# La Méthode d'Estimation des Moindres Carrés Modifiés ou Fully Modified*

Christophe Hurlin et Papa MB.P. N'Diaye [†]

Juin 1998

**Abstract**

Cet article présente la méthode d'estimation des Moindres Carrés Modifiés ou "Fully Modified" (FM) proposée par Phillips et Hansen (1990) et Phillips (1995).

A partir d'un exemple simple, nous examinons l'incidence des biais d'endogénéité de long terme sur les estimations par les Moindres Carrés Ordinaires des paramètres d'une relation de cointégration et nous proposons une interprétation simple de la correction FM.

La présentation générale de la méthode des FM étendue au cas vectoriel, est complétée par une étude des procédures non paramétriques d'estimation de la matrice de variance covariance de long terme.

Nous insistons sur les conditions générales de validité ainsi que sur les principales limites de la méthode Fully Modified. Nous démontrons en particulier la nécessité de tester au préalable l'ordre d'intégration des séries avant toute application de l'estimateur Fully Modified.

*Les auteurs tiennent à remercier P.-Y. Hénin, P. Fève et F. Collard pour leurs nombreuses remarques et corrections apportées à une précédente version de cet article. Nous remercions également les participants du séminaire Econométrie du M.A.D.

[†]CEPREMAP et M.A.D, Université de Paris I. 101-106 Boulevard de l'Hopital 75013 PARIS. e-mail : dia@univ-paris1.fr

L'apparition du concept de cointégration a fourni aux économistes un mode d'évaluation des relations d'équilibre de long terme pouvant exister entre différents agrégats macroéconomiques. Cette avancée s'est traduite par le développement de nombreuses procédures de test et d'estimation de relations de cointégration.

Parmi les différentes méthodes d'estimation proposées dans la littérature, celle de Engle et Granger (1987) constitue la méthode standard d'estimation des paramètres d'une relation de cointégration structurelle. Toutefois une des principales limites de cette méthode réside dans le fait que les distributions asymptotiques des estimateurs obtenues sous l'hypothèse nulle de cointégration dépendent potentiellement de paramètres de nuisance. La présence de ces derniers exclut la possibilité d'utiliser les procédures d'inférence standard. En particulier, il est alors impossible de tester les paramètres du vecteur de cointégration.

En effet, dès lors que le résidu de la relation de cointégration est corrélé avec les innovations des régresseurs, les estimateurs des Moindres Carrés Ordinaires (MCO) des paramètres du vecteur de cointégration sont biaisés à taille d'échantillon finie. La présence de ce biais, qualifié par Phillips (1995) de biais d'endogénéité de long terme, implique des distributions non standard pour les statistiques des principaux tests usuels.

La méthode des Moindres Carrés Modifiés ou "Fully Modified" (FM) est une des méthodes permettant de corriger ce biais d'endogénéité à taille d'échantillon finie. Introduite initialement par Phillips et Hansen (1990), puis étendue par Phillips (1995), cette procédure semi-paramétrique d'estimation est fréquemment utilisée dans la littérature empirique.

Son principe de base consiste à appliquer les Moindres Carrés Ordinaires sur des variables transformées, la transformation utilisée étant fondée sur une estimation préalable de la matrice de variance covariance de long terme. L'idée est de se ramener à une nouvelle représentation de la relation de cointégration dans laquelle les résidus vérifient les bonnes propriétés d'orthogonalité.

Plusieurs applications empiriques de cette méthode ont été proposées dans la littérature. En effet de nombreuses relations structurelles issues de modèles théoriques supposent que les taux de croissance des variables explicatives non stationnaires sont corrélés avec le résidu de la relation de cointégration. De plus, l'estimation d'une relation de long terme n'est souvent pertinente qu'à la condition qu'il soit possible de tester certaines contraintes structurelles portant sur les paramètres du vecteur de cointégration.

Parmi les plus récentes, on peut citer les travaux de Dutt et Gosh (1996) ou de Mc Donald et Mooore (1997) qui utilisent les FM pour tester l'hypothèse de Parité des Pouvoirs d'Achat. Une telle démarche suppose en effet de régresser les taux de changes sur les prix domestiques et étrangers et de tester *ex-post* l'égalité des coefficients. Or dans ce contexte, les "*régresseurs sont très probablement endogènes [..] et les erreurs sont très fortement autocorrélées*" (Mc Donald et Moore p. 17). C'est pourquoi, afin de tester les coefficients de la relation estimée de long terme, il convient d'utiliser les Fully Modified qui permettent de rendre les distributions asymptotiques des estimateurs indépendantes de tout paramètre de nuisance et de se ramener à des procédures d'inférence standard.

Dans un autre contexte, Otto et Voss (1997) ou Hénin et Hurlin (1998), utilisent les FM pour estimer la contribution productive du capital public à la croissance et mesurer l'importance des biais de simultanéité dans les résultats d'Aschauer (1989). En effet, l'estimation des rendements des facteurs privés et publics à partir d'une fonction de production, sous l'hypothèse de cointégration, suppose que l'on contrôle la corrélation susceptible d'apparaître entre le taux de croissance de ces facteurs et le résidu de Solow purgé de l'effet des externalités publics.

Enfin nous pouvons citer l'étude de Mamingi (1997) qui propose une évaluation du degré de mobilité des capitaux à partir de la corrélation entre l'épargne et l'investissement estimée dans 58 pays en voie de développement. Dans un tout autre domaine, sur la base d'un test de Wald fondé sur un estimateur FM, Farland, MacMahon et Ngama (1994) montrent que, pour certaines monnaies, les taux de change à terme constituent des indicateurs avancés des taux au comptant futurs.

Nous proposons dans cet article, une présentation de la méthode des Fully Modified étendue au cas vectoriel. Dans une première section, nous étudions l'incidence de l'endogénéité de long terme des régresseurs $I(1)$ dans une relation de cointégration. Nous illustrons cette étude par la dérivation des distributions asymptotiques de l'estimateur des MCO dans un modèle simple.

Dans une seconde section, nous présentons la méthode générale des Fully Modified, étendue par Phillips (1995), au cas multivarié, permettant l'introduction conjointe de variables $I(0)$ et $I(1)$ sans test à priori du degré d'intégration des séries. Nous étudions l'incidence des procédures d'estimation non-paramétriques de la matrice de variance covariance de long terme sur les propriétés de l'estimateur FM.

Nous proposons dans la troisième section une application de la méthode des FM sur données simulées mettant en évidence numériquement l'incidence du biais d'endogénéité de long terme et la capacité des FM à le corriger. L'utilisation des données simulées nous permet de d'approximer des distributions empiriques

Enfin, la dernière partie met l'accent sur les conditions de validité des FM dans l'approche de Phillips (1995) portant sur l'espace des régresseurs stationnaires. Nous montrons en effet qu'en présence de variables stationnaires, les estimations FM sont biaisées dès lors que ces variables sont corrélées au résidu de la relation de cointégration et aux innovations des variables non stationnaires. Ce résultat, contrairement aux indications de Phillips (1995), montre la nécessité de tester au préalable le degré d'intégration des séries, afin de vérifier les conditions de validité de l'estimateur FM portant en particulier sur les variables stationnaires. Cette étude est illustrée par l'application d'un test de l'hypothèse de revenu permanent.

# 1 Estimation des relations de cointégration

L'estimation d'une ou plusieurs relations de long terme entre des agrégats non stationnaires constitue le coeur de nombreuses études économiques appliquées. Plusieurs méthodes d'estimation des paramètres d'une relation de cointégration sont aujourd'hui à la disposition de l'économètre.

La plus simple et la plus fréquemment utilisée est celle proposée par Engle et Granger (1997). En présence de variables intégrées d'ordre 1 (notées $I(1)$), cette méthode consiste tout simplement à estimer les paramètres de la relation de cointégration en appliquant les Moindres Carrés Ordinaires (MCO) au modèle. Le vecteur de cointégration obtenu ne sera ensuite retenu que si, *ex-post*, l'hypothèse nulle de non cointégration est rejetée par le test de la non stationnarité des résidus. L'estimateur des MCO est alors asymptotiquement super convergent.

Cette méthode, contrairement à celle proposée par Johansen et Juselius (1988), permet d'imposer à priori le choix de la normalisation de la relation de cointégration. Il est ainsi possible d'estimer très facilement une relation structurelle issue d'un modèle théorique.

Toutefois, les distributions asymptotiques des estimateurs, sous l'hypothèse nulle de cointégration, dépendent dans le cas général de paramètres de nuisance. Les estimateurs des paramètres d'une relation de cointégration, obtenus par la méthode des MCO, sont alors biaisés à distance finie. Le biais est d'autant plus important que la taille d'échantillon est réduite.

De plus, la présence de ces paramètres de nuisance exclut la possibilité d'utiliser les procédures d'inférence standard. Les distributions associées aux statistiques des principaux tests usuels sont différentes de celles habituellement utilisées dans un univers stationnaire.

Ce résultat s'explique par la présence potentielle d'une corrélation (instantanée ou avancée/retardée) non nulle entre le résidu de la relation de cointégration et les innovations des régresseurs. En présence d'une telle corrélation apparaît un biais de second ordre dans la distribution asymptotique de l'estimateur des MCO pour des échantillons de taille finie. On qualifie ce biais (Phillips 1995), de *biais d'endogénéité de long terme.*

En effet, sous l'hypothèse de cointégration, la covariance empirique entre les régresseurs et le résidu de la relation de cointégration converge vers une intégrale stochastique faisant intervenir deux mouvements Browniens. Ces derniers représentent les processus limites de la somme des variables explicatives et la somme des résidus de la relation de cointégration.

