



HAL
open science

Modèles dynamiques à erreurs composées avec autocorrélation : bref aperçu

Serge-Alain Matondzi Ngouma

► **To cite this version:**

Serge-Alain Matondzi Ngouma. Modèles dynamiques à erreurs composées avec autocorrélation : bref aperçu. [Rapport de recherche] Laboratoire d'analyse et de techniques économiques(LATEC). 1997, 24 p., ref. bib.: 1 p. 1/2. hal-01526910

HAL Id: hal-01526910

<https://hal.science/hal-01526910>

Submitted on 23 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

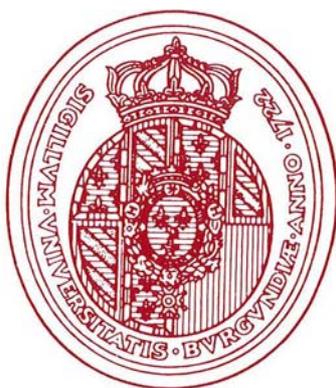
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LATEC

LABORATOIRE D'ANALYSE ET DE TECHNIQUES ÉCONOMIQUES

U.R.A. 342 C.N.R.S.

DOCUMENT de TRAVAIL



UNIVERSITE DE BOURGOGNE

FACULTE DE SCIENCE ECONOMIQUE ET DE GESTION

4, boulevard Gabriel - 21000 DIJON - Tél. 80 39 54 30 - Fax 80 39 56 48

ISSN : 0292-2002

n° 9702

**Modèles dynamiques à erreurs composées avec
autocorrélation : bref aperçu.**

Serge-Alain MATONDZI NGOUMA

janvier 1997

RESUME

L'étude des modèles dynamiques à erreurs composées avec autocorrélation est très récente en économétrie des données de panel. Aucun article ou ouvrage à ce jour ne traite de façon exhaustive cette question.

Nous proposons dans ce travail un bref aperçu de l'étude de ces modèles. Nous analysons notamment les estimateurs usuels, du maximum de vraisemblance et des variables instrumentales en faisant à chaque fois ressortir les particularités de ces modèles. Cette étude fait suite aux deux premières études.

SUMMARY

Studying dynamic error component models with autocorrelated disturbances has started very recently in econometrics analysis of panel data. There is no article or book which tackles this topic exhaustively.

In this study, a brief survey of these models is presented. The within, between, OLS, GLS, maximum likelihood and instrumental variable estimators are especially analysed, and the particular characteristics of these models are underlined each time.

MOTS CLES

Modèles dynamiques à erreurs composées, Autocorrélation, Conditions initiales, Propriétés asymptotiques, Estimation convergente, Limite en probabilité, Maximum de vraisemblance, Variable instrumentale, Conditions d'orthogonalité

SOMMAIRE

INTRODUCTION	3
I - LE MODELE	3
II - HYPOTHESES, NOTATION ET TRANSFORMATION DU MODELE	4
2.1 HYPOTHESES	
2.2 NOTATION	
2.3 TRANSFORMATION DU MODELE	
III - ETUDE DU COMPORTEMENT DES ESTIMATEURS USUELS DU MODELE AR(1) A ERREURS COMPOSEES AVEC AUTOCORRELATION	7
3.1 COMPORTEMENT ASYMPTOTIQUE	8
3.1.1 LES ESTIMATEURS WITHIN, MCG ET MCQG	8
3.1.2 LES ESTIMATEURS DES MCO ET BETWEEN	9
3.2 COMPORTEMENT SEMI-ASYMPTOTIQUE	10
IV - LA METHODE DU MAXIMUM DE VRAISEMBLANCE	10
4.1 MAXIMUM DE VRAISEMBLANCE : CAS 1	11
4.1.1 LES CONDITIONS DE PREMIER ORDRE	12
4.1.2 LES CONDITIONS DE DEUXIEME ORDRE	13
4.1 MAXIMUM DE VRAISEMBLANCE : CAS 2	16
V - LA METHODE DES VARIABLES INSTRUMENTALES	18
5.1 LES CONDITIONS D'ORTHOGONALITE	19
5.1.1 LA CONDITION DE LIVIATAN	19
5.1.2 LA CONDITION DE BALESTRA NERLOVE	19
5.1.3 LA CONDITION DE LIVIATAN 2	20
5.1.4 LA CONDITION DE AHN ET SCHMIDT	20
5.1.5 AUTRE CONDITION D'ORTHOGONALITE	21
5.2 ESTIMATION DU COEFFICIENT D'AUTOCORRELATION	21
5.2.1 LE MODELE DE BASE	21
5.2.2 LE MODELE TRANSFORME	22
CONCLUSION	23
BIBLIOGRAPHIE	23

INTRODUCTION

Si l'étude des modèles dynamiques à erreurs composées est devenue un classique en économétrie des données de panel, il n'en demeure pas moins que certains aspects de ce modèle restent encore à ce jour non explorés. L'un de ces aspects est la prise en compte de l'autocorrélation dans l'étude de ces modèles. Récemment, Baltagi et Li (1991) ont introduit l'autocorrélation dans le modèle simple à erreurs composées et ont souligné les problèmes que cette nouvelle structure pose dans l'estimation des paramètres.

