UNIVERSITE DE MONTREAL

LE TRAITEMENT DE LA PREMIERE OBSERVATION EN PRESENCE D'AUTOCORRELATION

PAR

ROBERT GOSSELIN

DEPARTEMENT DE SCIENCES ECONOMIQUES

FACULTE DES ARTS ET DES SCIENCES

MEMOIRE PRESENTE A LA FACULTE DES ETUDES SUPERIEURES
EN VUE DE L'OBTENTION DU GRADE DE
MAITRE ES SCIENCES (M.Sc.)

JANVIER 1985

TABLE DES MATIERES

LISTE DES TABLEAUX

SOMMAIRE

Les méthodes pour corriger l'autocorrélation du premier ordre peuvent être divisées en deux groupes : celles qui laissent tomber la première observation (ex. : Cochrane et Orcutt (1949)) et celles qui l'incluent avec un poids de $\sqrt{1-\rho^2}$ (ex. : Prais et Winsten (1954)). Cette deuxième méthode a toujours été considérée comme la plus efficace, cependant, il faut alors supposer que la variance de la première observation est $\sigma_e^2/(1-\rho^2)$, ce qui peut s'avérer inexact en pratique.

Nous démontrons ici que lorsque ρ est connu, il est plus efficace de laisser tomber l'observation initiale si sa variance est plus grande qu'une certaine borne. Ces résultats restent approximativement corrects pour le coefficient de la pente même si ρ doit être estimé. L'estimation de la constante pose cependant certains problèmes lorsqu'on laisse tomber la première observation et que ρ est grand, comme le montre l'étude de Monte-Carlo que nous avons réalisée.

INTRODUCTION

Dans un modèle simple tel que : $\underline{y} = X\underline{\beta} + \underline{u}$, $u_t = \rho\, u_{t-1} + e_t$, $\mathrm{Var}(e_t) = \sigma_e^2$; $\mathrm{Var}(y_1) = \sigma_1^2$, l'autocorrélation peut être corrigée de plusieurs façons. Les différentes méthodes d'estimations peuvent être regroupées en deux (2) classes : celles qui incluent la première observation et celles qui la laissent tomber.

De nombreuses études, autant théoriques qu'empiriques (e.g., Prais et Winsten (1954), Beach et MacKinnon (1978), etc.), arrivent à la conclusion qu'il est toujours préférable d'inclure les valeurs initiales avec un poids de $\sqrt{1 - \rho^2}$. Cette proposition n'est cependant valable que si la variance de la première variable dépendante est effectivement $(\sigma_e^2/(1-\rho^2))$. Comme le suggère Theil (1971, p. 253), il est possible que cette variance soit différente, notamment lorsque les données suivent immédiatement une période de conflit armé. Dans ce cas, Theil suggère d'employer l'autre méthode puisqu'elle est conditionnelle à la première observation, ce qui peut s'avérer plus efficace et donner des tests de Student valides. Le lecteur aura certainement reconnu les méthodes qu'il est maintenant convenu d'appeler Prais-Winsten (PW) et Cochrane-Orcutt (CORC).

Si CORC a longtemps été la technique universellement utilisée, de plus en plus, un certain choix est offert aux praticiens à l'intérieur des logiciels de régression. Entre autres, l'algorithme du maximum de vraisemblance présenté par Beach et MacKinnon (1978), est maintenant disponible sur TSP, un des logiciels les plus employés en économétrie. Comme cet estimateur inclut la première observation, il nous apparaît essentiel de mieux définir les domaines où chacun sera le plus efficace. En particulier, nous nous posons la question : quand est-il plus efficace de laisser tomber la première observation plutôt que de l'inclure avec un poids erroné?

Personne n'a vraiment abordé ce problème. Le seul texte qui fait référence à une mauvaise spécification des conditions initiales, Gurland (1954), se limite à étudier la perte d'efficacité de CORC par rapport aux MCG (moindres carrés généralisés) en supposant les matrices de covariance (et ρ) connues. Ses conclusions sont aussi très restrictives puisqu'elles ne s'appliquent qu'à des cas limites (par exemple : ρ tend vers 1). Tous les autres textes qui traitent de l'efficacité relative de différents estimateurs en présence d'autocorrélation du premier ordre se limitent au cas stationnaire, c'est-à-dire lorsque PW est l'équivalent des MCG (moindres carrés généralisés).

Après une brève revue de la littérature pertinente (chapitre I), et une discussion sur la spécification des conditions initiales (chapitre II), nous verrons dans le cas où ρ est connu, et pour le modèle le plus simple (estimation d'une moyenne), qu'il est toujours préférable de procéder à l'estimation conditionnelle si la variance de la première observation (σ_1^2) est plus grande qu'une certaine borne (chapitre III). Cette borne dépend de ρ, de σ_e^2 et de la taille d'échantillon.

Dans le chapitre IV, nous prouverons qu'il existe également une seule valeur critique pour σ_1^2 dans le cas général avec un nombre K (K < N) de variables explicatives. Lorsque σ_1^2 est plus grand (petit) que cette valeur, alors la différence entre la matrice de covariance de PW et celle de CORC est positive (négative) semi-définie. Ce résultat va beaucoup plus loin que ce qui a été fait jusqu'à présent : d'abord, nous n'imposons aucune restriction sur les conditions initiales, ensuite, nous utilisons un modèle parfaitement général quant au nombre et à la forme des variables explicatives, enfin nous sommes les premiers à utiliser la différence des matrices de covariance comme mesure d'efficacité en présence d'autocorrélation.

Dans les chapitres V et VI, nous éliminerons l'hypothèse que ρ est connu. Nous présenterons alors le plan et les résultats d'une vaste expérience de Monte-Carlo comparant la performance de différents estimateurs. Comme toujours, nous observons plusieurs différences entre la théorie et la pratique, en particulier pour la constante et (ou) lorsque ρ est très grand.

CHAPITRE I

Revue de la littérature

Les procédures de transformation de CORC et PW sont bien présentées dans tous les manuels de base et nous supposons que le lecteur est familier avec celles-ci. Précisons également que partout dans ce chapitre et les trois suivants, nous supposons que ρ est connu et que les variables explicatives sont parfaitement exogènes. Nous ne verrons ici que les auteurs qui font les mêmes hypothèses. Une revue des contributions où ρ est estimé sera présentée au début du chapitre V.

Toutes les études effectuées jusqu'à présent, sauf Gurland (1954), arrivent à la conclusion qu'il est toujours plus efficace d'inclure la première observation avec un poids de $(1-\rho^2)^{1/2}$. Prais et Winsten (1954), Kadiyala (1968), Maeshiro (1976, 1979), Taylor (1980) et enfin Doran (1980) nous suggèrent tous d'utiliser cette transformation.

Prais et Winsten (1954) ainsi que Maeshiro (1979) ont calculé les variances exactes des coefficients $(\hat{B})$ de CORC par rapport aux MCG (PW). Ils ont utilisé des modèles simples (deux variables explicatives dont une contante) en spécifiant directement les valeurs des variables exogènes. Dans certains cas (e.g., une tendance linéaire) ils ont trouvé que le ratio des variances était excessivement élevé pour le coefficient de la pente. Kadiyala (1968) et Maeshiro (1976) procèdent de la même façon que les deux premiers mais en comparant plutôt CORC aux MCO (moindres carrés ordinaires). Le premier (Kadiyala) montre que les MCO sont toujours plus efficaces que CORC pour l'estimation d'une moyenne échantillonnale (régression sur une constante). Maeshiro (1976) trouve un gain important en utilisant les MCO lorsque la variable explicative comporte une tendance.

Taylor (1980) choisit une approche différente. Avec une seule variable exogène (X) dont la forme reste assez générale, il trouve une approximation pour les variances des MCG (PW), des MCO et de CORC en fonction de la distribution de X. Selon lui, CORC serait généralement une bonne approximation des MCG sauf si on observe peu de variations de la variable explicative.

Le rapport des variances généralisées (CORC/MCG) a été étudié par Doran (1980) avec un modèle général, sans contrainte sur le nombre et la forme des variables explicatives. L'expression qu'il trouve dépend inversement de la taille d'échantillon et positivement du nombre de variables explicatives. CORC serait aussi relativement moins efficace en présence de multicollinéarité. Cette expression se révélera très intéressante parce que la borne que nous avons dégagée est directement reliée à celle-ci (c.f. chap. IV).

Si le modèle utilisé par Doran (1980) est le plus général que nous ayons rencontré, il repose cependant sur l'hypothèse que le processus générant les erreurs aléatoires se poursuit depuis un passé infini.

Gurland (1954) est le seul qui n'impose pas cette condition. Malheureusement, il se limite à un modèle simple avec deux variables explicatives. Le seul résultat qu'il obtient est que le ratio des variances généralisées (CORC/MCG) peut devenir infini lorsque ρ tend vers un ou que la première valeur de la variable explicative tend vers l'infini.

En fait, toutes les études précédentes ont considéré CORC comme une approximation des MCG. Une approche beaucoup plus fructueuse sera de le considérer comme un estimateur conditionnel à la première observation.

CHAPITRE II

Discussion des hypothèses

Plusieurs raisons ont été avancées pour justifier la pondération appliquée à la première observation. Nous verrons qu'aucune de ces hypothèses n'est pleinement satisfaisante et que de telles erreurs de spécification sont susceptibles de se produire plus souvent qu'on ne serait porté à le croire.

Lorsque $u_t = \rho\, u_{t-1} + e_t$, une condition suffisante pour que $Var(u_1) = \sigma_e^2/(1-\rho^2)$, est que le processus générant les erreurs aléatoires (u) ait débuté à moins l'infini. Cette formulation ne peut évidemment être retenue et on suppose souvent que l'erreur d'approximation est faible si le passé est suffisamment long. Combien de temps faut-il pour que l'erreur soit négligeable? La théorie ne répond pas à cette question, chaque cas est un cas d'espèce qui doit être jugé par le praticien. Theil (1971, p. 253) nous propose au moins un cas où l'on peut douter fortement de la pertinence d'une telle hypothèse : si les données disponibles suivent immédiatement une grande guerre, les changements structurels de l'économie affecteront sûrement les erreurs et l'approximation risque d'être très mauvaise. Les conflits armés ne sont qu'un exemple parmi tant d'autres où la structure économique est affectée. Comme une pratique courante en économétrie est justement de procéder à l'estimation d'un modèle entre deux périodes de changements, la formulation précédente risque d'être inexacte dans de très nombreux cas.

Comme on le sait, une source importante d'erreurs provient de l'observation des variables. Une modification des méthodes de collecte des données peut donc affecter le processus de la même façon qu'un changement structurel de l'économie. Ici encore, on peut douter de la validité de l'approximation si un changement des méthodes d'observation a eu lieu près de la période initiale.

Un des inconvénients majeurs de cette première formulation est qu'elle repose sur des valeurs passées qui ne sont pas observées. Pour éviter ceci, nous pouvons supposer directement que $VAR(u_1) = \sigma_e^2/(1-\rho^2)$ ou bien que le processus est stationnaire. Cette formulation reste pour le moins arbitraire et demanderait à être vérifiée empiriquement. Si la stationnarité est une qualité importante, elle n'est cependant pas nécessaire et ne devrait pas être imposée sans contrôle.

Prenons ici un exemple simple. Supposons que nous avons un modèle théorique simple (une équation) pour expliquer les quantités <u>vendues</u> de téléviseurs. Ce modèle nous suggère que les erreurs aléatoires sont de la forme bien connue :

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

Il est clair que les quantités vendues avant l'introduction du produit sur le marché sont identiquement nulles, les erreurs le sont donc forcément. Si nous disposons de données complètes sur les ventes, nous aurons pour u_1 : $u_1 = \rho\, u_0 + e_1 = e_1$, puisque u_0 est identiquement nul. Il est alors impossible de supposer en même temps que les innovations (e) sont homoscédastiques et que le processus est stationnaire ($Var(u_1) = \sigma_e^2/(1-\rho^2)$).

Nous verrons plus loin (chap.IV) que CORC admet certaines formes d'erreurs non stationnaires ce qui en fait une méthode plus souple et plus robuste que PW. De plus, les tests de Student appliqués aux coefficients ne sont valides que si les erreurs sont stationnaires, pour PW ou lorsque $\rho = 0$, pour les MCO. En utilisant CORC, ces tests restent valides pour toutes les valeurs de ρ et de σ_1^2.

CHAPITRE III

Régression sur une constante

Avant de s'attaquer au modèle général avec plusieurs variables explicatives, l'examen d'un problème plus simple peut être intéressant. La forme la plus simple de régression est évidemment le calcul d'une moyenne échantillonnale puisqu'il s'agit d'une régression sur une seule variable explicative constante. Comme l'a montré Kadiyala (1968), ce modèle est particulièrement défavorable à CORC qui est toujours moins efficace que les MCO si le processus est stationnaire. Les conditions sous lesquelles CORC sera meilleur que PW risquent donc d'être très sévères. Nous verrons pourtant qu'elles sont en fait assez peu restrictives. Nous nous attendons alors à trouver des conditions encore moins rigoureuses dans les cas où CORC est une bonne approximation des MCG.

Soit le modèle :

(1) $$y_t = B + u_t \qquad t = 1, \ldots, N$$

(2) $$u_t = \rho\, u_{t-1} + e_t \qquad t = 2, \ldots, N$$

(3) $$e_t \sim \text{N.i.d. } (0, \sigma_e^2) \qquad t = 2, \ldots, N$$

(4) $$u_1 \sim \text{N.i.d. } (0, \sigma_1^2)$$

On vérifie aisément que l'estimation de la moyenne par CORC se réduit à :

(5) $$\hat{B}(CORC) = \Sigma \frac{(y_t - \rho\, y_{t-1})}{(N-1)\ (1-\rho)}$$

(Le signe de sommation Σ sans indice doit être interprété comme : $\sum_{t=2}^{N}$).

L'estimateur pour PW sera plutôt

(6) $$\hat{B}(PW) = \frac{(1-\rho^2)y_1 + \Sigma(y_t - \rho\, y_{t-1})(1-\rho)}{(1-\rho^2) + (N-1)\ (1-\rho)^2}$$

Ce qui donne après substitution :

$$\hat{B}(CORC) = B + \Sigma\, e_t/[(N-1)\ (1-\rho)]$$

et

$$\hat{B}(PW) = B + [(1-\rho^2)u_1 + (1-\rho)\ \Sigma\, e_t]/[(1-\rho^2) + (N-1)\ (1-\rho)^2]$$

Ces estimateurs sont évidemment sans biais et leurs variances sont respectivement :

(7) $$\mathrm{Var}(\hat{B}(CORC)) = E(\hat{B}(CORC) - B)^2 = E\,\frac{(\Sigma\, e_t)^2}{(N-1)^2\ (1-\rho)^2} = \boxed{\sigma_e^2/[(N-1)\ (1-\rho)^2]}$$

et

(8) $$\mathrm{Var}(\hat{B}(PW)) = E(\hat{B}(PW) - B)^2 = E\,\frac{[(1-\rho^2)u_1 + (1-\rho)\ \Sigma\, e_t]^2}{[(1-\rho^2) + (N-1)\ (1-\rho)^2]^2}$$

$$= \boxed{\frac{(1-\rho^2)^2\ \sigma_1^2 + (N-1)\ (1-\rho)^2\ \sigma_e^2}{[(1-\rho^2) + (N-1)\ (1-\rho)^2]^2}}$$

A partir de ces deux formules nous pouvons maintenant énoncer et démontrer un théorème qui régit le domaine d'efficacité relative de ces estimateurs.

THEOREME 1 : Si les hypothèses (1) à (4) sont vérifiées, et si ρ est connu et $|\rho| < 1$ alors la variance de l'estimateur de CORC sera plus petite (grande) que celle de PW si et seulement si la variance de la première observation (σ_1^2) est plus grande (petite) que

$$\boxed{\left(2 + \frac{(1-\rho^2)}{(N-1)\ (1-\rho)^2}\right) \frac{\sigma_e^2}{1-\rho^2}} \tag{9}$$

Preuve :

La preuve est ici assez simple, soit :

$\mathrm{Var}(\hat{B}(\mathrm{CORC})) < \mathrm{Var}(\hat{B}(\mathrm{PW})) \implies$

$$\frac{\sigma_e^2}{(N-1)\ (1-\rho)^2} < \frac{[(1-\rho^2)^2\ \sigma_1^2 + (N-1)\ (1-\rho)^2\ \sigma_e^2]}{[(1-\rho^2) + (N-1)\ (1-\rho)^2]^2}$$

Comme les dénominateurs sont positifs, le signe d'inégalité ne change pas en écrivant :

$$\sigma_e^2[(1-\rho^2) + (N-1)\ (1-\rho)^2]^2 \leq [(1-\rho^2)^2\ \sigma_1^2 + (N-1)\ (1-\rho)^2\ \sigma_e^2]\ (N-1)\ (1-\rho)^2$$

En effectuant les multiplications, nous obtenons

$$\sigma_e^2[(1-\rho^2)^2 + 2(N-1)\ (1-\rho)^2\ (1-\rho^2) + (N-1)^2\ (1-\rho)^4]$$

$$< (N-1)\ (1-\rho)^2\ (1-\rho^2)^2\ \sigma_1^2 + (N-1)^2\ (1-\rho)^4\ \sigma_e^2$$

Les termes $(N-1)^2\ (1-\rho)^4\ \sigma_e^2$ s'annulent des deux côtés, donc :

$$\sigma_e^2[(1-\rho^2)^2 + 2(N-1)\ (1-\rho)^2\ (1-\rho^2)] < \sigma_1^2(N-1)\ (1-\rho)^2\ (1-\rho^2)^2$$

et

$$\frac{\sigma_e^2}{(1-\rho^2)}\ \frac{[(1-\rho^2)^2 + 2(N-1)\ (1-\rho)^2\ (1-\rho^2)]}{(N-1)\ (1-\rho)^2\ (1-\rho^2)} < \sigma_1^2$$

Donc,

$$\mathrm{Var}(\hat{B}(\mathrm{CORC})) < \mathrm{Var}(\hat{B}(\mathrm{PW})) \quad \text{ssi}$$

$$\sigma_1^2 > \frac{\sigma_e^2}{(1-\rho^2)}\left[2 + \frac{1-\rho^2}{(N-1)\ (1-\rho)^2}\right]$$

Q.E.D.

Le premier terme de droite, $\sigma_e^2/(1-\rho^2)$ représente la variance des "u" dans le modèle stationnaire. Ici, une nuance s'impose : nous verrons dans le chapitre suivant que la variance de chacune des erreurs aléatoires (σ_u^2) est affectée lorsque σ_1^2 est différent de $\sigma_e^2/(1-\rho^2)$. Cependant, lorsque la taille d'échantillon tend vers l'infini, alors $\mathrm{Var}(u_N)$ tend vers la valeur stationnaire. Cette expression doit donc être interprétée comme la variance de long terme.

Le second terme entre crochets $[(1-\rho^2)/(N-1)\ (1-\rho)^2]$ est toujours positif pour $|\rho| < 1$, il tend vers l'infini lorsque ρ tend vers 1, mais tend vers zéro lorsque N tend vers l'infini. La variance de la première observation doit donc être au moins le double de celle de long terme pour que CORC soit plus efficace que PW. Les cas les moins favorables à CORC sont ainsi ceux où ρ est grand et N est petit.

Le tableau suivant donne quelques exemples des valeurs qu'on trouve pour le rapport des variances sous différentes hypothèses pour ρ et N.

Tableau 1

Exemple de valeurs pour la borne

ρ	N	$2 + (1-\rho)^2/(N-1)\ (1-\rho)^2$
-0,99	50	2,00055
0,6	21	2,2
0,8	28	2,333
0,95	40	3
0,99	50	6,06

Pour toutes les valeurs négatives de ρ, et les valeurs positives moyennes de ρ et N, l'expression reste relativement faible. Cependant, lorsque ρ approche l'unité, il devient de moins en moins intéressant de laisser tomber la première observation.

Il est assez difficile d'évaluer si de telles conditions sont susceptibles de se produire souvent en pratique parce que la distribution des "u" ne retient jamais l'attention des analystes. Cependant, une simple transformation nous permettra de mieux juger la situation. L'expression (9) peut se réécrire :

$$\sigma_1^2(1-\rho^2) > \sigma_e^2[2 + (1-\rho^2)/(N-1)\ (1-\rho)^2]$$

Le terme de gauche est maintenant la variance de la première erreur transformée sous l'hypothèse de stationnarité $(Var(\sqrt{1-\rho^2}\ u_1))$. En général, σ_e^2 est estimé de façon assez précise, mais pas $\sigma_1^2(1-\rho^2)$. Un rapport de l'ordre de deux ou trois entre ces termes passera donc inaperçu la plupart du temps ce qui montre bien que la condition du théorème 1 est peu restrictive si ρ n'est pas trop grand.

L'estimation d'une moyenne échantillonnale en présence d'autocorrélation reste assez rare en économétrie. Ce modèle simple nous a cependant permis de prouver que dans certains cas, l'estimation conditionnelle peut être préférable. Nous pousserons maintenant cette étude vers un modèle parfaitement général.

CHAPITRE IV

Le modèle général[1]

[1]Rappelons que partout dans ce chapitre ρ est connu et $|\rho| < 1$.

S'il existe une borne unique lorsque le modèle comporte une seule variable explicative constante, rien ne nous permet de croire qu'il en sera de même avec plusieurs variables. Le théorème 2, que nous présenterons bientôt, nous garantit qu'une seule borne existe aussi dans le cas général. De plus, le corollaire 1 nous donne une interprétation beaucoup plus intuitive de cette borne. En premier lieu, il est évidemment nécessaire de présenter le modèle général.

IV.1. Le modèle

Soit :

$$\underline{y} = X\underline{B} + u \tag{1}$$

$\underline{y}$ est de dimension Nx1;

X est de dimension NxK (K < N);

$\underline{B}$ est de dimension Kx1;

$\underline{u}$ est de dimension Nx1.

Nous supposons évidemment que les hypothèses de base sont satisfaites :

- $\underline{y}$ est la variable aléatoire dépendante;
- X est non stochastique; rang (X) = K;
- $\underline{B}$ et $\underline{u}$ ne sont pas observés;

(2a) $$u_t = \rho\, u_{t-1} + e_t \qquad t = 2, \ldots, N$$

(2b) $$e_t \underset{\sim}{\text{ind}} N(0, \sigma_e^2) \qquad t = 2, \ldots, N$$

(2c) $$u_1 = e_1$$

(2d) $$e_1 \sim N(0, a^2 \sigma_e^2) \qquad \text{indépendant de } e_2, \ldots, e_N$$

La différence essentielle entre ce modèle et les autres est représentée par les équations (2c) et (2d). Nous aurons alors $\sigma_1^2 = a^2 \sigma_e^2$, et si $a^2 = 1/(1-\rho^2)$ nous retrouvons le modèle stationnaire qui est habituellement utilisé.

Les estimateurs de CORC et PW seront respectivement :

(3) $$\hat{\underline{B}}(\text{CORC}) = X'Q'QX)^{-1} X'Q'Q\underline{y}$$

où

(4) $$\underset{(N-1)\text{x}N}{Q} = \begin{bmatrix} -\rho & 1 & 0 & \ldots\ldots & 0 \\ 0 & -\rho & 1 & & 0 \\ 0 & 0 & -\rho & & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \ldots -\rho & 1 \end{bmatrix}$$

(5) $$\hat{\underline{B}}(\text{PW}) = (X'P'PX)^{-1} X'P'P\underline{y}$$

(6)
$$\underset{N\times N}{P} = \begin{bmatrix} \underline{p}_1' \\ Q \end{bmatrix}$$

où $\underline{p}_1' = (1-\rho^2)^{1/2}\ \underline{i}_1'$,

et $\underline{i}_1' = [1 \quad 0 \quad 0 \quad \dots \quad 0]$.

IV.2. Matrices de covariances

Avant de définir l'estimateur des MCG, nous devons obtenir la matrice de covariance des "u". Remarquons d'abord qu'en prémultipliant le vecteur $\underline{u}$ par une matrice R, nous obtenons le vecteur $\underline{e}$. On vérifie aisément que

(7)
$$\underset{N\times N}{R} = \begin{bmatrix} \underline{i}_1' \\ Q \end{bmatrix} .$$

Les équations (2a) à (2d) peuvent donc s'écrire sous forme matricielle :

(8a)
$$\underline{e} = R\ \underline{u}$$

(8b)
$$\underline{e} \sim N(0,\ \sigma_e^2\ A)$$

où

$$\underset{N\times N}{A} = \begin{bmatrix} a^2 & \underline{0}' \\ \underline{0} & I_{N-1} \end{bmatrix}$$

et

$$\underset{1\times(N-1)}{\underline{0}'} = [0 \quad 0 \quad 0 \quad \dots \quad 0]$$

De (8a) et (8b) nous obtenons la distribution de $\underline{u}$:

$$\underset{\sim}{u} = R^{-1} \underset{\sim}{e}$$

$$E(\underset{\sim}{u}) = R^{-1} E(\underset{\sim}{e}) = \underset{\sim}{0}$$

(9) $$\mathrm{Cov}(\underset{\sim}{u}) = E(\underset{\sim}{u}\underset{\sim}{u}') = R^{-1} E(\underset{\sim}{e}\underset{\sim}{e}')R'^{-1} = \sigma_e^2 R^{-1} AR'^{-1}$$

$$\underset{\sim}{u} \sim N(\underset{\sim}{0}, \sigma_e^2 R^{-1} AR'^{-1})$$

$$\underset{\sim}{u} \sim N(\underset{\sim}{0}, \sigma_e^2 V)$$

D'après Gurland (1954), R^{-1} est de la forme :

(9a) $$R^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \rho & 1 & 0 & \cdots & 0 \\ \rho^2 & \rho & 1 & & \\ \vdots & \vdots & & & \\ \rho^{N-1} & \rho^{N-2} & & \cdots & 1 \end{bmatrix}$$

Ce qui donne pour V :

(9b) $$V = \begin{bmatrix} a^2 & \rho a^2 & \rho^2 a^2 & & \cdots\cdots \\ \rho a^2 & \rho^2 a^2+1 & \rho^3 a^2+\rho & & \cdots\cdots \\ \rho^2 a^2 & \rho^3 a^2+\rho & \rho^4 a^2 + \frac{(1-\rho^4)}{1-\rho^2} & \ddots & \\ \rho^3 a^2 & \rho^4 a^2+\rho^2 & \vdots & & \\ \vdots & & \vdots & & \\ \rho^{N-1} a^2 & \rho^N a^2+\rho^{N-2} & \rho^{N+1} a^2+\rho^{N-3} \frac{(1-\rho^4)}{1-\rho^2} & \cdots & \rho^{2(N-1)} a^2 + \frac{(1-\rho^{2(N-1)})}{1-\rho^2} \end{bmatrix}$$

Remarquons ici que cette matrice de covariance n'est pas stationnaire excepté lorsque $a^2 = 1/(1-\rho^2)$. Cependant, lorsque N devient très grand, $Var(u_N)$ tend vers $\sigma_e^2/(1-\rho^2)$ et les éléments hors-diagonale tendent aussi vers leur valeur "stationnaire". La matrice V est donc la matrice de covariance d'un processus (AR1) qui retourne à la stabilité après avoir reçu un choc exogène. Plus ce choc est important (a^2) et (ou) plus ρ est grand, plus il faut de temps pour retrouver cette stabilité. $\sigma_e^2/(1-\rho^2)$ doit donc être interprété comme la variance de long terme.

L'estimateur des MCG sera alors :

$$\hat{\underline{B}}(MCG) = (X'V^{-1}X)^{-1}X'V^{-1}\underline{y}$$

où

$$V^{-1} = R^{-1}\ AR'^{-1}$$

On vérifie aisément que les trois estimateurs, CORC, PW et MCG sont sans biais. Par exemple :

$$\hat{\underline{B}}(PW) = (X'P'PX)^{-1}\ X'P'P\ \underline{y}$$

$$= (X'P'PX)^{-1}\ X'P'P\ (X\underline{B} + \underline{u})$$

$$= \underline{B} + (X'P'PX)^{-1}\ X'P'P\ \underline{u}$$

$$E(\hat{\underline{B}}(PW) - \underline{B}) = (X'P'PX)^{-1}\ X'P'P\ E(\underline{u}) = \underline{0}$$

Les matrices de covariance des paramètres estimés seront respectivement :

(10) $$\Sigma(CORC) = E\left[(\hat{B}(CORC) - B)\ (\hat{B}(CORC) - B)'\right]$$

$$= (X'Q'QX)^{-1}\ X'Q'QE(\underline{u}\underline{u}')\ Q'QX(X'Q'QX)^{-1}$$

$$= \sigma_e^2(X'Q'QX)^{-1}\ X'Q'QVQ'QX(X'Q'QX)^{-1}$$

(11) $$\Sigma(PW) = \sigma_e^2(X'P'PX)^{-1}\ X'P'PVP'PX(X'P'PX)^{-1}$$

(12) $$\Sigma(MCG) = \sigma_e^2(X'V^{-1}X)^{-1}$$

IV.3. Efficacité CORC vs PW

Théoriquement, le meilleur estimateur est celui des MCG, mais en pratique il est difficilement applicable parce qu'il faut connaître a^2. Il semble donc préférable de définir les domaines d'efficacité relative des estimateurs connus, en particulier CORC et PW.

THEOREME 2 : Si les hypothèses (1) à (2d) sont vérifiées, si ρ est connu et $|\rho| < 1$ alors : l'estimateur de CORC tel que défini en (3) sera plus (moins) efficace que PW (défini en (5)) au sens où la différence des matrices de covariance est positive semi-définie si et seulement si :

$$k > 2 + (1-\rho^2)\ \underline{x}_1'(X'Q'QX)^{-1}\ \underline{x}_1$$

où $k = \sigma_1^2/[\sigma_e^2/(1-\rho^2)]$

k est le rapport entre la variance de la première observation (u_1) et celle du long terme, et

(13) $$\underset{K\times 1}{\underline{x}_1} = X' \underline{i}_1 , \qquad \underset{N\times 1}{\underline{i}'_1} = [1 \quad 0 \quad 0 \quad \ldots \quad 0]$$

($\underline{x}_1$ est un vecteur contenant la première observation de chaque variable explicative).