Dès lors, que les innovations des régresseurs non stationnaires sont corrélés avec le résidu de la relation de cointégration, ces mouvements Browniens sont corrélés. Les moments empiriques croisés convergent à la vitesse $T$ vers une distribution non centrée, ce qui induit à taille d'échantillon finie un biais dans les estimations MCO des paramètres de la relation de long terme.

Ce biais d'endogénéité apparaît dans la distribution de l'estimateur des MCO dès lors que les taux de croissance des variables explicatives non stationnaires sont corrélés avec le résidu de la relation de cointégration. Sur le plan empirique cette hypothèse ne peut pas être testée *ex-post*, puisque si la corrélation est non nulle, l'estimateur de la population des résidus est biaisé. De plus, de nombreuses relations structurelles de long terme issues de modèles théoriques supposent implicitement l'existence d'une telle corrélation.

Afin d'illustrer ce problème et de mieux comprendre les enjeux associés à la méthode des Fully Modified, nous proposons l'étude d'un exemple simple. On considère la relation de cointégration suivante :

$$y_{1t} = \beta x_t + \mu_{1t} \tag{1}$$

$$\Delta x_t = \mu_{2t} \tag{2}$$

où $\epsilon_t = [\mu_{1t}\, \mu_{2t}]'$ et $\epsilon_t$ $i.i.d$ $(0, \Sigma)$ où $\Sigma$ est définie par :

$$E\left(\epsilon_t \epsilon_t'\right) = \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \tag{3}$$

Le biais d'endogénéité de long terme provient, dans ce contexte, de la non nullité du paramètre $\sigma_{12}$ correspondant à la covariance entre le résidu de la relation de cointégration $\mu_{1t}$ et l'innovation du régresseur non stationnaire $\mu_{2t.}$ Lorsque l'on suppose $\sigma_{12} \neq 0$, le biais associé à l'estimateur des MCO de $\beta$, noté $\widehat{\beta}_{MCO}$, suit la distribution suivante.

$$\begin{aligned} T\left(\widehat{\beta}_{MCO} - \beta\right) &= \frac{\frac{1}{T}\sum\limits_{t=0}^{T} x_t \mu_{1t}}{\frac{1}{T^2}\sum\limits_{t=0}^{T} x_t^2} \qquad (4) \\ &\underset{T\to\infty}{\overset{L}{\longrightarrow}} \frac{\sigma_1}{\sigma_2}\left\{ \frac{\tau}{2}\frac{\left[W_2(1)^2+1\right]}{\int_0^1 W_2(r)^2\,dr} + \left(1-\tau^2\right)^{1/2}\frac{\int_0^1 W_2(r)\,dW_1(r)}{\int_0^1 W_2(r)^2\,dr} \right\} \end{aligned}$$

où $\tau$ désigne la corrélation instantanée entre les résidus $\mu_{1t}$ et les innovations $\mu_{2t}$ ($\tau = \sigma_{12}/\sigma_1\sigma_2$). $W_1(.)$ et $W_2(.)$ désignent deux mouvements Browniens scalaires standards et indépendants. Cette distribution peut s'interpréter comme une moyenne pondérée de deux variables aléatoires, dont les poids dépendent du niveau de la corrélation $\tau$. La première, pondéré par $\tau/2$, est distribuée selon un $\chi^2(1)$ et possède une espérance strictement positive. Sa distribution est donc non symétrique et non centrée. Le second terme, pondéré par $\left(1-\tau^2\right)^{1/2}$ est distribué selon une loi normale centrée.

On vérifie ainsi qu'en présence d'une corrélation entre le résidu de la relation de cointégration et les innovations des régresseurs, la distribution de l'estimateur des MCO de $\beta$ présente un biais, qualifié de second ordre, à distance finie. Ce biais est $Op\,(1)$, puisque l'on conserve le résultat selon lequel :

$$\widehat{\beta}_{MCO} - \beta \underset{T\to\infty}{\overset{p}{\longrightarrow}} 0$$

A taille d'échantillon finie, la distribution asymptotique de l'estimateur MCO dépend du paramètre de nuisance $\tau$ et apparaît comme une moyenne pondérée de deux distributions. Dès lors :

- L'importance en valeur absolue du biais est une fonction croissante du poids accordé à la composante non centrée de la distribution, c'est à dire du niveau de la corrélation.
- Le signe du biais dépend du signe de la corrélation. Pour une corrélation positive (respectivement négative), la distribution de l'estimateur des MCO se déplace à droite (respectivement vers la gauche) de la vraie valeur du paramètre.
- Pour une corrélation non nulle, la distribution de l'estimateur des MCO est non symétrique.

Ainsi lorsque les innovations des régresseurs et le résidu de la relation de cointégration ne sont pas indépendants, la distribution de l'estimateur MCO est biaisée et non symétrique. Ce biais conduit à surestimer ou sous estimer la valeur du paramètre suivant le signe de la corrélation $\tau$.

Le caractère endogène du régresseur $x_t$ induit en outre des distributions non standard pour les différentes statistiques de tests usuels. Considérons le cas de la *t-stat* du test $\widehat{\beta}_{MCO} = \beta$, notée $t_{\widehat{\beta}}$. En présence d'une corrélation non nulle, la distribution de la *t-stat* converge[1] vers une combinaison linéaire de deux variables, la première suivant la distribution de la statistique de Dickey Fuller et la seconde suivant une loi normale centrée réduite. Les poids associés à ces deux distributions dépendent du paramètre de nuisance $\tau$.

$$t_{\widehat{\beta}} \xrightarrow{L} \frac{\tau}{2} \frac{\left[W_2\,(1)^2 + 1\right]}{\left[\int_0^1 W_2\,(r)^2\,dr\right]^{\frac{1}{2}}} + \left(1 - \tau^2\right)^{1/2} \frac{\int_0^1 W_2\,(r)\,dW_1\,(r)}{\left[\int_0^1 W_2\,(r)^2\,dr\right]^{\frac{1}{2}}} \tag{5}$$

[1]Cette distribution est obtenue en utilisant le résultat selon lequel l'estimateur des MCO est super-convergent. L'estimateur $_T$ est alors un estimateur convergent de $^2_1$

$${}^{2}_{T} = \frac{1}{-1} \sum_{t=1}^{T} \widehat{}^{2}_{1t} \xrightarrow{p} ({}_{1t}) = \left({}^{2}_{1t}\right) = {}^{2}_{1}$$

Etant donné que la statistique de Dickey et Fuller est non centrée (espérance positive) et non symétrique, le paramètre de nuisance $\tau$ affecte la distribution de la statistique de Student de trois manières :

- La distribution de $t_{\widehat{\beta}}$ est non centrée. Pour une corrélation positive, cette distribution se situe à droite de la distribution standard, ce qui induit un biais en faveur de l'hypothèse alternative de non nullité du paramètre testé.
- Plus le niveau en valeur absolue de la corrélation est important, plus le test est biaisé en faveur de l'hypothèse alternative.
- Dès lors que la corrélation est non nulle, la distribution de $t_{\widehat{\beta}}$ n'est plus symétrique.

Il apparaît ainsi important tant sur le plan de l'estimation que sur celui de l'inférence de corriger ce biais d'endogénéité de long terme. La méthode des Fully Modified est une des méthodes qui permet de corriger ce biais.

# 2 L'estimateur des Fully Modified

La méthode des Fully Modified (FM) proposée initialement par Phillips et Hansen (1990), puis étendue par Phillips (1995), est une procédure semi-paramétrique d'estimation des paramètres d'une relation de cointégration qui permet de corriger le biais d'endogénéité de long terme.

La distribution des estimateurs obtenus par cette méthode est indépendante des paramètres de nuisance présents dans la distribution des MCO. Dès lors, les statistiques des tests usuels appliqués aux estimateurs FM suivent des distribution standard identiques à celles utilisées dans un univers stationnaire.

Le principal avantage de cette méthode réside dans sa facilité de mise en oeuvre. Les Fully Modified consistent tout simplement à appliquer les MCO sur un modèle transformé. La transformation utilisée est obtenue à partir d'un estimateur convergent de la matrice de variance covariance de long terme des résidus et des innovations des variables non stationnaires.

L'intuition de la transformation retenue est très simple. Le but est d'orthogonaliser le résidu de la relation de cointégration par rapport aux innovations des variables non stationnaires. Une fois que le système a été réécrit de telle sorte que le résidu de la relation de cointégration soit orthogonal aux innovations des régresseurs $I(1)$, on peut alors appliquer les MCO. La distribution des estimateurs est dans ce cas indépendante des paramètres de nuisance, centrée et symétrique.

## 2.1 Principe général de la méthode d'estimation FM

La méthode des FM a été étendue par Phillips (1995) au cas vectoriel ($VAR$) intégrant à la fois des variables explicatives stationnaires et non stationnaires dans la relation de long terme. On considère ainsi le modèle suivant :

$$\underset{(n,1)}{y_t} = \beta \underset{(m,1)}{x_t} + \underset{(n,1)}{\mu_{0t}} \tag{6}$$

où $\beta$ est matrice $(n, m)$ et $m = m_1 + m_2$. On distingue maintenant deux types de régresseurs dans la relation principale de cointégration suivant leur degré respectif d'intégration.

$$\underset{(m_1,1)}{x_{1t}} = \mu_{1t} \tag{7}$$

$$\underset{(m_2,1)}{\Delta x_{2t}} = \mu_{2t} \tag{8}$$

La relation de cointégration (6) peut se réécrire en fonction de ces deux composantes de l'espace des régresseurs.

