$
- $\psi = 0.9$
- $\beta = 1.00$
- $\gamma = 0.8$
- variance de $x = 1.00$

Procédure qui génère les variables

- Calcul de la variance de epsilon à partir de $R_{\text{carré}}$ et de la variance de x puis de celle de u_t
- calcul de la variance de y_t
- dimensionner au préalable le vecteur des quatre (4) estimateurs pour m répétitions.
- générer u_t de la manière suivante: tirer la première observation de u_t d'une normale de moyenne 0 et de variance égale à celle de u_t . Obtenir les autres observations à partir de leur valeur passée immédiate et par itération successive.
- générer x_t de manière analogue à u_t : tirer sa première observations d'une normale de moyenne 0 et de variance égale à un (1). Obtenir les autres observations à partir de leur valeur passée immédiate et par itérations successives.
- générer y_t de la manière suivante : tirer sa première observation d'une normale de moyenne 0 et de variance égale à celle de y_t . Obtenir les autres observations à partir du modèle et par itération successive.
- ajouter des erreurs de mesure sur les variables de la manière suivante : calculer la variance de chaque erreur est égale à λ fois celle de la variable correspondante. générer les erreurs v_t et s_t d'une normale(0, variance). Additionner les variables v_t et s_t respectivement à x_t et y_t .

Calcul des estimateurs:

- Appliquer la formule usuelle des MCO avec les notations $X = (x_t, y_{t-1})$ et à chaque répétition.

- Calculer l'estimateur de Fuller de la manière suivante : Avec les notations $Z = (x_{t+1}, x_{t+2})$ et $YX = (y_t, x_t, y_{t-1})$, calculer s, \hat{y} et \hat{x} , prendre le minimum des valeurs propres de la matrice indiquée dans la méthode de Fuller (voir C-4). Et enfin appliquer la formule de l'estimateur de Fuller.

- Calculer l'estimateur de variables instrumentales utilisant comme instruments

$$Z = (x_{t+1}, x_{t+2})$$

- calculer l'estimateur de variable instrumentale avec l'instrument $Z_1 = (x_t, x_{t-1})$

- Calculer la moyenne des estimateurs

- calculer leur variance et leur biais

- Calculer la racine carré de l'erreur quadratique moyenne.

ANNEXE 4**% Simulation de Monte Carlo : Programme informatique avec le logiciel matlab.**

```

%
rcarré = 0.75;
beta = 1;
gamma = 0.8;
vraitheta = [ bêta gamma ]
psi = 0.9
rho = 0.7
sigmaX2 = 1;
lambda = 0.20
T = 151;
Nobs = T-1
m = 500;
Nfois = m
%
sigmaEpsil2 = (beta^2)*((1 + gamma * psi) / (1- gamma * psi)) * ((1 - rcarré) / (rcarré -
gamma^2)) * (1 - 2 * rho * psi + rho^2) * sigmaX2
%
% Variance de ut
sigmaU2 = (inv(1 - rho^2)) * sigmaEpsil2
%
% Variance de yt
sigmay2 = ((beta^2) * ((1 + gamma * psi) * inv((1 - gamma * psi) * (1 - gamma^2)))) * sigmaX2
+ (1 + gamma * rho) * inv((1 - gamma^2) * (1 - gamma * rho)) * sigmaU2)
%
% Dimensionner le vecteur des estimateurs
%
% thetaMCO = estimateur des MCO
thetaMCO = zeros(m,2);
% theta = estimateur proposé

```

```

theta = zeros(m,2);
%
% thetaz = 1er estimateur de variables instrumentales z
thetaz = zeros(m,2);
%
% thetaz1 = 2eme estimateur de variables instrumentales z1
thetaz1 = zeros(m,2);
%
% test pour un échantillon de Nobs. observations pour Nfois
q = 2;
gam = 1;
for k = 1:m
% dimensionner  $u_t = \rho * u_{t-1} + \text{Epsilt}$ 
u_t = zeros(T+2, 1);
Epsilt = sqrt(sigmaEpsil2) .* randn(T+2, 1);
%
u_0 = sqrt(sigmaU2) .* randn(1, 1);
u_t(1, 1) = u_0 ;
u_t(2, 1) = rho * u_t(1, 1) + Epsilt(2, 1);
%
% dimensionner  $x_t = \psi * x_{t-1} + w_t$  tel que  $\text{var}(x_t) = 1$ 
x_0 = randn(1, 1);
x_t(1, 1) = x_0 ;
x_t(2, 1) = psi .* x_t(1,1) + w_t(2, 1);
%
% dimensionner  $y_t = \beta * x_t + \gamma * y_{t-1} + u_t$ 
y_0 = sqrt(sigmay2) .* randn(1, 1);
y_t(1, 1) = y_0 ;
y_t(2, 1) = beta .* x_t(2, 1) + gamma .* y_t(1, 1) + u_t(2, 1);
%
for i = 3 : T+2