Preuve :

Nous devons dans un premier temps exprimer Σ(CORC) et Σ(PW) sous des formes plus maniables que les équations (10) et (11). Reprenons d'abord l'équation (10) :

$$\Sigma(\text{CORC}) = \sigma_e^2 (X'Q'QX)^{-1} X'Q'QVQ'QX(X'Q'QX)^{-1}$$

Le terme du milieu (QVQ') est une matrice unitaire d'ordre N-1 parce que :

$$QVQ' = QR^{-1} AR'^{-1} Q' \qquad \text{(d'après (9))}$$

$$I_N = RR^{-1} = \begin{bmatrix} \underline{i}'_1 \\ \\ Q \end{bmatrix} R^{-1} \qquad \text{(d'après (7))}$$

$$= \begin{bmatrix} \underline{i}'_1 R^{-1} \\ \\ Q R^{-1} \end{bmatrix} = \begin{bmatrix} 1 & \underline{0}' \\ \\ \underline{0} & I_{N-1} \end{bmatrix}$$

donc :

(14) $$QR^{-1} = [\underline{0} : I_{N-1}]$$

et :

$$QVQ' = QR^{-1} AR'^{-1} Q' = [\underset{\sim}{0} : I_{N-1}] \begin{bmatrix} a^2 & \underset{\sim}{0}' \\ \underset{\sim}{0} & I_{N-1} \end{bmatrix} \begin{bmatrix} \underset{\sim}{0}' \\ I_{N-1} \end{bmatrix} = I_{N-1}$$

(d'après (8b) et (14)). Par conséquent,

$$(15) \qquad \Sigma(CORC) = \sigma_e^2 (X'Q'QX)^{-1} X'Q'IQX(X'Q'QX)^{-1}$$

$$= \sigma_e^2 (X'Q'QX)^{-1}$$

$$= \sigma_e^2 W \qquad \text{(pour simplifier la notation).}$$

Nous exprimerons maintenant $\Sigma(PW)$ en fonction des mêmes éléments que $\Sigma(CORC)$, notamment :

$$(X'P'PX) = W^{-1} + (1-\rho^2)\, \underset{\sim}{x}_1 \underset{\sim}{x}_1'$$

et

$$X'P'PVP'PX = W^{-1} + a^2(1-\rho^2)^2\, \underset{\sim}{x}_1 \underset{\sim}{x}_1' \, .$$

On peut démontrer ces deux identités comme suit. D'abord,

$$(16) \qquad X'P'PX = X'[\underset{\sim}{p}_1 : Q'] \begin{bmatrix} \underset{\sim}{p}_1' \\ Q \end{bmatrix} X \qquad \text{(d'après (6))}$$

$$= X'(\underset{\sim}{p}_1 \underset{\sim}{p}_1' + Q'Q)\, X$$

$$= X'Q'QX + (1-\rho^2)\, X'\underset{\sim}{i}_1 \underset{\sim}{i}_1'\, X$$

$$= W^{-1} + (1-\rho^2)\, \underset{\sim}{x}_1 \underset{\sim}{x}_1' \qquad \text{(d'après (13) et (15)).}$$

en posant $\underline{z} = (1-\rho^2)^{1/2} \underline{x}_1$, nous avons

(17) $$X'P'PX = W^{-1} + \underline{z}\,\underline{z}' \ ,$$

et on vérifie aisément que : (voir Maddala (1977, p. 446)

$$(X'P'PX)^{-1} = (W^{-1} + \underline{z}\,\underline{z}')^{-1} = W - \frac{1}{1 + \underline{z}'\,W\underline{z}}\,(W\,\underline{z}\,\underline{z}'\,W)\ ,$$

d'où en définissant $h = 1/(1 + \underline{z}'\,W\,\underline{z})$, un scalaire,

(18) $$(X'P'PX)^{-1} = W - h\,W\underline{z}\,\underline{z}'\,W\ .$$

D'autre part :

$$PVP' = PR^{-1}\,AR'^{-1}\,P' \qquad \text{(d'après (9))}$$

(19) $$PR^{-1} = \begin{bmatrix} \underline{p}_1' \\ Q \end{bmatrix} R^{-1} = \begin{bmatrix} \underline{p}_1' R^{-1} \\ QR^{-1} \end{bmatrix} \qquad \text{(d'après (6))}$$

$$= \begin{bmatrix} (1-\rho^2)\underline{i}'_{\,1}R^{-1} \\ QR^{-1} \end{bmatrix} = \begin{bmatrix} (1-\rho^2)^{1/2} & \underline{0}' \\ \underline{0} & I_{N-1} \end{bmatrix} \qquad \text{(d'après (6), (9) et (14))}$$

Donc,

(20) $$PVP' = \begin{bmatrix} (1-\rho^2)^{1/2} & \underline{0}' \\ \underline{0} & I_{N-1} \end{bmatrix} \begin{bmatrix} a^2 & \underline{0}' \\ \underline{0} & I_{N-1} \end{bmatrix} \begin{bmatrix} (1-\rho^2)^{1/2} & \underline{0}' \\ \underline{0} & I_{N-1} \end{bmatrix}$$

$$= \begin{bmatrix} a^2(1-\rho^2) & \underline{0}' \\ \underline{0} & I_{N-1} \end{bmatrix} \qquad \text{(d'après (8b) et (19))}$$

et,

$$(21) \quad P'PVP'P = [\underline{p}_1 \; : \; Q'] \begin{bmatrix} a^2(1-\rho^2) & \underline{0}' \\ \underline{0} & I_{N-1} \end{bmatrix} \begin{bmatrix} \underline{p}_1' \\ Q \end{bmatrix} \qquad \text{(d'après (6)}$$

$$= [a^2(1-\rho^2) \; \underline{p}_1 \; : \; Q'] \begin{bmatrix} \underline{p}_1' \\ Q \end{bmatrix}$$

$$= a^2(1-\rho^2) \; \underline{p}_1 \, \underline{p}_1' + Q'Q \, .$$

Par suite,

$$(22) \quad X'P'PVP'PX = X'(a^2(1-\rho^2) \; \underline{p}_1 \, \underline{p}_1' + Q'Q) \; X \qquad \text{(d'après (21))}$$

$$= X'Q'QX + a^2(1-\rho^2)^2 X' \; \underline{p}_1 \, \underline{p}_1' \; X$$

$$= W^{-1} + a^2(1-\rho^2)^{\,2} X' \underline{i}_1 \, \underline{i}_1' \; X \qquad \text{(d'après (6) et (15))}$$

$$= W^{-1} + a^2(1-\rho^2)^2 \; \underline{x}_1 \, \underline{x}_1' \qquad \text{(d'après (13))}$$

$$= W^{-1} + k \, \underline{z} \, \underline{z}' \qquad \text{(d'après (17))}$$

$$(23) \qquad k = a^2(1-\rho^2)$$

En utilisant (18) et (23) nous obtenons

$$(24) \quad \Sigma(PW) = \sigma_e^2 \; (W^{-1} + h \; W \; \underline{z} \; \underline{z}' \; W) \; (W^{-1} + k \; \underline{z} \; \underline{z}') \; (W - h \; W \; \underline{z} \; \underline{z}' \; W)]$$

Nous pouvons maintenant chercher les conditions nécessaires et suffisantes pour que $\Sigma(PW) - \Sigma(CORC)$ soit positif semi-défini. En regroupant les équations (15) et (24), nous aurons :

$$\Sigma(PW) - \Sigma(CORC) = \sigma_e^2 M$$

$$= \sigma_e^2[(W - hW\underline{z}\,\underline{z}'W)(W^{-1} + k\underline{z}\,\underline{z}')(W - hW\underline{z}\,\underline{z}'W) - W]$$

En effectuant chacune des multiplications, nous obtenons :

$$M = (W + kW\underline{z}\,\underline{z}'W - 2hW\underline{z}\,\underline{z}'W - 2khW\underline{z}\,\underline{z}'W\underline{z}\,\underline{z}'W$$

$$+ h^2W\underline{z}\,\underline{z}'W\underline{z}\,\underline{z}'W + kh^2W\underline{z}\,\underline{z}'W\underline{z}\,\underline{z}'W\underline{z}\,\underline{z}'W - W)$$

En posant $g = z'Wz$ (un scalaire) et en regroupant les termes nous obtenons :

$$(25) \qquad M = (k - 2h - 2khg + h^2g + kh^2g^2)\,W\underline{z}\,\underline{z}'W$$

Pour que $\Sigma(PW) - \Sigma(CORC)$ soit positive semi-définie, il est nécessaire et suffisant que le terme entre parenthèses de l'équation (25) soit positif puisque σ_e^2 est positif et $W\underline{z}\,\underline{z}'W$ est une forme quadratique positive semi-définie; $W\underline{z}$ est un vecteur kx1 et $W = W'$. Pour que le produit d'un scalaire et d'une matrice positive semi-définie soit positif semi-défini, il est nécessaire et suffisant que le scalaire soit positif:

$$k - 2h - 2khg + h^2g + kh^2g^2 > 0$$

Rappelons que $h = 1/(1+\underline{z}'W\underline{z}) = 1/(1+g)$

$$k - \frac{2}{1+g} - \frac{2kg}{1+g} + \frac{g}{(1+g)^2} + \frac{kg^2}{(1+g)^2} > 0$$

En multipliant les deux côtés par $(1+g)^2$, on ne change pas l'inégalité :

$$k(1+g)^2 - 2(1+g) - 2kg(1+g) + g + kg^2 > 0$$

$$k + 2kg + kg^2 - 2 - 2g - 2kg - 2kg^2 + g + kg^2 > 0$$

$$k - 2 - g > 0$$

$$k > 2 + g$$

où

$$k = a^2(1-\rho^2) = \sigma_1^2(1-\rho^2)/\sigma_e^2$$

$$g = \underline{z}' W \underline{z} = (1-\rho^2)\ \underline{x}_1'(X'Q'QX)^{-1}\ \underline{x}_1 .$$

Donc, la différence des matrices de covariance est positive semi-définie si et seulement si :

$$\sigma_1^2(1-\rho^2)/\sigma_e^2 \geq 2 + (1-\rho^2)\ \underline{x}_1'(X'Q'QX)^{-1}\ \underline{x}_1$$

Cette différence sera négative semi-définie (PW plus efficace) si le signe d'inégalité est inversé et sera nulle si on a l'égalité stricte.

Q.E.D.

Nous obtenons ainsi une borne unique, même dans le cas de plusieurs variables explicatives. Cette borne délimite parfaitement le domaine d'efficacité relative de CORC et PW. Une des deux méthodes se révélera donc plus efficace que l'autre pour au moins un des paramètres et sera au moins aussi efficace pour les autres. S'il y a égalité stricte, les deux estimateurs ont la même variance.

IV.4. Interprétation

Présenté ainsi, le résultat apparaît assez peu intuitif. Le terme de gauche est le rapport entre la variance de la première observation et celle de long terme, mais le terme de droite n'a pas d'interprétation économétrique. Le corollaire suivant propose une formulation beaucoup plus claire.

Avant de présenter ce corollaire, quelques remarques et définitions sont nécessaires.

- Le prédicteur simple de y_1 par CORC est

$$\hat{y}_1 = \underline{x}_1' \, \hat{\underline{B}}(\mathrm{CORC}) \tag{26}$$

- sa variance est

$$\begin{aligned} \mathrm{Var}(\hat{y}_1) &= E(\hat{y}_1 - y_1)^2 = E(\underline{x}_1' \, \hat{\underline{B}}(\mathrm{CORC}) - y_1)^2 \\ &= E(\underline{x}_1'(\hat{\underline{B}}(\mathrm{CORC}) - \underline{B}) - u_1)^2 \end{aligned}$$

Puisque $\hat{\underline{B}}(\mathrm{CORC})$ et u_1 sont indépendants :

$$\begin{aligned} \mathrm{Var}(\hat{y}_1) &= \underline{x}_1' \, E\left[(\hat{\underline{B}}(\mathrm{CORC}) - \underline{B}) \; (\hat{\underline{B}}(\mathrm{CORC}) - \underline{B})'\right]\underline{x}_1 + E(u_1)^2 \\ &= \sigma_e^2 \, \underline{x}_1'(X'Q'QX)^{-1} \, \underline{x}_1 + \sigma_1^2 \qquad \text{(d'après (15))} \end{aligned} \tag{27}$$

- Si $\sigma_1^2 \geq \sigma_e^2/(1-\rho^2)$, le modèle présenté au début du chapitre ne peut être différencié d'un modèle où les hypothèses (2c) et (2d) sont remplacées par

(2e) $$u_1 = e_1/(1-\rho^2)^{1/2} + v$$

(2f) $$\underline{v} \sim N(\underline{0},\ \sigma_v^2) \quad \text{indépendant de } e_1,\ e_N$$

(2g) $$\underline{e} \sim N(\underline{0},\ \sigma_e^2\ I_N)$$

v représente un choc aléatoire exogène qui frappe le modèle stationnaire à la période initiale. Si $\sigma_v^2 = 0$, le processus reste stationnaire.

Nous aurons bien entendu :

(28) $$\sigma_1^2 = \sigma_e^2/(1-\rho^2) + \sigma_v^2$$

et

(29) $$\mathrm{Var}(\hat{y}_1) = \sigma_v^2 + \sigma_e^2\,[(1/(1-\rho^2) + \underline{x}_1'(X'Q'QX)^{-1}\ \underline{x}_1)]$$

La variance de $\hat{y}_1$ peut donc être expliquée par deux facteurs indépendants. $\sigma_v^2/\mathrm{Var}(\hat{y}_1)$ est la partie qui est due au choc, l'autre part étant due au processus stationnaire.

<u>COROLLAIRE 1</u> :

Si les hypothèses du théorème 2 sont satisfaites, et si $\sigma_1^2 \geq \sigma_e^2/(1-\rho^2)$ alors CORC sera toujours plus (moins) efficace que PW, au sens où la différence des matrices de variance est positive semi-définie, si et seulement si plus (moins) de 50% de la variance de $\hat{y}_1$ est due au choc exogène "v" (défini en (2e)), c'est-à-dire

(30) $$\sigma_v^2/\mathrm{Var}(\hat{y}_1) > 0,5$$

Preuve :

$$\sigma_v^2/\mathrm{Var}(\hat{y}_1) = \frac{\sigma_v^2}{\sigma_1^2 + \sigma_e^2(\underset{\sim}{x}_1'(X'Q'QX)^{-1}\ \underset{\sim}{x}_1)} \qquad \text{(d'après (27))}$$

La condition énoncée peut donc s'écrire :

$$2\ \sigma_v^2 > \sigma_1^2 + \sigma_e^2(\underset{\sim}{x}_1'(X'Q'QX)^{-1}\ \underset{\sim}{x}_1)$$

$$2(\sigma_1^2 - \sigma_e^2/(1-\rho^2)) > \sigma_1^2 + \sigma_e^2\ \underset{\sim}{x}_1'(X'Q'QX)^{-1}\ \underset{\sim}{x}_1 \qquad \text{(d'après (28))}$$

$$\sigma_1^2 > 2\ \sigma_e^2/(1-\rho^2) + \sigma_e^2\ \underset{\sim}{x}_1'(X'Q'QX)^{-1}\ \underset{\sim}{x}_1$$

$$\text{(31)} \quad \frac{\sigma_1^2(1-\rho^2)}{\sigma_e^2} > 2 + (1-\rho^2)\ \underset{\sim}{x}_1'(X'Q'QX)^{-1}\ \underset{\sim}{x}_1$$

Les conditions du théorème 2 et du corollaire 1 sont donc équivalentes.

Q.E.D.

De façon symétrique, si moins de 50% de la variance de $\hat{y}_1$ est due au terme <u>stationnaire</u>, alors CORC sera plus efficace. Il faut donc que la variance de $\hat{y}_1$ soit le double de la variance "stationnaire", $\sigma_e^2(1/(1-\rho^2) + (\underset{\sim}{x}_1(X'Q'QX)^{-1}\ \underset{\sim}{x}_1))$ pour que les deux estimateurs soient également efficaces. En fait, plus le choc (σ_v^2) est important plus PW se révélera inefficace.

IV.5. Evaluation

Que l'on évalue la variance de $\hat{y}_1$, où la borne de σ_1^2, le terme prédominant sera souvent $\underline{x}_1'(X'Q'QX)^{-1}\,\underline{x}_1$. Cette expression est évidemment positive et ne peut être nulle que si tous les éléments du vecteur $\underline{x}_1$ sont nuls. Doran (1980) trouve un terme similaire pour le rapport des variances généralisées. Il propose entre autres qu'il est d'autant plus petit que N est grand, mais qu'il peut devenir très grand en présence de multicollinéarité. Dans plusieurs cas, on peut aussi s'attendre à une augmentation lorsque le nombre de variables explicatives s'accroît. Cette dernière relation reste cependant subjective puisqu'il reste facile de trouver des contres-exemples dans des cas précis. Enfin, il semble que $\underline{x}_1'(X'Q'QX)^{-1}\,\underline{x}_1$ augmente avec ρ de sorte que CORC sera moins susceptible d'être préféré à PW lorsque ρ est grand.

IV.6. Validité des tests

Jusqu'à présent, nous nous sommes limités à étudier l'efficacité de CORC par rapport à PW. Un autre point tout aussi important s'ajoute en faveur de l'estimation conditionnelle. Lorsque $\sigma_1^2 \neq \sigma_e^2/(1-\rho^2)$, il est le seul (avec les MCG) dont les écarts-types estimés sont sans biais. Il est donc le seul à donner des tests de Student valides, indépendamment de la variance de la première observation.

En utilisant PW, la matrice de covariance est estimée par $(X'P'PX)^{-1}$. Celle-ci sera différente de la matrice véritable (équation

(11)) sauf si $a^2 = 1/(1-\rho^2)$, soit un processus stationnaire. De plus, l'estimateur de la variance (s^2) sera également biaisé puisque :

$$s^2(N-K) = \hat{\underline{e}}' \; \hat{\underline{e}}$$

où $\hat{\underline{e}} = P(\underline{y} - X\hat{\underline{B}})$

$$= P(I - X(X'P'PX)^{-1} X'P'P) \; \underline{y}$$

$$= P(I - X(X'P'PX)^{-1} X'P'P) \; \underline{u}$$

en nous servant des propriétés de la trace (tr), nous aurons

$$E(\hat{\underline{e}}' \; \hat{\underline{e}}) = E[\underline{u}'(I - P'PX(X'P'PX)^{-1} X') \; P'P(I - X(X'P'PX)^{-1} X'P'P) \; \underline{u}]$$

$$= E(\underline{u}'(P'P - P'PX(X'P'PX)^{-1} X'P'P) \; \underline{u})$$

$$= E(\text{tr } \underline{u}'P'P\underline{u}) - E(\text{tr } \underline{u}'P'PX(X'P'PX)^{-1} X'P'P\underline{u})$$

$$= E(\text{tr } P\underline{u}\underline{u}'P') - E(\text{tr } P\underline{u}\underline{u}'P'PX(X'P'PX)^{-1} X'P')$$

$$= \text{tr } P \; E(\underline{u}\underline{u}') \; P' - \text{tr } P \; E \; (\underline{u}\underline{u}') \; P'PX(X'P'PX)^{-1} X'P'$$

$$= \sigma_e^2(\text{tr } PVP' - \text{tr } PVP'PX(X'P'PX)^{-1} X'P')$$

$$= \sigma_e^2 \; \text{tr}(PVP'(I - PX(X'P'PX)^{-1} X'P'))$$

Posons $M = (I - PX(XP'PX)^{-1} X'P')$, une matrice idempotente; $\text{tr } M = N-K$. En partitionnant M et en nous servant de l'équation (20), nous obtenons

$$E(\hat{\underset{\sim}{e}}'\,\hat{\underset{\sim}{e}}) = \sigma_e^2 \,\mathrm{tr} \begin{bmatrix} a^2(1-\rho^2) & \underset{\sim}{0}' \\ \underset{\sim}{0} & I_{N-1} \end{bmatrix} \begin{bmatrix} m_{11} & \underset{\sim}{m}_{12} \\ \underset{\sim}{m}_{21} & M_{22} \end{bmatrix}$$

$$= \sigma_e^2(a^2(1-\rho^2)\; m_{11} + \mathrm{tr}\; M_{22})$$

Pour que s^2 soit sans biais, il faut que, $a^2(1-\rho^2)\; m_{11} + \mathrm{tr}\; M_{22} = N-K$. Cependant,

$$\mathrm{tr}\; M = m_{11} + \mathrm{tr}\; M_{22} = N-K$$

donc la condition n'est réalisée que si

$$a^2(1-\rho^2)\; m_{11} = m_{11}$$

c'est-à-dire $a^2 = 1/(1-\rho^2)$ si $m_{11} \neq 0$.

Un argument similaire peut être appliqué aux MCO. Ainsi, même si, dans certains cas, ils sont plus efficaces que CORC, comme le soulignent certains auteurs (e.g., Maeshiro (1976), ils ne peuvent être considérés comme une alternative valable puisque les écarts-types estimés sont biaisés et les tests de Student appliqués aux coefficients ne sont plus valides.

IV.7. Conclusion

Nous avons vu que lorsque ρ est connu, il existe une borne unique qui détermine le domaine d'efficacité de CORC et PW. De plus, seule l'estimation conditionnelle donne toujours des écarts-types corrects pour les coefficients de régression.

Cependant, en pratique, ρ et σ_1^2 ne sont jamais connu. Si la taille d'échantillon est finie, l'estimation de ρ peut affecter la distribution des coefficients de régression et les écarts-types estimés. Comme les distributions deviennent trop difficiles à analyser théoriquement, nous aurons maintenant recours aux techniques de Monte-Carlo pour étudier ce problème.

CHAPITRE V

Estimation de ρ

V.1. Les estimateurs

Plusieurs méthodes d'estimation sont disponibles lorsque ρ est inconnu. De façon générale, elles peuvent être divisées en cinq groupes :

1) les méthodes "stationnaires dérivées des MCG; on remplace ρ par un estimateur convergent et l'on utilise la matrice de transformation de PW (P);
2) les méthodes "conditionnelles" dérivées des MCG; on utilise ici la matrice de transformation de CORC (Q) où ρ est remplacé par un estimé convergent;
3) le maximum de vraisemblance "stationnaire"; il s'agit de maximiser la fonction de vraisemblance sous l'hypothèse de stationnarité;
4) le maximum de vraisemblance conditionnel; on maximise alors la fonction de vraisemblance conditionnelle à y_1;
5) toutes les autres méthodes; plusieurs autres méthodes ont été proposées pour tenir compte de l'autocorrélation, par exemple : estimation bayésienne, différences premières, etc., etc.;

(Notons que les estimateurs de vraisemblance sont des cas spéciaux des deux premiers groupes).

Pour chacun des groupes, plusieurs techniques différentes peuvent être appliquées. Pour ne pas alourdir inutilement la discussion, nous ne présenterons qu'un exemple concret pour chaque cas. Ce sont les techniques les plus utilisées où celles qui ont présenté les meilleures performances dans les études empiriques précédentes. Toutes ces méthodes sont bien connues et sont présentées dans tous les manuels de base (sauf le maximum de vraisemblance stationnaire) de sorte que nous nous contenterons d'une présentation très brève.

La procédure qui a été la plus utilisée est sûrement celle proposée par Cochrane et Orcutt (1949). C'est une méthode itérative où un premier estimé de ρ, obtenu des résidus des MCO, est utilisé dans la matrice Q (cf. chap. IV) pour obtenir un vecteur $(\hat{\underline{\beta}})$ de paramètres estimés. Une nouvelle valeur de $\hat{\rho}$ est calculée à partir des nouveaux résidus non transformés $(\underline{u})$. Ces deux étapes, calcul de $\hat{\rho}$ et calcul de $\hat{\underline{\beta}}$, sont itérées jusqu'à convergence. Comme on sait, à chaque itération,

$$\hat{\rho} = \sum_{t=2}^{N} \hat{u}_t \hat{u}_{t-1} \Big/ \sum_{t=2}^{N} \hat{u}_{t-1}^2$$

Cette méthode est encore très souvent considérée comme une approximation des MCG. Cependant, Cooper (1972) entre autres, remarque qu'il s'agit en fait du maximum de vraisemblance conditionnel (MVC) à la première observation. Nous conserverons ce terme (MVC) pour bien marquer la différence avec CORC qui suppose ρ connu.

Depuis son inclusion dans "TSP", l'algorithme du maximum de vraisemblance stationnaire (MVS), présenté par Beach et MacKinnon (1978), devient de plus en plus populaire. Selon les auteurs, cet estimateur devrait toujours être employé en pratique parce que :

- $\hat{\rho}$ est contraint dans le domaine d'invertibilité et de stationnarité. $|\hat{\rho}| < 1$;
- la convergence est assurée;
- l'algorithme converge très rapidement;
- les estimateurs sont plus efficaces que ceux de MVC, en particulier pour la constante.

C'est une méthode itérative qui utilise la matrice de transformation "P" pour obtenir le vecteur $\hat{\underline{\beta}}$. Son originalité provient de l'estimation de ρ. En maximisant le log de la fonction de vraisemblance par rapport à $\hat{\rho}$, nous obtenons un polynôme du 3^e degré en $\hat{\rho}$. Beach et MacKinnon (1978) ont démontré que ce polynôme ne possédait qu'une et une seule racine entre -1 et 1. Soit,

$$\partial L/\partial\hat{\rho} = \rho^3 + a\rho^2 + b\rho + c = 0$$

La racine désirée est obtenue directement par la formule :

$$\hat{\rho} = -2\sqrt{-p/3}\,\cos(d/3 + \pi/3) - a/3$$

où $p = b - a^2/3$

$q = c - ab/3 + 2a^3/27$

$a = -(T-2)\,\Sigma\,\hat{u}_t\,\hat{u}_{t-1}/(T-1)\,(\Sigma\,\hat{u}^2_{t-1} - \hat{u}^2_1)$

$b = ((T-1)\,\hat{u}^2_1 - T\,\Sigma\,\hat{u}^2_{t-1} - \Sigma\,\hat{u}^2_t)/(T-1)\,(\Sigma\,\hat{u}^2_{t-1} - \hat{u}^2_1)$

$c = T\,\Sigma\,\hat{u}_t\,\hat{u}_{t-1}/(T-1)\,(\Sigma\,\hat{u}^2_{t-1} - \hat{u}^2_1)$

$d = \cos^{-1}((q\sqrt{27})/(2\rho\sqrt{-p}))$

et les $\hat{u}$ sont les résidus non transformés obtenus de l'étape précédente.

Les méthodes d'estimation dérivées des MCG sont maintenant utilisées assez rarement. Néanmoins elles peuvent nous donner des points de comparaison intéressants.

La méthode conditionnelle qui semble avoir donné les meilleurs résultats est celle de Durbin (1960), dénotée DUR. $\hat{\rho}$ est d'abord obtenu du coefficient de y_{t-1} dans le modèle de régression :

$$y_t = \hat{\rho}\,y_{t-1} + \hat{\alpha} + \underline{x}'_t\,\hat{\underline{\beta}} + \underline{x}'_{t-1}\,\hat{\underline{\gamma}} + \hat{e}_t\;; \qquad t = 2, \ldots, N$$

Le vecteur de coefficients $\hat{\underline{\beta}}$ est ensuite obtenu de la façon habituelle en substituant $\hat{\rho}$ dans la matrice Q.

Selon Park et Mitchell (1980), l'estimateur le plus efficace dans le cas stationnaire est celui qui minimise (sur $\hat{\underline{\beta}}$ et $\hat{\rho}$) la somme des carrés des résidus[1] (SCR = $\sum_{t=1}^{N} \hat{e}_t^2$) de tout l'échantillon. C'est encore une procédure itérative qui ne diffère du MVS que par la formule d'estimation de ρ. Nous avons ici :

$$\hat{\rho} = \sum_{2}^{N} \hat{u}_t \hat{u}_{t-1} / \sum_{2}^{N-1} \hat{u}_t^2$$

Nous conserverons ici le terme utilisé par Park et Mitchell (1980) pour désigner cette méthode, soit : Prais-Winsten itéré (PW1). Rappelons qu'il faut généralement ajouter une contrainte pour s'assurer que $|\hat{\rho}| < 1$ sinon l'estimateur n'existe pas à cause du terme $(1-\rho^2)^{1/2}$.

Parmi les autres méthodes, la seule qui présente ici un certain intérêt est l'estimation en première différence. $\hat{\rho}$ est alors tout simplement contraint à l'unité et la constante n'est pas estimée.

En plus des MCO, ces cinq méthodes feront l'objet de l'étude de Monte-Carlo que nous présenterons après la revue des études empiriques précédentes.

V.2. Revue de la littérature

Plusieurs études de Monte-Carlo ont déjà été réalisées sur l'efficacité de différents estimateurs en présence d'autocorrélation du premier ordre. Une brève revue de celles-ci nous sera précieuse pour établir le plan de notre propre expérience et pourra éventuellement servir de point de comparaison pour nos résultats.

---

[1] Il s'agit évidemment des résidus transformés sous l'hypothèse de stationnarité ($\hat{e} = \hat{P}(\underline{y} - \underline{X}\hat{\beta})$).

Rao et Grilliches (1969), Pfaffenberger et Diellman (1979) ainsi que Spitzer (1979), ont tous utilisé un même modèle très simple avec une seule variable explicative ($x_t = \lambda x_{t-1} + v_t$; $v_t \sim N.i.d.(0, \sigma_v^2)$) sans terme constant. Les seules différences entre eux sont les estimateurs considérés et le nombre de reprises (50 pour Rao et Grilliches, 100 pour les deux autres). La taille de l'échantillon était fixée à 20 et les paramètres ρ et λ variaient par intervalles fixes entre -0,99 et 0,99. Les MCO se sont toujours révélés plus efficaces pour $|\rho| < 0,3$. Si Rao et Grilliches (1969) estiment que le gain d'efficacité dû à l'inclusion de la première observation n'est pas important, cette conclusion est cependant infirmée par les deux autres études, en particulier pour les méthodes itératives. Les premiers n'ont utilisé que des estimateurs en deux étapes.

Hong et L'Espérance (1973) ont étudié plus particulièrement les effets de la taille d'échantillon, de la collinéarité entre les variables explicatives et de la valeur des coefficients (β). Ils ajoutent au modèle précédent un terme constant et une autre variable présentant différents degrés de collinéarité avec la première. Comme Rao et Grilliches (1969) ils ne considèrent que des estimateurs en deux étapes. Ils observent d'abord que les MCO deviennent relativement plus efficaces lorsque la taille de l'échantillon est petite (N=10). Ni la valeur des paramètres, ni la présence de collinéarité ne semblent affecter l'efficacité relative des estimateurs considérés. Selon eux, il serait préférable de laisser tomber la première observation lorsque ρ est très grand.

L'étude de Beach et MacKinnon (1978) est sûrement une des plus importantes, même si elle se limite à deux estimateurs (MVC et MVS), trois valeurs de ρ (0,6, 0,8, 0,99) et deux variables explicatives : une constante et une autre variable qui prend trois formes différentes. L'avantage de cette simplicité est de présenter des conclusions claires et nettes. Dans tous les cas étudiés, le MVS était préférable au MVC, en particulier pour la constante.

La dernière étude connue est celle de Park et Mitchell (1980). Elle se distingue surtout par le nombre élevé de reprises, l'utilisation de variables explicatives réelles et les tests de Student appliqués aux paramètres. Généralement, les conclusions sont les mêmes que celles de Beach et MacKinnon (1978) mais de plus, l'erreur quadratique moyenne (EQM) de MVC devient quelquefois infiniment grande. D'autre part, les tests de Student appliqués au coefficient de la pente rejettent beaucoup trop souvent l'hypothèse nulle sauf si "X" est un bruit blanc. Malheureusement, ces tests n'ont été effectués que pour la pente, avec les estimateurs qui incluent la première observation.

Evidemment, toutes ces expériences ont été faites avec le modèle stationnaire ($a^2 = 1/(1-\rho^2)$). Il est donc normal que l'estimation conditionnelle soit moins efficace.

V.3. Plan

Pour obtenir le maximum d'information de cette étude empirique, nous avons choisi d'utiliser la même spécification que celle de Beach et MacKinnon (1978) en la complétant de différentes façons. Comme nous l'avons dit, cette étude a l'avantage d'être claire, elle est aussi bien connue et les modèles utilisés semblent assez réalistes. Nous présenterons d'abord le plan complet de Beach et MacKinnon (1978) puis nous verrons comment notre étude se distingue de celle-ci.

Soit le modèle simple :

$$y_t = \beta_0 + \beta_1 x_t + u_t \qquad t = 1, \ldots, N$$

$$u_t = \rho u_{t-1} + e_t \qquad t = 2, \ldots, N$$

$$u_1 = e_1/(1-\rho^2)^{1/2}$$

$$e_t \underset{\sim}{\text{ind}} N(0,\ 0{,}0036) \qquad t = 1, \ldots, N$$

$$\beta_0 = \beta_1 = 1$$

$$\rho = 0{,}6;\ 0{,}8;\ 0{,}99.$$

Trois séries de variables explicatives différentes ont été utilisées, les deux premières correspondent à une taille d'échantillon de 20 et la troisième de dimension 50. Dans le premier cas, x comportait une tendance telle que :

$$x_t = \exp(0{,}04t) + w_t \qquad t = 1, \ldots, 20$$

et $$w_t \underset{\sim}{\text{ind}} N(0,\ 0{,}009) \qquad t = 1, \ldots, 20$$

Dans le second cas, il s'agissait d'un bruit blanc :

$$x_t \underset{\sim}{\text{ind}} N(0,\ 0{,}0625) \qquad t = 1, \ldots, 20$$

La troisième spécification était équivalente à la première (tendance) sauf pour le nombre d'observations (50).

Enfin, pour chaque paramètre estimé ($\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\rho}$) ils ont calculé le biais, la racine de l'erreur quadratique moyenne (REQM) et le nombre de fois que l'estimateur du MVS était plus près de la vraie valeur que celui du MVC.

L'aspect le plus original de notre étude est évidemment l'hypothèse retenue sur la première observation, qui n'a encore fait l'objet d'aucune recherche. Pour vérifier objectivement nos résultats théoriques, nous devons considérer autant de valeurs de part et d'autre de la borne en plus de cette dernière. Pour chaque valeur de ρ, cinq spécifications différentes seront retenues pour l'écart-type de la première observation :

$$\sigma_1 = 0{,}5\ \sigma_e/\sqrt{1-\rho^2}$$

$$= \sigma_e/\sqrt{1-\rho^2} \qquad \text{(hypothèse de stationnarité)}$$

$$= \sigma^*\text{, tel que } \operatorname{VAR}(\hat{\underline{\beta}}(\text{CORC})) = \operatorname{VAR}(\hat{\underline{\beta}}(\text{PW}))$$

$$= 1{,}5\ \sigma^*$$

$$= 2\ \sigma^*$$

L'intervalle ainsi couvert nous apparaît plus que suffisant pour dégager des conclusions intéressantes sur le comportement des différents paramètres estimés.

Nous couvrirons également trois autres sujets qui n'ont été abordés dans aucune étude précédente.

1. Comme nous le disions au début de ce chapitre, nous inclurons l'estimation en première différence (DIFF) puisque la pratique est assez courante lorsque ρ est grand.

2. Parce que la distribution de $\hat{\alpha} = \hat{\beta}_0(1-\hat{\rho})$ est au moins aussi importante que celle de $\hat{\beta}_0$ dans les modèles destinés à la prévision $(y_{t+1} = \hat{\alpha} + \hat{\rho}\ y_t + \hat{\beta}_1(x_{t+1} - \hat{\rho}\ x_t))$ les statistiques habituelles seront compilés pour ce paramètre.