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \mu_{0t} \tag{9}$$

On suppose que $\mu_t' = (\mu_{0t}, \mu_{1t}, \mu_{2t})$ suit un processus $VMA$ inversible défini par

$$\underset{(m+n,1)}{\mu_t} = \underset{(m+n,m+n)}{C\,(L)} \underset{(m+n,1)}{\epsilon_t} \tag{10}$$

où les $\epsilon_t$ sont $i.i.d.$ $(0, \Sigma)$ et les coefficients de la matrice $C\,(L)$ satisfont les hypothèses standard de sommabilité[2]. On suppose en outre que les variables explicatives stationnaires $\mu_{1t}$ sont indépendantes à toute date du résidu de la relation de cointégration $\mu_{0t}$.

La procédure d'estimation des Fully Modified du vecteur de paramètres $\beta$ peut se décomposer en deux étapes :

### 2.1.1 *Etape 1: Estimation des matrices et* $\Delta$

La matrice de variance covariance de $\mu_t$ peut se réécrire comme la somme de deux composantes. La première, notée , exprimant les variances covariances de long terme

$$= C\,(1)\,\Sigma C\,(1)' \tag{11}$$

La seconde, notée $\Delta$, correspondant aux variances covariances avancées

$$\Delta = \sum_{j=1}^{\infty} E\left(u_j u_0'\right) = \sum_{j=1}^{\infty} \Gamma\,(j) \tag{12}$$

On subdivise alors les matrices et $\Delta$ de la façon suivante :

---

[2] En particulier $\sum_{j=1}^{\infty} |_j| \;\; \infty \;|\;(1)| \neq 0 \;\nabla\;\; 1$

$$= \begin{bmatrix} \underset{(n,n)}{_{yy}} & \underset{(n,m)}{_{yx}} \\ \underset{(m,n)}{'_{yx}} & \underset{(m,m)}{_{xx}} \end{bmatrix} \qquad \Delta = \begin{bmatrix} \underset{(n,n)}{\Delta_{yy}} & \underset{(n,m)}{\Delta_{yx}} \\ \underset{(m,n)}{\Delta_{xy}} & \underset{(m,m)}{\Delta_{xx}} \end{bmatrix}$$

La méthode des Fully Modified repose sur l'estimation préalable de  et $\Delta$. L'estimation de ces deux matrices requiert une approximation de la population des résidus $\mu_{0t}$. Pour ce faire, on utilise le résultat selon lequel l'estimateur des $MCO$ est asymptotiquement convergent et l'on utilise pour le calcul des covariances empiriques les résidus $\widehat{\mu}_{0t} = y_t - \widehat{\beta} x_t$ obtenus par les MCO.

A partir de cette population empirique des résidus estimés, on construit les estimateurs à noyau (ou estimateur *kernel*) des matrices  et $\Delta$. Cette estimation est obtenue par un lissage des autocovariances empiriques du vecteur $\widehat{\mu}'_t = (\widehat{\mu}_{0t}, \mu_{1t}, \mu_{2t})$ (voir encadré 1).

Cette étape préliminaire constitue une des faiblesses de la méthode des Fully Modified, puisque elle suppose que les principales hypothèses de convergence de l'estimateur des $MCO$ soient satisfaites. On retrouve ainsi la critique qui peut être adressée plus généralement à l'ensemble des méthodes d'estimation en deux étapes.

### 2.1.2 Etape 2 : Correction des FM

Dans le cas général, l'endogénéité de long terme des variables explicatives non stationnaires se traduit de deux manières :

- le premier problème réside dans la non nullité du bloc $_{yx}$ de la matrice de variance covariance de long terme qui traduit une corrélation instantanée entre $\mu_{0t}$ et $\mu_{2t}$[3].
- le deuxième problème réside dans la non nullité du bloc $\Delta_{yx}$, qui traduit une corrélation entre le résidu de la relation de cointégration et le passé des innovations $\mu_{2t}$. Etant donnée la persistance de ces chocs, une telle corrélation conduit à un biais dans la distribution de l'estimateur des MCO.

Le premier problème est résolu en rééxprimant le résidu de la relation de cointégration en deux composantes. Une composante correspondant à la projection de $\mu_{0t}$ sur l'espace engendré par les innovations $\mu_{2t}$ et une composante $\mu_{0t}^+$ orthogonale à $\mu_{2t}$. Analytiquement, cela revient à redéfinir le modèle de la façon suivante (cf encadré 2) :

$$y_t^+ = \beta x_t + \mu_{0t}^+ \tag{13}$$

$$y_t^+ = y_t - \widehat{}_{yx} \widehat{}_{xx}^{\,-1} \Delta x_t \tag{14}$$

$$\mu_{0t}^+ = \mu_{0t} - \widehat{}_{yx} \widehat{}_{xx}^{\,-1} \Delta x_t \tag{15}$$

[3] Dans le cas de l'exemple précédent, le bloc $_{yx}$ correspond à la covariance $_{12}$.

où $\widehat{\ }_{yx}$ et $\widehat{\ }_{xx}$ sont les estimateurs convergent de $\ _{yx}$ et $\ _{xx}$ obtenus dans la première étape. De par sa construction, le résidu $\mu_{0t}^{+}$ est asymptotiquement orthogonal aux innovations de la composante non stationnaire de $x_t$.

Phillips (1995) présente sa méthode comme une méthode d'estimation très générale qui ne requiert aucun test à priori sur le degré d'intégration des séries. On suppose donc qu'il est impossible de distinguer $x_{1t}$ de $x_{2t}$.

En effet, étant donné que dans la correction tous les régresseurs, y compris les régresseurs stationnaires $x_{1t}$, sont différenciés, ces derniers n'auront aucune incidence asymptotiquement. Quand bien même, les éléments de $x_{1t}$ seraient spécifiés en niveau, leur caractère stationnaire impliquerait le même résultat à long terme (les éléments correspondant de la matrice de variance covariance de long terme sont nuls). C'est pourquoi les FM ne supposent pas de tester au préalable le degré d'intégration de chaque série.

Ainsi, un des principaux avantages de l'approche de Phillips (1995) réside dans le fait qu'il n'est pas nécessaire de tester à priori le degré d'intégration des variables explicatives avant d'appliquer l'estimateur FM.

Cependant, la méthode d'estimation des FM est une méthode en deux étapes, qui requiert l'utilisation d'un estimateur convergent de la matrice de variance covariance de long terme. Or les résultats de convergence des estimateurs standards de cette matrice reposent sur l'hypothèse que seules les séries non stationnaires puissent être corrélées avec le résidu de la relation de cointégration. Les variables stationnaires doivent en effet être orthogonales au résidu, afin que l'estimateur des $MCO$, nécessaire dans la première étape (estimation de et $\Delta$), soit convergent. On rappelle que le biais d'endogénéité de second ordre ne disparaît qu'en raison de la non stationnarité des régresseurs.

L'hypothèse fondamentale de la convergence de l'estimateur des $MCO$, qui ne peut être satisfaite que sous certaines conditions sur les variables stationnaires, implique ainsi des conditions plus ou moins restrictives que l'on doit tester à priori (intégration, orthogonalité). Ainsi, en contradiction avec l'approche de Phillips (1995), l'utilisation des FM nécessite de tester au préalable le degré d'intégration des variables explicatives et de distinguer les séries stationnaires des séries non stationnaires.

La seconde correction des FM concerne la matrice $\Delta_{yx}$. La correction de l'autocorrélation est obtenue en calculant le terme (cf. Annexe B1):

$$\widehat{\Delta}_{yx}^{+} = \widehat{\Delta}_{yx} - \widehat{\ }_{yx}\widehat{\ }_{xx}^{-1}\widehat{\Delta}_{xx} \tag{16}$$

Dès lors, si l'on note $Y' = (y_1, ..., y_T)$, la forme générale de l'estimateur des Fully Modified est donnée par

$$\widehat{\beta}_{FM} = \left(Y^{+'}X - T\widehat{\Delta}_{yx}^{+}\right)(X'X)^{-1} \tag{17}$$

La distribution asymptotique de cet estimateur, contrairement à celui des MCO, ne dépend pas des paramètres de nuisance $\ _{yx}$ et $\Delta_{yx}$.

### 2.1.3 Distribution asymptotique de l'estimateur FM

On suppose que l'on peut décomposer la matrice $\beta$ en deux sous matrices correspondant respectivement aux matrices de coefficients pour les variables stationnaires et non stationnaires.

$$
\begin{aligned}
& y_t = \beta_1 x_{1t} + \beta x_{2t} + \mu_{0t} \\
& \beta_i = \beta H_i \quad i = 1, 2 \\
& H_1^{'} x_t = x_{1t} \\
& H_2' \Delta x_t = \Delta x_{2t}
\end{aligned}
\tag{18}
$$

Phillips (1995) établit les distributions asymptotiques des deux composantes $\beta_1$ et $\beta_2$ sous les hypothèses décrites précédemment sur la matrice $C(L)$, sous l'hypothèse d'exogénéité des variables stationnaires, et sous les hypothèses relatives à l'estimation kernel des matrices de variance covariance.