Nous proposons dans ce travail d'introduire l'autocorrélation dans le modèle dynamique à erreurs composées et de présenter brièvement les propriétés des estimateurs habituels utilisés sur données de panel. Il s'agira ici d'analyser les estimateurs within, between, MCO, MCG, MCQG, maximum de vraisemblance et les estimateurs des variables instrumentales.

Cet article est tiré d'un travail plus important portant sur les modèles dynamiques à erreurs composées avec autocorrélation [MATONDZI (1997b)]. Tous les résultats présentés ici sont démontrés dans ce travail.

I LE MODELE

Dans ce travail, nous allons considérer un modèle dynamique à erreurs composées avec autocorrélation contenant deux variables exogènes. La première x_{it} varie à la fois selon les individus et les périodes, et la deuxième z_i ne possède qu'une variabilité individuelle.

Ce modèle s'écrit :

$$\begin{cases} y_{it} = \delta y_{it-1} + \beta x_{it} + \phi z_i + \alpha_i + v_{it} \\ v_{it} = \rho v_{it-1} + \varepsilon_{it} \end{cases} \quad (1.1)$$

La descente par récurrence jusqu'à l'instant initial nous permet d'écrire encore ce modèle comme suit :

$$y_{it} = \delta^t y_{i0} + \beta \sum_{j=0}^{t-1} \delta^j x_{it-j} + \frac{1-\delta^t}{1-\delta} \phi z_i + \frac{\rho}{\rho-\delta} (\rho^t - \delta^t) v_{i0} + \frac{1-\delta^t}{1-\delta} \alpha_i + \frac{1}{\rho-\delta} \sum_{j=0}^{t-1} (\rho^{j+1} - \delta^{j+1}) \varepsilon_{it-j} \quad (1.2)$$

Ainsi, nous pouvons écrire :

$$y_{it} = f(y_{i0}, x_{it}, z_i, v_{i0}, \alpha_i, \varepsilon_{it}) \quad (1.3)$$

Il apparaît de cette dernière écriture que les observations initiales influencent de façon non négligeable le comportement asymptotique des estimateurs notamment lorsque la dimension temporelle de l'échantillon est finie.

Le problème qui se pose alors est de savoir quel est le véritable statut de ces observations. Les hypothèses qui sont faites sur le statut de ces observations peuvent être regroupées en deux familles.

La première est composée des hypothèses qui supposent que les observations initiales sont exogènes c'est à dire que ces observations sont indépendantes du processus (1.1).

Dans ce cadre, nous pouvons considérer que ces observations sont soit des constantes fixes, (hypothèse 1) soit engendrées par un processus aléatoire d'espérance nulle non corrélées avec les erreurs du modèle (1.1) (hypothèse 2) c'est à dire :

$$P \lim \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{i0} (\alpha_i + v_{i0}) = 0 \quad (1.4)$$

avec

$$H_2 \begin{cases} y_{io} = c + \pi_i \\ E(\pi_i) = 0 \\ E(\pi_i \pi_{i'}) = \begin{cases} \sigma_\pi^2 & \text{si } i = i' \\ 0 & \text{sinon} \end{cases} \end{cases}$$

La deuxième est celle des hypothèses qui acceptent l'existence d'une corrélation entre ces observations et les erreurs du modèle ou de façon générale entre le modèle et les observations initiales.

Dans cette situation, nous pouvons considérer que les observations initiales dépendent de l'effet individuel α_i , des valeurs passées de la variable exogène x_{it-j} , de la variable exogène z_i et de l'erreur initiale v_{io} .

$$y_{io} = f(\alpha_i, x_{it-j}, z_i, v_{io}) \quad (1.5)$$

Soit de façon générale :

$$P \lim \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{io} y_{it} \neq 0 \quad (1.6)$$

L'hypothèse générale qui regroupe tous les cas importants étudiés par les économètres est :

$$H_3 : y_{io} = c + k_1 \alpha_i + k_2 v_{io}$$

Cette hypothèse peut être élargie à :

$$H_4 : y_{io} = c + k_1 \alpha_i + k_2 v_{io} + k_3 x_{io} + k_4 z_i$$

Une bonne présentation des hypothèses sous cette deuxième famille pour les différentes valeurs de k_1 et k_2 est faite par ANDERSON.T.W et HSIAO.C(1981 ET 1982) et SEVESTRE .P (1983).