3. Enfin, personne n'a jugé bon d'étudier la distribution de l'estimateur de la variance (s^2). Nous n'insisterons pas sur l'importance de cette statistique, mais son inclusion apparaît primordiale.

D'autres extensions seront encore apportées au plan de Beach et MacKinnon (1978). D'abord, nous ajouterons une série de variables explicatives sans tendance avec une taille d'échantillon de 50 ($x_t \underset{\sim}{\text{ind}} N(0,\ 0{,}0625)$). En plus de MVC, MVS et DIFF, trois autres méthodes d'estimation seront considérées : MCO, DUR, PWI (cf. section V.1). Le critère de convergence pour les méthodes itératives est le même que Beach et MacKinnon (1×10^{-5}). Pour PWI, nous contraignons

$\hat{\rho} = 0{,}99999$ si $\hat{\rho} \geq 1{,}0$ comme l'ont fait Park et Mitchell (1980). Le nombre de reprises sera porté à 1 000 plutôt que 200. Enfin, la validité des tests de Student (sur $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\alpha}$) sera étudiée avec des niveaux de signification de 5, 10 et 20%.

Aucune expérience empirique sur l'autocorrélation n'a encore présenté une telle envergure. Seuls Park et Mitchell (1980) ont un nombre aussi élevé de reprises. Nous comptons 60 modèles différents qui varient selon la forme de la variable explicative, le nombre d'observations, la valeur de ρ et celle de σ_1^2. Dans la plupart des cas (MVC, DUR, MVS, PWI) nous avons cinq paramètres à estimer ($\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\rho}$, $\hat{\alpha}$, s^2) pour lesquels nous compilerons le biais, la racine de l'erreur quadratique moyenne (REQM) et le nombre de fois que l'estimateur considéré est plus près de la vraie valeur que MVS (NF).

CHAPITRE VI

Résultats empiriques

VI.1. Covariances théoriques

La première étape dans la mise en oeuvre du plan a été de générer les quatre séries de variables explicatives[1]. Ensuite, pour chacune d'elles, nous avons calculé les cinq valeurs de σ_1, associées à chaque valeur de ρ. Nous pouvions dès lors calculer les matrices de covariance théoriques (ρ connu) pour chacun des 60 modèles d'après les formules du chapitre IV, soit :

$$\Sigma(MCO) = \sigma_e^2 (X'X)^{-1} X'VX (X'X)^{-1}$$

$$\Sigma(CORC) = \sigma_e^2 (X'Q'QX)^{-1}$$

$$\Sigma(PW) = \sigma_e^2 (X'P'PX)^{-1} X'P'PVP'PX (X'P'PX)^{-1}$$

$$\Sigma(MCG) = \sigma_e^2 (X'V^{-1}X)^{-1}$$

Le tableau 2 présente les variances des paramètres, ainsi que les valeurs de l'écart-type de la première observation (σ_1) associées à chaque modèle. Si l'estimation de ρ n'affecte pas l'efficacité des estimateurs, les REQM devraient être sensiblement égales à ces valeurs théoriques. Nous reviendrons un peu plus loin à ces comparaisons. Auparavant, quelques remarques peuvent s'avérer intéressantes pour relier cette partie du travail aux chapitres précédents.

[1]La liste complète est donnée au tableau 21, annexe C.

Tableau 2

Variances de $\hat{\beta}_0$ et $\hat{\beta}_1$ (ρ connu)

		$\rho = 0,6$					$\rho = 0,8$					$\rho = 0,99$				
X1 (avec tendance, N = 0)	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
	σ_1	.038	.075	.134	.201	.269	.050	.100	.226	.340	.453	.213	.425	2.958	4.438	5.917
	V(B0) MCO	.0122	.0152	.0242	.0404	.0632	.0267	.0386	.1044	.2065	.3495	.1136	.2873	11.2619	25.2698	44.8808
	CORC	.0273	.0273	.0273	.0273	.0273	.1056	.1056	.1056	.1056	.1056	10.0388	10.0388	10.0388	10.0388	10.0388
	PW	.0083	.0131	.0273	.0532	.0893	.0143	.0283	.1056	.2257	.3938	.0772	.2325	10.0388	22.5554	40.0786
	MCQ	.0058	.0131	.0202	.0236	.0251	.0112	.0283	.0670	.0839	.0921	.0747	.2325	5.1356	7.0457	8.1020
	V(B1) MCO	.0049	.0058	.0085	.0134	.0202	.0112	.0144	.0319	.0592	.0973	.0454	.0481	.2209	.4414	.7501
	CORC	.0091	.0091	.0091	.0091	.0091	.0288	.0288	.0288	.0288	.0288	.1122	.1122	.1122	.1122	.1122
	PW	.0036	.0050	.0091	.0165	.0270	.0069	.0103	.0288	.0576	.0979	.0258	.0272	.1122	.2206	.3725
	MCQ	.0029	.0050	.0070	.0080	.0084	.0062	.0103	.0195	.0236	.0256	.0258	.0272	.0697	.0862	.0954
X2, (bruit blanc, N = 20)	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
	σ_1	.038	.075	.114	.171	.228	.050	.100	.159	.238	.318	.213	.425	1.503	2.254	3.005
	V(B0) MCO	.0009	.0010	.0011	.0014	.0018	.0030	.0035	.0045	.0066	.0095	.0565	.1693	1.8966	4.2438	7.5300
	CORC	.0012	.0012	.0012	.0012	.0012	.0047	.0047	.0047	.0047	.0047	1.8949	1.8949	1.8949	1.8949	1.8949
	PW	.0009	.0010	.0012	.0016	.0023	.0025	.0032	.0047	.0079	.0122	.0524	.1654	1.8949	4.2452	7.5356
	MCQ	.0007	.0010	.0011	.0011	.0012	.0018	.0032	.0040	.0044	.0045	.0445	.1654	1.0302	1.3800	1.5662
	V(B1) MCO	.0047	.0055	.0069	.0101	.0146	.0075	.0095	.0135	.0219	.0337	.0220	.0231	.0392	.0612	.0919
	CORC	.0032	.0032	.0032	.0032	.0032	.0027	.0027	.0027	.0027	.0027	.0022	.0022	.0022	.0022	.0022
	PW	.0028	.0030	.0032	.0038	.0046	.0025	.0026	.0027	.0028	.0031	.0022	.0022	.0022	.0022	.0022
	MCQ	.0026	.0030	.0031	.0032	.0032	.0025	.0026	.0026	.0026	.0026	.0022	.0022	.0022	.0022	.0022

d : $\sigma_1 (1 - \rho^2)^{1/2}/\sigma_e$, d* : valeur à la borne,

$V(\beta_0)$: $Var(\hat{\beta}_0)$,

$V(\beta_1)$: $Var(\hat{\beta}_1)$.

Tableau 2 (suite)

Variances de $\hat{\beta}_0$ et $\hat{\beta}_1$ (ρ connu)

		ρ = 0,6					ρ = 0,8					ρ = 0,99				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X3, (avec tendance, N = 50)	σ_1	.038	.075	.111	.167	.223	.050	.100	.158	.237	.316	.213	.425	1.638	2.457	3.276
	V(B0) MCO	.0016	.0017	.0019	.0022	.0028	.0053	.0059	.0073	.0100	.0139	.0719	.2033	2.6261	5.8735	10.4199
	CORC	.0020	.0020	.0020	.0020	.0020	.0078	.0078	.0078	.0078	.0078	2.5578	2.5578	2.5578	2.5578	2.5578
	PW	.0014	.0016	.0020	.0027	.0038	.0040	.0053	.0078	.0131	.0204	.0588	.1874	2.5578	5.7350	10.1831
	MCG	.0011	.0016	.0018	.0019	.0019	.0028	.0053	.0066	.0072	.0075	.0515	.1874	1.3726	1.8482	2.1034
	V(B1) MCO	.0001	.0001	.0001	.0001	.0002	.0004	.0004	.0004	.0006	.0007	.0037	.0042	.0131	.0250	.0416
	CORC	.0001	.0001	.0001	.0001	.0001	.0004	.0004	.0004	.0004	.0004	.0071	.0071	.0071	.0071	.0071
	PW	.0001	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0006	.0008	.0023	.0026	.0071	.0131	.0215
	MCG	.0001	.0001	.0001	.0001	.0001	.0002	.0003	.0004	.0004	.0004	.0023	.0026	.0048	.0057	.0062
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X4, (bruit blanc, N = 50)	σ_1	.038	.075	.108	.163	.217	.050	.100	.148	.222	.296	.213	.425	1.047	1.571	2.094
	V(B0) MCO	.0004	.0004	.0004	.0005	.0005	.0016	.0016	.0018	.0020	.0024	.0692	.1539	.7259	1.5821	2.7807
	CORC	.0005	.0005	.0005	.0005	.0005	.0018	.0018	.0018	.0018	.0018	.7347	.7347	.7347	.7347	.7347
	PW	.0004	.0004	.0005	.0005	.0007	.0014	.0016	.0018	.0025	.0034	.0578	.1452	.7347	1.6172	2.8527
	MCG	.0003	.0004	.0004	.0005	.0005	.0011	.0016	.0017	.0018	.0018	.0426	.1452	.4399	.5661	.6293
	V(B1) MCO	.0015	.0015	.0015	.0015	.0015	.0020	.0021	.0022	.0025	.0029	.0037	.0038	.0042	.0047	.0054
	CORC	.0008	.0008	.0008	.0008	.0008	.0007	.0007	.0007	.0007	.0007	.0006	.0006	.0006	.0006	.0006
	PW	.0008	.0008	.0008	.0009	.0009	.0007	.0007	.0007	.0007	.0007	.0006	.0006	.0006	.0006	.0006
	MCG	.0008	.0008	.0008	.0008	.0008	.0007	.0007	.0007	.0007	.0007	.0006	.0006	.0006	.0006	.0006

Comme il se doit, PW est toujours équivalent aux MCG lorsque $d = 1$ et les variances de CORC et PW sont égales, à la borne (d^*). On vérifie aisément que la différence des matrices de covariance est positive semi-définie conformément au théorème 2. Notons qu'il y a très peu de différences entre CORC, PW et MCG pour la pente lorsque la variable explicative est un bruit blanc.

Le comportement des MCO est beaucoup plus instable. Dans la majorité des cas, l'estimateur de la pente ($\hat{\beta}_1$) est moins efficace que ceux de CORC et PW, mais la constante se situe plutôt entre les deux. En général, l'efficacité des MCO augmente lorsque ρ, d ou N baissent.

Si nous portons notre attention sur la valeur de la borne (d^*) nous voyons qu'elle augmente avec ρ, mais baisse avec N. Elle est aussi relativement plus grande lorsque la variable explicative comporte un trend. Elle atteint un maximum de 6,9 pour la première variable explicative (trend, N=20) lorsque ρ est très grand (0,99).

Après ces quelques observations sur les variances théoriques, nous pouvons maintenant concentrer nos efforts sur les résultats obtenus lorsque ρ est estimé.

VI.2. Analyse des résultats empiriques

Etant donné le grand nombre de statistiques qui ont été compilées, l'analyse des résultats devra se faire de façon très systématique, en nous limitant aux questions les plus importantes, de façon à réduire le plus possible toute source de confusion. Parce que MVS et MVC sont les méthodes les plus utilisées, leur comportement sera étudié de plus près que les autres même si elles ne constituent pas toujours le meilleur choix. Ceci nous permettra de simplifier énormément l'analyse sans affecter vraiment la portée pratique des résultats.

Comme la distribution des coefficients de régression $(\hat{\beta}_0, \hat{\beta}_1)$ est le problème qui nous intéresse plus particulièrement, nous lui consacrons la plus grande partie de ce chapitre. Nous examinerons ensuite plus brièvement la distribution des autres paramètres estimés $(\hat{\rho}, \hat{\alpha}, s^2)$, puis la validité des tests et enfin les performances en terme de coût de traitement.

VI.3. Distribution de $\hat{\beta}_0$ et $\hat{\beta}_1$

Dans les chapitres précédents, nous avons soulevé plusieurs questions concernant l'efficacité de $\hat{\beta}_0$ et $\hat{\beta}_1$.

- L'estimation de ρ affecte-t-elle la valeur de la borne?
- La variance des estimateurs conditionnels reste-t-elle indépendante de σ_1^2?
- Quand est-il préférable de laisser tomber la première observation?
- L'estimation de ρ affecte-t-elle la variance des coefficients?

Avant de répondre à ces questions, nous devons nous demander si les estimateurs sont bien centrés.

VI.3.1. Biais de $\hat{\beta}_0$ et $\hat{\beta}_1$

Même si toutes les études précédentes concluent que le biais des coefficients $\hat{\beta}_0$ et $\hat{\beta}_1$ est nul, une vérification empirique s'impose.

En examinant les tableaux 3 et 4, nous remarquons que certaines valeurs du biais sont très élevées, en particulier pour la constante. Dans un cas, il s'élève même jusqu'à -8,1 (variable sans tendance, N=50), $\rho=0,99$, $d=2d^*$, MVC, $\hat{\beta}_0$). Si on se contente de cette statistique simple, il nous faut admettre que l'estimateur n'est pas toujours centré. Par contre, en tenant compte de la variance les conclusions sont totalement différentes.

En fait, le biais qui est calculé est tout simplement la moyenne échantillonnale d'une population dont chacune des observations $(\hat{\beta} - \beta)$ est indépendante pour un estimateur (MCO, MVC, etc.) et un coefficient (β_0, β_1) donné. Pour chacune des valeurs calculées, nous pouvons effectuer un test simple, soit :

$$H0 : \overline{\beta} - \beta = 1/1000 \sum_{i=1}^{1000} (\hat{\beta}_i - \beta) = 0$$

$$H1 : \overline{\beta} - \beta \neq 0$$

Sous l'hypothèse nulle, l'application du théorème limite-centrale nous donne :

Tableau 3

Biais de $\hat{\beta}_0$

	D=	ρ = 0,6 .5	ρ = 0,6 1.0	ρ = 0,6 D*	ρ = 0,6 1.5D*	ρ = 0,6 2D*	ρ = 0,8 .5	ρ = 0,8 1.0	ρ = 0,8 D*	ρ = 0,8 1.5D*	ρ = 0,8 2D*	ρ = 0,99 .5	ρ = 0,99 1.0	ρ = 0,99 D*	ρ = 0,99 1.5D*	ρ = 0,99 2D*
X1, (avec tendance, N = 20)	MCO	.003 (.93)	.003 (.71)	-.006 (1.21)	.006 (.92)	.001 (.12)	.002 (.44)	.008 (1.20)	.014 (1.34)	.008 (.58)	-.013 (.69)	-.003 (.30)	.003 (.19)	.051 (.47)	.087 (.57)	.180 (.87)
	DUR	.005 (.97)	.002 (.34)	.000 (.06)	.005 (.75)	-.005 (.83)	.003 (.32)	.003 (.31)	.022 (1.84)	.012 (.90)	.154 (1.08)	.002 (.10)	-.012 (.32)	.067 (.65)	.154 (.98)	.105 (.51)
	MVC	-.001 (.11)	.003 (.34)	.004 (.49)	.001 (.09)	.002 (.18)	-.005 (.32)	.012 (.84)	.019 (1.11)	-.008 (.56)	.017 (1.07)	.033 (1.39)	.030 (.87)	.085 (.88)	.122 (.92)	.086 (.52)
	PWI	.002 (.55)	.003 (.92)	-.005 (.91)	.010 (1.24)	.007 (.61)	.003 (.73)	.006 (1.00)	.011 (1.06)	.001 (.05)	-.007 (.33)	-.002 (.26)	.007 (.45)	.054 (.52)	.088 (.59)	.185 (.91)
	MVB	.002 (.61)	.003 (.91)	-.005 (1.03)	.009 (1.21)	.006 (.56)	.003 (.68)	.006 (1.09)	.012 (1.11)	.002 (.14)	-.008 (.41)	-.003 (.31)	.006 (.39)	.053 (.51)	.087 (.57)	.182 (.89)
X2, (bruit blanc, N = 20)	MCO	-.003 (2.81)	-.000 (.15)	-.001 (1.37)	.002 (1.33)	-.000 (.04)	.001 (.70)	-.003 (1.45)	.001 (.54)	.008 (3.09)	.004 (1.18)	-.000 (.06)	-.009 (.68)	.063 (1.46)	.021 (.32)	.142 (1.71)
	DUR	-.003 (2.74)	-.001 (.88)	-.001 (.72)	.001 (1.14)	.002 (.92)	.001 (.27)	-.001 (.45)	.007 (1.62)	-.026 (.49)	.010 (1.97)	-.109 (1.93)	-.004 (.19)	.216 (1.89)	.044 (.31)	-.313 (.88)
	MVC	-.004 (3.07)	-.001 (.78)	-.001 (.77)	.001 (1.26)	.001 (.75)	-.001 (.35)	.003 (.44)	.004 (1.37)	.035 (1.00)	-.009 (.90)	-.008 (.17)	.157 (1.13)	.014 (.21)	-.142 (1.00)	.097 (.27)
	PWI	-.003 (3.07)	.000 (.26)	-.003 (2.12)	.001 (.82)	.000 (.06)	.000 (.02)	-.003 (1.42)	.001 (.59)	.011 (3.12)	.005 (.92)	-.002 (.23)	-.008 (.61)	.066 (1.51)	.021 (.32)	.142 (1.71)
	MVB	-.003 (3.02)	.000 (.16)	-.002 (1.99)	.002 (1.11)	.000 (.09)	.000 (.21)	-.003 (1.49)	.001 (.54)	.010 (3.14)	.004 (.98)	-.001 (.16)	-.008 (.62)	.065 (1.49)	.021 (.32)	.142 (1.71)
X3, (avec tendance, N = 50)	MCO	.003 (2.03)	-.001 (.78)	.001 (.37)	.001 (.36)	-.000 (.26)	.002 (.82)	.002 (.97)	.001 (.40)	-.001 (.17)	-.001 (.39)	.006 (.74)	-.010 (.66)	.033 (.62)	.135 (1.87)	.094 (.92)
	DUR	.003 (2.41)	-.001 (.85)	.001 (.59)	.001 (.70)	-.002 (1.46)	.001 (.57)	.000 (.16)	.000 (.05)	.000 (.02)	-.001 (.43)	.013 (.94)	-.011 (.53)	.082 (1.22)	.234 (2.16)	.130 (.54)
	MVC	.003 (2.42)	-.001 (.85)	.001 (.60)	.001 (.71)	-.002 (1.49)	.001 (.54)	.001 (.35)	.000 (.05)	.000 (.13)	-.006 (1.08)	.037 (.53)	.007 (.20)	-.023 (.31)	.122 (1.15)	.167 (1.44)
	PWI	.003 (2.08)	-.000 (.33)	.001 (.43)	.000 (.26)	.001 (.52)	.001 (.43)	.002 (.92)	.000 (.08)	-.000 (.08)	-.002 (.35)	.001 (.11)	-.012 (.82)	.035 (.66)	.133 (1.86)	.091 (.90)
	MVB	.003 (2.08)	-.000 (.36)	.001 (.43)	.001 (.29)	.001 (.42)	.001 (.47)	.002 (.94)	.000 (.15)	-.000 (.09)	-.002 (.35)	.001 (.17)	-.012 (.81)	.034 (.65)	.134 (1.86)	.091 (.90)
X4, (bruit blanc, N = 50)	MCO	-.001 (1.23)	.000 (.44)	-.000 (.01)	.001 (1.12)	.001 (.99)	-.001 (.68)	.001 (.58)	-.002 (1.55)	-.001 (.55)	.003 (1.88)	.009 (1.05)	.001 (.07)	.033 (1.25)	-.005 (.14)	.040 (.75)
	DUR	-.001 (1.44)	.000 (.50)	-.000 (.16)	.001 (1.26)	.000 (.57)	-.002 (1.30)	-.001 (.57)	-.001 (.97)	-.000 (.21)	.002 (1.64)	-.292 (.93)	.033 (1.08)	.339 (1.01)	-.383 (.71)	.341 (.53)
	MVC	-.001 (1.45)	.000 (.47)	-.000 (.18)	.001 (1.26)	.000 (.55)	-.003 (1.39)	-.000 (.13)	-.001 (.86)	-.000 (.16)	.002 (1.63)	.244 (1.50)	-.216 (1.01)	.177 (1.87)	.158 (.84)	-8.129 (1.13)
	PWI	-.001 (1.53)	.000 (.48)	-.000 (.01)	.001 (.96)	.001 (.54)	-.001 (.54)	.001 (.43)	-.002 (1.28)	-.001 (.50)	.006 (1.71)	.006 (.75)	.003 (.23)	.027 (1.02)	-.010 (.26)	.038 (.71)
	MVB	-.001 (1.51)	.000 (.48)	-.000 (.01)	.001 (1.06)	.001 (.74)	-.001 (.57)	.001 (.48)	-.002 (1.36)	-.001 (.44)	.005 (1.72)	.007 (.85)	.002 (.17)	.028 (1.07)	-.009 (.23)	.039 (.73)

La valeur absolue de la statistique "z" est entre parenthèses.

Ta[illegible]au 4

Biais de $\hat{\beta}_1$

		ρ = 0,6					ρ = 0,8					ρ = 0,99				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	MCO	-.000 (.23)	-.001 (.61)	.004 (1.28)	-.004 (.99)	-.002 (.37)	-.002 (.73)	-.005 (1.35)	-.008 (1.33)	-.005 (.73)	.005 (.53)	.003 (.55)	.004 (.62)	-.008 (.54)	-.013 (.64)	-.011 (.40)
	DUR	-.001 (.37)	-.001 (.39)	-.000 (.02)	-.001 (.45)	.001 (.34)	-.003 (.58)	-.003 (.67)	-.010 (1.83)	-.004 (.59)	-.004 (.67)	-.005 (.66)	-.000 (.03)	-.011 (.84)	-.017 (.96)	-.008 (.39)
	MVC	.001 (.19)	-.001 (.37)	-.002 (.52)	.000 (.03)	-.001 (.15)	.000 (.08)	-.006 (.93)	-.010 (1.57)	.004 (.58)	-.006 (.95)	-.009 (1.10)	.003 (.36)	-.017 (1.31)	-.022 (1.38)	-.002 (.13)
	PWI	.000 (.22)	-.002 (.81)	.003 (.86)	-.006 (1.28)	-.005 (.79)	-.003 (1.03)	-.004 (1.16)	-.006 (1.17)	-.000 (.07)	.001 (.10)	.002 (.45)	.002 (.34)	-.011 (.82)	-.014 (.77)	-.014 (.64)
	MVS	.000 (.16)	-.002 (.81)	.003 (1.01)	-.005 (1.26)	-.004 (.75)	-.003 (.99)	-.004 (1.26)	-.007 (1.19)	-.001 (.19)	.002 (.21)	.003 (.52)	.003 (.41)	-.010 (.71)	-.013 (.69)	-.012 (.51)
	DIF	.002 (.65)	.003 (.67)	.000 (.10)	-.007 (1.15)	-.009 (1.26)	-.002 (.49)	.000 (.07)	-.004 (.78)	.006 (.82)	.001 (.10)	-.000 (.06)	.002 (.30)	.000 (.02)	-.012 (.84)	-.011 (.58)
X2 (bruit blanc, (N = 20)	MCO	.001 (.65)	-.003 (1.38)	-.002 (.84)	.002 (.63)	-.002 (.45)	-.002 (.74)	-.001 (.19)	.001 (.34)	.012 (2.55)	.002 (.30)	-.007 (1.51)	-.003 (.63)	.003 (.47)	-.002 (.32)	.011 (1.19)
	DUR	.000 (.02)	-.004 (2.24)	.003 (1.35)	.001 (.36)	-.001 (.30)	-.002 (.90)	.000 (.20)	.000 (.07)	-.002 (1.23)	-.001 (.50)	-.003 (1.77)	-.000 (.03)	-.001 (.37)	.002 (1.22)	-.001 (.49)
	MVC	.000 (.23)	-.004 (2.18)	.002 (1.33)	.001 (.43)	-.001 (.34)	-.002 (1.02)	.000 (.21)	.000 (.01)	-.002 (1.14)	-.001 (.54)	-.003 (1.87)	-.000 (.08)	-.001 (.35)	.002 (1.20)	-.001 (.59)
	PWI	.001 (.83)	-.003 (1.82)	-.000 (.01)	-.000 (.10)	-.000 (.04)	-.001 (.83)	-.000 (.17)	.000 (.05)	-.001 (.65)	.001 (.41)	-.003 (1.96)	.001 (.36)	.000 (.03)	.002 (1.27)	-.000 (.33)
	MVS	.001 (.80)	-.003 (1.80)	-.000 (.08)	-.000 (.05)	-.000 (.05)	-.001 (.87)	-.000 (.21)	.000 (.08)	-.001 (.62)	.001 (.44)	-.003 (1.96)	.001 (.41)	.000 (.07)	.002 (1.29)	-.000 (.28)
	DIF	.001 (.62)	-.004 (2.07)	.001 (.30)	.001 (.45)	.000 (.00)	-.001 (.50)	-.000 (.24)	.000 (.22)	-.001 (.73)	.001 (.32)	-.003 (2.16)	-.000 (.04)	.000 (.10)	.002 (1.20)	-.001 (.46)
X3, (avec tendance, N = 50)	MCO	-.001 (1.74)	.000 (.81)	-.000 (.56)	.000 (.25)	-.000 (.17)	-.001 (1.44)	-.001 (1.95)	.000 (.08)	-.000 (.33)	.000 (.41)	-.003 (1.44)	-.001 (.26)	.000 (.07)	-.009 (1.94)	-.005 (.83)
	DUR	-.001 (2.25)	.000 (.74)	-.000 (.86)	-.000 (.01)	.000 (.78)	-.001 (1.05)	-.001 (1.25)	.000 (.48)	-.001 (.98)	.000 (.46)	-.001 (.43)	-.001 (.55)	-.003 (1.22)	-.005 (1.59)	-.003 (.73)
	MVC	-.001 (2.25)	.000 (.73)	-.000 (.86)	-.000 (.02)	.000 (.80)	-.001 (1.02)	-.001 (1.37)	.000 (.48)	-.001 (.99)	.000 (.67)	-.001 (.71)	-.001 (.46)	.000 (.03)	-.004 (1.49)	-.004 (1.09)
	PWI	-.001 (1.92)	.000 (.29)	-.000 (.76)	.000 (.23)	-.000 (.80)	-.000 (.97)	-.001 (1.84)	.000 (.30)	-.000 (.59)	.000 (.32)	-.001 (.51)	-.000 (.26)	-.001 (.20)	-.007 (2.08)	-.004 (.88)
	MVS	-.001 (1.91)	.000 (.32)	-.000 (.75)	.000 (.22)	-.000 (.55)	-.001 (1.02)	-.001 (1.89)	.000 (.30)	-.000 (.59)	.000 (.38)	-.001 (.60)	-.000 (.19)	-.000 (.11)	-.008 (2.06)	-.004 (.85)
	DIF	-.001 (.69)	-.001 (.71)	-.000 (.59)	-.001 (1.14)	-.001 (1.41)	-.001 (.66)	-.001 (.69)	.000 (.40)	-.001 (.53)	.001 (.65)	.001 (.61)	-.001 (.39)	-.000 (.07)	-.007 (2.09)	-.003 (.68)
X4, (bruit blanc, N = 50)	MCO	.001 (1.15)	-.002 (1.74)	.003 (2.20)	-.001 (.48)	.002 (1.34)	.002 (1.46)	-.002 (1.44)	.000 (.13)	.001 (.87)	.001 (.49)	-.003 (1.44)	-.004 (2.29)	-.003 (1.50)	-.002 (.94)	.002 (.68)
	DUR	-.000 (.06)	-.000 (.28)	.002 (2.37)	-.000 (.27)	.001 (.65)	.001 (.65)	.000 (.05)	-.002 (1.83)	-.001 (1.06)	-.001 (.75)	-.000 (.46)	-.001 (.75)	-.001 (.81)	-.000 (.16)	.000 (.20)
	MVC	-.000 (.08)	-.000 (.24)	.002 (2.36)	-.000 (.26)	.001 (.63)	.001 (.64)	.000 (.07)	-.002 (1.86)	-.001 (1.07)	-.001 (.72)	-.000 (.45)	-.001 (.73)	-.001 (.80)	-.000 (.16)	.000 (.21)
	PWI	-.000 (.11)	-.000 (.17)	.002 (2.31)	-.000 (.31)	.000 (.18)	.000 (.58)	.000 (.09)	-.002 (1.85)	-.001 (1.07)	-.001 (.77)	-.000 (.37)	-.000 (.66)	-.001 (.75)	-.000 (.08)	.000 (.23)
	MVS	-.000 (.09)	-.000 (.19)	.002 (2.31)	-.000 (.32)	.000 (.19)	.000 (.58)	.000 (.08)	-.002 (1.84)	-.001 (1.06)	-.001 (.77)	-.000 (.36)	-.000 (.66)	-.001 (.74)	-.000 (.07)	.000 (.24)
	DIF	-.000 (.34)	.000 (.29)	.002 (2.32)	-.000 (.38)	.000 (.18)	.000 (.56)	.000 (.15)	-.002 (1.89)	-.001 (1.13)	-.001 (.84)	-.000 (.37)	-.000 (.65)	-.001 (.76)	-.000 (.11)	.000 (.24)

La valeur absolue de la statistique "z" est entre parenthèses.

$$z = \frac{\sqrt{1000}\ (\bar{\beta} - \beta)}{\left[1/1000 \sum_{i=1}^{1000} (\hat{\beta}_i - \bar{\beta})^2\right]^{1/2}} \sim \text{approx } N(0,\ 1)$$

Le dénominateur est la variance estimée soit la différence entre l'erreur quadratique moyenne (EQM) et le biais au carré.

En fixant le niveau du test à 1% nous devons rejeter l'hypothèse nulle si $|z| > 2,576$.

Les valeurs de "z" sont données entre parenthèses au tableau 3. Nous retrouvons maintenant deux modèles où le biais est statistiquement différent de zéro. Ceci se produit dans les deux cas pour la constante, lorsque la taille d'échantillon est petite (20) et que la variable explicative est un bruit blanc (ρ=0,6; d=0,5 et ρ=0,8; d=1,5d*). Si l'hypothèse nulle doit être rejetée, nous croyons qu'il s'agit plutôt de l'effet du hasard puisque les biais restent très faibles en valeur absolue. Notons enfin que les cas où le biais est important en valeur absolue, sont facilement rejetés puisque la variance est aussi plus grande.

Les paramètres estimés $\hat{\beta}_0$ et $\hat{\beta}_1$ semblent donc effectivement centrés.

VI.3.2. Efficacité de $\hat{\beta}_0$ et $\hat{\beta}_1$

Le but premier des études empiriques précédentes était d'évaluer les performances de divers estimateurs lorsque ρ doit être estimé. Pour cela, les auteurs ont eu recours à diverses statistiques telles que les EQMs (ou REQMs), leurs ratios, ou le nombre de fois (NF) qu'un estimateur était plus près de la vraie valeur qu'un autre. Une lacune importante de cette approche est qu'elle ne nous permet pas de distinguer les causes des différences d'efficacité.

En effet, chacune des EQMs des paramètres $(\hat{\beta}_0, \hat{\beta}_1)$ peut être vu comme la somme de deux termes soit la variance théorique (ρ connu) plus un terme (k) qui est dû uniquement à l'estimation de ρ; $EQM(\hat{\beta}_{\hat{\rho}}) = Var(\hat{\beta}_{\rho}) + k$. Les différences entre deux estimateurs peuvent donc être dues à l'un ou l'autre de ces termes ou à une combinaison des deux.

Nous examinerons dans un premier temps le rapport entre les EQMs et les variances théoriques (ρ connu). Lorsque ces rapports sont près de l'unité, nous pourrons conclure que les variances sont indépendantes de l'estimation de ρ. Les résultats théorique (chapitre IV) seront directement confirmés puisque les variances empiriques ne sont pas différentes des variables théoriques avec lesquelles ces résultats ont été obtenus. Une analyse plus approfondie ne se révélera donc nécessaire que dans les cas contraire.

Les tableaux 5 à 8 donnent les REQMs de chaque estimateur avec le ratio entre l'EQM et la variance théorique et le nombre de fois que cet estimateur est plus près de la vraie valeur que le MVS. Rappelons ici que les covariances théoriques de DUR et MVC sont celles de CORC alors que pour PWI et MVS nous utilisons PW.

Comme nous nous y attendions, les EQMs des MCO sont toujours très près des variances théoriques. Etant donné que ρ n'est pas estimé, les variances empiriques ne sont affectées que par la génération des variables aléatoires. Seul le hasard ou une erreur de programmation auraient pu expliquer une différence appréciable entre ces valeurs. Nous pourrons donc porter toute notre attention sur les méthodes qui corrigent l'autocorrélation d'une façon ou d'autre autre.