```

```

% générer  $u_t$  (3)
 $u_t(i, 1) = \rho \cdot u_t(i-1, 1) + \text{Epsilt}(i, 1);$ 
%
% générer  $x_t$  (3)
 $x_t(i, 1) = \psi \cdot x_t(i-1, 1) + w_t(i, 1);$ 
%
% générer  $y_t$  (3)
 $y_t(i, 1) = \beta \cdot x_t(i, 1) + \gamma \cdot y_t(i-1, 1) + u_t(i, 1);$ 
end
%
% En tenant compte des erreurs de mesure sur les variables
%
%  $x_{\text{td}} = x_t + v_t$  tel que  $\text{var}(v_t) = \lambda \times \text{var}(x_t)$ 
 $x_{\text{td}} = \text{zeros}(T+2, 1);$ 
 $\text{sigmav2} = \lambda \cdot \text{sigmaX2};$ 
 $v_t = \text{sqrt}(\text{sigmav2}) \cdot \text{randn}(T+2, 1);$ 
 $x_{\text{td}}(:, 1) = x_t(:, 1) + v_t(:, 1);$ 
%
%  $y_{\text{td}} = y_t + s_t$  tel que  $\text{var}(s_t) = \lambda \times \text{var}(y_t)$ 
 $y_{\text{td}} = \text{zeros}(T+2, 1);$ 
 $\text{sigmas2} = \lambda \cdot \text{sigmay2};$ 
 $s_t = \text{sqrt}(\text{sigmas2}) \cdot \text{randn}(T+2, 1);$ 
 $y_{\text{td}}(:, 1) = y_t(:, 1) + s_t(:, 1);$ 
%
% calcul de l'estimateur MCO
% En posant :
 $X = [x_{\text{td}}(2:T, 1) \ y_{\text{td}}(1:T-1, 1)];$ 
 $Y = y_{\text{td}}(2:T, 1);$ 
 $\text{estMCO} = (\text{inv}(X' \cdot X)) \cdot (X' \cdot Y);$ 
 $\text{thetaMCO}(k, :) = (\text{estMCO})';$ 
%
```

```

% Calcul de l'estimateur proposé
% Calcul de f
% En posant :
Z = [xitd(3: T+1, 1) xitd(4: T+2, 1)];
Z1 = [xitd(2: T, 1) xitd(1: T-1, 1)];
YX = [yitd(2: T, 1) xitd(2: T, 1) yitd(1: T-1, 1)];
%
f = ((YX' * YX) - (YX' * (Z * inv(Z' * Z) * Z') * YX)) ./ (Nobs-q);
%
% Notation : est = estimé de (yitd, xitd)
est = (Z * inv(Z' * Z) * Z') * YX;
% Trouver les valeurs propres telles que det ((est'est) -vf) = 0
c = inv(f) * (est' * est);
v = eig(c);
v1 = min(v);
% est1 = estimateur des paramètres à chaque itération
est1 = inv((est(:, 2:3)' * est(:, 2:3)) - ((v1 - gam) * f(2:3, 2:3))) * ((est(:, 2:3)' * est(:, 1)) - ...
((v1 - gam) * f(2:3, 1)));
theta(k,:) = (est1)';
% Calcul de l'estimateur de variables instrumentales Z
estz = (inv(Z' * X)) * (Z' * Y);
thetaz(k, :) = (estz)';
% Calcul de l'estimateur de variables instrumentales z1
estz1 = (inv(Z1' * X)) * (Z1' * Y);
thetaz1(k, :) = (estz1)';
%
clear ut u0 Epsilt xt x0 wt yt y0 xitd vt yitd st X Y estMCO Z Z1 YX f est c v v1 est1 estz
estz1
end
% Calcul de la moyenne des estimateurs
% % mcomoy = moyenne des estimateurs MCO

```

```
% % estmoy = moyenne de l'estimateur proposé
% % estzmoy = moyenne de l'estimateur Z
% % estz1moy = moyenne de l'estimateur Z1
%
mcomoy = mean(thetaMCO)
estmoy = mean(theta)
estzmoy = mean(thetaz)
estz1moy = mean(thetaz1)
% Calcul de la variance des estimateurs
varmco = (std(thetaMCO)).^2
varest = (std(theta)).^2
varestz = (std(thetaz)).^2
varestz1 = (std(thetaz1)).^2
% Calcul du biais
biaismco = vraitheta - mcomoy
biaisest = vraitheta - estmoy
biaisestz = vraitheta - estzmoy
biaisestz1 = vraitheta - estz1moy
% Racine carré de l'erreur quadratique moyenne = REQM
REQMmco = sqrt((varmco + (biaismco.^2)))
REQMest = sqrt((varest + (biaisest.^2)))
REQMestz = sqrt((varestz + (biaisestz.^2)))
REQMestz1 = sqrt((varestz1 + (biaisestz1.^2)))
```

REFERENCES

- Dagenais, M.G. et D.L. Dagenais, " GMM estimators for linear regression models with errors in the variables ", Centre de Recherche et Developpement en économique (CRDE) Université de Montréal, février 1994.
- Dagenais, M.G., " Parameter estimation in regression models with errors in the variables and autocorrelated disturbances ", journal of econometrics 64, 1994, 145-163.
- Dagenais, M.G., " Consistent estimators for regression models with errors in the variables and autocorrelated disturbances ", CRDE Université de Montréal, janvier 1995.
- Fomby, T.B., R.C. Hill, and S.R. Johnson, " advanced econometrics methods (Springer-verlag, New York, NY), 1984, 205-277.
- Hatanaka, M., " An efficient two step estimator for the dynamic adjustment model with autoregressive errors ", journal of econometrics 2, 1974, 199-220.
- Hamilton, D.J., " Time series analysis ", (Princeton University press, Princeton, New Jersey), 1994, 1-59 et 180-230.
- Malinvaud, E., " Methodes statistique de l'économetrie ", 3^e édition (Dunod, Paris) ,420-458.
- Johnston, J., " Methodes économetriques ", 3^e édition, tome 2 (Economica, 1988), 407-515
- The mathwork, Inc., " Pro-matlab for unix computers ", (Cochituate place, 24 prime park way, South malick, MA 01760), march 1, 1991