**Proposition 1** *La distribution asymptotique de l'estimateur FM associé aux coefficients des variables stationnaires est :*

$$\sqrt{T}\left(\widehat{\beta}_1 - \beta_1\right) H_1 \underset{T \to \infty}{\overset{L}{\longrightarrow}} N(0, \Theta)$$

*avec* $\Theta = \begin{pmatrix} I_n & \Sigma_{11}^{-1} \end{pmatrix} \begin{pmatrix} \Sigma_{00} & \Sigma_{11} \end{pmatrix} \begin{pmatrix} I_n & \Sigma_{11}^{-1} \end{pmatrix}$, *où* $\Sigma_{ii} = E\left(\mu_i \mu_i'\right)$ *(*$i = 0, 1, 2$*) désigne le bloc ii de la matrice de variance covariance des résidus i.i.d.* $\epsilon_t$.

*Concernant les coefficients associés aux variables non stationnaires, on montre que*

$$T\left(\widehat{\beta}_2 - \beta_2\right) H_2 \underset{T \to \infty}{\overset{L}{\longrightarrow}} \left(\int_0^1 dW_{y.x_2} B_{x_2}'\right) \left(\int_0^1 dW_{x_2} W_{x_2}'\right)^{-1}$$

*Le terme* $W_{y.x}$ *désigne un mouvement brownien vectoriel de dimension* $(n, 1)$ *défini par* $W_{y.x_2} = W_y - {}_{yx}\ {}_{xx}^{-1} W_{x_2}$ *où* $W_y$ *et* $W_{x_2}$ *désigne deux mouvements Browniens vectoriels de dimension respectives* $(n, 1)$ *et* $(m_1, 1)$ *non indépendants. La variance du mouvement* $W_{y.x}$ *est égale à* ${}_{yy.x_2} = {}_{yy} - {}_{yx_2}\ {}_{x_2 x_2}^{-1}\ {}_{x_2 y}$.

La démonstration de cette proposition figure dans l'article de Phillips (1995). L'idée générale consiste à prémultiplier le vecteur des paramètres de la relation cointégration par deux matrices choisies de telle sorte que l'on obtienne deux sous vecteurs de paramètres, l'un étant associé aux variables stationnaires, l'autre aux variables $I(1)$. En utilisant le théorème central limite et le théorème central limite fonctionnel, on peut alors retrouver ces distributions asymptotiques.

Dans le cas où tous les régresseurs sont stationnaires ($m_2 = 0$, $H_1 = I_{m_1}$), l'estimateur des FM est convergent et possède la même distribution limite que les MCO, à la condition toutefois que les hypothèses nécessaires à la convergence de l'estimateur des MCO soient satisfaites.

Dans le cas non stationnaire ($m_2 = 0, H_2 = I_{m_2}$), la correction semi-paramétrique des Fully Modified nous permet de retrouver une distribution standard faisant intervenir un mouvement Brownien vectoriel dont la variance $_{yy.x_2}$ correspond à celle de la composante des résidus $\mu_{0t}$ orthogonale à l'espace engendré par les innovations $\mu_{2t}$.

On peut alors se ramener à des distributions standard pour les principales statistiques des tests usuels. En particulier pour $m_1 = 0$ ou $m_2 = 0$, la statistique du test de Wald associé à $r$ contraintes a pour distribution asymptotique un $\chi^2$ à $r$ degrés de liberté.

# 3 Application

Afin de mieux illustrer les enjeux associés à la correction du biais d'endogénéité de long terme, nous proposons dans cette section une application de la méthode des Fully Modified sur des données simulées.

On considère les processus décrit par les équations (1) et (2), avec $\beta = 1$. Les innovations $\mu_{2t}$ et le résidu de la relation de cointégration $\mu_{1t}$ sont distribués suivant des lois normales[4].

L'influence du biais d'endogénéité sur la distribution de l'estimateur des MCO, est illustrée par les résultats présentés dans le tableau 1, obtenus à partir de pseudo échantillons simulés par la méthode de Monte Carlo. Dans ce tableau figurent d'une part le biais moyen $(10^{-2})$ et d'autre part la variance moyenne $(10^{-4})$ associé à $\widehat{\beta}_{MCO}$ . Ces moments ont été calculés à partir de 5000 réplications du modèle.

On vérifie tout d'abord que pour une corrélation donnée (positive ou nulle), lorsque la taille d'échantillon tend vers l'infini, le biais et la variance de l'estimateur diminuent, conformément au résultat théorique selon lequel le biais d'endogénéité de long terme est $Op\,(1)$.

De plus, à taille d'échantillon donnée, le biais et la variance sont des fonctions croissantes du niveau de la corrélation en valeur absolue entre $\mu_{1t}$ et $\mu_{2t}$, conformément au résultat décrit précédemment. Ce résultat s'explique par la déviation de la fonction de densité de l'estimateur des MCO par rapport à la vraie valeur du paramètre. On vérifie en outre que pour une corrélation positive le biais est positif, ce qui indique un déplacement vers la droite de la densité de l'estimateur des MCO et que pour une corrélation négative le biais devient négatif en raison du déplacement vers la gauche de cette distribution. Les biais sont en outre symétriques par rapport au cas $\tau = 0$.

Sur la figure (1) sont représentés les approximations des fonctions de densité empiriques obtenues pour différentes valeurs positives de la corrélation $\tau$ avec $T$ fixé à 50. Lorsque le régresseur $x_t$ est exogène, la distribution de l'estimateur des MCO est parfaitement centrée sur la vraie valeur du paramètre. Plus le niveau de cette

[4] Les variances des lois normales ont été choisies arbitrairement et fixées repspectivement à $^2_1 =$ 025 et $^2_2 = 081$.

Table 1: Biais d'Endogènéité

| Corrélation Positive | | | | | | |
|---|---|---|---|---|---|---|
| T | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 |
| 50 | 002 (7.09) | 036 (7.00) | 115 (7.61) | 209 (8.30) | 280 (8.62) | 373 (10.25) |
| 100 | 0002 (1.80) | 023 (1.82) | 062 (1.75) | 104 (2.08) | 142 (2.14) | 191 (2.76) |
| 500 | 0002 (0.069) | 004 (0.068) | 012 (0.073) | 020 (0.08) | 029 (0.09) | 038 (0.11) |
| 1000 | 0002 (0.017) | 002 (0.017) | 006 (0.018) | 010 (0.02) | 014 (0.02) | 019 (0.03) |

| Corrélation Négative | | | | | | |
|---|---|---|---|---|---|---|
| T | 0 | -0.1 | -0.3 | -0.5 | -0.7 | -0.9 |
| 50 | 002 (7.09) | −036 (7.29) | −120 (7.11) | −200 (7.44) | −291 (9.34) | −366 (9.98) |
| 100 | 0002 (1.80) | −018 (1.70) | −071 (1.69) | −100 (1.94) | −141 (2.30) | −187 (2.66) |
| 500 | 0002 (0.069) | −004 (0.07) | −012 (0.07) | −020 (0.08) | −029 (0.09) | −038 (0.12) |
| 1000 | 0002 (0.017) | −002 (0.02) | −006 (0.02) | −010 (0.02) | −014 (0.02) | −019 (0.03) |

corrélation augmente, plus la fonction de densité se déplace vers la droite et s'éloigne de cette vraie valeur.

Ainsi, conformément aux résultats décrits précédemment dès lors que les taux de croissance des variables explicatives du modèle sont corrélés avec le résidu de la relation de cointégration, il existe un biais à taille d'échantillon finie dans les estimations du paramètre $\beta$ par les MCO.

Nous allons voir que la méthode des Fully Modified permet de corriger ce biais[5]. Dans le tableau (2) sont reportés les biais moyens ($10^{-2}$) ainsi que les variances moyennes ($10^{-4}$) associés à l'estimateur des FM du paramètre $\beta$ obtenus à partir du exercice de simulation que celui décrit précédemment. Pour ces simulations, nous avons distingué deux cas. Dans le premier cas, nous supposons que la matrice de variance covariance est connue. La correction FM est alors basée sur les vraies valeurs de la covariance entre le résidu $\mu_{1t}$ et les innovations $\mu_{2t}$. Dans le second cas, la correction FM est basée sur l'estimateur des MCO de la matrice de variance covariance.

Lorsque l'on suppose que la matrice de variance covariance $\Sigma$ est connue, en

[5] La procedure informatique utilisée a été programmée sous Matlab 4.0. Une procédure d'estimation par les FM est disponible sous GAUSS. Les auteurs ont en outre programmé les FM sous TSP 4.4 dans le cas vectoriel à partir du programme de P. Fève développé dans le cas univarié.

Figure 1: Fonction de Densité Empirique de l'Estimateur des MCO $\left(\widehat{\beta}-\beta_0\right)/T$ $(T=50)$

présence d'une corrélation nulle on retrouve exactement les résultats des MCO puisque dans ce cas la correction des FM est nulle. En revanche pour $\tau \neq 0$, pour des petites tailles d'échantillons ($< 500$), le biais et les variances associés aux FM sont négligeables comparés à ceux des MCO.

Cette conclusion demeure valable la correction FM est fondée sur l'estimateur des MCO de la matrice $\Sigma$. On vérifie, à corrélation donné, que plus la taille d'échantillon augmente plus le biais des FM diminue. Ce résultat s'explique par le fait que le biais associé à l'estimateur des MCO est $Op\left(1\right)$, donc plus la taille d'échantillon augmente plus l'estimateur des résidus $\widehat{\mu}_{1t}$ converge vers la vraie population $\mu_{1t}$. Dès lors, la matrice de variance covariance des résidus estimés converge vers sa vraie valeur $\Sigma$ et la correction des FM devient de plus en plus efficace (au sens où le terme de correction tend vers celui obtenu en posant $\Sigma$ connu).