VI.3.3. La pente $(\hat{\beta}_1)$

Observons d'abord le comportement des ratios pour la pente ($\hat{\beta}_1$, tableau 5). Lorsque la variable explicative est un bruit blanc (X2 et X4), les différences entre les variances théoriques et les EQMs sont négligeables, peu importe la méthode d'estimation. Par contre, avec une tendance (X1, X3), l'estimation de ρ affecte assez considérablement les variances empiriques de MVC et DUR et dans une moindre mesure celles de PWI et MVS. Cet effet est d'autant plus important que ρ est grand et (ou) la taille d'échantillon (N) est petite. Par exemple, lorsque N=50, aucun rapport n'est différent de 1 sauf pour ρ=0,99; mais si N=20 l'EQM de MVC est toujours différente de la variance théorique

Tableau 5

Efficacité de $\hat{\beta}_1$

	D=	REQM .5	REQM 1.0	REQM D*	REQM 1.5D*	REQM 2D*	EQM/COV .5	EQM/COV 1.0	EQM/COV D*	EQM/COV 1.5D*	EQM/COV 2D*
X1 (avec tendance, N = 20)	RO=.60										
	MCO	.069	.076	.092	.116	.144	.964	.989	.986	1.001	1.024
	DUR	.090	.096	.099	.096	.104	.893	1.023	1.083	1.020	1.181
	MVC	.111	.117	.123	.107	.132	1.361	1.501	1.665	1.267	1.903
	PWI	.062	.073	.095	.142	.188	1.085	1.071	.997	1.218	1.306
	MVS	.062	.073	.094	.137	.180	1.084	1.057	.964	1.127	1.203
	DIF	.110	.127	.149	.189	.230	ND	ND	ND	ND	ND
	RO=.80										
	MCO	.106	.124	.180	.234	.318	1.005	1.067	1.013	.927	1.037
	DUR	.146	.154	.175	.188	.196	.742	.822	1.063	1.221	1.330
	MVC	.191	.193	.201	.199	.208	1.263	1.293	1.404	1.380	1.494
	PWI	.092	.111	.174	.228	.311	1.227	1.200	1.054	.902	.991
	MVS	.093	.111	.175	.230	.316	1.241	1.198	1.058	.917	1.023
	DIF	.114	.126	.175	.227	.302	ND	ND	ND	ND	ND
	RO=.99										
	MCO	.196	.220	.481	.639	.847	.847	1.001	1.048	.925	.955
	DUR	.227	.249	.426	.544	.651	.461	.553	1.618	2.634	3.783
	MVC	.264	.275	.422	.517	.605	.623	.673	1.591	2.384	3.263
	PWI	.169	.189	.422	.559	.716	1.109	1.312	1.585	1.416	1.378
	MVS	.173	.192	.437	.585	.765	1.153	1.363	1.700	1.552	1.569
	DIF	.157	.164	.336	.449	.590	ND	ND	ND	ND	ND
X2 (bruit blanc, N = 20)	RO=.60										
	MCO	.070	.074	.082	.098	.121	1.042	1.004	.971	.958	1.010
	DUR	.059	.061	.059	.059	.058	1.067	1.158	1.078	1.073	1.054
	MVC	.058	.061	.059	.059	.058	1.060	1.150	1.070	1.075	1.044
	PWI	.054	.057	.057	.060	.061	1.030	1.085	1.008	.942	.811
	MVS	.054	.057	.057	.061	.062	1.029	1.086	1.018	.966	.844
	DIF	.055	.057	.054	.056	.059	ND	ND	ND	ND	ND
	RO=.80										
	MCO	.085	.099	.113	.146	.182	.966	1.035	.948	.973	.981
	DUR	.054	.055	.053	.052	.052	1.076	1.146	1.054	1.014	1.007
	MVC	.053	.055	.053	.052	.052	1.060	1.131	1.042	1.004	1.001
	PWI	.051	.054	.052	.052	.052	1.014	1.138	1.015	.951	.860
	MVS	.051	.054	.052	.052	.052	1.016	1.146	1.028	.968	.880
	DIF	.050	.052	.050	.050	.051	ND	ND	ND	ND	ND
	RO=.99										
	MCO	.148	.151	.196	.243	.298	.995	.981	.976	.961	.967
	DUR	.050	.050	.047	.050	.049	1.137	1.120	1.003	1.148	1.077
	MVC	.050	.049	.047	.050	.048	1.124	1.102	.991	1.118	1.050
	PWI	.049	.048	.046	.048	.047	1.079	1.069	.951	1.063	.997
	MVS	.049	.049	.046	.048	.047	1.092	1.078	.961	1.066	1.001
	DIF	.048	.048	.045	.048	.046	ND	ND	ND	ND	ND

REQM : Racine de l'erreur quadratique moyenne.

EQM/COV : Ratio de l'erreur quadratique moyenne sur la variance théorique (ρ connu).

ND : Non disponible.

Tableau 5 (suite)

Efficacité de $\hat{\beta}_1$

	D=	REQM .5	REQM 1.0	REQM D*	REQM 1.5D*	REQM 2D*	EQM/COV .5	EQM/COV 1.0	EQM/COV D*	EQM/COV 1.5D*	EQM/COV 2D*
X3 (avec tendance, N = 50)	RO=.60										
	MCO	.011	.011	.011	.012	.013	1.014	.973	1.048	1.012	.987
	DUR	.011	.011	.011	.011	.011	1.017	.968	1.037	.951	.999
	MVC	.011	.011	.011	.011	.011	1.013	.965	1.035	.949	.999
	PWI	.010	.010	.011	.013	.015	1.044	.973	1.054	1.174	1.215
	MVS	.010	.010	.011	.013	.015	1.044	.973	1.046	1.143	1.174
	DIF	.029	.029	.026	.028	.027	ND	ND	ND	ND	ND
	RO=.80										
	MCO	.018	.019	.021	.023	.027	.941	.942	1.009	.985	1.027
	DUR	.019	.020	.020	.021	.021	.841	.943	.978	1.108	1.067
	MVC	.019	.020	.020	.021	.022	.840	.935	.976	1.104	1.145
	PWI	.016	.018	.021	.024	.029	.908	.960	1.071	.972	.982
	MVS	.016	.018	.021	.024	.029	.915	.957	1.058	.969	1.016
	DIF	.030	.030	.030	.031	.031	ND	ND	ND	ND	ND
	RO=.99										
	MCO	.057	.065	.119	.150	.206	.887	.992	1.087	.903	1.023
	DUR	.057	.060	.085	.102	.120	.454	.513	1.012	1.463	2.026
	MVC	.058	.062	.082	.094	.109	.477	.544	.961	1.243	1.671
	PWI	.049	.055	.092	.113	.149	1.041	1.185	1.205	.969	1.026
	MVS	.050	.056	.098	.121	.162	1.058	1.218	1.363	1.123	1.226
	DIF	.048	.051	.082	.101	.139	ND	ND	ND	ND	ND
X4 (bruit blanc, N = 50)	RO=.60										
	MCO	.039	.041	.040	.038	.039	1.022	1.090	1.058	.944	.983
	DUR	.029	.030	.029	.029	.029	1.023	1.037	1.010	.963	1.009
	MVC	.029	.029	.029	.028	.029	1.017	1.029	1.007	.960	1.008
	PWI	.029	.029	.029	.029	.030	1.015	1.019	1.010	.982	1.003
	MVS	.029	.029	.029	.029	.030	1.016	1.020	1.010	.983	1.004
	DIF	.029	.029	.029	.029	.030	ND	ND	ND	ND	ND
	RO=.80										
	MCO	.043	.045	.046	.050	.053	.932	.970	.968	.997	.966
	DUR	.026	.027	.026	.025	.026	.969	1.019	1.005	.895	.975
	MVC	.026	.027	.026	.025	.026	.967	1.019	1.004	.895	.973
	PWI	.026	.027	.026	.025	.026	.976	1.019	1.006	.883	.984
	MVS	.026	.027	.026	.025	.026	.976	1.019	1.007	.884	.985
	DIF	.026	.027	.026	.025	.026	ND	ND	ND	ND	ND
	RO=.99										
	MCO	.062	.061	.066	.069	.072	1.016	.991	1.055	1.009	.946
	DUR	.023	.024	.024	.024	.025	.948	.992	.988	1.032	1.055
	MVC	.023	.024	.024	.024	.025	.947	.991	.988	1.032	1.054
	PWI	.023	.024	.024	.024	.025	.941	.987	.979	1.028	1.047
	MVS	.023	.024	.024	.024	.025	.941	.988	.980	1.027	1.047
	DIF	.023	.024	.024	.024	.025	ND	ND	ND	ND	ND

Tableau 6

Nombre de fois qu'un estimateur de β_1 est plus près de la vraie valeur que MVS

D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
	X1 (avec tendance, N = 20)					X3 (avec tendance, N = 50)				
RO=.60										
MCO	.383	.429	.498	.607	.659	.428	.472	.488	.587	.578
DUR	.268	.369	.503	.610	.689	.349	.443	.504	.571	.607
MVC	.267	.365	.492	.598	.685	.348	.444	.505	.572	.609
PWI	.576	.518	.480	.382	.338	.541	.500	.481	.382	.404
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.309	.327	.318	.301	.312	.210	.216	.249	.275	.304
RO=.80										
MCO	.340	.369	.438	.441	.500	.376	.415	.480	.504	.531
DUR	.265	.334	.549	.620	.732	.331	.411	.517	.557	.651
MVC	.244	.307	.509	.583	.703	.329	.408	.520	.557	.654
PWI	.602	.579	.554	.567	.558	.569	.526	.481	.487	.490
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.432	.439	.514	.540	.568	.305	.339	.344	.392	.493
RO=.99										
MCO	.354	.302	.259	.228	.215	.351	.356	.303	.268	.188
DUR	.367	.354	.503	.539	.612	.461	.477	.688	.705	.806
MVC	.324	.319	.454	.458	.529	.459	.471	.676	.699	.812
PWI	.629	.651	.742	.767	.802	.599	.596	.701	.747	.827
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.528	.593	.742	.789	.834	.518	.557	.705	.757	.830

	X2 (bruit blanc, N = 20)					X4 (bruit blanc, N = 50)				
RO=.60										
MCO	.373	.366	.327	.272	.280	.361	.346	.343	.386	.390
DUR	.393	.426	.493	.505	.547	.481	.468	.477	.499	.500
MVC	.394	.443	.485	.501	.552	.480	.471	.491	.504	.508
PWI	.496	.520	.545	.577	.616	.522	.520	.513	.506	.505
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.464	.476	.518	.560	.599	.486	.499	.490	.512	.507
RO=.80										
MCO	.313	.287	.251	.174	.152	.305	.301	.306	.277	.281
DUR	.446	.463	.490	.517	.524	.503	.506	.482	.487	.540
MVC	.448	.465	.493	.528	.514	.507	.511	.481	.485	.540
PWI	.511	.506	.541	.580	.597	.495	.517	.500	.525	.557
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.493	.494	.514	.556	.582	.492	.502	.493	.506	.547
RO=.99										
MCO	.166	.161	.141	.111	.106	.214	.231	.203	.217	.211
DUR	.507	.470	.493	.461	.471	.494	.475	.468	.502	.488
MVC	.514	.488	.497	.480	.480	.494	.478	.470	.506	.491
PWI	.569	.548	.556	.517	.545	.495	.499	.493	.476	.463
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.548	.531	.538	.502	.537	.492	.492	.491	.476	.475

Tous les chiffres sont en proportion de l'échantillon total.

alors que ceux de MVS sont très semblables sauf lorsque ρ est très grand (0,99).

Le point le plus étonnant est sûrement le ratio des variances pour les estimateurs conditionnels qui est beaucoup plus petit que l'unité lorsque ρ est grand et $\sigma_1^2(1-\rho^2)/\sigma_e^2$ est petit. Il semble ici préférable d'estimer ρ plutôt que d'utiliser sa vraie valeur. Ceci paraît contre-intuitif, mais puisque MVC n'est pas le meilleur estimateur, il peut exister des valeurs de $Q_1 \neq Q$ telles que :

$$\sigma_e^2(X'Q_1'Q_1X)^{-1}\ X'Q_1'Q_1VQ_1'Q_1X(X'Q_1'Q_1X)^{-1} < \sigma_e^2\ (X'Q'QX)^{-1}$$

Si les valeurs de $\hat{\rho}$ sont concentrés autour de la valeur de Q_1, il est probable que les variances empiriques seront plus petites que celles calculées avec la vraie valeur de ρ. Notons cependant que cette réduction de la variance n'est pas suffisante pour que l'estimation conditionnelle devienne préférable à l'estimation stationnaire. Les conclusions théoriques restent donc valides : il est préférable d'inclure la première observation si $d < d^*$. Seule la différence entre les deux méthodes est moins importante que prévue.

Examinons maintenant plus en détail les autres cas où l'estimation de ρ affecte les EQMs. Toujours pour ρ=0,99, et X avec tendance, les REQMs de MVC, DUR, PWI et MVS nous semblent approximativement égales lorsque $d = d^*$. Notons que la taille d'échantillon affecte toutes les méthodes de la même façon : lorsqu'elle est grande (50), une seule EQM est différente de la variance théorique (MVS), lorsqu'elle est petite (20) toutes les EQM augmentent d'environ 25% par rapport à leurs valeurs théoriques. Ainsi, toutes les méthodes restent aussi efficaces l'une que l'autre avec une légère préférence pour MVC.

Pour terminer avec les grandes valeurs de ρ (0,99), il reste les cas où $d > d^*$. Nous nous attendons alors à une meilleure performance des estimateurs conditionnels et c'est effectivement ce que nous observons. Le MVC offre encore la plus faible REQM, suivi de près par DUR.

Pour les petites valeurs de ρ (0,6, 0,8), les EQMs ne sont éloignées des variances théoriques que lorsque N est petit (20). Comme prévu, il est préférable d'inclure la première observation lorsque $d < d^*$. PWI et MVS sont alors équivalents mais DUR surpasse clairement MVC. Si $d = d^*$, DUR, PWI et MVS sont très peu différents, mais MVC reste moins efficace que les autres. Enfin, lorsque $d > d^*$ les résultats théoriques sont respectés puisque DUR est le meilleur estimateur suivi de MVC.

Avant de passer à l'efficacité de la constante, quelques remarques s'imposent sur l'estimation en premières différences (DIFF). Cette méthode est toujours la plus efficace lorsque ρ est près de l'unité, en particulier si la taille d'échantillon est petite (20). De plus, il se révèle au moins aussi efficace que les autres méthodes pour les plus faibles valeurs de ρ (0,6, 0,8) lorsque la variable explicative est un bruit blanc (X2, X4). Par exemple, c'est lui qui présente la plus faible REQM pour X2, N=20 , ρ=0,6 et $d=d^*$.

VI.3.4. La constante ($\hat{\beta}_0$)

Si l'estimation de ρ affecte somme toute assez peu les variances de la pente et que les résultats théoriques du chapitre IV sont généralement confirmés, il en est tout autrement pour la constante. Comme le montre le tableau 7, seule l'EQM des MCO reste toujours égale à la variance théorique. Pour les méthodes stationnaires (MVS et PWI), les ratios sont relativement importants lorsque ρ est petit (0,6, 0,8) et d est grand. Pour les méthodes conditionnelles (MVC, DUR), elles ne sont vraiment équivalentes à CORC que si N est grand (50) et ρ est très petit (0,6).

Il ne reste donc pratiquement aucun cas qui se comporte identiquement au modèle théorique pour toutes les valeurs de d. Une analyse comparative de toutes les spécifications serait ici beaucoup trop longue et n'apporterait pas nécessairement les éclaircissements souhaités. Nous tenterons plutôt de dégager les points les plus importants en laissant au lecteur l'initiative d'établir les comparaisons qui l'intéresse plus particulièrement.

De façon globale, les estimateurs stationnaires offrent une meilleure performance que les méthodes conditionnelles par rapport à leurs variances théoriques. Nous observons moins de cas où les ratios sont différents de 1 et ces différences sont en elles-mêmes moins importantes. Par exemple, l'EQM de PWI est au maximum 3,2 fois plus élevée que la variance théorique (X4, ρ=0,8, d=2d*) alors que celles de DUR et MVC

Tableau 7

Efficacité de $\hat{\beta}_0$

	D=	REQM .5	REQM 1.0	REQM D*	REQM 1.5D*	REQM 2D*	EQM/COV .5	EQM/COV 1.0	EQM/COV D*	EQM/COV 1.5D*	EQM/COV 2D*
X1 (avec tendance, N = 20)	RO=.60										
	MCO	.107	.124	.154	.200	.253	.941	1.008	.982	.988	1.014
	DUR	.154	.167	.199	.190	.183	.869	1.023	1.446	1.319	1.225
	MVC	.236	.257	.284	.236	.303	2.039	2.424	2.951	2.029	3.366
	PWI	.095	.119	.166	.261	.361	1.085	1.071	1.011	1.276	1.460
	MVS	.095	.118	.162	.247	.339	1.092	1.060	.965	1.146	1.282
	RO=.80										
	MCO	.162	.202	.323	.437	.604	.985	1.056	1.000	.923	1.044
	DUR	.293	.288	.373	.424	4.527	.810	.785	1.318	1.705	193.992
	MVC	.524	.469	.531	.435	.511	2.603	2.079	2.671	1.792	2.476
	PWI	.136	.183	.335	.462	.659	1.295	1.185	1.059	.946	1.102
	MVS	.138	.183	.330	.457	.650	1.325	1.178	1.033	.924	1.073
	RO=.99										
	MCO	.328	.540	3.378	4.874	6.582	.947	1.013	1.013	.940	.965
	DUR	.564	1.198	3.269	4.957	6.511	.032	.143	1.065	2.448	4.223
	MVC	.744	1.086	3.048	4.194	5.293	.055	.117	.925	1.752	2.791
	PWI	.289	.502	3.302	4.759	6.397	1.081	1.085	1.086	1.004	1.021
	MVS	.294	.507	3.323	4.801	6.469	1.118	1.107	1.100	1.022	1.044
X2 (bruit blanc, N = 20)	RO=.60										
	MCO	.032	.031	.034	.036	.042	1.087	.965	1.025	.929	.978
	DUR	.036	.035	.041	.035	.058	1.105	1.032	1.426	1.022	2.796
	MVC	.039	.036	.038	.035	.038	1.292	1.070	1.183	1.053	1.227
	PWI	.031	.031	.038	.054	.080	1.133	1.006	1.214	1.740	2.804
	MVS	.031	.031	.036	.045	.063	1.127	.986	1.087	1.251	1.720
	RO=.80										
	MCO	.053	.059	.068	.085	.098	.920	1.003	1.021	1.097	1.011
	DUR	.108	.104	.130	1.678	.155	2.473	2.289	3.563	593.606	5.085
	MVC	.103	.207	.084	1.118	.331	2.232	9.068	1.483	263.510	23.084
	PWI	.050	.057	.074	.115	.154	1.003	1.008	1.154	1.676	1.939
	MVS	.050	.057	.070	.103	.134	.993	.991	1.040	1.347	1.464
	RO=.99										
	MCO	.236	.400	1.373	2.072	2.628	.989	.943	.994	1.012	.917
	DUR	1.782	.714	3.620	4.504	11.300	1.676	.269	6.916	10.705	67.385
	MVC	1.483	4.421	2.096	4.461	11.140	1.161	10.314	2.318	10.502	65.490
	PWI	.231	.395	1.375	2.071	2.631	1.021	.942	.998	1.010	.919
	MVS	.232	.396	1.374	2.070	2.630	1.028	.948	.996	1.009	.918

Tableau 7 (suite)

Efficacité de $\hat{\beta}_0$

	D=	REQM .5	1.0	D*	1.5D*	2D*	EQM/COV .5	1.0	D*	1.5D*	2D*
X3 (avec tendance, N = 50)	RO=.60										
	MCO	.040	.042	.045	.047	.053	1.008	1.021	1.083	.969	1.002
	DUR	.044	.045	.046	.043	.044	.998	1.018	1.095	.918	.969
	MVC	.044	.045	.046	.042	.044	.994	1.013	1.092	.916	.968
	PWI	.039	.041	.046	.059	.080	1.037	1.021	1.095	1.277	1.703
	MVS	.039	.041	.046	.057	.075	1.040	1.021	1.083	1.193	1.473
	RO=.80										
	MCO	.070	.076	.086	.102	.118	.935	.959	1.015	1.038	1.007
	DUR	.081	.086	.086	.110	.092	.830	.947	.937	1.549	1.084
	MVC	.080	.085	.085	.104	.182	.827	.916	.932	1.385	4.260
	PWI	.062	.072	.093	.125	.174	.957	.983	1.110	1.200	1.487
	MVS	.063	.072	.092	.121	.165	.968	.981	1.074	1.127	1.326
	RO=.99										
	MCO	.266	.468	1.665	2.296	3.230	.984	1.075	1.056	.898	1.001
	DUR	.445	.649	2.126	3.436	7.560	.077	.164	1.767	4.616	22.345
	MVC	2.210	1.023	2.351	3.356	3.661	1.910	.409	2.161	4.403	5.240
	PWI	.253	.453	1.645	2.270	3.195	1.088	1.097	1.058	.899	1.002
	MVS	.253	.455	1.651	2.278	3.206	1.089	1.104	1.066	.905	1.009
X4 (bruit blanc, N = 50)	RO=.60										
	MCO	.020	.021	.021	.023	.023	.988	1.005	1.012	1.048	.985
	DUR	.021	.021	.021	.022	.022	.990	.986	1.003	1.009	1.024
	MVC	.021	.021	.021	.022	.022	.991	.987	1.003	1.009	1.024
	PWI	.020	.021	.022	.029	.043	.995	.996	1.039	1.576	2.802
	MVS	.020	.021	.022	.027	.033	.996	.996	1.032	1.308	1.709
	RO=.80										
	MCO	.038	.041	.041	.047	.051	.947	1.013	.937	1.065	1.060
	DUR	.047	.064	.042	.045	.045	1.184	2.230	.964	1.092	1.126
	MVC	.068	.049	.042	.044	.045	2.508	1.304	.970	1.077	1.120
	PWI	.037	.040	.044	.072	.105	.985	1.009	1.078	2.072	3.225
	MVS	.037	.039	.043	.061	.083	.982	1.002	.994	1.481	2.008
	RO=.99										
	MCO	.260	.388	.827	1.261	1.657	.974	.976	.942	1.005	.987
	DUR	9.883	.973	10.660	17.010	20.270	132.945	1.288	154.671	393.824	559.244
	MVC	5.154	6.730	2.998	5.977	228.200	36.2	61.6	12.2	48.6	70880.2
	PWI	.242	.377	.830	1.277	1.682	1.013	.980	.937	1.008	.992
	MVS	.244	.378	.828	1.272	1.678	1.030	.986	.933	1.000	.987

Tableau 8

Nombre de fois qu'un estimateur de β_0 est plus près de la vraie valeur que celui du MVS

D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
	X1 (avec tendance, N = 20)					X3 (avec tendance, N = 50)				
RO=.60										
MCO	.348	.431	.533	.645	.702	.409	.461	.525	.637	.687
DUR	.230	.366	.491	.617	.710	.296	.420	.485	.585	.659
MVC	.225	.364	.482	.609	.700	.296	.421	.487	.589	.659
PWI	.598	.519	.428	.342	.293	.567	.507	.439	.336	.276
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
RO=.80										
MCO	.296	.378	.497	.534	.614	.369	.437	.557	.618	.704
DUR	.199	.326	.548	.635	.740	.274	.397	.525	.603	.703
MVC	.189	.300	.508	.607	.715	.271	.397	.528	.603	.704
PWI	.656	.570	.470	.453	.403	.580	.533	.404	.362	.295
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
RO=.99										
MCO	.361	.388	.313	.269	.246	.407	.416	.450	.454	.434
DUR	.312	.389	.546	.567	.635	.303	.415	.678	.717	.792
MVC	.277	.355	.527	.525	.574	.301	.407	.666	.710	.808
PWI	.615	.600	.696	.750	.772	.557	.558	.541	.576	.628
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	X2 (bruit blanc, N = 20)					X4 (bruit blanc, N = 50)				
RO=.60										
MCO	.423	.457	.521	.624	.687	.430	.482	.541	.586	.642
DUR	.306	.419	.527	.597	.652	.357	.446	.524	.572	.627
MVC	.314	.416	.519	.593	.652	.356	.447	.522	.568	.626
PWI	.556	.528	.433	.335	.278	.555	.507	.425	.382	.318
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
RO=.80										
MCO	.422	.417	.508	.622	.706	.419	.434	.525	.667	.686
DUR	.268	.398	.505	.633	.676	.321	.429	.497	.626	.653
MVC	.268	.384	.504	.625	.673	.313	.431	.492	.625	.656
PWI	.575	.545	.438	.331	.251	.551	.536	.419	.302	.258
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
RO=.99										
MCO	.425	.464	.493	.505	.497	.396	.423	.504	.551	.557
DUR	.307	.403	.655	.669	.697	.253	.375	.574	.680	.681
MVC	.292	.401	.654	.676	.726	.253	.376	.578	.687	.693
PWI	.560	.538	.481	.470	.472	.572	.560	.487	.441	.437
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

sont souvent deux fois plus élevées que si ρ était connu. On trouve même un ratio de 70 880 pour MVC (X4, ρ=0,99, d=2d*). En général, les différences observées entre les valeurs théoriques et empiriques sont de moins de 20% pour PWI et MVS.

Entre les deux méthodes stationnaires, MVS nous semble légèrement préférables sauf pour la première variable explicative (avec tendance, N=20). Les différences entre ces deux estimateurs sont toujours assez faibles.

A l'opposé, MVC et DUR présentent un comportement assez imprévisible. On note souvent un écart très appréciable entre les deux. Il semble que plus ρ et d sont grands, plus leurs REQMs deviennent erratiques. Par contre, la présence d'une tendance atténue ces problèmes. En fait, aucune des deux méthodes n'est systématiquement préférable à l'autre.

Si nous comparons directement les REQMs des estimateurs conditionnels et stationnaires, ces derniers sont nettement préférables lorsque d n'est pas trop grand. Même lorsque d est grand, PWI et MVS sont quelquefois supérieurs à DUR et (ou) MVC. Ceci se produit lorsque N et ρ sont grands pour les deux formes de variables explicatives et aussi lorsque N est petit et X est un bruit blanc (X2) pour les valeurs de ρ de 0,8 et 0,99. Il ne reste alors que peu de place aux méthodes conditionnelles soit lorsque ρ est très petit (0,6) et d est grand ou pour ρ=0,8, N=50 et d est grand.

Comme l'ont noté Beach et MacKinnon (1978) ainsi que Park et Mitchell (1980), les estimateurs conditionnels produisent un petit nombre de très mauvaises valeurs pour la constante. Ces valeurs affectent les REQMs de façon dramatique de sorte qu'elles donnent une image déformée de la dispersion véritable. Une autre forme de mesure est donc souhaitable.

En considérant le nombre de fois qu'un estimateur est plus près de la vraie valeur que MVS (NF) en proportion du nombre de reprises (1 000), nous devons admettre que les conclusions tirées des REQMs sont pour le moins douteuses. (tableau 8).

Avec un intervalle de confiance de 5%, des valeurs entre 0,469 et 0,531 ne sont pas différentes de 0,5 pour une binomiale (1 000, 0,5). Pour tous les modèles ou d=d*, le NF des estimateurs conditionnels n'est jamais statistiquement plus petit que 0,5. Les conclusions théoriques sont aussi confirmées dans tous les autres cas : lorsque d est plus grand (petit) que d*, MVC et DUR sont plus (moins) souvent près de vraie valeur que MVS, leur dispersion est donc plus faible (grande). Il semblerait alors que la distribution de la constante des estimateurs conditionnels soit plus concentrée au centre et dans les queues, au détriment des valeurs moyennes. Soulignons enfin que PWI apparaît très souvent préférable à MVS pour X1 et X3 (X avec tendance). Quelquefois le NF de PWI est même plus grand que ceux de DUR et MVC mais ceci ne signifie pas qu'il leur est préférable.

Avant de passer en revue la distribution des autres statistiques, un bref résumé s'impose pour éclaircir la situation. Nous avons d'abord vu que les EQM des estimateurs conditionnels dépendent dans une large mesure de l'estimation de ρ. Lorsque ρ est grand, les variances empiriques ne sont plus indépendantes de σ_1^2, mais les autres résultats théoriques sont respectés pour $\hat{\beta}_1$. La constante pose certains problèmes puisque des statistiques différentes donnent des conclusions différentes, en particulier lorsque ρ est grand. Par contre, lorsque ρ est très grand l'estimation en premières différences est toujours celle qui est préférable. Les problèmes affectant la constante sont donc indirectement résolus puisque ce paramètre n'est pas estimé.

VI.4. Distribution des autres paramètres

En plus des coefficients de régression, les statistiques sur trois autres paramètres ont été compilées soit $\hat{\rho}$, $\hat{\alpha}$ et s^2. Nous examinerons d'abord le comportement de $\hat{\rho}$ pour lequel nous avons certains points de comparaison dans l'étude de Beach et MacKinnon (1978). Nous verrons ensuite que la distribution de $\hat{\alpha}$ est très similaire à celle de $\hat{\rho}$. Nous terminerons enfin par quelques observations sur s^2.

VI.4.1. Distribution de $\hat{\rho}$ (tableau 9)

Toutes les études empiriques précédentes concluent que le biais de $\hat{\rho}$ est négatif. C'est effectivement ce que nous retrouvons ici dans les modèles stationnaires (d=1). Les valeurs calculées sont aussi directement comparables à celles de Beach et MacKinnon si on tient

Tableau 9

Biais et efficacité de $\hat{\rho}$

	D=	BIAIS .5	1.0	D*	1.5D*	2D*	REQM .5	1.0	D*	1.5D*	2D*	NF .5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	RO=.60															
	DUR	-.245	-.250	-.233	-.215	-.191	.332	.344	.325	.306	.290	.536	.514	.469	.461	.499
	MVC	-.244	-.250	-.230	-.211	-.183	.332	.343	.326	.301	.287	.594	.574	.498	.480	.508
	PWI	-.229	-.227	-.187	-.133	-.067	.324	.335	.315	.293	.295	.795	.760	.711	.652	.528
	MVS	-.249	-.248	-.212	-.163	-.103	.330	.339	.318	.293	.286	ND	ND	ND	ND	ND
	RO=.80															
	DUR	-.320	-.303	-.280	-.264	-.222	.393	.382	.356	.349	.305	.535	.517	.434	.417	.400
	MVC	-.318	-.300	-.273	-.249	-.209	.396	.383	.355	.342	.302	.516	.495	.443	.472	.474
	PWI	-.289	-.262	-.216	-.185	-.120	.373	.357	.325	.315	.273	.909	.881	.810	.741	.635
	MVB	-.319	-.293	-.252	-.222	-.163	.388	.372	.339	.326	.282	ND	ND	ND	ND	ND
	RO=.99															
	DUR	-.429	-.416	-.410	-.388	-.354	.484	.470	.472	.445	.416	.521	.487	.546	.597	.633
	MVC	-.428	-.418	-.411	-.394	-.359	.488	.475	.487	.475	.454	.527	.514	.549	.581	.621
	PWI	-.381	-.364	-.374	-.361	-.333	.449	.429	.451	.438	.421	.984	.989	.983	.983	.983
	MVB	-.418	-.402	-.415	-.407	-.386	.473	.455	.476	.466	.449	ND	ND	ND	ND	ND
X2 (bruit blanc, N = 20)	RO=.60															
	DUR	-.166	-.172	-.154	-.143	-.118	.278	.270	.256	.250	.231	.464	.402	.434	.455	.489
	MVC	-.141	-.148	-.133	-.123	-.100	.265	.255	.243	.236	.218	.556	.527	.509	.560	.550
	PWI	-.128	-.115	-.070	-.020	.043	.262	.248	.243	.258	.258	.675	.691	.610	.489	.402
	MVB	-.153	-.143	-.103	-.061	-.005	.265	.251	.239	.243	.233	ND	ND	ND	ND	ND
	RO=.80															
	DUR	-.214	-.198	-.191	-.158	-.145	.301	.285	.277	.246	.241	.420	.384	.348	.418	.448
	MVC	-.189	-.175	-.167	-.135	-.130	.283	.267	.254	.224	.223	.620	.551	.500	.547	.535
	PWI	-.164	-.128	-.092	-.031	.010	.265	.249	.234	.216	.217	.814	.728	.659	.512	.405
	MVB	-.198	-.167	-.136	-.079	-.040	.280	.261	.242	.215	.208	ND	ND	ND	ND	ND
	RO=.99															
	DUR	-.267	-.265	-.227	-.186	-.173	.330	.326	.298	.267	.255	.312	.299	.273	.251	.211
	MVC	-.240	-.238	-.204	-.162	-.153	.306	.302	.276	.243	.233	.517	.505	.417	.409	.350
	PWI	-.181	-.176	-.133	-.103	-.091	.260	.257	.224	.205	.192	.992	.997	.957	.885	.816
	MVS	-.228	-.223	-.180	-.146	-.131	.288	.285	.250	.227	.213	ND	ND	ND	ND	ND

NF : Nombre de fois (sur 1 000) qu'un estimateur est plus près de la vraie valeur que MVS.