Si l'on raisonne à taille d'échantillon donnée, on observe que pour des corrélations inférieures à 0.5, le biais tend à diminuer avec $\tau$ et qu'au contraire il tend à augmenter avec $\tau$ pour des valeurs de $\tau$ supérieures à 0.5. En effet, pour de faibles corrélations, le biais associé à l'estimateur MCO des résidus de la relation de cointégration est négligeable, et la correction des FM est fondée sur un estimateur relativement peu biaisé de la matrice de variance covariance de long terme. En revanche pour des

Table 2: Estimation FM

| Correction FM avec $\Sigma$ connue | | | | | | |
|---|---|---|---|---|---|---|
| T | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 |
| 50 | 002 (7.29) | 002 (6.98) | $-002$ (6.55) | $-002$ (5.14) | 001 (3.54) | 003 (1.31) |
| 100 | 0002 (1.80) | $-0004$ (1.72) | $-0001$ (1.53) | 0005 (1.24) | 0008 (0.88) | $-0006$ (0.34) |
| 500 | 0002 (0.069) | 0003 (0.07) | $-0004$ (0.06) | $-0002$ (0.05) | $-00006$ (0.03) | $-0002$ (0.01) |
| 1000 | 0002 (0.017) | $-00007$ (0.02) | $-00007$ (0.02) | 0002 (0.01) | 00007 (0.01) | $-00001$ (0.005) |

| Correction FM avec $\widehat{\Sigma}_{MCO}$ | | | | | | |
|---|---|---|---|---|---|---|
| T | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 |
| 50 | 003 (7.85) | 007 (7.61) | 006 (7.08) | 017 (5.85) | 025 (4.24) | 035 (1.86) |
| 100 | 0003 (1.87) | 0002 (1.80) | 002 (1.62) | 004 (1.31) | 007 (0.92) | 008 (0.38) |
| 500 | $-0003$ (0.07) | 0004 (0.07) | $-0002$ (0.06) | $-0003$ (0.05) | 0002 (0.03) | 00009 (0.01) |
| 1000 | $-0002$ (0.02) | $-0003$ (0.02) | $-00004$ (0.02) | 0002 (0.01) | 0001 (0.01) | 00009 (0.006) |

corrélations élévées, la correction FM est fondée sur un estimateur fortement biaisé de la matrice $\Sigma$, ce qui explique que les biais sont alors plus importants.

Comme pour toute méthode d'estimation en deux étapes, on vérifie que lorsque les Fully Modified reposent sur une estimation préalable de la matrice de variance covariance de long terme, les résultats dépendent fondamentalement de la convergence de l'estimateur de $\Sigma$. C'est pourquoi sur des petits échantillons, des méthodes d'estimation non paramétriques comme la méthode des Moindres Carrés Dynamiques (*DOLS*) de Stock et Watson (1993) peuvent apporter, sous certaines hypothèses, une meilleure correction du biais d'endogénéité de long terme que la méthode des Fully Modified (Montalvo 1995).

Dans tous les cas, on observe en outre que la variance de l'estimateur est d'autant plus faible que la corrélation est forte. En effet, nous avons vu que la variance de la composante orthogonale aux innovations est une fonction décroissante de la corrélation ($\sigma_1^{+2} = \sigma_1^2\left(1-\tau^2\right)$). Dans notre expérience étant donné que $\sigma_1^2$ et $\sigma_2^2$ sont fixes, plus la corrélation est forte, plus la variance de l'estimateur $\widehat{\beta}_{FM}$ est faible.

On peut vérifier l'importance de la correction des FM en comparant pour une taille d'échantillon relativement faible ($T = 100$) et une corrélation importante ($\tau = 0.9$), les distributions approximées de l'estimateur des MCO et de l'estimateur FM obtenu à partir d'une estimation préalable par les MCO de la matrice $\Sigma$ (figure 2).

Figure 2: Fonctions de Densité Empiriques des Biais des Estimateurs FM et MCO ($T = 100$, $\tau = 0.9$)

# 4 Limites de la méthode des estimations Fully Modified

La méthode d'estimation des Fully Modified, telle qu'elle a été proposée par Phillips (1995), est une méthode d'estimation très générale pouvant être appliquée avec succès dans de nombreuses situations d'inférence comportant des séries persistantes.

En effet, dans l'approche de Phillips (1995), les FM permettent d'estimer les paramètres d'une ou plusieurs relations de cointégration sans avoir à priori à tester le degré d'intégration des séries ou l'hypothèse de cointégration. De plus, à distance finie, contrairement aux Moindres Carrés Ordinaires, ces estimations sont non biaisées lorsque les innovations des variables explicatives $I(1)$ sont corrélées avec les résidus des relations de long terme (biais d'endogénéité de long terme).

Enfin, cette méthode permet de se ramener à des distributions standards pour les principaux tests usuels. Contrairement à la méthode d'Engle et Granger (1987), il est alors possible de tester les paramètres de la relation de cointégration. Toutes les techniques traditionnelles d'inférence sont valides asymptotiquement.

Mais la principale limite des Fully Modified réside dans les conditions de validité d'une procédure d'estimation en deux étapes. Comme nous l'avons vu précédemment, la première étape des Fully Modified consiste en l'estimation kernel des matrices de variance covariance de long terme. Toute la correction du biais d'endogénéité repose sur l'obtention d'un estimateur convergent des matrices et $\Delta$.

Or, l'estimation par la fonction kernel de ces deux matrices est réalisée à partir des covariances empiriques calculées, en particulier, à partir des résidus des relations de long terme obtenus par la méthode des MCO.

Nous avons vu que lorsque toutes les variables explicatives sont $I(1)$, l'estimateur des MCO est super-convergent, puisque le biais d'endogénéité de long terme est $Op(1)$. Ce résultat demeure valable lorsque l'on introduit des variables explicatives $I(0)$ orthogonales aux résidus des relations de long terme.

Cependant, lorsque les variables stationnaires sont endogènes (au sens traditionnel du terme), les conditions nécessaires à la convergence de l'estimateur des MCO ne sont plus satisfaites. La correction des FM sera alors fondée sur une mauvaise approximation de la véritable population des résidus $\mu_{0t}$ et sur une estimation biaisée de la matrice de variance covariance de long terme.

**Résultat 1** *En présence de variables stationnaires dans la relation principale les estimateurs des Fully Modified des paramètres associés aux variables $I(0)$ sont asymptotiquement biaisés dès lors que ces variables sont corrélées au résidu de la relation de cointégration (biais standard de simultanéité).*

En effet, la transformation semi-paramétrique retenue ne permet pas d'orthogonaliser les variables stationnaires du système. Mais de plus, l'introduction de variables $I(0)$ endogènes peut affecter les estimations des paramètres associés aux variables non stationnaires.

**Résultat 2** *Lorsque les variables stationnaires et les innovations des variables non stationnaires sont corrélées au résidu de la relation de cointégration, les estimateurs des Fully Modified des paramètres associés aux variables $I(1)$ sont biaisés à taille d'échantillon finie dès lors que les variables $I(0)$ sont corrélées aux innovations des variables $I(1)$ (biais d'endogénéité de long terme résiduel).*

La démonstration de ces deux résultats figure en annexe B.2. L'intuition est très simple. Lorsque la relation de cointégration comporte des variables explicatives stationnaires qui sont corrélées avec le résidu de la relation de cointégration, les estimateurs des MCO des paramètres associés à ces variables sont biaisés (bais d'endogénéité standard). Dans ce cas, les MCO ne fournissent pas une bonne estimation de la population des résidus de la première étape est biaisé. La correction semi-paramètrique des Fully Modified est alors fondée sur une mesure biaisée de la matrice de variance covariance de long terme des résidus. Elle peut, dans ce contexte, laisser subsister une partie du biais d'endogénéité de long terme et les estimateurs de tous les paramètres (y compris ceux des variables $I(1)$) sont alors biaisés.

Considérons par exemple le cas de l'estimation d'une fonction de consommation sous l'hypothèse de revenu permanent. Sous cette hypothèse, l'épargne est définie comme une proportion constante de la variation de la richesse totale. Supposons que le revenu courant et la consommation soient des variables non stationnaires. Etant donnée la contrainte de ressource des agents, le caractère stationnaire de l'épargne implique l'existence d'une relation de cointégration entre la consommation, $c_t$, et le revenu courant, $y_t$, de vecteur $(1, -1)$.

Dès lors, une manière possible de tester l'hypothèse de revenu permanent consiste à estimer une relation de long terme entre $c_t$ et $y_t$ et à tester le vecteur de cointégration. La méthode des Fully Modified est particulièrement appropriée à ce type d'exercice, puisque, entre autres, elle permet l'utilisation des procédures d'inférence standard sur les paramètres de la relation de cointégration.

Supposons par ailleurs que la relation structurelle à estimer tienne compte des mécanismes d'épargne de précaution et que la consommation soit définie comme une fonction décroissante de la variation du taux de chômage, notée $\widetilde{u}_t$. Le taux de chômage étant une variable $I(1)$, son taux de croissance est stationnaire. Bien que la relation de cointégration théorique dérivée de l'hypothèse de revenu permanent n'implique pas la présence du taux de chômage, on peut penser que la présence de ce dernier permet d'améliorer le contenu informationel du système. Le modèle à estimer est alors le suivant :

$$\begin{aligned} c_t &= \alpha_0 + \alpha_1 y_t + \alpha_2 \widetilde{u}_t + \epsilon_{c,t} \quad \alpha_1 = 1,\ \alpha_2 < 0 \\ \Delta y_t &= \epsilon_{y,t} \\ \widetilde{u}_t &= \epsilon_{u,t} \end{aligned} \tag{19}$$

avec $\epsilon_t' = (\epsilon_{c,t}, \epsilon_{y,t}, \epsilon_{u,t})$, $E(\epsilon_t) = 0$, $E(\epsilon_t \epsilon_t') = \Sigma$. Si l'on s'en tient à l'approche de Phillips (1995), il n'est pas nécessaire de tester à priori le degré d'intégration des variables explicatives avant d'appliquer la méthode d'estimation des Fully Modified (sous réserve qu'il y ait bien entendu au moins une variable $I(1)$). Dans la première étape d'estimation de la matrice de variance covariance de long terme, toutes les variables explicatives, y compris le taux de croissance du chômage, sont différenciées.