Tableau 9 (suite)

Biais et efficacité de $\hat{\rho}$

	D=	BIAIS .5	BIAIS 1.0	BIAIS D*	BIAIS 1.5D*	BIAIS 2D*	REQM .5	REQM 1.0	REQM D*	REQM 1.5D*	REQM 2D*	NF .5	NF 1.0	NF D*	NF 1.5D*	NF 2D*
X3 (avec tendance, N = 50)	RO=.60															
	DUR	-.090	-.087	-.080	-.084	-.074	.160	.154	.153	.150	.139	.682	.607	.538	.525	.537
	MVC	-.098	-.094	-.087	-.090	-.080	.164	.160	.156	.154	.143	.685	.582	.506	.480	.513
	PWI	-.093	-.083	-.063	-.038	-.000	.162	.155	.151	.151	.160	.790	.738	.681	.605	.511
	MVS	-.103	-.093	-.075	-.052	-.016	.166	.159	.153	.151	.153	ND	ND	ND	ND	ND
	RO=.80															
	DUR	-.109	-.117	-.113	-.095	-.095	.156	.166	.160	.148	.147	.705	.610	.529	.495	.524
	MVC	-.114	-.122	-.118	-.100	-.099	.160	.170	.164	.153	.151	.671	.571	.505	.472	.499
	PWI	-.105	-.103	-.084	-.045	-.018	.154	.158	.146	.139	.144	.840	.839	.772	.627	.532
	MVS	-.119	-.118	-.100	-.063	-.039	.162	.166	.153	.141	.142	ND	ND	ND	ND	ND
	RO=.99															
	DUR	-.178	-.171	-.144	-.140	-.119	.206	.199	.176	.175	.158	.549	.539	.476	.416	.337
	MVC	-.181	-.174	-.146	-.142	-.119	.209	.202	.179	.178	.160	.531	.499	.460	.408	.383
	PWI	-.153	-.144	-.111	-.105	-.080	.185	.177	.149	.149	.133	.999	.999	.972	.939	.860
	MVS	-.172	-.163	-.132	-.126	-.101	.200	.191	.164	.162	.145	ND	ND	ND	ND	ND
X4 (bruit blanc, N = 50)	RO=.60															
	DUR	-.054	-.059	-.047	-.057	-.048	.133	.140	.128	.134	.125	.529	.464	.450	.475	.536
	MVC	-.050	-.054	-.043	-.054	-.045	.130	.137	.124	.131	.123	.600	.527	.511	.509	.534
	PWI	-.046	-.042	-.020	.001	.035	.129	.134	.123	.141	.155	.629	.598	.542	.502	.426
	MVS	-.057	-.054	-.032	-.014	.018	.132	.136	.123	.137	.144	ND	ND	ND	ND	ND
	RO=.80															
	DUR	-.074	-.070	-.068	-.063	-.054	.133	.129	.124	.122	.110	.538	.465	.459	.502	.545
	MVC	-.072	-.068	-.066	-.061	-.053	.130	.127	.121	.119	.108	.638	.524	.518	.538	.565
	PWI	-.065	-.051	-.032	-.000	.034	.125	.118	.112	.119	.124	.737	.679	.612	.483	.380
	MVS	-.080	-.066	-.049	-.020	.012	.132	.124	.115	.116	.114	ND	ND	ND	ND	ND
	RO=.99															
	DUR	-.100	-.100	-.079	-.069	-.057	.132	.133	.111	.105	.091	.447	.447	.363	.308	.320
	MVC	-.099	-.099	-.079	-.069	-.057	.130	.131	.110	.104	.090	.495	.464	.405	.334	.341
	PWI	-.071	-.069	-.048	-.037	-.028	.108	.107	.086	.081	.067	.952	.934	.804	.672	.574
	MVS	-.091	-.089	-.068	-.055	-.044	.121	.120	.098	.091	.077	ND	ND	ND	ND	ND

compte des différences dans la génération des variables exogènes (voir tableau 17, Annexe A).

La différence entre la moyenne empirique et la vraie valeur augmente (vers zéro) à mesure que N augmente, ou que ρ baisse ou si X est un bruit blanc. Cependant, une différence importante distingue notre étude des précédentes. Disons d'abord que le biais tend à augmenter avec la variance de la première observation; il augmente tellement qu'il devient quelquefois positif pour les estimateurs stationnaires (PWI, MVS). Par exemple, pour X4 (N=50, bruit blanc), ρ=0,6 et d=2d*, le biais pour PWI est de 0,035. Ces observations contredisent nettement les conclusions des travaux précédents qui ne sont valables que si le modèle est purement stationnaire. Ce résultat peut avoir plusieurs implications, par exemple sur les tests appliqués au coefficient d'autocorrélation et une étude plus approfondie s'imposerait.

En général, la moyenne des estimateurs conditionnels semble moins sensible à la spécification de d. Le biais reste plus stable, mais aussi plus élevé en valeur absolue que celui des méthodes stationnaires.

Les REQMs suivent approximativement la même direction que la valeur absolue du biais excepté par rapport à ρ : elles baissent lorsque d ou N augmentent ou si X est un bruit blanc, mais elles augmentent avec ρ. Une augmentation de la variance de la première observation améliore donc l'estimation de ρ ce qui est normal puisqu'asymptotiquement, la variance de $\hat{\rho}$ est inversement proportionnelle à la somme des u_t^2 qui augmente avec d.

Le meilleur estimateur de ρ, au sens des REQM semble être celui de PWI. Viendraient ensuite dans l'ordre : MVS, MVC et DUR (DUR est cependant préférable à MVC lorsque X comporte une tendance et $N = 50$). Nous constatons ici qu'une estimation plus précise du coefficient d'autocorrélation n'est pas suffisante pour améliorer l'efficacité des coefficients de régression. Dans plusieurs cas, le meilleur estimateur de ρ ne correspond pas au meilleur pour $\hat{\beta}_0$ et $\hat{\beta}_1$ même en comparant DUR et MVC ou PWI et MVS entre eux.

Bien que les méthodes conditionnelles soient moins efficaces au sens des REQMs, dans plusieurs cas, elles sont plus près de la vraie valeur que MVS pour la majorité des reprises. Ceci se produit lorsque d est grand et ρ est relativement petit (0,6, 0,8).

VI.4.2. Distribution de $\hat{\alpha}$

Les statistiques sur $\hat{\alpha}$ (tableau 10) sont pratiquement équivalentes à celles de $\hat{\rho}$ et paraissent indépendantes de $\hat{\beta}_0$. Dans tous les cas, les biais de $\hat{\alpha}$ et $\hat{\rho}$ sont égaux à la deuxième décimale sauf pour le signe qui est inversé. On ne retrouve que deux exceptions à cette règle soit les modèles où $d > d^*$ lorsque X comporte une tendance, $N = 20$ et $\rho=0,99$. Les différences sont alors de l'ordre de 0,1 (entre $\hat{\alpha}$ et $\hat{\rho}$).

Tableau 10

Biais et efficacité de $\hat{\alpha}$

	D=	BIAIS .5	BIAIS 1.0	BIAIS D*	BIAIS 1.5D*	BIAIS 2D*	REQM .5	REQM 1.0	REQM D*	REQM 1.5D*	REQM 2D*	NF .5	NF 1.0	NF D*	NF 1.5D*	NF 2D*
X2 (bruit blanc, N = 20)	RO=.60															
	DUR	.247	.255	.231	.215	.189	.347	.365	.336	.320	.308	.512	.478	.500	.492	.534
	MVC	.247	.253	.228	.211	.181	.348	.362	.336	.317	.302	.567	.510	.518	.521	.544
	PWI	.231	.229	.183	.135	.066	.333	.347	.324	.310	.311	.779	.747	.696	.645	.522
	MVS	.251	.251	.208	.165	.102	.339	.352	.328	.311	.305	ND	ND	ND	ND	ND
	RO=.80															
	DUR	.322	.307	.286	.265	.222	.415	.410	.393	.388	.348	.574	.509	.482	.490	.472
	MVC	.319	.305	.279	.248	.212	.420	.413	.392	.382	.343	.565	.533	.493	.539	.484
	PWI	.292	.267	.219	.185	.118	.387	.379	.359	.360	.311	.906	.867	.779	.716	.667
	MVS	.322	.299	.255	.223	.160	.402	.394	.375	.376	.329	ND	ND	ND	ND	ND
	RO=.99															
	DUR	.429	.419	.473	.428	.376	.513	.550	1.558	1.995	2.344	.555	.533	.573	.610	.650
	MVC	.428	.423	.511	.434	.431	.520	.563	1.742	2.446	2.979	.562	.543	.536	.581	.634
	PWI	.377	.368	.465	.403	.449	.463	.493	1.604	2.261	2.802	.984	.984	.973	.978	.982
	MVS	.414	.406	.506	.451	.503	.488	.522	1.699	2.413	3.019	ND	ND	ND	ND	ND
X1 (avec tendance, N = 20)	RO=.60															
	DUR	.164	.172	.154	.143	.118	.276	.270	.256	.250	.232	.464	.415	.434	.455	.489
	MVC	.139	.147	.132	.124	.100	.264	.255	.243	.236	.219	.566	.530	.507	.565	.552
	PWI	.126	.115	.068	.020	-.043	.260	.248	.243	.257	.259	.673	.692	.609	.493	.396
	MVS	.151	.143	.102	.062	.005	.264	.251	.239	.242	.234	ND	ND	ND	ND	ND
	RO=.80															
	DUR	.214	.197	.190	.159	.146	.303	.285	.275	.249	.244	.429	.390	.340	.412	.449
	MVC	.189	.174	.166	.136	.131	.284	.267	.252	.226	.227	.617	.545	.494	.544	.542
	PWI	.164	.127	.091	.032	-.009	.267	.249	.231	.218	.220	.811	.722	.656	.511	.418
	MVS	.199	.166	.135	.080	.041	.282	.261	.239	.218	.212	ND	ND	ND	ND	ND
	RO=.99															
	DUR	.267	.258	.246	.184	.189	.339	.336	.496	.461	.491	.315	.296	.309	.304	.263
	MVC	.240	.230	.223	.153	.171	.315	.311	.457	.404	.440	.502	.506	.453	.464	.430
	PWI	.181	.169	.142	.093	.100	.266	.263	.367	.312	.328	.987	.985	.954	.970	.970
	MVS	.227	.216	.193	.138	.144	.295	.293	.408	.357	.376	ND	ND	ND	ND	ND

NF : Nombre de fois (sur 1 000) qu'un estimateur est plus près de la vraie valeur que MVS.

Tableau 10 (suite)

Biais et efficacité de $\hat{\alpha}$

	D*	BIAIS .5	BIAIS 1.0	BIAIS D*	BIAIS 1.5D*	BIAIS 2D*	REQM .5	REQM 1.0	REQM D*	REQM 1.5D*	REQM 2D*	NF .5	NF 1.0	NF D*	NF 1.5D*	NF 2D*
X3 (avec tendance, N = 50)	RO=.60															
	DUR	.091	.086	.080	.085	.073	.162	.156	.153	.151	.140	.654	.611	.538	.518	.542
	MVC	.099	.094	.087	.091	.078	.167	.161	.157	.155	.143	.642	.568	.496	.491	.525
	PWI	.094	.083	.063	.038	.000	.164	.157	.152	.152	.160	.774	.733	.671	.604	.513
	MVS	.105	.093	.075	.051	.016	.168	.160	.154	.151	.154	ND	ND	ND	ND	ND
	RO=.80															
	DUR	.109	.117	.113	.096	.094	.159	.168	.162	.152	.147	.695	.605	.544	.507	.529
	MVC	.114	.122	.118	.101	.098	.163	.173	.166	.156	.151	.664	.584	.510	.484	.506
	PWI	.105	.104	.084	.045	.017	.156	.160	.149	.142	.145	.839	.843	.761	.619	.527
	MVB	.119	.119	.100	.063	.037	.164	.169	.156	.145	.143	ND	ND	ND	ND	ND
	RO=.99															
	DUR	.180	.169	.154	.152	.122	.218	.219	.318	.363	.382	.561	.543	.526	.509	.457
	MVC	.182	.171	.154	.153	.126	.221	.222	.321	.368	.381	.544	.517	.516	.506	.480
	PWI	.154	.141	.114	.112	.084	.194	.192	.266	.297	.303	.994	.984	.977	.981	.989
	MVB	.173	.160	.136	.136	.107	.209	.208	.293	.330	.340	ND	ND	ND	ND	ND
X4 (bruit blanc, N = 50)	RO=.60															
	DUR	.053	.059	.048	.058	.048	.133	.140	.129	.135	.126	.524	.464	.459	.476	.538
	MVC	.049	.054	.043	.054	.045	.130	.137	.125	.132	.124	.600	.516	.507	.503	.540
	PWI	.046	.042	.020	-.000	-.034	.129	.134	.124	.142	.155	.632	.604	.545	.506	.428
	MVB	.057	.054	.032	.014	-.018	.132	.136	.124	.138	.145	ND	ND	ND	ND	ND
	RO=.80															
	DUR	.074	.070	.068	.063	.055	.133	.130	.124	.124	.111	.530	.460	.461	.497	.542
	MVC	.072	.068	.066	.061	.054	.130	.127	.122	.121	.109	.637	.524	.514	.527	.558
	PWI	.065	.051	.032	.001	-.034	.126	.119	.113	.121	.125	.732	.676	.611	.481	.378
	MVB	.079	.066	.049	.020	-.011	.132	.124	.116	.118	.114	ND	ND	ND	ND	ND
	RO=.99															
	DUR	.101	.102	.081	.067	.059	.138	.145	.138	.138	.131	.466	.434	.423	.398	.394
	MVC	.100	.101	.081	.067	.059	.135	.144	.136	.136	.131	.503	.452	.449	.416	.399
	PWI	.071	.069	.051	.036	.029	.109	.117	.107	.102	.093	.936	.907	.840	.833	.805
	MVB	.092	.089	.071	.054	.046	.123	.130	.123	.117	.108	ND	ND	ND	ND	ND

Les REQMs de ces deux paramètres sont aussi à peu près identiques, mais les différences sont ici beaucoup plus évidentes. Contrairement à $\hat{\rho}$, les REQMs de $\hat{\alpha}$ tendent à augmenter avec d lorsque ρ est très grand (0,99). Ce sont les seuls cas (ρ=0,99 et d assez grand) où l'on peut noter une différence entre les deux distributions empiriques. Rappelons qu'il s'agit des mêmes cas où les variances empiriques de la constante étaient très élevées, en particulier pour les méthodes conditionnelles, mais aussi pour les estimateurs stationnaires. Il faut donc que la variance de $\hat{\beta}_0$ soit très grande pour influencer celle de $\hat{\alpha}$.

Les meilleurs estimateurs pour $\hat{\alpha}$ restent les mêmes que ceux de $\hat{\rho}$: les méthodes stationnaires sont légèrement supérieures aux autres.

VI.4.3. Distribution de s^2

La variance estimée (s^2) est sûrement la statistique la plus importante dans un modèle de régression. C'est en grande partie parce que son estimation est faussée que les MCO ne sont pas employés en présence d'autocorrélation. Les méthodes stationnaires risquent également d'être biaisées puisque le poids attribué à la première observation est incorrect lorsque $d \neq 1$.

Les statistiques habituelles ont été compilées pour s^2, elles sont présentées aux tableaux 11 et 12. Précisons ici que pour les MCO les statistiques ont été compilées par rapport à $\sigma_e^2/(1-\rho^2)$ soit la variance

Tableau 11

Biais et efficacité de s^2

	D=	BIAIS .5	1.0	D*	1.5D*	2D*	REQM .5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	RO=.60										
	MCO	-.00157	-.00144	-.00106	-.00022	.00113	.0019	.0020	.0024	.0037	.0058
	DUR	-.00035	-.00034	-.00038	-.00030	-.00032	.0012	.0012	.0012	.0012	.0013
	MVC	-.00036	-.00036	-.00041	-.00033	-.00037	.0012	.0012	.0012	.0012	.0012
	PWI	-.00039	-.00030	-.00014	.00025	.00067	.0012	.0012	.0013	.0015	.0020
	MVS	-.00039	-.00030	-.00014	.00026	.00068	.0012	.0012	.0013	.0016	.0020
	DIF	.00093	.00104	.00109	.00139	.00168	.0018	.0019	.0019	.0022	.0026
	RO=.80										
	MCO	-.00519	-.00483	-.00394	-.00282	-.00047	.0029	.0033	.0046	.0066	.0108
	DUR	-.00048	-.00041	-.00041	-.00039	-.00036	.0013	.0013	.0012	.0013	.0013
	MVC	-.00051	-.00045	-.00046	-.00045	-.00045	.0013	.0013	.0012	.0013	.0012
	PWI	-.00042	-.00030	-.00008	.00013	.00062	.0012	.0012	.0012	.0015	.0021
	MVS	-.00041	-.00030	-.00006	.00014	.00064	.0012	.0012	.0013	.0015	.0021
	DIF	.00047	.00055	.00069	.00083	.00119	.0015	.0015	.0016	.0018	.0023
	RO=.99										
	MCO	-.17531	-.17510	-.17492	-.17472	-.17365	.0348	.0351	.0353	.0352	.0364
	DUR	-.00070	-.00068	-.00051	-.00034	-.00001	.0013	.0013	.0013	.0014	.0017
	MVC	-.00076	-.00074	-.00065	-.00057	-.00037	.0013	.0013	.0013	.0013	.0014
	PWI	-.00054	-.00048	-.00033	-.00016	.00030	.0012	.0012	.0013	.0014	.0019
	MVS	-.00053	-.00047	-.00031	-.00014	.00033	.0012	.0012	.0013	.0014	.0019
	DIF	.00008	.00009	.00025	.00039	.00080	.0013	.0012	.0014	.0015	.0020
X2 (bruit blanc, N = 20)	RO=.60										
	MCO	-.00101	-.00088	-.00043	.00050	.00197	.0023	.0025	.0029	.0041	.0067
	DUR	-.00020	-.00021	-.00024	-.00017	-.00016	.0012	.0013	.0012	.0012	.0012
	MVC	-.00022	-.00024	-.00026	-.00020	-.00020	.0012	.0012	.0012	.0012	.0012
	PWI	-.00031	-.00019	-.00002	.00035	.00075	.0012	.0012	.0012	.0015	.0020
	MVS	-.00030	-.00019	-.00001	.00036	.00078	.0012	.0012	.0012	.0016	.0020
	DIF	.00082	.00088	.00094	.00122	.00148	.0018	.0018	.0019	.0021	.0024
	RO=.80										
	MCO	-.00388	-.00329	-.00243	.00007	.00333	.0043	.0049	.0061	.0101	.0151
	DUR	-.00025	-.00027	-.00024	-.00026	-.00022	.0012	.0012	.0012	.0012	.0012
	MVC	-.00028	-.00030	-.00028	-.00031	-.00027	.0012	.0012	.0012	.0012	.0012
	PWI	-.00033	-.00022	-.00005	.00021	.00067	.0012	.0012	.0012	.0014	.0018
	MVS	-.00032	-.00021	-.00003	.00025	.00074	.0012	.0012	.0012	.0015	.0019
	DIF	.00033	.00036	.00046	.00060	.00099	.0014	.0015	.0014	.0016	.0019
	RO=.99										
	MCO	-.16975	-.16973	-.16446	-.15709	-.14685	.0407	.0406	.0480	.0580	.0732
	DUR	-.00042	-.00046	-.00044	-.00035	-.00031	.0012	.0012	.0012	.0013	.0012
	MVC	-.00048	-.00051	-.00050	-.00042	-.00038	.0012	.0012	.0012	.0013	.0012
	PWI	-.00026	-.00029	-.00004	.00025	.00062	.0012	.0012	.0013	.0016	.0020
	MVS	-.00023	-.00025	.00003	.00035	.00075	.0012	.0012	.0013	.0016	.0021
	DIF	.00002	-.00002	.00017	.00041	.00076	.0012	.0012	.0013	.0016	.0020

Pour MCO, il s'agit du biais et REQM par rapport à $\sigma_e^2/(1-\rho^2)$. Pour les autres, les statistiques sont compilées par rapport à σ_e^2.

Tableau 11 (suite)

Biais et efficacité de s^2

		BIAIS					REQM				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X3 (avec tendance, N = 50)	RO=.60										
	MCO	-.00076	-.00066	-.00042	.00002	.00061	.0019	.0020	.0022	.0028	.0036
	DUR	-.00011	-.00012	-.00010	-.00009	-.00010	.0007	.0007	.0007	.0008	.0007
	MVC	-.00011	-.00012	-.00010	-.00009	-.00010	.0007	.0007	.0007	.0008	.0007
	PWI	-.00015	-.00011	-.00002	.00019	.00038	.0007	.0007	.0007	.0009	.0011
	MVS	-.00015	-.00011	-.00002	.00019	.00038	.0007	.0007	.0007	.0009	.0011
	DIF	.00095	.00094	.00100	.00114	.00122	.0014	.0013	.0014	.0015	.0016
	RO=.80										
	MCO	-.00278	-.00270	-.00234	-.00100	.00047	.0045	.0048	.0052	.0068	.0093
	DUR	-.00010	-.00014	-.00017	-.00012	-.00013	.0008	.0008	.0007	.0007	.0007
	MVC	-.00010	-.00014	-.00017	-.00013	-.00013	.0008	.0008	.0007	.0007	.0007
	PWI	-.00012	-.00012	-.00007	.00010	.00028	.0007	.0008	.0007	.0008	.0009
	MVS	-.00012	-.00011	-.00007	.00011	.00028	.0007	.0008	.0007	.0008	.0010
	DIF	.00047	.00046	.00046	.00055	.00068	.0010	.0010	.0009	.0010	.0011
	RO=.99										
	MCO	-.16844	-.16757	-.16361	-.16084	-.15339	.0399	.0408	.0454	.0498	.0584
	DUR	-.00031	-.00032	-.00028	-.00023	-.00023	.0008	.0008	.0008	.0008	.0008
	MVC	-.00031	-.00032	-.00029	-.00025	-.00026	.0008	.0008	.0008	.0008	.0008
	PWI	-.00023	-.00023	-.00009	.00001	.00016	.0007	.0008	.0008	.0008	.0009
	MVS	-.00023	-.00022	-.00008	.00003	.00018	.0007	.0008	.0008	.0008	.0009
	DIF	.00003	.00002	.00009	.00019	.00029	.0007	.0008	.0008	.0008	.0009
X4 (bruit blanc, N = 50)	RO=.60										
	MCO	-.00046	-.00031	-.00013	.00035	.00092	.0022	.0023	.0024	.0031	.0041
	DUR	-.00010	-.00006	-.00008	-.00006	-.00009	.0008	.0008	.0007	.0007	.0008
	MVC	-.00010	-.00006	-.00008	-.00007	-.00009	.0008	.0008	.0007	.0007	.0008
	PWI	-.00015	-.00006	-.00002	.00020	.00033	.0008	.0007	.0007	.0008	.0010
	MVS	-.00015	-.00006	-.00002	.00020	.00033	.0008	.0007	.0007	.0008	.0011
	DIF	.00082	.00091	.00088	.00104	.00108	.0013	.0013	.0013	.0014	.0015
	RO=.80										
	MCO	-.00174	-.00139	-.00085	.00060	.00232	.0059	.0062	.0068	.0086	.0111
	DUR	-.00008	-.00007	-.00008	-.00005	-.00009	.0007	.0007	.0008	.0007	.0007
	MVC	-.00009	-.00008	-.00008	-.00005	-.00010	.0007	.0007	.0008	.0007	.0007
	PWI	-.00013	-.00007	.00001	.00017	.00028	.0007	.0007	.0008	.0008	.0009
	MVS	-.00012	-.00007	.00001	.00017	.00029	.0007	.0007	.0008	.0008	.0009
	DIF	.00038	.00041	.00044	.00053	.00057	.0009	.0009	.0010	.0010	.0010
	RO=.99										
	MCO	-.15470	-.15312	-.14040	-.12169	-.09515	.0544	.0564	.0736	.0988	.1345
	DUR	-.00018	-.00019	-.00017	-.00013	-.00014	.0007	.0007	.0007	.0007	.0008
	MVC	-.00018	-.00019	-.00018	-.00013	-.00014	.0007	.0007	.0007	.0007	.0008
	PWI	-.00011	-.00010	-.00002	.00012	.00023	.0007	.0007	.0008	.0008	.0009
	MVS	-.00009	-.00008	.00000	.00016	.00027	.0007	.0007	.0008	.0008	.0009
	DIF	.00002	.00002	.00007	.00019	.00029	.0007	.0007	.0008	.0008	.0009

Tableau 12

Nombre de fois qu'un estimateur de s^2 est plus près de la vraie valeur (σ_e^2) que MVS

D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
	X1 (avec tendance, N = 20)					X3 (avec tendance, N = 50)				
RO=.60										
MCO	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DUR	.526	.540	.490	.526	.533	.543	.513	.511	.533	.581
MVC	.528	.529	.492	.519	.533	.543	.511	.509	.531	.583
PWI	.328	.353	.392	.508	.590	.409	.427	.462	.553	.601
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.464	.436	.393	.335	.277	.333	.325	.282	.248	.218
RO=.80										
MCO	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DUR	.501	.460	.474	.488	.556	.522	.482	.489	.503	.554
MVC	.487	.449	.464	.480	.552	.521	.483	.489	.503	.554
PWI	.319	.365	.410	.476	.561	.399	.413	.439	.522	.578
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.525	.507	.454	.439	.406	.439	.441	.410	.354	.365
RO=.99										
MCO	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DUR	.475	.464	.476	.483	.508	.476	.460	.458	.489	.524
MVC	.460	.441	.475	.490	.504	.473	.458	.457	.490	.517
PWI	.299	.295	.356	.386	.492	.366	.357	.436	.470	.562
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.603	.611	.581	.564	.539	.577	.564	.543	.581	.610
	X3 (bruit blanc, N = 20)					X4 (bruit blanc, N = 50)				
RO=.60										
MCO	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DUR	.519	.494	.465	.531	.563	.554	.507	.499	.522	.591
MVC	.513	.491	.458	.526	.552	.555	.509	.496	.520	.590
PWI	.354	.390	.435	.546	.620	.396	.446	.467	.560	.604
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.471	.434	.406	.366	.382	.386	.336	.315	.226	.279
RO=.80										
MCO	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DUR	.512	.495	.474	.531	.567	.536	.519	.510	.533	.550
MVC	.509	.497	.473	.529	.565	.536	.515	.511	.532	.548
PWI	.357	.393	.435	.498	.619	.408	.432	.460	.547	.598
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.530	.486	.492	.494	.558	.448	.434	.431	.373	.398
RO=.99										
MCO	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DUR	.454	.420	.484	.517	.573	.456	.464	.474	.509	.522
MVC	.452	.418	.470	.500	.569	.453	.462	.475	.507	.520
PWI	.385	.379	.435	.525	.570	.412	.424	.459	.554	.571
MVS	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
DIF	.555	.549	.586	.627	.681	.530	.517	.539	.571	.610

de long terme des "u". Les comparaisons avec les autres estimateurs doivent donc être évitées parce que ceux-ci sont basés sur σ_e^2.

Comme nous nous y attendions, les estimateurs conditionnels sont ceux qui dépendent le moins de la variance de la première observation. Leurs biais et leurs REQMs varient très peu avec d sauf lorsque ρ est grand et N est petit. Nous remarquons qu'ils sous-estiment systématiquement la variance véritable et ce, d'autant plus que N est petit ou ρ est grand. Le biais (en valeur absolue) est aussi plus grand lorsque X est un bruit blanc. Lorsque N est petit, la différence entre la moyenne empirique et la vraie valeur est de l'ordre de 10 à 20%. Une correction supplémentaire serait donc très profitable (en plus de l'ajustement pour le nombre de degrés de liberté).

Entre les deux estimateurs conditionnels, DUR est généralement mieux centré que MVC, mais ce dernier présente les plus faibles REQMs.

Le biais et les REQMs de tous les autres estimateurs dépendent dans une large mesure de σ_1^2. Dans tous les cas, sauf DIFF, le biais est négatif pour les faibles valeurs de "d" mais devient positif pour les valeurs plus élevées. Par contre, DIFF surestime constamment σ_e^2, et d'autant plus que d est grand. Les REQMs augmentent elles aussi avec d, elles sont souvent presque deux fois plus grandes entre les valeurs extrêmes de "d".

Evidemment, les méthodes stationnaires sont les plus efficaces lorsque d = 1, mais aussi lorsque d = d*. Pour les valeurs plus faibles ou plus grandes de d, il est généralement préférable d'employer un estimateur conditionnel.

VI.4.4. Validité des tests

Nous ne reviendrons pas sur l'importance qu'il faut attribuer à la validité des tests de Student. Rappelons tout de même que CORC est le seul estimateur dont les tests sont corrects pour toutes les valeurs de "d" (si ρ est connu).

En pratique, la situation est très différente. Dans la grande majorité des cas, l'hypothèse nulle ($\hat{\beta} = \beta$) est rejetée beaucoup trop souvent comme le montrent les tableaux 13 à 15. Ces résultats confirment ceux de Park et Mitchell (1980) qui, comme nous, trouvent que les tests ne sont valides que pour le coefficient de la pente ($\hat{\beta}_1$) lorsque X est un bruit blanc. Lorsque N est petit, les estimateurs conditionnels nous semblent nettement préférables puisque la proportion de rejets de PWI, MVS et DIFF est trop grande si d et ρ sont assez grands (β_1, X sans tendance). Lorsque N est plus grand (50), les différences entre ces estimateurs sont moins grandes, mais alors même MVC et DUR tendent à ne pas rejeter l'hypothèse nulle assez souvent.

Dans tous les autres cas, l'hypothèse nulle est rejetée beaucoup plus souvent que ne le laisse supposer le niveau du test; DIFF est l'exception. Notons qu'il est normal de trouver une proportion de rejets très élevée pour $\hat{\alpha}$ (tableau 15) puisque ce coefficient est biaisé. Sa variance a été estimée en utilisant la matrice de covariance $(X'\hat{Q}'\hat{Q}X)^{-1}$ où la première colonne de $\hat{Q}X$ est remplacée par un vecteur unitaire.

Tableau 13

Proportions de rejets de l'hypothèse nulle : $\hat{\beta}_0 = \beta_0$
(distribution de Student)

NIVEAU=		20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	RO=.60															
	MCO	.486	.509	.599	.666	.715	.346	.391	.492	.566	.621	.246	.299	.406	.471	.551
	DUR	.369	.371	.399	.369	.370	.237	.268	.284	.252	.269	.153	.197	.195	.182	.201
	MVC	.384	.395	.427	.389	.390	.260	.284	.295	.263	.281	.160	.198	.217	.186	.210
	PWI	.289	.349	.469	.568	.666	.166	.243	.341	.442	.519	.101	.173	.257	.339	.400
	MVS	.300	.368	.480	.583	.674	.180	.260	.358	.463	.551	.107	.184	.273	.358	.439
	RO=.80															
	MCO	.633	.654	.787	.832	.893	.523	.577	.726	.771	.842	.415	.505	.662	.715	.811
	DUR	.475	.485	.531	.542	.555	.361	.360	.409	.424	.436	.276	.274	.340	.341	.361
	MVC	.523	.525	.554	.547	.541	.389	.407	.444	.417	.434	.295	.313	.358	.343	.348
	PWI	.342	.430	.626	.709	.801	.233	.325	.542	.607	.730	.171	.246	.454	.530	.651
	MVS	.372	.464	.661	.735	.826	.262	.350	.568	.643	.755	.182	.264	.488	.566	.696
	RO=.99															
	MCO	.790	.851	.985	.983	.990	.719	.817	.976	.978	.985	.662	.776	.974	.977	.978
	DUR	.627	.704	.941	.957	.966	.520	.624	.929	.942	.953	.439	.563	.913	.927	.938
	MVC	.686	.728	.899	.915	.923	.574	.649	.871	.889	.894	.480	.589	.852	.869	.876
	PWI	.562	.732	.961	.949	.914	.468	.655	.950	.943	.907	.400	.606	.945	.932	.898
	MVS	.611	.762	.973	.977	.981	.517	.683	.962	.969	.976	.431	.634	.957	.960	.966
X2 (bruit blanc, N = 20)	RO=.60															
	MCO	.539	.544	.535	.547	.563	.437	.429	.429	.427	.461	.348	.344	.336	.333	.357
	DUR	.348	.341	.328	.315	.309	.246	.239	.229	.211	.214	.179	.169	.156	.135	.151
	MVC	.338	.323	.315	.304	.303	.226	.227	.210	.200	.202	.164	.164	.145	.122	.136
	PWI	.307	.295	.285	.259	.246	.209	.195	.184	.156	.161	.152	.142	.138	.111	.103
	MVS	.329	.316	.312	.296	.277	.222	.216	.205	.182	.188	.163	.150	.148	.121	.126
	RO=.80															
	MCO	.669	.690	.724	.776	.795	.583	.618	.632	.695	.720	.519	.555	.555	.627	.661
	DUR	.438	.451	.438	.438	.430	.333	.348	.333	.322	.331	.251	.261	.256	.261	.247
	MVC	.423	.432	.416	.416	.421	.300	.318	.303	.306	.317	.235	.235	.238	.232	.233
	PWI	.338	.365	.335	.332	.255	.235	.258	.251	.256	.201	.173	.194	.189	.200	.147
	MVS	.391	.401	.395	.421	.396	.278	.310	.291	.307	.265	.202	.227	.229	.248	.201
	RO=.99															
	MCO	.916	.952	.984	.985	.992	.887	.939	.978	.981	.989	.865	.930	.976	.976	.988
	DUR	.776	.853	.924	.918	.941	.705	.804	.905	.893	.921	.638	.768	.884	.875	.904
	MVC	.756	.837	.897	.876	.905	.683	.779	.866	.853	.880	.624	.739	.838	.832	.855
	PWI	.560	.670	.641	.518	.472	.504	.621	.631	.500	.462	.457	.579	.621	.486	.454
	MVS	.702	.830	.944	.956	.972	.629	.783	.928	.935	.963	.580	.743	.914	.916	.959

Tableau 13 (suite)

Proportions de rejets de l'hypothèse nulle : $\hat{\beta}_0 = \beta_0$
(distribution de Student)

	NIVEAU=	20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X3 (avec tendance, N = 50)	RO=.60															
	MCO	.536	.550	.544	.553	.602	.415	.447	.454	.438	.475	.332	.347	.358	.348	.391
	DUR	.293	.297	.299	.275	.284	.187	.189	.196	.168	.171	.119	.123	.121	.098	.105
	MVC	.300	.303	.304	.280	.290	.192	.193	.202	.174	.179	.122	.124	.127	.102	.109
	PWI	.272	.294	.313	.322	.366	.164	.185	.199	.207	.217	.115	.115	.126	.138	.142
	MVS	.283	.301	.328	.336	.397	.174	.190	.206	.219	.234	.120	.117	.128	.147	.158
	RO=.80															
	MCO	.657	.688	.724	.724	.774	.567	.614	.635	.650	.703	.508	.546	.562	.589	.661
	DUR	.330	.363	.360	.355	.372	.229	.247	.250	.246	.244	.146	.170	.174	.186	.179
	MVC	.336	.375	.373	.363	.379	.234	.253	.252	.257	.250	.152	.175	.181	.189	.181
	PWI	.286	.362	.410	.429	.479	.168	.236	.288	.304	.315	.119	.161	.217	.210	.236
	MVS	.308	.378	.433	.461	.548	.183	.253	.308	.343	.382	.128	.180	.234	.235	.281
	RO=.99															
	MCO	.874	.939	.973	.980	.995	.835	.925	.969	.978	.991	.804	.905	.966	.973	.986
	DUR	.661	.781	.877	.902	.922	.575	.708	.830	.879	.901	.516	.646	.801	.851	.876
	MVC	.675	.782	.869	.890	.876	.594	.710	.827	.856	.844	.523	.651	.787	.831	.804
	PWI	.609	.765	.844	.818	.725	.525	.705	.821	.793	.715	.442	.647	.796	.775	.707
	MVS	.653	.813	.931	.949	.966	.576	.755	.914	.930	.956	.505	.704	.894	.910	.954
X4 (bruit blanc, N = 50)	RO=.60															
	MCO	.529	.512	.528	.539	.552	.404	.410	.418	.428	.425	.321	.335	.329	.331	.337
	DUR	.245	.246	.254	.249	.247	.158	.145	.155	.151	.152	.102	.090	.102	.104	.098
	MVC	.246	.240	.253	.249	.247	.153	.145	.153	.153	.151	.100	.088	.095	.101	.097
	PWI	.240	.228	.241	.229	.222	.138	.136	.144	.140	.116	.088	.088	.091	.084	.072
	MVS	.245	.237	.252	.241	.240	.149	.144	.152	.152	.134	.101	.090	.099	.087	.080
	RO=.80															
	MCO	.640	.684	.655	.704	.697	.553	.596	.580	.627	.623	.484	.534	.508	.566	.550
	DUR	.292	.329	.295	.306	.295	.201	.223	.206	.204	.188	.145	.144	.138	.144	.135
	MVC	.287	.324	.301	.296	.300	.197	.223	.201	.199	.186	.142	.150	.136	.144	.133
	PWI	.257	.293	.256	.256	.217	.175	.193	.155	.164	.132	.130	.129	.115	.120	.092
	MVS	.281	.317	.283	.299	.283	.193	.215	.178	.190	.155	.143	.144	.130	.142	.108
	RO=.99															
	MCO	.923	.938	.968	.982	.977	.905	.916	.956	.978	.973	.876	.903	.949	.973	.968
	DUR	.698	.735	.786	.835	.808	.606	.665	.729	.781	.761	.548	.616	.676	.736	.716
	MVC	.687	.737	.785	.822	.794	.607	.660	.717	.770	.747	.545	.603	.667	.724	.703
	PWI	.472	.566	.557	.484	.427	.400	.504	.523	.450	.402	.353	.456	.491	.426	.378
	MVS	.601	.706	.827	899	918	.515	.645	.768	.871	.891	.452	.581	.732	.835	.858

Tableau 14

Proportion de rejets de l'hypothèse nulle : $\hat{\beta}_1 = \beta_1$
(distribution de Student)

	NIVEAU=	20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	RO=.60															
	MCO	.489	.498	.581	.641	.682	.365	383	464	.530	.598	.273	296	387	442	.521
	DUR	.378	.368	.390	.359	.364	.247	.269	.275	.247	258	161	.194	199	.179	191
	MVC	.395	.389	.405	.376	.391	.258	.281	.290	262	.277	.177	.203	.214	.191	.205
	PWI	.306	.357	.440	.524	.593	.190	.242	.329	.387	.446	.125	.180	.236	.294	.324
	MVB	.312	.371	.456	.537	.618	.202	.246	.345	.400	.476	.130	185	.254	309	.361
	DIF	.021	.059	.102	.166	.238	.003	.009	.037	.078	.119	0.000	.002	004	.034	.059
	RO=.80															
	MCO	.624	.642	.761	.803	.857	.507	.549	.678	.745	.813	.417	.482	.619	.679	778
	DUR	.463	.485	.509	.514	.511	.345	.366	373	.408	.403	.269	.279	.300	.323	.319
	MVC	.500	.516	.539	.519	.511	.385	.411	.397	.407	.385	.298	.314	.319	.328	308
	PWI	.366	.429	.576	.649	.735	.267	.307	.465	.545	.656	.202	.251	.391	.443	.575
	MVB	.392	.456	.610	.680	.769	.285	.332	.489	.576	.694	.217	.265	.416	.490	.625
	DIF	.042	.063	.168	.273	.406	.007	.015	.075	.155	.258	.002	.004	.037	.074	.150
	RO=.99															
	MCO	.759	.808	.910	.927	.934	.703	.728	.880	.905	.917	.635	.672	.849	.883	.905
	DUR	.585	.618	.748	.755	.785	.472	.536	.677	700	.725	.404	.457	.617	.652	.685
	MVC	.647	.646	.711	.692	.676	.522	.563	.640	.612	.610	.442	.483	.575	.563	.563
	PWI	.535	.584	.808	.850	.877	.437	.499	.751	.805	.850	.360	.427	.700	.754	824
	MVB	.574	.605	.827	.868	.892	.486	.532	.780	.829	.861	.399	.459	.738	.787	843
	DIF	174	.173	.526	.594	.704	.074	.083	.389	.493	.615	.031	.039	.278	.411	.542
X2 (bruit blanc, N = 20)	RO=.60															
	MCO	.218	.235	.254	.306	.360	.100	.107	.134	.167	.196	.050	.052	.068	.078	.092
	DUR	.213	.228	.198	.204	.210	.104	.107	.107	.095	.110	.044	.062	.054	.053	.056
	MVC	.214	.230	.208	.206	.212	.108	.108	.108	.104	.111	.048	.063	.055	.055	.057
	PWI	.218	.214	.217	.233	.231	.096	.110	.110	.126	.115	.046	.064	.061	.054	.052
	MVB	.218	.211	.217	.233	.227	.089	.110	.106	.124	.113	.044	.062	.059	.054	.050
	DIF	.234	.238	.219	.226	.236	.122	.124	.119	.125	.130	.059	.077	.059	.063	.058
	RO=.80															
	MCO	.245	.289	.333	402	.461	.110	144	.171	.216	.264	.049	.075	.079	.096	.096
	DUR	.183	.212	.192	.195	.180	.091	.116	.087	.085	.089	.052	.057	044	.046	.049
	MVC	194	.213	.197	.195	194	.094	.117	.100	.090	.096	.053	.058	.048	.046	.049
	PWI	199	.221	.205	.188	.185	.098	.123	.095	.101	.084	.051	.059	.055	.046	.037
	MVB	.195	.217	201	.183	.178	093	112	.089	.094	.083	.051	.060	.053	.041	.033
	DIF	.216	.229	.205	.194	190	.110	.126	.105	.100	.092	.055	.063	.056	.043	.040
	RO=.99															
	MCO	390	.406	474	538	585	199	196	237	245	.243	077	.068	060	077	055
	DUR	192	197	.181	227	.202	.105	112	103	110	104	052	053	048	058	056
	MVC	197	.207	186	224	.206	109	.118	110	115	108	059	057	.050	059	058
	PWI	201	213	174	.190	164	105	.114	089	.102	084	050	048	036	049	036
	MVB	196	204	.163	.179	154	.097	103	083	.095	075	.047	050	.033	044	032
	DIF	213	214	.182	196	.165	.117	.119	099	.100	.089	055	.054	039	055	.035

Tableau 14 (suite)

Proportion de rejets de l'hypotèse nulle : $\hat{\beta}_0 = \beta_1$
(distribution de Student)

	NIVEAU=	20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X3 (avec tendance, N = 50)	RO=.60															
	MCO	.534	.505	.521	.523	.554	.417	.400	.427	.430	.435	.331	.335	.354	.342	.352
	DUR	.286	.263	.279	.268	.273	.193	.180	.174	.168	.164	.132	.110	.111	.102	.105
	MVC	.292	.272	.284	.270	.278	.195	.185	.184	.175	.167	.135	.116	.116	.107	.105
	PWI	.288	.269	.280	.298	.287	.183	.161	.173	.176	.170	.121	.106	.112	.113	.098
	MVB	.295	.276	.289	.312	.306	.190	.175	.182	.186	.176	.131	.110	.123	.121	.110
	DIF	.006	.004	.002	.001	.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	RO=.80															
	MCO	.659	.658	.670	.674	.728	.575	.572	.576	.589	.651	.496	.501	.500	.530	.595
	DUR	.326	.347	.348	.322	.340	.204	.240	.235	.227	.236	.143	.177	.165	.159	.173
	MVC	.334	.353	.356	.327	.345	.211	.247	.243	.231	.244	.144	.181	.176	.161	.174
	PWI	.288	.334	.346	.340	.367	.189	.223	.247	.214	.242	.123	.165	.180	.141	.164
	MVB	.310	.349	.366	.369	.410	.201	.241	.262	.240	.281	.131	.175	.192	.156	.196
	DIF	.008	.012	.006	.019	.010	0.000	0.000	0.000	.004	0.000	0.000	0.000	0.000	0.000	0.000
	RO=.99															
	MCO	.851	.857	.926	.940	.960	.818	.814	.908	.919	.944	.775	.776	.893	.909	.936
	DUR	.558	.584	.635	.658	.676	.483	.492	.544	.568	.593	.400	.420	.485	.512	.537
	MVC	.571	.594	.639	.646	.644	.491	.505	.532	.555	.566	.409	.428	.478	.496	.503
	PWI	.543	.571	.710	.747	.813	.467	.468	.625	.671	.752	.387	.394	.554	.592	.692
	MVB	.577	.619	.771	.807	.868	.509	.512	.700	.732	.822	.424	.439	.639	.674	.773
	DIF	.139	.167	.366	.451	.631	.063	.081	.256	.348	.528	.022	.033	.174	.257	.439
X4 (bruit blanc, N = 50)	RO=.60															
	MCO	.163	.183	.154	.125	.139	.073	.082	.067	.049	.060	.037	.040	.032	.018	.023
	DUR	.206	.182	.193	.174	.199	.093	.106	.108	.093	.102	.042	.054	.044	.048	.044
	MVC	.208	.182	.190	.173	.199	.093	.104	.107	.093	.103	.041	.055	.045	.049	.044
	PWI	.218	.186	.198	.183	.201	.095	.108	.109	.096	.105	.041	.046	.050	.053	.055
	MVB	.214	.185	.197	.179	.197	.094	.107	.107	.096	.100	.041	.043	.048	.051	.051
	DIF	.253	.228	.251	.220	.251	.141	.128	.133	.122	.126	.061	.077	.081	.066	.068
	RO=.80															
	MCO	.123	.122	.137	.119	.116	.037	.049	.040	.048	.029	.012	.021	.010	.009	.009
	DUR	.188	.197	.187	.153	.187	.093	.091	.091	.077	.100	.052	.040	.039	.037	.038
	MVC	.188	.197	.187	.152	.190	.092	.092	.091	.079	.099	.051	.042	.039	.036	.042
	PWI	.190	.198	.192	.159	.189	.097	.094	.084	.080	.093	.055	.041	.045	.034	.043
	MVB	.189	.192	.190	.154	.185	.095	.092	.081	.074	.089	.052	.039	.041	.031	.042
	DIF	.213	.215	.216	.181	.206	.122	.117	.108	.088	.101	.069	.060	.055	.043	.050
	RO=.99															
	MCO	.068	.067	.072	.052	.033	.019	.017	.013	.008	.009	.004	.003	.002	.002	.002
	DUR	.179	.191	.193	.200	.201	.084	.095	.087	.093	.116	.039	.043	.048	.049	.060
	MVC	.181	.192	.194	.201	.201	.084	.095	.089	.093	.115	.039	.044	.049	.049	.060
	PWI	.186	.194	.183	.189	.182	.083	.092	.088	.086	.099	.038	.038	.047	.045	.046
	MVB	.184	.184	.174	.186	.181	.077	.091	.085	.085	.095	.037	.038	.044	.043	.044
	DIF	.197	.204	.192	.198	.189	.090	.102	.093	.087	.101	.041	.041	.052	.049	.048

Tableau 15

Proportion de rejets de l'hypothèse nulle : $\hat{\alpha} = \alpha$
(distribution de Student)

	NIVEAU=	20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	RO=.60															
	DUR	.799	.791	.793	.740	.726	.740	.728	.727	.689	.661	.691	.686	.671	.637	.602
	MVC	.811	.810	.770	.757	.739	.756	.748	.717	.693	.665	.710	.711	.675	.643	.616
	PWI	.779	.761	.759	.749	.782	.723	.701	.693	.688	.727	.674	.648	.645	.622	.660
	MVB	.798	.771	.767	.753	.747	.743	.710	.700	.682	.681	.689	.664	.653	.618	.624
	RO=.80															
	DUR	.843	.819	.821	.788	.727	.797	.771	.768	.724	.660	.761	.738	.713	.677	.614
	MVC	.858	.844	.833	.782	.712	.826	.804	.778	.711	.658	.790	.759	.726	.664	.603
	PWI	.794	.758	.767	.751	.741	.751	.711	.694	.672	.671	.710	.655	.633	.622	.606
	MVB	.855	.808	.765	.759	.723	.807	.747	.693	.685	.649	.761	.707	.628	.623	.597
	RO=.99															
	DUR	.938	.905	.933	.947	.966	.921	.861	.902	.939	.953	.904	.833	.889	.927	.941
	MVC	.924	.892	.900	.913	.923	.908	.854	.861	.891	.892	.884	.822	.835	.869	.865
	PWI	.959	.882	.909	.925	.896	.917	.829	.883	.909	.879	.866	.785	.867	.890	.865
	MVB	.982	.925	.930	.958	.972	.963	.886	.917	.948	.962	.935	.850	.906	.934	.954
X2 (bruit blanc, N = 20)	RO=.60															
	DUR	.942	.946	.925	.927	.925	.921	.931	.908	.902	.906	.902	.916	.888	.889	.885
	MVC	.945	.952	.938	.908	.925	.924	.926	.919	.886	.902	.903	.901	.898	.872	.878
	PWI	.940	.940	.927	.942	.950	.924	.920	.908	.921	.933	.909	.899	.892	.909	.919
	MVB	.939	.941	.937	.929	.921	.919	.929	.916	.908	.899	.902	.911	.896	.890	.877
	RO=.80															
	DUR	.941	.938	.934	.922	.916	.926	.917	.913	.894	.889	.906	.894	.894	.877	.866
	MVC	.938	.936	.922	.897	.895	.921	.913	.893	.876	.861	.901	.885	.866	.850	.845
	PWI	.931	.930	.919	.933	.947	.914	.904	.895	.921	.936	.891	.884	.866	.903	.909
	MVB	.935	.916	.908	.927	.938	.917	.892	.889	.899	.917	.906	.869	.862	.876	.894
	RO=.99															
	DUR	.970	.951	.924	.931	.946	.951	.941	.899	.901	.917	.944	.935	.878	.886	.895
	MVC	.952	.955	.891	.886	.908	.936	.934	.869	.862	.879	.923	.916	.845	.840	.860
	PWI	.775	.750	.590	.490	.440	.760	.737	.574	.474	.422	.743	.713	.565	.462	.413
	MVB	.970	.960	.902	.932	.946	.956	.938	.862	.909	.922	.944	.919	.844	.879	.886

Tableau 15 (suite)

Proportion de rejets de l'hypothèse nulle : $\hat{\alpha} = \alpha$
(distribution de Student)

	NIVEAU=	20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X3 (avec tendance, N = 50)	RO=.60															
	DUR	.884	.878	.882	.858	.840	.845	.844	.840	.816	.802	.810	.812	.802	.780	.770
	MVC	.884	.883	.880	.861	.849	.861	.857	.836	.819	.806	.837	.830	.803	.786	.780
	PWI	.887	.887	.874	.881	.890	.859	.854	.839	.825	.856	.826	.820	.805	.792	.817
	MVS	.889	.885	.885	.861	.874	.860	.854	.842	.824	.832	.838	.823	.811	.791	.801
	RO=.80															
	DUR	.853	.890	.870	.841	.829	.816	.849	.838	.795	.769	.779	.808	.804	.761	.737
	MVC	.869	.895	.880	.842	.830	.825	.855	.842	.801	.784	.793	.817	.818	.776	.741
	PWI	.852	.858	.855	.851	.884	.805	.817	.809	.808	.836	.770	.791	.779	.779	.802
	MVS	.862	.894	.856	.855	.866	.825	.854	.822	.801	.833	.790	.816	.787	.761	.795
	RO=.99															
	DUR	.958	.943	.895	.922	.926	.948	.926	.862	.898	.904	.938	.910	.839	.873	.886
	MVC	.966	.946	.884	.905	.887	.951	.929	.855	.875	.848	.938	.911	.826	.844	.830
	PWI	.935	.890	.783	.763	.665	.918	.867	.758	.727	.639	.901	.836	.733	.710	.602
	MVS	.984	.938	.907	.934	.964	.974	.926	.879	.911	.949	.959	.909	.858	.890	.932
X4 (bruit blanc, N = 50)	RO=.60															
	DUR	.918	.928	.921	.922	.918	.900	.898	.893	.902	.898	.883	.885	.878	.880	.885
	MVC	.922	.929	.917	.920	.924	.908	.899	.898	.900	.906	.894	.880	.881	.880	.885
	PWI	.929	.929	.917	.926	.938	.911	.917	.900	.908	.920	.901	.903	.885	.899	.904
	MVS	.926	.934	.918	.938	.930	.899	.921	.893	.914	.907	.876	.900	.878	.899	.890
	RO=.80															
	DUR	.923	.931	.916	.916	.908	.904	.902	.892	.893	.871	.885	.892	.863	.870	.841
	MVC	.916	.920	.915	.900	.900	.891	.900	.894	.874	.869	.876	.887	.869	.854	.850
	PWI	.916	.915	.917	.934	.924	.889	.894	.884	.903	.905	.878	.863	.860	.881	.885
	MVS	.917	.920	.899	.913	.931	.893	.897	.877	.891	.906	.881	.873	.848	.868	.896
	RO=.99															
	DUR	.932	.908	.849	.836	.826	.909	.892	.815	.796	.777	.895	.861	.783	.768	.732
	MVC	.938	.913	.846	.835	.820	.918	.890	.812	.787	.759	.900	.860	.777	.764	.725
	PWI	.764	.719	.600	.482	.425	.723	.673	.526	.413	.366	.701	.652	.502	.397	.357
	MVS	.922	.880	.783	.787	.756	.898	.857	.741	.712	.673	.883	.836	.699	.647	.587

Tableau 16

Nombre d'itérations et temps de traitement

			X1 (avec trend, N=20)			X2 (bruit blanc, N=20)		
ρ	"d"		MVC	PWI	MVS	MVC	PWI	MVS
0.6	0.5	N ITERATION=	5.0130	3.8220	3.5950	4.3280	4.0540	3.9880
	1.0	N ITERATION=	5.5270	4.1440	3.8270	4.4860	4.1840	4.0780
	d*	N ITERATION=	6.5090	4.5480	4.1940	4.7220	4.4150	4.2860
	1.5d*	N ITERATION=	6.3580	5.1310	4.6270	4.8620	5.2990	4.7090
	2d*	N ITERATION=	7.7000	5.9440	5.1550	5.0690	5.2010	4.8820
0.8	0.5	N ITERATION=	8.3660	4.7510	4.3600	5.5240	4.2460	4.1190
	1.0	N ITERATION=	9.3500	5.0540	4.5780	6.2550	4.5470	4.2790
	d*	N ITERATION=	10.6240	5.4780	4.7770	6.5850	4.7240	4.3010
	1.5d*	N ITERATION=	11.9300	5.6410	4.8680	6.9550	5.4080	4.3300
	2d*	N ITERATION=	13.1010	6.3460	5.1400	8.0170	7.6110	4.3740
0.99	0.5	N ITERATION=	12.0460	5.7600	4.9810	14.6600	5.9700	4.0720
	1.0	N ITERATION=	11.8000	5.9150	5.1100	16.6550	5.8100	4.0560
	d*	N ITERATION=	18.3160	7.4150	5.8640	22.5380	9.3370	3.8220
	1.5d*	N ITERATION=	23.0130	8.8320	6.5220	29.8270	14.1740	3.6780
	2d*	N ITERATION=	28.8570	10.2090	7.6350	32.0520	17.8840	3.5960
0.6	0.5	N SECONDES=	5.0010	3.6050	4.8180	4.1980	4.2310	5.1250
	1.0	N SECONDES=	5.1400	4.5150	5.8370	4.4740	4.7890	4.8960
	d*	N SECONDES=	6.0480	5.0490	5.6040	5.2170	5.0440	5.3180
	1.5d*	N SECONDES=	5.7620	4.8720	5.8040	4.5780	5.0260	5.2040
	2d*	N SECONDES=	6.7540	5.4750	6.1940	4.6140	5.0130	7.0870
0.8	0.5	N SECONDES=	7.0800	5.3950	5.6860	5.1290	4.1150	4.7340
	1.0	N SECONDES=	7.6840	4.7200	5.6640	5.7290	4.6070	5.3270
	d*	N SECONDES=	8.6960	5.9120	6.2230	5.5380	4.6510	5.1600
	1.5d*	N SECONDES=	9.1660	5.6270	6.4790	5.9240	5.0960	5.6360
	2d*	N SECONDES=	10.2350	6.6890	5.6110	6.8230	7.1550	6.0100
0.99	0.5	N SECONDES=	9.4540	5.2730	6.5900	11.8740	5.6480	5.1760
	1.0	N SECONDES=	9.6860	5.3700	6.8360	13.5940	5.4460	6.0400
	d*	N SECONDES=	14.8550	6.8750	6.9010	16.7880	8.5470	4.9230
	1.5d*	N SECONDES=	17.1760	8.5510	7.8150	23.0390	11.4520	5.3210
	2d*	N SECONDES=	21.8180	9.7500	7.9190	24.2370	14.2640	5.0210

Les mêmes tests ont été effectués en utilisant la distribution normale plutôt que celle de Student. Comme le lecteur pourra le constater à l'Annexe B, les deux tests sont pratiquement identiques et ne doivent alors être utilisés qu'avec la plus grande prudence.

VI.5. Coûts de traitement

Beach et MacKinnon (1978) mentionnaient que leur algorithme (MVS) était plus rapide que MVC et que cet avantage devait être pris en compte dans l'évaluation de ces deux estimateurs. Cette observation n'était basée que sur le nombre d'itérations. En plus de calculer le nombre moyen d'itérations, nous avons aussi considéré les temps requis par MVC, PWI et MVS. Ces résultats sont présentés au tableau 16.

Pour tous les estimateurs, le temps et le nombre d'itérations augmentent avec ρ et d. S'il est vrai que MVS peut être jusqu'à cinq fois plus rapide que MVC, la différence en secondes pour une estimation est infime (moins de 0,03) et peut être considérée comme nulle.

CONCLUSION

Après avoir vu qu'en présence d'autocorrélation, la première observation pouvait être incluse avec un poids erroné. Nous avons démontré que dans un modèle simple, l'estimation conditionnelle était préférable si le rapport entre la variance de la première observation et celle de long terme était plus grande que la valeur critique :

$$2 + (1-\rho^2)/(1-\rho)^2 \ (N-1)$$

Nous avons ensuite étendu ce résultat au modèle général avec un nombre indéterminé de variables explicatives. La valeur critique devient alors :

$$2 + (1-\rho^2) \ \underline{x}_1'(X'Q'QX)^{-1} \ \underline{x}_1$$

si le rapport est plus grand que cette borne, CORC est plus efficace que PW au sens où la différence de leur matrice de covariance est positive semi-définie.

Nous avons ensuite élaboré une vaste étude de Monte-Carlo pour vérifier si les résultats précédents sont affectés par l'estimation de ρ. Les estimateurs conditionnels (MVC et DUR) sont alors ceux qui sont les plus sensible à cette nouvelle inconnue, et d'autant plus que ρ est grand. En fait, la borne théorique reste valable pour le coefficient de la pente (β_1) mais non pour la constante (β_0) d'après les erreurs quadratiques moyennes (EQM). Par contre, tous les résultats théoriques sont confirmés si l'on se fie au nombre de fois où un estimateur est plus près de la vraie valeur que le maximum de vraisemblance stationnaire (MVS). Ceci est dû à un petit nombre de très mauvaises valeurs qui gonflent l'EQM de $\hat{\beta}_0$ (MVC et DUR).

Plusieurs autres statistiques ont été compilées, notamment pour $\hat{\rho}$, $(\hat{\beta}_0(1-\hat{\rho}))$ et s^2. Dans la majorité des cas, les méthodes stationnaires présentent les meilleurs estimés de ρ et α mais sont surpassées par MVC et DUR pour s^2.

Enfin, les tests de Student sur les coefficients de régression ne sont valides que pour la pente, lorsque la variable explicative est un bruit blanc. Dans tous les autres cas, l'hypothèse nulle est rejetée beaucoup trop souvent.

Toutes les études effectuées jusqu'à présent concluaient qu'il est toujours préférable d'inclure la première observation avec un poids de $\sqrt{1-\rho^2}$. Nous ne prétendrons pas ici qu'il est toujours préférable de la laisser tomber, mais nous incitons le lecteur à faire preuve de prudence avant d'effecteur cette transformation; le résultat peut être l'inverse de celui recherché.

Si nous avons pu résoudre certains problèmes dans ce travail, plusieurs autres questions ont été soulevées en cours de route :

- comment vérifier en pratique si la variance de la première observation est plus grande que la borne?
- existe-t-il une méthode plus efficace pour toutes les valeurs de σ_1^2?
- comment doit-on effectuer les tests sur les coefficients de régression?
- le biais de s^2 peut-il être corrigé? Si oui, comment?
- pourquoi l'estimation conditionnelle de la constante est-elle si mauvaise? Son espérance mathématique existe-t-elle?

Ces quelques questions ne sont qu'un bref aperçu du travail qui reste à accomplir.

ANNEXE A

Résultats de Beach et MacKinnon (1978)

Tableau 17

Résultats de Beach et MacKinnon (1978)

BIAS AND ROOT MEAN SQUARE ERROR OF FML AND CONVENTIONAL PROCEDURES[a]

Parameter	True ρ	Trending X, $T = 20$ Bias	RMSE	No.	Trending X, $T = 50$ Bias	RMSE	No.	Nontrending X, $T = 20$ Bias	RMSE	No.
ρ	.6	−.250	.328	96	−.075	.143	100	−.125	.249	94
		−.251	.335		−.076	.143		−.132	.252	
	.8	−.320	.386	104	−.099	.148	91	−.163	.263	96
		−.326	.396		−.102	.150		−.174	.272	
	99	−.442	.497	104	−.153	.178	97	−.219	.281	97
		−.460	.514		−.161	.189		−.234	.300	
β_1	.6	.010	.138	121*	.002	.043	119*	−.001	.035	115*
		.072	.551		.000	.049		.000	.041	
	.8	.009	.215	123*	.005	.076	121*	−.002	.062	126*
		.086	.578		.002	.098		−.047	.495	
	.99	−.025	.597	110	.053	.436	127*	−.018	.450	109
		.047	1.283		.056	1.682		.107	3.078	
β_2	.6	−.006	.088	125*	−.001	.011	106	.002	.041	101
		−.021	.134		.000	.012		.002	.041	
	.8	−.006	.134	119*	−.001	.018	109	.001	.037	98
		−.024	.183		−.001	.022		.001	.037	
	.99	.005	.235	113	−.005	.059	111	.001	.033	107
		−.011	.266		−.004	.065		.001	.033	

[a] The first element in each cell refers to the FML procedure, and the second refers to the conventional ML procedure.