Mais, nous avons montré qu'une telle approche suppose que les variables stationnaires du modèle vérifient les bonnes propriétés d'orthogonalité à la fois par rapport au résidu de la relation de cointégration et par rapport aux innovations des régresseurs $I(1)$.

Afin d'illustrer l'incidence de ces propriétés d'orthogonalité, nous avons simulé 5000 pseudo-échantillons à partir du modèle (19) par la méthode de Monte-Carlo. Dans un premier groupe d'expériences, les paramètres du modèle ont été étalonnés sur donnés semestrielles françaises (source : *Perspectives Economiques de l'OCDE*) sur la période 1960:1-1997:2. Dans ce cas, le taux de croissance du chômage est corrélé négativement au taux de croissance du revenu ($corr(\Delta y_t, \widetilde{u}_t) = -0.25$) et positivement au résidu de la relation de cointégration[6] ($corr(\widehat{\epsilon}_{c,t}, \widetilde{u}_t) = -0.49$). Pour le second groupe de réalisations, nous avons supposé que le taux de croissance du chômage vérifiait les bonnes propriétés d'orthogonalité ($\sigma_{c,u} = \sigma_{y,u} = 0$).

[6] La relation de cointégration a été estimée par les Moindres Carrés Dynamiques afin d'obtenir un estimateur non biaisé de la population des résidus. On trouve $_2 = -013$ $_0 = -051$

Sur la figure (3), le premier graphique indique clairement que lorsque le taux de croissance du chômage (variable $I(0)$) est corrélé avec les innovations du produit (variable $I(1)$), la distribution de l'estimateur des Fully Modified du paramètre $\alpha_1$ est biaisée. Ce biais disparaît dès lors que $\sigma_{yu} = 0$, puisque dans ce cas le terme de correction des FM est fondé sur un estimateur convergent et non biaisé de la matrice de variance covariance de long terme.

En revanche, le biais dans la distribution du paramètre $\alpha_2$ associé à la variable stationnaire demeure présent même lorsque $\sigma_{yu} = 0$. En effet, ce biais correspond à un biais d'endogénéité standard, et est par là même indépendant de la covariance entre les taux de croissance du chômage et de la production.

Concernant le test de l'hypothèse de revenu permanent , on constate que dans le cas où la corrélation $\sigma_{yu}$ est nulle, la distribution de la t-statistique associée au test $\alpha_1 = 1$ est parfaitement centrée sur la valeur nulle. Mais lorsque la correction des FM est fondée sur un estimateur biaisé de la matrice de variance covariance de long terme ($\sigma_{yu} \neq 0$), le test de Student est biaisé en faveur de l'hypothèse alternative. La correction imparfaite du biais d'endogénéité de long terme tend peut conduire à rejeter de manière fallacieuse l'hypothèse nulle de revenu permanent. Au seuil standard à 5%, l'hypothèse de revenu permanent est rejetée dans près de 11% des cas lorsque $\sigma_{yu} \neq 0$.

Ainsi les Fully Modified peuvent être appliqués dans un cadre très général, intégrant à la fois des variables stationnaires et non stationnaires. Cependant, contrairement à ce que préconise Phillips (1995), il convient cependant de tester à priori le degré d'intégration des séries afin de vérifier d'une part que les variables $I(0)$ satisfont les bonnes propriétés d'orthogonalité et que d'autre part les innovations des variables $I(1)$ et les variables $I(0)$ ne sont pas corrélées. Sans ces précautions, les estimations FM peuvent d'une part être affectées d'un biais standard de simultanéité et d'autre part ne corriger que partiellement le biais d'endogénéité de long terme.

Il est de plus évident que si la correction du biais d'endogénéité de long terme n'est que partielle, les distributions asymptotiques des tests usuels ne seront plus standard.

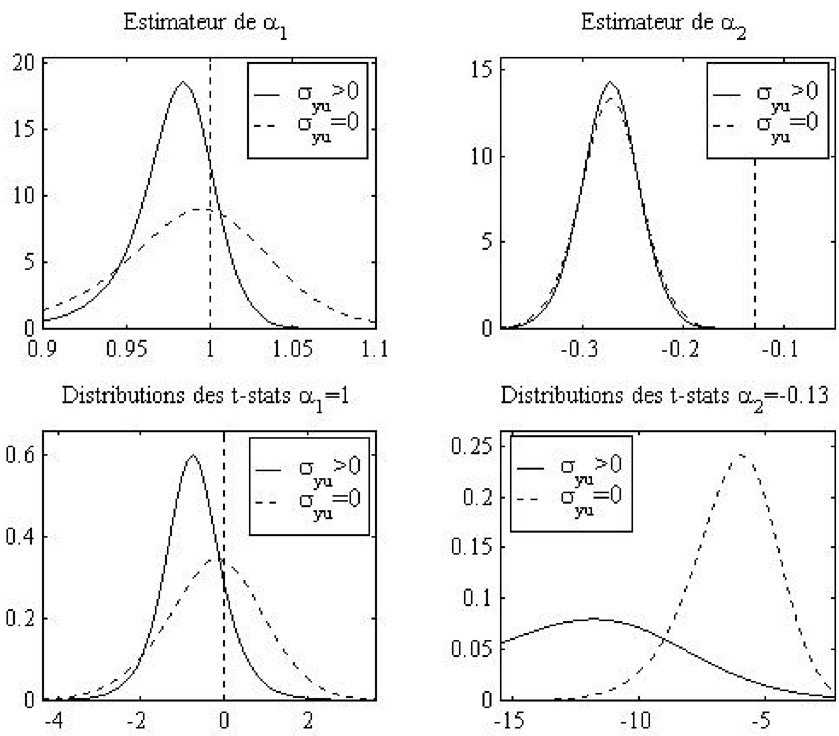
# 5 Conclusion

La méthode d'estimation Fully Modified est une méthode très générale d'estimation des paramètres d'une ou plusieurs relations de cointégration. Elle permet de corriger les biais d'endogénéité de long terme et de se ramener à des distributions standard pour les statistiques de tests usuels.

Sa mis en oeuvre est très simple[7] et consiste à appliquer les MCO à un modèle transformé. La transformation est fondée sur une estimation convergente de la matrice de variance covariance de long terme. L'idée étant d'orthogonaliser les

[7] Des procédures informatiques d'estimation par les Fully Modified sont disponibles dans plusieurs logiciels (Gauss). Les simulations de cette étude ont été réalisée à partir de programmes réalisés sous TSP et sous Matlab.

Figure 3: Estimation d'une Fonction de Consommation




résidus de la relation de cointégration par une projection sur l'espace engendré par les innovations des variables explicatives.

Cependant, du fait de sa structure en deux étapes, son efficacité est soumise à certaines conditions notamment en présence de variables stationnaires. Nous avons montré que la correction FM du biais d'endogénéité de long terme n'était que partielle lorsque les variables $I(0)$ étaient d'une part endogènes et d'autre part corrélées avec les innovations des variables $I(1)$.

Ainsi, contrairement à ce qu'affirme Phillips (1995), l'application des FM nécessite la connaissance préalable de l'ordre d'intégration des séries utilisées. Dès lors que le système comprend des variables stationnaires, celles-ci doivent vérifier les bonnes propriétés d'orthogonalité.

# A Encadrés

## A.1 Encadré 1 : Estimation de  et $\Delta$

Afin d'estimer les matrices de variance covariance de long terme, on utilise généralement la méthode des estimateurs à noyau ou estimateurs kernel (Andrews 1991). Cette méthode nécessite le choix d'une fonction kernel et d'un paramètre de troncature. L'idée de base de cette estimation non paramétrique repose sur le fait que la matrice de variance covariance de long terme  peut être approximée par $2\pi$ fois une estimation de la matrice de densité spectrale considérée à la fréquence nulle.

Afin d'estimer la matrice de densité spectrale considérée à la fréquence nulle, on utilise une moyenne pondérée du périodogramme de la série considérée calculée pour des fréquences voisines de la fréquence nulle. En effet, le périodogramme calculé pour une fréquence particulière nous donne un estimateur de la densité au voisinage de cette fréquence. Puisque l'on sait que pour des fréquences très proches, les densités spectrales sont elles aussi très proches, il suffit alors de calculer une moyenne pondérée des valeurs du périodogramme obtenues autour de la fréquence nulle.

$$\begin{aligned}
\widehat{\ } &= \sum_{j=-T+1}^{T+1} w\left(\frac{j}{K}\right)\widehat{\Gamma}(j) \quad \widehat{\Delta} = \sum_{j=0}^{T+1} w\left(\frac{j}{K}\right)\widehat{\Gamma}(j) \\
\widehat{\Gamma}(j) &= \frac{1}{T}\sum_{j=1}^{T}\widehat{u}_{t+j}\widehat{u}'_t \quad \widehat{u}'_t = (\widehat{u}_{0t}, u_{1t}, u_{2t})
\end{aligned}$$

La fonction kernel $w(.)$ détermine le poids à affecter à chaque fréquence. Le paramètre de troncature (*bandwidth parameter*) $K$ détermine la bande de fréquence qui est nécessaire pour obtenir une estimation convergente de la matrice de densité spectrale considérée à la fréquence nulle.

$$w\left(\frac{j}{K}\right) = 0 \quad \nabla |j| \geq K$$

Dans ce contexte, il est possible d'utiliser l'ensemble des fonctions kernel présentées par Andrew (1991) : kernel tronqué, kernel de Bartlett, kernel de Parzen, kernel de Tukey-Hanning, ou kernel spectrale quadratique.