ANNEXE B

Validité des tests avec la distribution normale

Tableau 18

Proportion de rejets de l'hypothèse nulle : $\hat{\beta}_0 = \beta_0$
(distribution normale)

NIVEAU=		20%					10%					5%				
D=		.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	RO=.60															
	MCO	.502	.520	.616	.679	.724	.370	.413	.519	.584	.637	.278	.333	.432	.499	.583
	DUR	.384	.392	.419	.390	.387	.271	.288	.311	.275	.292	.180	.224	.233	.206	.224
	MVC	.401	.417	.441	.403	.408	.294	.305	.324	.288	.301	.189	.240	.248	.209	.235
	PWI	.298	.375	.486	.584	.677	.196	.265	.373	.469	.557	.124	.197	.288	.380	.446
	MVS	.319	.388	.499	.603	.685	.210	.275	.386	.483	.578	.131	.201	.295	.394	.479
	RO=.80															
	MCO	.643	.664	.793	.837	.896	.541	.588	.740	.785	.851	.460	.535	.688	.740	.821
	DUR	.497	.500	.553	.558	.570	.392	.381	.437	.451	.463	.308	.308	.372	.370	.391
	MVC	.544	.536	.575	.565	.552	.423	.425	.472	.443	.452	.337	.350	.385	.374	.376
	PWI	.354	.444	.642	.721	.816	.255	.345	.554	.627	.745	.188	.269	.484	.559	.680
	MVS	.386	.480	.672	.743	.835	.286	.370	.581	.660	.770	.209	.295	.518	.598	.724
	RO=.99															
	MCO	.794	.853	.986	.984	.992	.734	.828	.977	.978	.987	.678	.793	.976	.977	.979
	DUR	.646	.712	.944	.957	.966	.551	.645	.932	.946	.957	.465	.587	.921	.933	.943
	MVC	.696	.735	.903	.916	.925	.604	.663	.880	.896	.902	.521	.613	.856	.876	.882
	PWI	.580	.744	.963	.951	.915	.495	.670	.952	.943	.910	.423	.621	.947	.937	.900
	MVS	.623	.776	.975	.978	.981	.539	.698	.963	.971	.978	.453	.656	.958	.966	.971
X2 (bruit blanc, N = 20)	RO=.60															
	MCO	.550	.561	.551	.560	.579	.458	.460	.447	.449	.486	.379	.371	.372	.368	.404
	DUR	.359	.365	.352	.325	.322	.265	.262	.251	.240	.233	.204	.195	.183	.159	.174
	MVC	.347	.339	.329	.314	.315	.246	.247	.234	.215	.225	.186	.189	.173	.148	.166
	PWI	.323	.311	.302	.272	.254	.225	.214	.204	.178	.172	.176	.162	.158	.125	.123
	MVS	.344	.338	.327	.316	.293	.245	.232	.223	.199	.198	.185	.174	.165	.140	.149
	RO=.80															
	MCO	.680	.701	.734	.786	.804	.601	.631	.651	.713	.737	.545	.575	.589	.651	.684
	DUR	.453	.478	.451	.452	.453	.357	.367	.351	.344	.358	.278	.293	.285	.286	.283
	MVC	.434	.453	.430	.424	.436	.329	.347	.330	.323	.343	.259	.270	.253	.256	.263
	PWI	.356	.380	.346	.340	.273	.253	.286	.273	.267	.212	.196	.214	.208	.222	.164
	MVS	.404	.416	.412	.438	.420	.300	.332	.316	.324	.282	.230	.247	.249	.269	.224
	RO=.99															
	MCO	.919	.953	.984	.986	.992	.895	.943	.978	.982	.990	.870	.931	.977	.977	.989
	DUR	.787	.857	.927	.919	.944	.724	.816	.911	.897	.926	.665	.782	.893	.883	.909
	MVC	.763	.840	.900	.881	.908	.694	.791	.874	.862	.887	.652	.754	.851	.840	.866
	PWI	.569	.672	.641	.519	.473	.516	.636	.633	.500	.465	.474	.595	.622	.492	.455
	MVS	.709	.834	.947	.957	.974	.644	.793	.931	.939	.966	.597	.755	.923	.923	.960

Tableau 18 (suite)

Proportion de rejets de l'hypothèse nulle : $\hat{\beta}_0 = \beta_0$
(distribution normale)

NIVEAU=		20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X3 (avec tendance, N = 50)	RO=.60															
	MCO	.541	.557	.549	.560	.609	.422	.460	.465	.442	.487	.344	.366	.372	.369	.400
	DUR	.297	.298	.306	.281	.292	.196	.200	.202	.175	.182	.126	.130	.126	.109	.109
	MVC	.311	.305	.310	.286	.294	.199	.201	.207	.182	.189	.130	.133	.137	.111	.113
	PWI	.277	.297	.322	.329	.372	.172	.189	.208	.215	.225	.122	.122	.133	.147	.151
	MVS	.288	.304	.335	.341	.408	.184	.197	.214	.224	.247	.123	.127	.140	.157	.168
	RO=.80															
	MCO	.660	.688	.727	.729	.776	.576	.618	.643	.657	.710	.520	.554	.569	.598	.672
	DUR	.336	.371	.363	.357	.380	.234	.257	.258	.257	.252	.164	.175	.187	.194	.190
	MVC	.346	.382	.380	.369	.385	.238	.262	.262	.265	.259	.166	.178	.192	.198	.194
	PWI	.290	.366	.414	.436	.484	.175	.242	.295	.317	.328	.125	.173	.227	.218	.246
	MVS	.316	.384	.439	.472	.556	.192	.259	.325	.356	.398	.137	.189	.244	.251	.293
	RO=.99															
	MCO	.875	.939	.973	.980	.995	.839	.925	.971	.978	.991	.812	.908	.967	.975	.988
	DUR	.664	.785	.878	.907	.923	.587	.717	.831	.883	.901	.521	.658	.805	.858	.880
	MVC	.679	.789	.875	.892	.878	.602	.715	.828	.859	.847	.534	.663	.794	.836	.809
	PWI	.612	.767	.844	.821	.725	.528	.708	.822	.796	.715	.457	.652	.801	.779	.708
	MVS	.655	.814	.931	.949	.967	.586	.759	.917	.931	.956	.518	.716	.898	.911	.954
X4 (bruit blanc, N = 50)	RO=.60															
	MCO	.538	.520	.533	.547	.559	.413	.421	.430	.435	.433	.334	.348	.335	.340	.347
	DUR	.249	.250	.259	.257	.255	.163	.156	.163	.159	.156	.108	.091	.110	.110	.108
	MVC	.249	.247	.258	.254	.251	.159	.150	.157	.160	.158	.108	.092	.106	.108	.107
	PWI	.243	.234	.245	.232	.231	.145	.144	.149	.147	.124	.101	.093	.101	.087	.078
	MVS	.247	.243	.255	.245	.249	.154	.150	.158	.157	.140	.103	.096	.104	.093	.084
	RO=.80															
	MCO	.642	.689	.659	.706	.699	.562	.600	.589	.635	.627	.492	.546	.522	.577	.558
	DUR	.299	.333	.307	.307	.300	.206	.228	.209	.209	.197	.153	.158	.149	.151	.146
	MVC	.292	.327	.309	.298	.306	.204	.230	.208	.203	.194	.146	.157	.144	.151	.144
	PWI	.262	.296	.257	.266	.224	.178	.198	.164	.171	.135	.135	.136	.120	.125	.095
	MVS	288	.323	.290	.305	.290	.198	.223	.186	.199	.165	.148	.151	.131	.150	.114
	RO=.99															
	MCO	.924	.939	.968	.982	.977	.905	.919	.956	.978	.973	.881	.904	.951	.973	.968
	DUR	.703	.739	.790	.838	.808	.612	.674	.729	.785	.765	.555	.624	.685	.742	.723
	MVC	.692	.741	.787	.822	.797	.614	.669	.723	.776	.752	.559	.612	.672	.730	.709
	PWI	.476	.572	.561	.484	.427	.405	.510	.525	.453	.402	.357	.466	.496	.428	.379
	MVS	.604	.708	.828	.900	.921	.520	.649	.771	.873	.895	.460	.596	.736	.840	.863

Tableau 19

Proportion de rejets de l'hypothèse nulle : $\hat{\beta}_1 = \beta_1$
(distribution normale)

	NIVEAU=	20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	RO=.60															
	MCO	.515	.507	.602	.652	.696	.390	.415	.490	.558	.612	.305	.323	.417	.469	.549
	DUR	.393	.387	.406	.378	.379	.267	.291	.305	.274	.283	.197	.224	.221	.203	.214
	MVC	.414	.407	.431	.396	.402	.274	.306	.318	.279	.293	.206	.238	.236	.213	.235
	PWI	.324	.373	.459	.545	.616	.209	.257	.355	.411	.474	.147	.203	.266	.325	.369
	MVB	.335	.389	.475	.565	.629	.221	.263	.363	.425	.504	.154	.208	.282	.337	.398
	DIF	.028	.072	.111	.181	.254	.004	.017	.047	.096	.139	.002	.003	.008	.044	.076
	RO=.80															
	MCO	.634	.653	.775	.814	.861	.535	.563	.692	.754	.818	.453	.505	.641	.703	.788
	DUR	.480	.503	.524	.525	.520	.372	.402	.406	.431	.426	.310	.310	.325	.350	.359
	MVC	.515	.537	.561	.532	.527	.403	.446	.431	.427	.413	.338	.346	.350	.356	.342
	PWI	.389	.446	.590	.659	.748	.280	.331	.477	.566	.669	.228	.270	.423	.481	.604
	MVB	.410	.469	.629	.691	.777	.303	.359	.520	.602	.707	.243	.284	.444	.523	.649
	DIF	.047	.073	.178	.286	.425	.011	.022	.093	.171	.285	.003	.009	.051	.103	.183
	RO=.99															
	MCO	.765	.814	.916	.929	.937	.718	.754	.887	.906	.921	.660	.688	.862	.890	.908
	DUR	.599	.630	.757	.760	.792	.496	.548	.694	.709	.736	.431	.483	.638	.668	.700
	MVC	.658	.655	.720	.698	.683	.548	.577	.661	.631	.621	.470	.513	.592	.588	.583
	PWI	.555	.590	.817	.855	.880	.462	.517	.763	.817	.854	.389	.445	.717	.773	.839
	MVB	.589	.619	.831	.872	.893	.505	.547	.790	.835	.867	.427	.490	.758	.804	.853
	DIF	.191	.192	.537	.604	.718	.088	.099	.423	.509	.631	.044	.052	.326	.441	.569
X2 (bruit blanc, N = 20)	RO=.60															
	MCO	.231	.254	.273	.327	.385	.124	.130	.158	.190	.227	.067	.069	.088	.122	.139
	DUR	.231	.240	.219	.220	.224	.129	.125	.124	.118	.125	.060	.078	.067	.069	.068
	MVC	.232	.252	.225	.230	.231	.127	.124	.127	.119	.133	.064	.078	.070	.075	.069
	PWI	.237	.228	.233	.249	.248	.115	.127	.126	.145	.135	.057	.076	.077	.077	.071
	MVB	.235	.230	.230	.248	.240	.112	.124	.120	.142	.131	.057	.074	.074	.074	.068
	DIF	.253	.258	.241	.236	.254	.140	.143	.134	.141	.154	.081	.085	.084	.082	.076
	RO=.80															
	MCO	.270	.314	.357	.423	.485	.139	.166	.207	.257	.311	.060	.095	.102	.142	.162
	DUR	.205	.230	.208	.208	.201	.109	.127	.114	.103	.106	.062	.079	.060	.059	.059
	MVC	.212	.232	.212	.209	.208	.114	.130	.120	.107	.112	.060	.080	.065	.062	.064
	PWI	.214	.234	.222	.202	.203	.113	.139	.112	.117	.099	.064	.080	.062	.061	.050
	MVB	.211	.230	.215	.193	.195	.109	.136	.109	.110	.095	.061	.077	.062	.059	.044
	DIF	.227	.248	.219	.208	.208	.127	.144	.124	.125	.109	.072	.083	.067	.055	.055
	RO=.99															
	MCO	.406	.446	.500	.572	.609	.223	.231	.287	.308	.325	.120	.114	.120	.116	.096
	DUR	.211	.216	.203	.243	.216	.126	.135	.120	.134	.120	.063	.076	.068	.075	.069
	MVC	.217	.218	.202	.242	.222	.125	.140	.126	.133	.126	.070	.079	.070	.077	.070
	PWI	.216	.225	.191	.203	.179	.124	.134	.106	.114	.091	.062	.070	.054	.072	.050
	MVB	.212	.218	.180	.192	.169	.114	.123	.097	.104	.087	.058	.059	.049	.066	.042
	DIF	.228	.225	.199	.210	.185	.136	.138	.117	.115	.098	.074	.076	.058	.071	.052

Tableau 19 (suite)

Proportion de rejets de l'hypothèse nulle : $\hat{\beta}_1 = \beta_1$
(distribution normale)

	NIVEAU=	20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X3 (avec tendance, N = 50)	RO=.60															
	MCO	.538	.511	.523	.526	.559	.422	.407	.431	.440	.445	.348	.344	.364	.360	.364
	DUR	.290	.275	.281	.274	.279	.198	.189	.184	.180	.171	.136	.121	.118	.109	.112
	MVC	.298	.279	.290	.277	.285	.201	.192	.190	.185	.175	.138	.124	.124	.118	.113
	PWI	.294	.274	.288	.304	.294	.191	.174	.180	.184	.174	.136	.112	.124	.122	.105
	MVS	.302	.283	.292	.318	.314	.198	.184	.191	.194	.186	.144	.115	.130	.130	.118
	DIF	.007	.005	.003	.002	.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	RO=.80															
	MCO	.662	.659	.670	.674	.732	.581	.584	.583	.600	.657	.506	.511	.512	.535	.604
	DUR	.329	.357	.356	.324	.346	.215	.247	.243	.232	.244	.153	.186	.183	.169	.181
	MVC	.342	.361	.365	.331	.351	.219	.254	.254	.233	.250	.159	.193	.190	.171	.185
	PWI	.296	.337	.354	.349	.369	.197	.232	.253	.221	.257	.130	.175	.190	.144	.176
	MVS	.319	.353	.370	.377	.418	.209	.246	.272	.257	.295	.143	.188	.200	.163	.206
	DIF	.009	.012	.008	.019	.010	0.000	0.000	0.000	.005	0.000	0.000	0.000	0.000	0.000	0.000
	RO=.99															
	MCO	.857	.859	.928	.942	.961	.819	.818	.908	.922	.944	.780	.778	.895	.909	.937
	DUR	.567	.586	.637	.665	.678	.497	.502	.548	.575	.604	.412	.431	.494	.517	.545
	MVC	.576	.599	.640	.650	.648	.504	.517	.544	.562	.572	.423	.438	.486	.504	.508
	PWI	.552	.578	.712	.750	.819	.473	.476	.633	.675	.755	.395	.409	.566	.602	.699
	MVS	.579	.622	.774	.809	.869	.512	.523	.704	.735	.826	.436	.450	.649	.689	.783
	DIF	.143	.174	.380	.456	.638	.073	.086	.266	.353	.537	.027	.039	.185	.267	.456
X4 (bruit blanc, N = 50)	RO=.60															
	MCO	.167	.185	.162	.133	.140	.082	.085	.070	.055	.062	.041	.046	.033	.022	.028
	DUR	.216	.189	.198	.180	.204	.101	.112	.114	.076	.114	.045	.061	.053	.053	.052
	MVC	.213	.186	.196	.178	.203	.099	.112	.113	.096	.114	.046	.061	.054	.054	.051
	PWI	.226	.191	.204	.190	.209	.105	.114	.114	.098	.109	.049	.057	.058	.055	.056
	MVS	.222	.193	.202	.188	.204	.103	.113	.112	.098	.107	.046	.055	.055	.054	.056
	DIF	.260	.232	.261	.224	.255	.151	.136	.142	.129	.133	.067	.085	.089	.074	.074
	RO=.80															
	MCO	.127	.126	.148	.122	.124	.044	.055	.044	.052	.037	.013	.023	.012	.013	.012
	DUR	.195	.202	.193	.160	.195	.097	.096	.096	.081	.104	.059	.048	.041	.040	.046
	MVC	.194	.201	.193	.157	.198	.097	.096	.096	.082	.103	.059	.048	.041	.041	.047
	PWI	.196	.202	.194	.164	.198	.103	.101	.093	.084	.095	.059	.051	.048	.041	.050
	MVS	.193	.199	.192	.160	.191	.099	.100	.091	.083	.094	.058	.048	.046	.037	.046
	DIF	.219	.223	.222	.192	.209	.132	.124	.115	.092	.108	.072	.064	.067	.054	.057
	RO=.99															
	MCO	.074	.070	.075	.055	.034	.021	.020	.018	.008	.010	.006	.004	.004	.002	.003
	DUR	.189	.197	.199	.202	.204	.090	.099	.096	.102	.118	.042	.045	.053	.052	.069
	MVC	.188	.196	.197	.203	.203	.090	.099	.096	.102	.118	.041	.045	.053	.053	.069
	PWI	.193	.203	.194	.195	.186	.093	.102	.098	.093	.104	.040	.041	.052	.049	.053
	MVS	.186	.195	.182	.189	.183	.087	.096	.091	.087	.099	.040	.040	.050	.046	.050
	DIF	.200	.214	.197	.204	.192	.101	.108	.103	.099	.109	.045	.049	.055	.056	.056

Tableau 20

Proportion de rejets de l'hypothèse nulle : $\hat{\alpha} = \alpha$
(distribution normale)

	NIVEAU=	20%					10%					5%				
	D=	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X1 (avec tendance, N = 20)	RO=.60															
	DUR	.805	.800	.799	.749	.737	.747	.742	.737	.698	.673	.712	.702	.696	.666	.627
	MVC	.822	.819	.779	.767	.752	.766	.760	.733	.710	.682	.731	.726	.697	.668	.635
	PWI	.785	.765	.765	.754	.789	.732	.712	.714	.704	.739	.700	.666	.675	.644	.684
	MVS	.803	.781	.775	.761	.755	.762	.722	.717	.692	.699	.711	.687	.670	.635	.652
	RO=.80															
	DUR	.849	.825	.822	.793	.734	.813	.778	.776	.736	.682	.776	.752	.743	.691	.631
	MVC	.863	.848	.837	.792	.718	.837	.811	.791	.721	.671	.806	.780	.748	.681	.619
	PWI	.804	.764	.774	.761	.751	.759	.720	.709	.692	.686	.727	.679	.657	.648	.626
	MVS	.860	.819	.776	.772	.733	.818	.760	.712	.709	.657	.782	.720	.643	.639	.620
	RO=.99															
	DUR	.941	.911	.934	.948	.972	.928	.871	.907	.940	.955	.910	.839	.896	.931	.945
	MVC	.925	.894	.901	.916	.923	.916	.861	.872	.899	.901	.897	.834	.846	.877	.874
	PWI	.961	.884	.910	.926	.896	.925	.841	.886	.910	.883	.888	.805	.876	.898	.871
	MVS	.984	.928	.931	.960	.972	.969	.894	.923	.954	.964	.955	.861	.908	.939	.955
X2 (bruit blanc, N = 20)	RO=.60															
	DUR	.945	.947	.927	.930	.927	.926	.934	.912	.909	.908	.911	.921	.892	.893	.891
	MVC	.948	.953	.942	.911	.928	.931	.932	.925	.894	.911	.909	.909	.907	.875	.886
	PWI	.941	.941	.930	.944	.953	.931	.928	.911	.928	.936	.915	.909	.897	.914	.927
	MVS	.941	.941	.938	.931	.925	.923	.929	.919	.913	.903	.910	.921	.906	.895	.882
	RO=.80															
	DUR	.941	.939	.937	.924	.916	.932	.923	.919	.895	.895	.913	.905	.901	.885	.875
	MVC	.941	.937	.925	.899	.902	.924	.919	.895	.880	.873	.910	.894	.877	.864	.852
	PWI	.937	.931	.919	.935	.950	.916	.912	.901	.922	.937	.901	.893	.876	.910	.916
	MVS	.937	.921	.910	.927	.942	.920	.898	.893	.905	.922	.910	.879	.870	.880	.902
	RO=.99															
	DUR	.970	.955	.927	.935	.950	.956	.943	.906	.908	.921	.946	.938	.888	.892	.904
	MVC	.954	.956	.896	.887	.913	.939	.939	.873	.866	.889	.929	.927	.855	.846	.866
	PWI	.777	.751	.593	.494	.443	.761	.739	.575	.477	.428	.752	.725	.567	.468	.415
	MVS	.971	.964	.907	.935	.949	.958	.941	.868	.913	.929	.950	.928	.850	.894	.898

Tableau 20 (suite)

Proportion de rejets de l'hypothèse nulle : $\hat{\alpha} = \alpha$
(distribution normale)

NIVEAU=		20%					10%					5%				
D=		.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*	.5	1.0	D*	1.5D*	2D*
X3 (avec tendance, N = 50)	RO=.60															
	DUR	.887	.879	.884	.859	.844	.850	.845	.843	.823	.809	.816	.818	.809	.784	.774
	MVC	.885	.885	.882	.864	.849	.862	.861	.841	.821	.807	.841	.832	.810	.788	.785
	PWI	.887	.887	.876	.882	.890	.863	.855	.842	.829	.858	.829	.825	.813	.797	.820
	MVS	.889	.887	.886	.862	.874	.863	.857	.848	.827	.835	.841	.830	.819	.794	.807
	RO=.80															
	DUR	.854	.896	.871	.844	.832	.818	.855	.840	.799	.775	.782	.814	.810	.763	.742
	MVC	.869	.898	.882	.845	.833	.827	.862	.846	.804	.786	.800	.822	.822	.780	.745
	PWI	.857	.861	.859	.852	.884	.809	.820	.812	.815	.840	.776	.792	.786	.782	.807
	MVS	.863	.894	.860	.856	.868	.826	.861	.825	.803	.833	.795	.822	.793	.766	.799
	RO=.99															
	DUR	.959	.943	.896	.922	.926	.950	.926	.869	.901	.908	.941	.913	.843	.877	.887
	MVC	.967	.947	.885	.905	.887	.951	.932	.859	.875	.853	.943	.911	.829	.850	.833
	PWI	.935	.890	.785	.763	.665	.921	.869	.759	.728	.641	.904	.843	.738	.710	.608
	MVS	.984	.938	.908	.934	.966	.975	.927	.880	.914	.951	.960	.913	.864	.893	.935
X4 (bruit blanc, N = 50)	RO=.60															
	DUR	.919	.928	.922	.924	.919	.902	.899	.895	.902	.899	.886	.886	.883	.883	.887
	MVC	.923	.930	.917	.921	.924	.908	.903	.898	.900	.908	.899	.883	.884	.885	.890
	PWI	.930	.929	.919	.928	.938	.912	.918	.903	.908	.923	.906	.905	.888	.899	.907
	MVS	.930	.935	.921	.938	.930	.901	.921	.895	.914	.910	.879	.901	.881	.902	.893
	RO=.80															
	DUR	.925	.931	.917	.916	.909	.904	.907	.895	.894	.873	.889	.893	.867	.876	.844
	MVC	.919	.921	.916	.903	.900	.893	.903	.899	.876	.872	.879	.889	.873	.856	.851
	PWI	.917	.918	.917	.936	.927	.895	.898	.889	.906	.907	.881	.864	.863	.885	.888
	MVS	.919	.920	.900	.914	.932	.897	.899	.879	.893	.909	.884	.876	.852	.872	.897
	RO=.99															
	DUR	.933	.909	.851	.839	.831	.910	.893	.816	.797	.783	.899	.865	.788	.772	.736
	MVC	.939	.914	.848	.836	.822	.918	.894	.816	.792	.763	.904	.863	.782	.769	.730
	PWI	.770	.726	.607	.496	.440	.724	.673	.526	.415	.367	.704	.657	.506	.402	.359
	MVS	.923	.881	.785	.789	.762	.901	.862	.743	.719	.683	.887	.840	.706	.655	.608

ANNEXE C

Liste des variables explicatives

Tableau 21

Valeurs des variables explicatives

X1, tel que $x_t = \exp(0,04t) + v_t$, $v_t \underset{\sim}{\text{ind}} N(0, \sigma_v^2)$ $t = 1, \ldots, 20$

.911297778900988	1.049896921156382	1.160335855416342	1.142831310452848	1.216939108374994
1.274231212786763	1.361117496835220	1.396553261675265	1.469973621276715	1.467260919019552
1.583163821398749	1.655794538665326	1.649329055088401	1.716713642004272	1.792947108642451
1.885268293830507	1.983117777784592	2.051785392738637	2.136629765039260	2.241324516416398

X2, tel que $x_t = v_t$, $v_t \underset{\sim}{\text{ind}} N(0, \sigma_v^2)$ $t = 1, \ldots, 20$

.391311613560260	-.077978738633302	.016424009114587	.114677741401911	.080884484710341
.043969019135161	.132331639340332	.095600888470163	-.075596923417061	-.023128753678908
-.371842617545379	.051714778022373	.107575518435878	-.372958158693720	.143134932360647
.051805433930495	-.420319652561757	-.353825324807628	-.226096974168000	.341765296101601

X3, tel que $x_t = \exp(0,04t) + v_t$, $v_t \underset{\sim}{\text{ind}} N(0, \sigma_v^2)$ $t = 1, \ldots, 50$

1.080650002450668	1.056713744288103	1.128233376281969	1.134598544843158	1.210568244989581
1.279134612746311	1.350764605811293	1.354736536163635	1.417721455893293	1.475313842593728
1.547835629819595	1.613251037678467	1.674903816973931	1.720278557660386	1.819602887977730
1.927586065575468	1.976986191461968	1.967470751266688	2.137083424502450	2.286793236506412
2.372497438884920	2.411242673242612	2.478412955783014	2.597076949361181	2.752276920164107
2.790585890696562	2.954150895314243	3.081149365198755	3.140650124062986	3.374449633398953
3.470394425115543	3.562551193177327	3.787890290716106	3.861174515597298	4.083190608256359
4.280650439914098	4.374006229014270	4.584399180682993	4.760900096171298	4.973572698363824
5.152533078102380	5.298360163302078	5.628320513868374	5.843043980373693	6.084707304177329
6.323342140461762	6.526027398607226	6.849991153616031	7.077953043500713	7.392066515749860

X4, tel que $x_t = v_t$, $v_t \underset{\sim}{\text{ind}} N(0, \sigma_v^2)$ $t = 1, \ldots, 50$

-.207755728104888	-.005875600013835	.369655234721911	.297242287221774	.196651692960262
.199983344549842	-.503798101983694	.195876238755965	.003716306140006	-.240520262569296
.096820716746450	.178110329413923	-.113844259822651	.006409343399031	-.076138951315485
.050559621506121	-.112374638808976	-.333693208872306	-.499931479867893	.069988667856472
-.083927365335601	-.275111424050770	.005011592723880	.323805555901135	-.152766205798844
.069862191416932	-.311256310111707	.488704176258107	.163417554381279	-.146235537444252
-.346901764200423	.156958262958950	-.216914551806569	-.181519700360020	-.356103057643240
.304545441397259	.144037467352087	-.309897347669501	.231063712578328	.197863160059344
.055887870356865	-.485611312298362	.117873477949231	.031251098531405	-.012800397426668
-.101673932959375	.387041167669883	-.145587088854974	.168994818222761	-.209196551298735

De gauche à droite, nous avons $t = 1, \ldots, 50$

ANNEXE D

Principaux programmes utilisés

Programme EXPL : génération des variables explicatives, des valeurs de $\sigma_1^2(1-\rho^2)/\sigma_e^2$ et des matrices de covariances

```
A1BRG,CM100000,T200.
*CODE
IMSL4.
FTN(PMD,OPT=2,R=3)
LGO.
REWIND,TAPE1.
REWIND,TAPE2.
LIB,C,TAPE1,N=VAR.
MONTE,GROUPE1.
REQUEST,CORC2,*PF,SN=GROUPE1.
LIB,C,TAPE2,N=COV,F=CORC2.
CATALOG(CORC2,XR=C2X2)
RETURN,CORC2.
*WEOR
*FTN
*NUM
      PROGRAM EXPL (INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT,TAPE1,TAPE2)
*
* CE PROGRAMME GENERE 4 SERIES DE VARIABLES EXPLICATIVES.
* 2 DE DIMENSION 20 ET 2 DE DIMENSION 50.
* DANS CHAQUE CAS, UNE SERIE COMPORTE UN TREND X(T)=EXP(.04T)+E(T)
* L'AUTRE EST UN BRUIT BLANC.
* DANS UN DEUXIEME TEMPS, NOUS CALCULONS LES COVARIANCES THEORIQUES
* DES ESTIMATEURS MCO,CORC,PW ET MCG AVEC DIFFERENTES HYPOTHESES
* POUR RHO ET LA VARIANCE DE U(1).
*
      DIMENSION E(50),X(50,4),XXINV(3),XGXIN(3)
      DIMENSION XVX(3),XPV(3),Y(50,2),XVXIN(3)
      DIMENSION COVMCO(3),XPXIN(3),COVPW(3)
      REAL NM,NP,NG
      DOUBLE PRECISION DSEED
      WRITE(6,900)
  900  FORMAT(1H1)
*
      N=20
      J=0
      T=.04
      DSEED=1.01D0
   10  CONTINUE
*
*CONSTITUONS D'ABORD LES SERIES DE DIMENSION 20
*
      DO 40 I=1,2
      L=0
      J=J+1
      CALL GGNML(DSEED,N,E)
      S=.25
      IF(I.EQ.2)GOTO 20
*
*LA SERIE SANS TREND A UN ECART-TYPE DE .25, L'AUTRE .03
*
      L=1
      S=.03
   20  CONTINUE
      DO 30 K=1,N
   30  X(K,J)=L*EXP(T*K)+E(K)*S
   40  CONTINUE
      IF (N.EQ.50)GOTO 50
      N=50
      GOTO 10
*
* ON REVIENT EFFECTUER LES SERIES DE DIM 50
*
   50  CONTINUE
      WRITE(1,91)
   91  FORMAT(1X,5HDSEED)
      WRITE(1,9)DSEED
    9  FORMAT(1X,D40.30)


***** CALCUL DES VARIANCES THEORIQUES ******


      DO 1000 I=1,4
      WRITE(6,900)
      N=20
      IF (I.GT.2)N=50
      WRITE(1,999)
  999  FORMAT(//)
      WRITE(1,92)I
   92  FORMAT(1X,7HEXPL    ,I5)
      WRITE(1,99)(X(J,I),J=1,N)
   99  FORMAT(1X,5F20.15)
      WRITE(1,990)
  990  FORMAT(1H0,4HDEL2)
      C=X(1,I)
      CALL SUM(N,X(1,I),S,S2,S21)

      XN=N
      CALL INVR(XN,S2,S,XXINV)

***** MATRICE DE COV DE CORC ****

      RHO=0.6
      DO 1000 J=1,3
      WRITE(6,900)
      WRITE(6,999)
      IF(J.EQ.2)RHO=.8
      IF(J.EQ.3)RHO=.99
       RHO2=RHO**2
       NG=(N-1)*((1.0-RHO)**2)
       SG2=S2*(1.+RHO2)-2.*RHO*S21-C**2-(X(N,I)*RHO)**2
       SG=(1.-RHO)*((S-C)-RHO*(S-X(N,I)))
       CALL INVR(NG,SG2,SG,XGXIN)

**** CALCUL DE XPPX INVERSE *****

       NP=NG+(1.-RHO2)
       SP2=SG2+(1.-RHO2)*C**2
       SP=SG+(1.-RHO2)*C
       CALL INVR(NP,SP2,SP,XPXIN)

**** CALCUL DES DELTAS****

       DRHO=1./(1.-RHO2)
       BORNE=2.*DRHO+XGXIN(1)+2*C*XGXIN(3)+XGXIN(2)*C**2
       BRN=SQRT(BORNE*(1-RHO2))
       DO 1000 K=1,5
       GOTO (81,82,83,84,85),K
   81   DSIG=.5
       GOTO 90
   82   DSIG=1.0
       GOTO 90
   83   DSIG=BRN
       GOTO 90
   84   DSIG=1.5*BRN
       GOTO 90
   85   DSIG=2*BRN
   90   CONTINUE
       DSIG2=DSIG**2
       DEL2=DSIG2*DRHO
```

```
      WRITE(1,99)DEL2
      DEL=SQRT(DEL2)

*** CALCUL DE X'VX ***

      DO 110 IX=1,N
      Y(IX,1)=Y(IX,2)=0.0
      DO 110 JX=IX,N
      KX=JX-IX
      IF(KX.NE.0)GOTO 100
      Y(IX,2)=Y(IX,2)+X(JX,I)
      Y(IX,1)=Y(IX,1)+1.
      GOTO 101
 100  CONTINUE
      RHOK=RHO**KX
      Y(IX,2)=Y(IX,2)+X(JX,I)*RHOK
      Y(IX,1)=Y(IX,1)+RHOK
 101  CONTINUE
 110  CONTINUE
      Y(1,1)=Y(1,1)*DEL
      Y(1,2)=Y(1,2)*DEL
      XVX(1)=XVX(2)=XVX(3)=0.0
      DO 120 IX=1,N
      XVX(1)=XVX(1)+Y(IX,1)**2
      XVX(2)=XVX(2)+Y(IX,2)**2
      XVX(3)=XVX(3)+Y(IX,1)*Y(IX,2)
 120  CONTINUE

*** MATRICE DE COV MCO ***
      CALL COV(XXINV,XVX,COVMCO)
      D=DEL2*(1.-RHO2)**2

*** MAT COV PW ***

      XPV(1)=NQ+D
      XPV(2)=SQ2+D*(C**2)
      XPV(3)=SQ+D*C
      CALL COV (XPXIN,XPV,COVPW)

*** COV MCG ***

      NM=NQ+1./DEL2
      SM=SQ+C/DEL2
      SM2=SQ2+(C**2)/DEL2
      CALL INVR(NM,SM2,SM,XVXIN)

      PRINT 999
      PRINT *,"EXPL=    ",I,"      DEL2=        ",DEL2
      PRINT*," VAR(U1)/VAR(U) =   ",DSIG2
      PRINT*," S.D(U1)/S.D(U) =   ",DSIG
      PRINT*," RHO=               ",RHO
      PRINT 999
      DO 200 ID=1,3
      COVMCO(ID)=COVMCO(ID)*.0036
      E(ID)=XGXIN(ID)*.0036
      COVPW(ID)=COVPW(ID)*.0036
      XVXIN(ID)=XVXIN(ID)*.0036
      WRITE(2,909)COVMCO(ID),E(ID),COVPW(ID),XVXIN(ID)
      WRITE(6,909)COVMCO(ID),E(ID),COVPW(ID),XVXIN(ID)
 200  CONTINUE
 909  FORMAT(1HO,4F20.15)
 500  CONTINUE
1000  CONTINUE
      STOP
      END

      SUBROUTINE SUM(N,X,S,S2,S21)
      DIMENSION X(N)

      PRINT*,"X(1)=    ",X(1),"      X(N)=    ",X(N)
      S=S2=S21=0.0
      DO 10 I=1,N
      I1=I-1
      S=S+X(I)
      S2=S2+X(I)**2
      IF(I.EQ.1)GOTO 10
      S21=S21+X(I)*X(I1)
  10  CONTINUE
      RETURN
      END

      SUBROUTINE INVR(X1,X2,X3,XIN)
      DIMENSION XIN(3)
      D=X1*X2-X3**2
      XIN(1)=X2/D
      XIN(2)=X1/D
      XIN(3)=(-1.)*X3/D
      RETURN
      END

      SUBROUTINE COV(XX,XVX,COVA)
      DIMENSION XX(3),XVX(3),COVA(3)
      COVA(1)=XVX(1)*XX(1)**2+2.*XX(1)*XX(3)*XVX(3)+XVX(2)*XX(3)**2
      COVA(2)=XVX(2)*XX(2)**2+2.*XX(2)*XX(3)*XVX(3)+XVX(1)*XX(3)**2
      COVA(3)=XX(1)*XX(3)*XVX(1)+XX(2)*XX(3)*XVX(2)+
     + XX(1)*XX(2)*XVX(3)+XVX(3)*XX(3)**2
      RETURN
      END
```