Concernant le paramètre de troncature, il existe deux types de choix possibles. Le premier consiste à retenir un paramètre de troncature fixe, qui soit inférieur au taux de convergence du paramètre de troncature optimal. La seconde possibilité, préconisée par Andrews (1991), consiste à utiliser un paramètre de troncature déterminé automatiquement[8].

[8] Dans ce cas on considère un modèle (1) univarié pour toutes les composantes de $\vec{\ }_t = (\widehat{\ }_{0t}\ {}_{1t}\ {}_{2t})$ Soient $\widehat{\ }_i$ et $\widehat{\ }_i$ les racines et les variances des résidus associées à la i ème composante de $\widehat{\ }_t$ Le paramètre de troncature automatique proposé par Andrews (1991) dans le cas de la fonction kernel de Bartlett

## A.2 Encadré 2 : Intuition dans un modèle simple

Reprenons l'exemple décrit par les équations (1) et (2). On suppose que l'on connaît un estimateur convergent $\widehat{\Sigma}$ de la matrice $\Sigma$ (on notera par la suite $\sigma_{ij} = \widehat{\sigma}_{ij}$). Ce point sera discuté par la suite. La méthode d'estimation des Fully Modified consiste alors simplement à appliquer les MCO au modèle transformé suivant

$$y_t^+ = \beta x_t + \mu_{1t}^+ \tag{21}$$

$$y_t^+ = y_t - \frac{\sigma_{12}}{\sigma_2^2} \Delta x_t \tag{22}$$

$$\mu_{1t}^+ = \mu_{1t} - \frac{\sigma_{12}}{\sigma_2^2} \Delta x_t \tag{23}$$

La transformation des FM est ici très simplement interprétable. Elle consiste simplement à régresser $y_t$ sur $x_t$ en ayant au préalable orthogonaliser les deux résidus $\mu_{1t}$ et $\mu_{2t}$. Le résidu de la relation de cointégration $\mu_{1t}$ peut en effet se réécrire sous la forme suivante :

$$\mu_{1t} = \gamma \mu_{2t} + \mu_{1t}^+ \tag{24}$$

avec

$$\gamma = \frac{cov\left(\mu_{1t}, \mu_{2t}\right)}{var(\mu_{2t})} = \frac{\sigma_{12}}{\sigma_2^2}$$

Le résidu $\mu_{1t}$ correspond à la somme d'une composante de projection sur $\mu_{2t}$ $(\sigma_{12}\mu_{2t}/\sigma_2^2)$ et d'une composante orthogonale à $\mu_{2t}$ de variance égale à $\sigma_1^2\left(1-\tau^2\right)$. Le modèle initial peut alors être exprimé sous la forme :

$$y_t = \beta x_t + \mu_{1t} \Longleftrightarrow y_t - \frac{\sigma_{12}}{\sigma_2^2}\mu_{2t} = \beta x_t + \mu_{1t}^+ \qquad cov\left(\mu_{1t}^+, \mu_{2t}\right) = 0$$

La distribution asymptotique de l'estimateur FM, noté $\widehat{\beta}_{FM}$, est alors la suivante

$$T\left(\widehat{\beta}_{FM} - \beta\right) = \frac{\frac{1}{T}\sum\limits_{t=0}^{T} x_t \mu_{1t}^+}{\frac{1}{T^2}\sum\limits_{t=0}^{T} x_t^2} \underset{T\to\infty}{\xrightarrow{L}} \frac{\sigma_1^+}{\sigma_2} \frac{\int\limits_0^1 W_2\left(r\right) dW_1\left(r\right)}{\int\limits_0^1 W_2\left(r\right)^2 dr} \tag{25}$$

avec

$$\sigma_1^{+2} = var\left(\mu_{1t}^+\right) = \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2} = \sigma_1^2\left(1-\tau^2\right).$$

---

est alors défini par

$$\widehat{\ } = 11147\left(\sum_{i=1}^{n+m} \frac{4\widehat{\ }_i^2 \widehat{\ }_i^2}{(1-\widehat{\ }_i)^6 (1+\widehat{\ }_i)^2} \Bigg/ \sum_{i=1}^{n+m} \frac{\widehat{\ }_i^2}{(1-\widehat{\ }_i)^4}\right)^{\frac{1}{3}} \tag{20}$$

On retrouve ainsi une distribution qui, conditionnellement au processus $W_2(.)$, est gaussienne et dont la variance est une fonction décroissante de la corrélation $\tau$.

Quel que soit le niveau de la corrélation, la statistique de Student converge en loi vers une variable aléatoire distribuée suivant une loi normale centrée réduite. En effet, d'après le résultat (25) on sait que :

$$s_T^{+2} = \frac{1}{T-1} \sum_{t=1}^{T} \widehat{\mu}_{1t}^{+2} \underset{T\to\infty}{\overset{p}{\longrightarrow}} E\left(\mu_{1t}^{+2}\right) = \sigma_1^{+2}$$

Dès lors, on montre que

$$\forall \tau \quad t_{\widehat{\beta}} \underset{T\to\infty}{\overset{L}{\longrightarrow}} \frac{\int_0^1 W_2(r)\, dW_1(r)}{\left[\int_0^1 W_2(r)^2\, dr\right]^{\frac{1}{2}}} \rightsquigarrow N(0,1) \tag{26}$$

De la même façon, on montre que la statistique du test de Fisher suit un $\chi_2(m)$ où $m$ désigne le nombre de contraintes testées (ici en l'occurrence $m = 1$).

# B Annexes

## B.1 Terme de correction FM

Ce terme correspond à la covariance entre les innovations $\mu_{1t}$ et $\mu_{2t}$ et la composante du résidu $\mu_{0t}^{+}$ orthogonalisée. Pour $m_1 = 0$ (c'est à dire lorsque toutes les variables explicatives sont non stationnaires) on a

$$\begin{aligned}
\widehat{\Delta}_{yx}^{+} &= \sum_{j=0}^{\infty} E\left(u_{2t-j} \;\; \mu_{0t}'^{+}\right) \\
&= \sum_{j=0}^{\infty} E\left(u_{2t-j} \;\; [u_{0t}\, \mu_{2t}]'\right) \begin{bmatrix} I \\ -\widehat{\ }_{yx} \widehat{\ }_{xx}^{-1} \end{bmatrix} \\
&= \sum_{j=0}^{\infty} E\left(u_{2t-j} u_{0t}'\right) - \widehat{\ }_{yx} \widehat{\ }_{xx}^{-1} \sum_{j=0}^{\infty} E\left(u_{2t-j} u_{2t}'\right)
\end{aligned}$$

Dans l'exemple décrit dans la section précédente, ce terme était nul en raison de l'absence de corrélation retardées des innovations et des résidus.

## B.2 Estimateur FM en présence de variables $I(0)$ endogènes

Afin de démontrer ces deux propositions, considérons un modèle simple avec une variable $I(1)$ et une variable $I(0)$ dans la relation principale ($n = m_1 = m_2 = 1$).

$$
\begin{aligned}
y_t &= \beta_1 x_{1t} + \beta_2 x_{2t} + \mu_{0t} && (27)\\
x_{1t} &= \mu_{1t} && (28)\\
\Delta x_{2t} &= \mu_{2t} && (29)
\end{aligned}
$$

On se place dans le cas où la variable explicative stationnaire $x_{1t}$ est endogène au sens traditionnel, c'est à dire corrélée avec le résidu $\mu_{0t}$. On suppose que les innovations des variables $I(1)$, $\mu_{2t}$, peuvent être corrélées avec le niveau des variables stationnaires, c'est à dire avec $\mu_{1t}$. Pour simplifier les calculs on suppose en outre que les résidus ne sont pas autocorrélés (mais la démonstration peut très facilement étendue au cas autocorrélé[9]).

$$
E\left(u_t u_t'\right) = \Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} \\ \sigma_{01} & \sigma_1^2 & \sigma_{12} \\ \sigma_{02} & \sigma_{12} & \sigma_2^2 \end{bmatrix} \qquad \mu_t' = \left(\mu_{0t}\, \mu_{1t}\, \mu_{2t}\right)
$$

Si la matrice de variance covariance de long terme (qui correspond ici à $\Sigma$) n'est pas connue, la procédure d'estimation des Fully Modified se décompose en deux étapes. La première étape consiste à estimer la population des résidus $\mu_{0t}$ par les MCO afin d'obtenir un estimateur de $\Sigma$.

$$
\widehat{\mu}_{0t} = \left(\beta_1 - \widehat{\beta}_1\right) x_{1t} + \left(\beta_2 - \widehat{\beta}_2\right) x_{2t} + \mu_{0t} \tag{30}
$$

Sachant que le biais de simultanéité associé à l'estimateur MCO de $\beta_1$ est $Op(0)$ et que le biais d'endogénéité de long terme associé à l'estimateur MCO de $\beta_2$, on a :

$$
\widehat{\beta}_2 - \beta_2 \underset{T\to\infty}{\overset{p}{\longrightarrow}} 0 \qquad \widehat{\beta}_1 - \beta_1 \underset{T\to\infty}{\overset{p}{\longrightarrow}} \frac{\sigma_{01}}{\sigma_1^2}
$$

Conformément aux résultats standard, l'estimateur MCO, noté $\widehat{\Sigma}$, de la matrice de variance covariance $\Sigma$ est alors asymptotiquement biaisé. Dès lors, la correction de l'estimateur FM est fondée sur un estimateur non convergent de la matrice de variance covariance de long terme. Le modèle transformé s'écrit :

$$
\begin{aligned}
y_t^+ &= y_t - \widehat{\Sigma}_{yx}\widehat{\Sigma}_{xx}^{-1}\Delta x_t && (31)\\
\mu_{0t}^+ &= \mu_{0t} - \widehat{\Sigma}_{yx}\widehat{\Sigma}_{xx}^{-1}\Delta x_t && (32)
\end{aligned}
$$

Le terme de correction obtenu à partir de l'estimateur des MCO de la matrice $\Sigma$ converge vers une mesure biaisée des coefficients issus de l'orthogonalisation de la matrice $\Sigma$.