Programme CARL : estimations

```
*WEOR
*FTN
*NUM
      PROGRAM CARL(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT,TAPE1,TAPE2,
     ,TAPE3,TAPE4)

      DOUBLE PRECISION DSEED
      COMMON/REPL/ID,ITMAX(15,3),ICON(15,3)
      COMMON/CONVER/CONV(15,2)
       COMMON/SQ/X1,X2,XN,Y1,Y2,YN
       COMMON/RS/Y(50)
       COMMON/XY/SXY(3),RHO
       COMMON/EXPL/X(50),N
       COMMON/QSTAT/BIAS(15,5,5),REGM(15,5,5),BEST(15,4,5),STATT(15,5,9
     ,),STATN(15,5,9),TN(3),TN1(3),ANOR(3)

       DIMENSION DEL2(15),PAR(8,5),XX(3),XXIN(3)
       DIMENSION E(50),U(50)
       DIMENSION SITER(15,3),SEC(15,3)
       DIMENSION RO(3)

       DATA RO/.6,.8,.99/
       DATA SIQ/.06/
       DATA SITER/45*0.0/,SEC/45*0.0/
       DATA BIAS/375*0.0/,REGM/375*0.0/,BEST/300*0.0/,STATT/675*0.0/,STA
     ,TN/675*0.0/
       DATA ITMAX/45*0/,ICON/45*0/,CONV/30*0.0/
       DATA ANOR/1.282,1.645,1.960/
       DATA TN1/1.333,1.740,2.110/
       DATA TN/1.330,1.734,2.101/
       DATA PAR/40*0.0/

       PRINT 9999
 9999   FORMAT(1H1)
       READ(1,*)N
       NR=0
       IF(N.EQ.20)GOTO 10
       TN1(1)=1.3001
       TN1(2)=1.6776
       TN1(3)=2.0116
       TN(1)=1.300
       TN(2)=1.6774
       TN(3)=2.0114
   10   CONTINUE

       NREPL=1000
       READ(2,9)DSEED
       WRITE(6,9199)DSEED
 9199   FORMAT(1X,D40.30)
    9   FORMAT(D40.30)
       READ(1,*) (X(I),I=1,N)
       READ(1,*) (DEL2(I),I=1,15)
       X1=X(1)
       X2=X(2)
       XN=X(N)
       PRINT*," X1=  ",X1,"  XN=  ",XN
       PRINT*," DEL2= ",(DEL2(I),I=1,15)

*** ON CALCULE IMMEDIATEMENT X'X ET (X'X)-1
*** LES 2 PREMIERS ELEMENTS SONTS LES TERMES DIAGONAUX

       CALL SUMXX(N,X(1),XX(1))
       SXY(1)=N
       SXY(2)=XX(2)
       SXY(3)=XX(1)
       CALL INVR(SXY,XXIN)

****** DEBUT DES ITERATIONS ********

       DO 1000 IREPL=1,NREPL
       ID=IB=0

**** DEBUT DE LA BOUCLE POUR RHO *****

       DO 1000 IRO=1,3
       RHO=RO(IRO)

**** BOUCLE POUR LES 5 VALEURS DE DELTA (SIGMA1) ****

       DO 1000 IDEL=1,5
       ID=ID+1
       DEL=DEL2(ID)

   30   CONTINUE

       SDEL=SQRT(DEL)
***** CALCULONS LES VARIABLES ALEATOIRES ET DEPENDANTES ***

       CALL GGNML(DSEED,N,E)
       E(1)=E(1)*SIG
       U(1)=SDEL*E(1)
       Y(1)=1.0+X(1)+U(1)
       DO 40 I=2,N
       E(I)=E(I)*SIG
       I1=I-1
       U(I)=RHO*U(I1)+E(I)
       Y(I)=1.0+X(I)+U(I)
   40   CONTINUE

       Y1=Y(1)
       Y2=Y(2)
       YN=Y(N)
   99   FORMAT(1X,3F20.15)

****** ESTIMATIONS ********

       CALL MCO(N,PAR(1,1),XXIN,XX,X)
       CALL DUR(N,PAR(1,2),XXIN,XX,X)
       CALL CORC(N,PAR(1,3),XXIN,XX,X,ITR,SECD)
       SITER(ID,1)=SITER(ID,1)+.001*ITR
       SEC(ID,1)=SEC(ID,1)+SECD
       CALL PW(N,PAR(1,4),XXIN,XX,X,ITR,SECD)
       SITER(ID,2)=SITER(ID,2)+.001*ITR
       SEC(ID,2)=SEC(ID,2)+SECD
       CALL BMK(N,PAR(1,5),XXIN,XX,X,ITR,SECD)
       SITER(ID,3)=SITER(ID,3)+.001*ITR
       SEC(ID,3)=SEC(ID,3)+SECD
       CALL SSTATS(ID,PAR,N,RHO)
  999   FORMAT(//)

 1000   CONTINUE


       PRINT*," NOMBRE DE REPRISES =  ",IREPL
       WRITE(6,939)((SITER(I,J),J=1,3),I=1,15)
  939   FORMAT(1HO,15H N ITERATION=   ,3F10.4)
       WRITE(6,949)((SEC(I,J),J=1,3),I=1,15)
  949   FORMAT(1HO,15H N SECONDES=    ,3F10.4)
       WRITE(6,959)((ITMAX(I,J),J=1,3),I=1,15)
  959   FORMAT(1HO,15H ITER MAXI=     ,3I10)
       WRITE(6,969)((ICON(I,J),J=1,3),I=1,15)
```

```
*FTN ,
*NUM
***
      FUNCTION SUM2(A,B,N)
      DIMENSION A(N),B(N)
      DOUBLE PRECISION S2
      S2=0.0
      DO 10 I=1,N
   10 S2=S2+A(I)*B(I)
      SUM2=SNGL(S2)
      RETURN
      END

      FUNCTION RHOCO(E,N)
      DIMENSION E(N)
      N1=N-1
      S12=SUM2(E(1),E(2),N1)
      S2=SUM2(E,E,N1)
      RHOCO=S12/S2
      RETURN
      END


      FUNCTION RHOPW(E,N)
      DIMENSION E(N)
      N1=N-1
      N2=N-2
      RH=SUM2(E(1),E(2),N1)/SUM2(E(2),E(2),N2)
      RHOPW=RH
      RETURN
      END


      FUNCTION RHOBMK(E,N)
      DIMENSION E(N)
      N1=N-1
      N2=N-2
      N11=N+1
      Z=N1*SUM2(E(2),E(2),N2)
      Y=SUM2(E(1),E(2),N1)
      A=-N2*Y/Z
      B=(-1.0)*((E(1)**2+N11*Z/N1)+E(N)**2)/Z
      C=N*Y/Z
      P=B-(A**2)/3.
      Q=C-(A*B)/3.+(2.*A**3)/27.
      P27=(-27.)/P
      PHI=(SQRT(P27)*Q)/(2.*P)
      PHI=ACOS(PHI)
      P3=-P/3.
      RHOBMK=(-2.)*SQRT(P3)*COS((PHI+3.141592654)/3.)-A/3.
      RETURN
      END


      SUBROUTINE INVR(X,XIN)
      DIMENSION X(3),XIN(3)
      DOUBLE PRECISION D
      D=X(1)*X(2)-X(3)**2
      XIN(1)=X(2)/D
      XIN(2)=X(1)/D
      XIN(3)=(-1.)*X(3)/D
      RETURN
```

```
  969 FORMAT(1H0,15H RO .GE. 1.0=   ,3I10)
      WRITE(6,979)((CONV(I,J),J=1,2),I=1,15)
  979 FORMAT(1H0,15H BO NON CONV =   ,2F20.4)
      WRITE(6,9199)DSEED


      DO 1500 I=1,15
      DO 1500 J=1,5
      DO 1500 K=1,5
 1500 REQM(I,J,K)=SQRT(REQM(I,J,K))

      WRITE(3,9)DSEED
      DO 5000 IFLAG=1,3
      ID=0
      DO 5000 IR=1,3
      WRITE(4,999)
      DO 5000 IPAR=1,5
      WRITE(4,9999)
      WRITE(4,9919)RO(IR),IPAR
 9919 FORMAT(1X,6H RHO= ,F6.4,10H   PAR =   ,I5,//)
      DO 5000 IDEL=1,5
      ID=(IR-1)*5+IDEL

      NEST=5
      IF(IFLAG.EQ.3)NEST=4
      DO 2000 IEST=1,NEST
 2000 E(IEST)=STATW(IFLAG,ID,IEST,IPAR)
      IA=1
      IB=5
      IF((IPAR.EQ.3).OR.(IPAR.EQ.4))IA=2
      IF((IPAR.NE.2).AND.(IPAR.NE.5))GOTO 2500
      IB=6
      IF(IPAR.EQ.2)E(NEST+1)=STATW(IFLAG,ID,1,3)
      IF(IPAR.EQ.5)E(NEST+1)=STATW(IFLAG,ID,1,4)
 2500 CONTINUE
      IF(IFLAG.EQ.3)IB=IB-1
      WRITE(4,989)(E(I),I=IA,IB)
  989 FORMAT(1H0,6E12.4)


 5000 CONTINUE

      DO 7000 IFLAG=1,2
      DO 7000 IR=1,3
      DO 7000 IPAR=1,3
      WRITE(4,9999)
      WRITE(4,9919)RO(IR),IPAR
      IA=1
      IB=5
      IF(IPAR.EQ.2)IB=6
      IF(IPAR.EQ.3)IA=2
      DO 7000 IT=1,3
      WRITE(4,999)
      DO 7000 IDEL=1,5
      ID=(IR-1)*5+IDEL
      DO 6000 IEST=1,5
 6000 E(IEST)=STW(IFLAG,ID,IEST,IT,IPAR)
      IF(IPAR.EQ.2)E(6)=STW(IFLAG,ID,1,IT,3)
      WRITE(4,987)(E(I),I=IA,IB)
  987 FORMAT(1H0,6F10.3)
 7000 CONTINUE
      STOP
      END
```

```
      END


      SUBROUTINE RES(B,N,E,IFLAG,X)
      DIMENSION B(3),X(N),E(N)
      COMMON/RS/Y(50)
      IF(IFLAG-1)10,20,30
   10  CONTINUE
      DO 50 I=1,N
      E(I)=Y(I)-B(1)-B(2)*X(I)
   50  CONTINUE
      RETURN
   20  CONTINUE
      K=1
   25  CONTINUE
      DO 100 I=2,N
      IK=I-K
      I1=I-1
  100  E(IK)=Y(I)-B(3)*Y(I1)-B(1)*(1.-B(3))-B(2)*(X(I)-B(3)*X(I1))
      RETURN

   30  CONTINUE
      ROSQ=SQRT(1.-B(3)**2)
      E(1)=ROSQ*(Y(1)-B(1)-B(2)*X(1))
      K=0
      GOTO 25
      END


      SUBROUTINE XQQX(RHO,N,XQX,XX)
      DIMENSION XQX(3),XX(3)
      COMMON/SQ/X1,X2,XN,Y1,Y2,YN
      S=XX(1)
      S2=XX(2)
      S21=XX(3)
       RHO2=RHO**2
       XQX(1)=(N-1)*(1.-RHO)**2
       XQX(2)=S2*(1.+RHO2)-2.*RHO*S21-X1**2-(XN*RHO)**2
       XQX(3)=(1.-RHO)*((S-X1)-RHO*(S-XN))
       RETURN
       END


      SUBROUTINE XPPX(RHO,N,XPX,XX)
      DIMENSION XPX(3),XX(3),XQX(3)
      COMMON/SQ/X1,X2,XN,Y1,Y2,YN
      RHO2=1.-RHO**2
      CALL XQQX(RHO,N,XQX,XX)
      XPX(1)=XQX(1)+RHO2
      XPX(2)=XQX(2)+RHO2*X1**2
      XPX(3)=XQX(3)+RHO2*X1
      RETURN
      END


      SUBROUTINE SUMXX(N,X,XX)
      DIMENSION XX(3),X(N)
      DOUBLE PRECISION S,S2,S21
      S=S2=S21=0.0
      DO 10 I=1,N
      I1=I-1
      S=S+X(I)
      S2=S2+X(I)**2
      IF(I.EQ.1)GOTO 10
      S21=S21+X(I)*X(I1)
   10  CONTINUE
      XX(1)=SNGL(S)
      XX(2)=SNGL(S2)
      XX(3)=SNGL(S21)
      RETURN
      END


      SUBROUTINE SUMXY(N,X)
      DIMENSION X(N)
      COMMON/XY/XY(3),DUM
      COMMON/RS/Y(50)
      COMMON/SQ/X1,X2,XN,Y1,Y2,YN
      XY(1)=XY(2)=XY(3)=0.0
      DO 10 I=1,N
      I1=I-1
      XY(1)=XY(1)+Y(I)
      XY(2)=XY(2)+X(I)*Y(I)
      IF(I.EQ.1)GOTO 10
      XY(3)=XY(3)+X(I)*Y(I1)+Y(I)*X(I1)
   10  CONTINUE
      RETURN
      END


      SUBROUTINE SUMXYQ(RHO,N,SXYQ)
      DIMENSION SXYQ(3)
      COMMON/XY/SXY(3),DUM
      COMMON/SQ/X1,X2,XN,Y1,Y2,YN
      RHO2=RHO**2
      SXYQ(1)=(1.-RHO)*(SXY(1)-Y1-RHO*(SXY(1)-YN))
      SXYQ(2)=(1.+RHO2)*SXY(2)-RHO*SXY(3)-X1*Y1-RHO2*XN*YN
      RETURN
      END


      SUBROUTINE SUMXYP(RHO,N,SXYP)
      DIMENSION SXYP(2),SXYQ(2)
      COMMON/XY/SXY(3),DUM
      COMMON/SQ/X1,X2,XN,Y1,Y2,YN
      CALL SUMXYQ(RHO,N,SXYQ)
      RHO2=RHO**2
      SXYP(1)=SXYQ(1)+(1.-RHO2)*Y1
      SXYP(2)=SXYQ(2)+(1.-RHO2)*X1*Y1
      RETURN
      END


      FUNCTION BETA(X,Y,B)
      DIMENSION X(3),Y(2)
      BETA=Y(1)*X(3)+X(2)*Y(2)
      B=X(1)*Y(1)+Y(2)*X(3)
      RETURN
      END


      SUBROUTINE MCO(N,PAR,XIN,XX,X)
      DIMENSION XIN(3),XX(3),X(50)
      DIMENSION PAR(8),U(50)
      DIMENSION XQX(3),XQY(3)
      COMMON/COR/RHOO(3)
      COMMON/XY/SXY(3),VRHO
      N1=N-1
      CALL SUMXY(N,X)
      PAR(2)=BETA(XIN,SXY,PAR(1))
      CALL RES(PAR,N,U,0,X)
      RHOO(1)=RHOCO(U,N)
      RHOO(2)=RHOPW(U,N)
      RHOO(3)=RHOBMK(U,N)
      S=SUM2(U,U,N)
      PAR(5)=S/(N-2)
```

```
      N2=N-2
      PAR(6)=(PAR(1)-1.)/(SQRT(XIN(1)*S/N2))
      PAR(7)=(PAR(2)-1.)/(SQRT(XIN(2)*S/N2))
      CALL  XQQX(1.,N,XQX,XX)
      CALL  SUMXYQ(1.0,N,XQY)
      PAR(3)=XQY(2)/XQX(2)
      XQY(1)=1.
      XQY(2)=PAR(3)
      XQY(3)=1.0
      CALL  RES(XQY,N,U,1,X)
      S=SUM2(U,U,N1)/(N-2)
      PAR(4)=S
      PAR(8)=(PAR(3)-1.)/SQRT(S/XQX(2))
      RETURN
      END
      SUBROUTINE DUR(N,PAR,XXIN,XX,X)
      DIMENSION XXIN(3),XX(3),X(50)
      DIMENSION PAR(8),W(49,4),E(50),WK(20)
      DIMENSION XQX(4),XQIN(3),SXYQ(3)
      COMMON/RS/Y(50)
      COMMON/XY/SXY(3),VRHO
      COMMON/REPL/ID,ITMAX(15,3),ICON(15,3)
      N1=N-1
      DO 10 I=2,N
      I1=I-1
      W(I1,1)=X(I)
      W(I1,2)=X(I1)
      W(I1,3)=Y(I1)
      W(I1,4)=1.0
   10  CONTINUE
      CALL  OFIMA3(W,49,Y(2),49,N1,1,4,PAR,5,WK,IER)
      CALL  XQQX(PAR(3),N,XQX,XX)
      CALL  INVR(XQX,XQIN)
      CALL  SUMXYQ(PAR(3),N,SXYQ)
      PAR(2)=BETA(XQIN,SXYQ,PAR(1))
      CALL  RES(PAR,N,E,1,X)
      S=SUM2(E,E,N1)/(N-3)
      PAR(5)=S
      PAR(4)=PAR(1)*(1.-PAR(3))
      PAR(6)=(PAR(1)-1.)/SQRT(S*XQIN(1))
      PAR(7)=(PAR(2)-1.)/SQRT(S*XQIN(2))
      XQX(1)=N-1.
      XQX(3)=XQX(3)/(1.-PAR(3))
      CALL  INVR(XQX,XQIN)
      PAR(8)=(PAR(4)-(1.-VRHO))/SQRT(S*XQIN(1))
      IF(PAR(3).GE.1.0)ICON(ID,1)=ICON(ID,1)+1
      RETURN
      END
      SUBROUTINE CORC(N,PAR,XXIN,XX,X,ITER,TIME)
      COMMON/COR/RHOO(3)
      COMMON/XY/SXY(3),VRHO
      COMMON/RS/Y(50)
      COMMON/REPL/ID,ITMAX(15,3),ICON(15,3)
      COMMON/CONVER/CONV(15,2)
      DIMENSION XXIN(3),XX(3),X(N)
      DIMENSION PAR(8),XQX(3),XQIN(3),XQY(3),U(50)
      TIME=0.0
      TIME=SECOND(CPU)
      N1=N-1
      ITER=0
      PAR(3)=RHOO(1)
   10  CONTINUE
      RO=PAR(3)
      ITER=ITER+1
      CALL  XQQX(PAR(3),N,XQX,XX)
      CALL  INVR(XQX,XQIN)
   11  CONTINUE
      CALL  SUMXYQ(PAR(3),N,XQY)
      PAR(2)=BETA(XQIN,XQY,PAR(1))
      CALL  RES(PAR,N,U,0,X)
      PAR(3)=RHOCO(U,N)
      IF(ITER.LE.100)GOTO 20
      ITMAX(ID,1)=ITMAX(ID,1)+1
    9  FORMAT(1X,3F20.15)
      GOTO 30
   20  CONTINUE
      IF(ABS(PAR(3)-RO).GT.0.00001)GOTO 10
   30  CONTINUE
   40  CONTINUE
      BO=PAR(1)
      IF(PAR(3).GE.1.0)ICON(ID,2)=ICON(ID,2)+1
      CALL  XQQX(PAR(3),N,XQX,XX)
      CALL  INVR(XQX,XQIN)
      CALL  SUMXYQ(PAR(3),N,XQY)
      PAR(2)=BETA(XQIN,XQY,PAR(1))
      IF(ABS(PAR(1)-BO).LT.1.0)GOTO 50
      CONV(ID,1)=CONV(ID,1)+(PAR(1)-BO)**2
      CONV(ID,2)=CONV(ID,2)+1
   50  CONTINUE
      CALL  RES(PAR,N,U,1,X)
      S=SUM2(U,U,N1)/(N-3)
      PAR(5)=S
      PAR(4)=PAR(1)*(1.-PAR(3))
      PAR(6)=(PAR(1)-1.)/SQRT(S*XQIN(1))
      PAR(7)=(PAR(2)-1.)/SQRT(S*XQIN(2))
      XQX(1)=N-1.
      XQX(3)=XQX(3)/(1.-PAR(3))
      CALL  INVR(XQX,XQIN)
      PAR(8)=(PAR(4)-(1.-VRHO))/SQRT(S*XQIN(1))
      TIME=SECOND(CPU)-TIME
      RETURN
      END
      SUBROUTINE PW(N,PAR,XXIN,XX,X,ITER,TIME)
      DIMENSION XXIN(3),XX(3),X(50)
      DIMENSION PAR(8),XPX(3),XPIN(3),XPY(3),U(50)
      COMMON/COR/RHOO(3)
      COMMON/XY/SXY(3),VRHO
      COMMON/SQ/X1,X2,XN,Y1,Y2,YN
      COMMON/REPL/ID,ITMAX(15,3),ICON(15,3)
      TIME=SECOND(CPU)
      ITER=0
      PAR(3)=RHOO(2)
   10  CONTINUE
      RO=PAR(3)
      ITER=ITER+1
      CALL  XPPX(PAR(3),N,XPX,XX)
      CALL  INVR(XPX,XPIN)
      CALL  SUMXYP(PAR(3),N,XPY)
      PAR(2)=BETA(XPIN,XPY,PAR(1))
      CALL  RES(PAR,N,U,0,X)
      PAR(3)=RHOPW(U,N)
      IF(ITER.LE.100)GOTO 20
      ITMAX(ID,2)=ITMAX(ID,2)+1
      GOTO 30
   20  CONTINUE
```

```
      IF(ABS(PAR(3)-RO).GT.0.00001)GOTO 10
   30  CONTINUE
      IF(PAR(3).GE.1.0)ICON(ID,3)=ICON(ID,3)+1
      IF(ABS(PAR(3)).GE.1.)PAR(3)=SIGN(.99999,PAR(3))
      CALL XPPX(PAR(3),N,XPX,XX)
      CALL INVR(XPX,XPIN)
      CALL SUMXYP(PAR(3),N,XPY)
      PAR(2)=BETA(XPIN,XPY,PAR(1))
      CALL RES(PAR,N,U,2,X)
      S=SUM2(U,U,N)/(N-2)
      PAR(5)=S
      PAR(4)=PAR(1)*(1.-PAR(3))
      PAR(6)=(PAR(1)-1.)/SQRT(S*XPIN(1))
      PAR(7)=(PAR(2)-1.)/SQRT(S*XPIN(2))
      XPX(1)=N
      XPX(3)=(XPX(3)/(1.-PAR(3)))-X1*(1.+PAR(3)-SQRT(1.-PAR(3)**2))
      CALL INVR(XPX,XPIN)
      PAR(8)=(PAR(4)-(1.-VRHO))/SQRT(S*XPIN(1))
      TIME=SECOND(CPU)-TIME
      RETURN
      END

      SUBROUTINE BMK(N,PAR,XXIN,XX,X,ITER,TIME)
      DIMENSION XXIN(3),XX(3),X(50)
      DIMENSION PAR(8),XPX(3),XPIN(3),XPY(3),U(50)
      COMMON/SQ/X1,X2,XN,Y1,Y2,YN
      COMMON/XY/SXY(3),VRHO
      COMMON/REPL/ID,ITMAX(15,3),ICON(15,3)
      COMMON/COR/RHOO(3)
      TIME=SECOND(CPU)
      ITER=0
      PAR(3)=RHOO(3)

   10  CONTINUE
      RO=PAR(3)
      ITER=ITER+1
      CALL XPPX(PAR(3),N,XPX,XX)
      CALL INVR(XPX,XPIN)
      CALL SUMXYP(PAR(3),N,XPY)
      PAR(2)=BETA(XPIN,XPY,PAR(1))
      CALL RES(PAR,N,U,0,X)
      PAR(3)=RHOBMK(U,N)
      IF(ITER.LE.100)GOTO 20
      ITMAX(ID,3)=ITMAX(ID,3)+1
      GOTO 30
   20  CONTINUE
      IF(ABS(PAR(3)-RO).GT.0.00001)GOTO 10
   30  CONTINUE

      CALL XPPX(PAR(3),N,XPX,XX)
      CALL INVR(XPX,XPIN)
      CALL SUMXYP(PAR(3),N,XPY)
      PAR(2)=BETA(XPIN,XPY,PAR(1))
      CALL RES(PAR,N,U,2,X)
      S=SUM2(U,U,N)/(N-2)
      PAR(5)=S
      PAR(4)=PAR(1)*(1.-PAR(3))
      PAR(6)=(PAR(1)-1.)/SQRT(S*XPIN(1))
      PAR(7)=(PAR(2)-1.)/SQRT(S*XPIN(2))
      XPX(1)=N
      XPX(3)=(XPX(3)/(1.-PAR(3)))-X1*(1.+PAR(3)-SQRT(1.-PAR(3)**2))
      CALL INVR(XPX,XPIN)
      PAR(8)=(PAR(4)-(1.-VRHO))/SQRT(S*XPIN(1))
      TIME=SECOND(CPU)-TIME
      RETURN
      END

      SUBROUTINE SSTATS(ID,PAR,N,RHO)
      COMMON/QSTAT/BIAS(15,5,5),REQM(15,5,5),BEST(15,4,5),STATT(15,5,9
     ,),STATN(15,5,9),TN(3),TN1(3),ANOR(3)
      DIMENSION PAR(8,5),VPAR(5),BMK(5)

      VPAR(1)=VPAR(2)=1.0
      VPAR(3)=RHO
      VPAR(4)=1.-RHO
      VPAR(5)=.0036
      DO 5 I=1,5
    5  BMK(I)=ABS(PAR(I,5)-VPAR(I))

      DO 100 I=1,5
      DO 20 J=1,5
      JJ=J
      IF((I.EQ.1).AND.(J.EQ.3))JJ=2
      IF((I.EQ.1).AND.(J.EQ.4))JJ=5

      BIAS(ID,I,J)=BIAS(ID,I,J)+.001*(PAR(J,I)-VPAR(JJ))
      REQM(ID,I,J)=REQM(ID,I,J)+.001*(PAR(J,I)-VPAR(JJ))**2
      IF (I.EQ.5)GOTO 20
      IF(ABS(PAR(J,I)-VPAR(JJ)).LE.BMK(JJ))BEST(ID,I,J)=BEST(ID,I,J)+
     , .001
   20  CONTINUE

      DO 50 J=1,3
      T=TN(J)
      SN=ANOR(J)
      IF((I.NE.2).AND.(I.NE.3))GOTO 30
      T=TN1(J)
   30  CONTINUE
      DO 50 K=1,3
      K1=K+5
      IF(ABS(PAR(K1,I)).GT.T)STATT(ID,I,ICAL(J,K))=STATT(ID,I,
     , ICAL(J,K))+.001
      IF(ABS(PAR(K1,I)).GT.SN)STATN(ID,I,ICAL(J,K))=STATN(ID,I,
     , ICAL(J,K))+.001
   50  CONTINUE

  100  CONTINUE
      RETURN
      END

      FUNCTION STATW(IFLAG,I,J,K)
      COMMON/QSTAT/BIAS(15,5,5),REQM(15,5,5),BEST(15,4,5),STATT(15,5,9
     ,),STATN(15,5,9),TN(3),TN1(3),ANOR(3)
      IF(IFLAG-2)10,20,30
   10  STATW=BIAS(I,J,K)
      RETURN
   20  STATW=REQM(I,J,K)
      RETURN
   30  STATW=BEST(I,J,K)
      RETURN
      END

      FUNCTION STW(IFLAG,I,J,K,L)
      COMMON/QSTAT/BIAS(15,5,5),REQM(15,5,5),BEST(15,4,5),STATT(15,5,9
     ,),STATN(15,5,9),TN(3),TN1(3),ANOR(3)
      IF(IFLAG.EQ.1)STW=STATT(I,J,ICAL(K,L))
      IF(IFLAG.EQ.2)STW=STATN(I,J,ICAL(K,L))
      RETURN
      END

      FUNCTION ICAL(I,J)

      ICAL=3*(I-1)+J
      RETURN
      END
```

REMERCIEMENTS

Je ne suis pas ce qu'on peut appeler un élève docile. Ceci n'a sûrement pas facilité la tâche de mon directeur. Malgré tout, nous sommes parvenus à la dernière étape de ce travail et une bonne part du mérite doit lui être attribuée. Je remercie donc Monsieur Dufour pour ses précieux conseils, son soutien moral et financier, et même pour les quelques "coups de pied au derrière" qui ont été nécessaires à la réalisation de ce mémoire.

Je remercie également M. Dagenais et Mme Salvas qui ont accepté de lire ce texte et y apporter leurs commentaires. Je veux souligner spécialement la contribution de M. Dagenais, pour lequel j'ai eu le privilège de travailler comme assistant de recherche. L'expérience acquise restera toujours un atout précieux.

Comme la rédaction de ce travail a été plus longue que prévue, le nombre de mes confrères et amis s'est considérablement accru avec le temps. Les limites qui me sont imposées m'empêchent de rendre à chacun le tribut qui lui est dû. Je veux tout de même mentionner les membres du "Tea Time Club" et du Club du "9-5-2"; André, Clément, Arnold, Jean-Olivier, Martine, Daniel et tous les autres. S'ils sont tous des Montréalais convaincus, je leur pardonne volontiers ce défaut et je les assure qu'ils seront toujours bien reçus à Québec. Je salue aussi Marc et Chantale, Jean-Denis, Etienne, Raymond et Marie-Josée. Les heures passées en leur compagnie ont été parmi les plus heureuses.

Oups! Bonjour Mme Psarianos! Je suis un peu pressé de terminer, mais nous nous reverrons sûrement.

Enfin, un gros merci à Suzanne pour son travail soigné.

Le mot de la fin sera pour mon père et ma mère, même si le premier n'est plus là. Je leur dirai simplement que je les aime.

BIBILOGRAPHIE

BEACH, C.M. et J.G. MacKINNON, "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors", Econometrica, Vol. 46, 1978, pp. 51-58.

COCHRANE, D. et G.H. ORCUTT, "Application ot Least Squares Regression to Relationships Containing Autocorrelated Error Terms", Journal of the American Statistical Association, Vol. 44, 1949, pp. 32-61.

COOPER, J.P., "Assymptotic Covariance Matrix of Procedures for Linear Regression in the Presence of First-Order Autoregressives Disturbances", Econometrica, Vol. 40, 1972, pp. 305-310.

DHRYMES, P.J., Econometrics : Statistical Foundations and Applications, New York, Springer-Verlag, 1980.

DORAN, H.E., "Omission of an Observation from a Regression Analysis", Journal of Econometrics, Vol. 61, 1981, pp. 367-374.

DURBIN, J., "Estimation of Parameters in Time-Series Regression Models", Journal of the Royal Statistical Society Series B, Vol. 22, 1960, pp. 139-153.

GURLAND, J., "An Example of Autocorrelated Disturbances in Linear Regression", Econometrica, Vol. 22, 1954, pp. 218-227.

HONG, D.M. et L. L'ESPERANCE, "Effects of Autocorrelated Errors on Various Least-Squares Estimators : A Monte-Carlo Study", Communications in Statistics, Vol. 2, 1973, pp. 507-523.

JOHNSTON, J., Econometric Methods, 2e édition, New York, McGraw-Hill, 1971.

KADIYALA, K.R., "A Transformation Used to Circumvent the Problem of Autocorrelation", Econometrica, Vol. 36, 1968, pp. 93-96.

MADDALA, G.S., Econometrics, New York, McGraw-Hill, 1977.

MAESHIRO, A., "Autoregressive Transformation, Trented Independent Variables, and Autocorrelated Disturbance Terms", Review of Economics and Statistics, Vol. 58, 1976, pp. 497-500.

MAESHIRO, A., "On the Retention of the First Observation in Serial Correlation Adjustments of Regression Models", International Economic Review, Vol. 20, 1979, pp. 259-265.

PARK, R.E. et B.M. MITCHELL, "Estimating the Autocorrelated Error Model With Trended Data", Journal of Econometrics, Vol. 13, 1980, pp. 185-201.

PFAFFENBERGER, R. et T. DIELMAN, "Small-Sample Properties of the Bayesian Estimator in the Regression Model with Autocorrelated Errors", Non publié, Texas Christian University, Decision Sciences Working Paper No. 10, 1979, 16 p.

PRAIS, S.J. et C.B. WINSTEN, "Trend Estimators and Serial Correlation", Non Publié, Cowles Commission Discussion Paper No. 383, Chicago, 1954.

RAO, P. et Z. GRILICHES, "Small-Sample Properties of Several Two-Stage Regression Method in the Context of Auto-Correlated Errors", Journal of the American Statistical Association, Vol. 64, 1969, pp. 253-272.

SPITZER, J.J., "Small-Sample Properties of Nonlinear Least Squares and Maximum Likelihood Estimators in the Context of Autocorrelated Errors", Journal of the American Statistical Association, Vol. 74, 1979, pp. 41-47.

TAYLOR, W.E., "On the Efficiency of the Cochrane-Orcutt Estimator", Journal of Econometrics, Vol. 17, 1981, pp. 67-82.

THEIL, H., Principles of Econometrics, New York, John Wiley and Son, 1971.