---

[9] En présence d'autocorrélation des résidus et des innovations, le cadre de validité de la seconde proposition doit cependant être étendu au cas où il existe une corrélation avancée ou retardée entre $_{1t}$ et $_{2t}$

$$\widehat{\Sigma}_{yx}\widehat{\Sigma}_{xx}^{-1} \underset{T\to\infty}{\overset{p}{\longrightarrow}} \left[\begin{array}{cc} 0 & \frac{1}{\sigma_2^2}\left(\sigma_{02} - \frac{\sigma_{01}\sigma_{12}}{\sigma_1^2}\right) \end{array}\right] \tag{33}$$

Les distributions des estimateurs FM des deux paramètres $\beta_1$ et $\beta_2$ respectivement associés aux variables $I(0)$ et $I(1)$ sont biaisés. Concernant le paramètre $\beta_1$ on retrouve le biais traditionnel d'endogénéité.

$$\widehat{\beta}_1^{FM} - \beta_1 \underset{T\to\infty}{\overset{p}{\longrightarrow}} \frac{\sigma_{01}^{+}}{\sigma_1^2} = \frac{1}{\sigma_1^2}\left(\sigma_{01} + \frac{\sigma_{01}\sigma_{12}^2}{\sigma_1^2\sigma_2^2} - \frac{\sigma_{02}\sigma_{12}}{\sigma_2^2}\right) \tag{34}$$

Dès lors que la variable stationnaire est endogène ($\sigma_{01} \neq 0$), l'estimateur FM du paramètre $\beta_1$ est biaisé même si les innovations des régresseurs ne sont pas corrélées, c'est à dire $\sigma_{12} = 0$.

En revanche, la distribution de l'estimateur FM du paramètre associé à la variable non-stationnaire, $\beta_2$, ne sera biaisé que si d'une part la variable $I(0)$ est endogène ($\sigma_{01} \neq 0$) et si d'autre part les innovations des régresseurs sont corrélés ($\sigma_{12} \neq 0$), comme on peut l'observer à partir de la distribution suivante.

$$\begin{aligned} T\left(\widehat{\beta}_2^{FM} - \beta_2\right) \underset{T\to\infty}{\overset{L}{\longrightarrow}} \frac{\sigma_0^{+}}{\sigma_2}\frac{\int_0^1 W_2(r)\,dW_0(r)}{\int_0^1 W_2(r)^2\,dr} & -\frac{\sigma_1^{+}\sigma_{01}^{+}}{\sigma_1^2\sigma_2}\frac{\int_0^1 W_2(r)\,dW_1(r)}{\int_0^1 W_2(r)^2\,dr} \\ & -\frac{\sigma_{12}\sigma_{01}^{+}}{\sigma_1^2\sigma_2^2}\frac{\left\{\frac{1}{2}\left[W_2^2(1)+1\right] - \frac{\sigma_{01}}{\sigma_{01}^{+}}\right\}}{\int_0^1 W_2(r)^2\,dr} \end{aligned} \tag{35}$$

avec

$$\begin{aligned} \sigma_0^{+2} &= E\left(\mu_{0t}^{+2}\right) = \sigma_0^2 - \frac{\sigma_{02}^2}{\sigma_2^2} + \left(\frac{\sigma_{01}\sigma_{12}}{\sigma_2\sigma_1^2}\right)^2 \\ \sigma_1^{+2} &= \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2} \end{aligned}$$

Le biais dans la distribution provient du dernier terme qui est distribué suivant un $\chi^2(1)$. Le biais dans l'estimateur de $\beta_2$ n'apparaît donc que si $\sigma_{12}$ est non nul. En effet, dans le cas où $\sigma_{12} = 0$, on obtient une distribution centrée, même si $\sigma_{01}$ est non nul.

$$T\left(\widehat{\beta}_2^{FM} - \beta_2\right) \underset{T\to\infty}{\overset{L}{\longrightarrow}} \frac{\sigma_0^{+}}{\sigma_2}\frac{\int_0^1 W_2(r)\,dW_0(r)}{\int_0^1 W_2(r)^2\,dr} - \frac{\sigma_{01}}{\sigma_1\sigma_2}\frac{\int_0^1 W_2(r)\,dW_1(r)}{\int_0^1 W_2(r)^2\,dr}$$

Dès lors que $\sigma_{12} \neq 0$, la correction FM est en partie inefficace. On retrouve en effet dans la distribution de l'estimateur de $\beta_2$ une composante résiduelle du biais d'endogénéité de long terme qui affecte la distribution des MCO. L'amplitude de ce biais dépend du niveau des corrélations $\sigma_{12}$ et $\sigma_{01}$.

## C Références


Andrews D.W.K (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica* 59, pp 817-858.

Cifarelli G. (1995), "Fundamentals, Regime Shifts, and Dollar Behavior in the 1980s", *Open Economies Review,* 6(1), pp 29-48.

Engle R.F. et Granger C.W.J. (1987), "Co-Integration and Error Correction : Representation, Estimation and Testing", *Econometrica* 55, pp 251-276.

Dutt S. et Gosh D. (1995),"Are Forward Rates Free of the Risk Premium ? An Empirical Examination". *Internaional Economic Journal,* 9(3), pp 49-60.

Dutt S. et Gosh D. (1996), "Purchasing Power Parity Doctrine : An Unrestricted Cointegration Test", *Studies in Economics and Finance,* 16(2), pp 22-45.

Hansen B.E. (1992), "Efficient Estimation and Testing of Cointegrating Vectors in the Presence of Deterministic Trends", 53(1-3), pp 87-121.

Hansen B.E. and Phillips P.C.B.(1990), "Estimation and Inference in Models of Cointegration : a Simulation Study", *Advances in econometrics,* 8, pp 225-248.

Hénin P.Y. et Hurlin C. (1997), "L'Evaluation de la Contribution Productive des Investissements Publics", *Rapport de contrat finalisé 1996 pour le Commissariat Général du Plan*, CEPREMAP

Kitamura Y. et Phillips P. (1997), "Fully Modified IV, GIVE and GMM Estimation with Possibly Non-Stationnary Regressors and Instruments", *Journal of Econometrics*, 80(1), pp 85-123.

Kostia K. (1995),"The Fully Modified OLS Estimator as a System Estimator: A Monte Carlo Analysis", *Working Paper 95/8,* European University Institute, Florence.

Li Y., Maddala G.S. et Rush M. (1995), "New Small Sample Estimators for Cointegrating Regression : Low-Pass Spectral Filter Method", *Economics Letters,* 47(2), pp 123-129.

MacDonald R. et Moore M.J. (1996), "Long-Run Purchasing Power Parity and Structural Change", *Economie Appliquée*, 49(3), pp 11-48.

MacFarland J.W., McMahon P.C. et Ngama Y. (1994), "Forward Exchange Rate and Expectations during the 1920s : A Re-examination of the Evidence", *Journal of International Money and Finance,* 13(6), pp 627-636.

Mamingi N. (1997), "Saving-Investment Correlations and Capital Mobility : The Experience of Developing Countries", *Journal of Policy Modeling, vol 19, pp 605-626.*

Montalvo J.G. (1995) , "Comparing Cointegrating Regression estimators : Somme Additional Monte Carlo Results", *Economics Letters*, 48, pp 229-234

Pesaran M.H. et Shin Y. (1995), "An Autoregressive Distributed Lag Modeling Approach to Cointegration Analysis", *Department of Applied Economics Working Paper,* 9514 University of Cambridge.

Otto G.D. et Voss G.M. (1997), Public Capital and Private Production in Australia", *Southern Journal of Economics",* 3, pp 723-738,

Phillips, P.C.B. (1988) "Weak Convergence of Sample Covariance Matrices to Stochastic Integrals via Martingale Approximations", *Econometric Theory* 4, pp 528-533.

Phillips, P.C.B. (1993), "Robust Nonstationary Regression", *Yale Cowles Foundation Discussion Paper*, 1064.

Phillips, P.C.B. (1995), "Fully Modified Least Squares and Vector Autoregression", *Econometrica*, vol 63, 5, pp 1023-1078.

Stock J.H. et Watson M.W. (1993), "A simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems", *Econometrica,* vol 61, 4, pp 783-820.

Toda H.Y. et Yamada H. (1997), "A Note on Hypothesis Testing Based on the Fully Modified Vector Autoregression", *Economic Letters,* 56(1) pp 27-39.