II - HYPOTHESES, NOTATION ET TRANSFORMATION DU MODELE

2.1 HYPOTHESES

Dans le cadre de ce travail, nous ferons les hypothèses suivantes sur les erreurs :

$$H_1 \quad |\delta| < 1, |\rho| < 1$$

Ces deux conditions ne sont pas nécessaires si T est fini.

$$H_2 \quad E(\alpha_i / y_{it-1}, x_{it}, z_i) = 0$$

$$H_3 \quad E(\alpha_i \alpha_{i'} / y_{it-1}, x_{it}, z_i) = \begin{cases} \sigma_\alpha^2 & \text{si } i = i' \\ 0 & \text{sinon} \end{cases}$$

$$H_4 \quad E(\varepsilon_{it} / y_{it-1}, x_{it}, z_i) = 0$$

$$H_5 \quad E(\varepsilon_{it} \varepsilon_{i't'} / y_{it-1}, x_{it}, z_i) = \begin{cases} \sigma_\varepsilon^2 & \text{si } i = i' \text{ et } t = t' \\ 0 & \text{sinon} \end{cases}$$

$$H_6 \quad E(\alpha_i \varepsilon_{it}) = E(\varepsilon_{it} \alpha_i) = 0$$

H_7 , Nous supposons en outre que v_{io} suit un processus stationnaire :

$$v_{io} \rightarrow \text{iid} \left(0, \frac{\sigma_\varepsilon^2}{1-\rho^2} \right)$$

Cette hypothèse peut être élargie comme le font BALTAGI, CHANG et LI (1992) qui posent :

$$v_{io} \rightarrow \text{iid} \left(0, \frac{\sigma_\varepsilon^2}{v} \right)$$

où v est un paramètre à estimer.

2.2 NOTATION

Sous forme matricielle le modèle étudié s'écrit :

$$\begin{cases} Y = \delta Y_{-1} + \beta X + \phi(I_N \otimes S_T)Z + U \\ U = (I_N \otimes S_T)\alpha + V \end{cases} \quad (2.1)$$

soit de façon compacte :

$$Y = M \phi + U \quad (2.2)$$

avec

$$M = (Y_{-1} \ X \ Z^*) \quad \phi' = (\delta \ \beta \ \phi) \quad Z^* = (I_N \otimes S_T)Z$$

$$\begin{aligned} \underset{(NTX1)}{Y} &= \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} & \underset{(TX1)}{Y_i} &= \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{iT} \end{bmatrix} & \underset{(NTX1)}{Y_{-1}} &= \begin{bmatrix} Y_{1-1} \\ Y_{2-1} \\ \vdots \\ Y_{N-1} \end{bmatrix} & \underset{(TX1)}{Y_{i-1}} &= \begin{bmatrix} Y_{i0} \\ Y_{i1} \\ \vdots \\ Y_{iT-1} \end{bmatrix} & \underset{(NTX1)}{X} &= \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} & \underset{(TX1)}{X_i} &= \begin{bmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{iT} \end{bmatrix} \\ \underset{(NX1)}{Z} &= \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} & \underset{(NX1)}{\alpha} &= \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} & \underset{(TX1)}{S_T} &= \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} & \underset{(NTX1)}{V} &= \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{bmatrix} & \underset{(TX1)}{v_i} &= \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{iT} \end{bmatrix} \end{aligned}$$

I_N est la matrice unitaire d'ordre N .

La matrice des variances-covariances de ce modèle est définie par :

$$E(UU') = \Omega = I_N \otimes \Sigma \quad (2.3)$$

avec

$$\Sigma = \sigma_\alpha^2 S_T S_T' + \sigma_\varepsilon^2 V_p \quad (2.4)$$

$$V_p = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & \rho & \dots & \rho^{T-1} \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho^{T-1} & \dots & \rho & 1 \end{bmatrix}$$

2.3 TRANSFORMATION DU MODELE

Lorsque le coefficient d'autocorrélation ρ est connu ou estimé de façon convergente dans une première étape nous pouvons transformer le modèle en le multipliant par la matrice de transformation C suivante :

$$C = \begin{bmatrix} (1-\rho^2)^{\frac{1}{2}} & 0 & \dots & 0 \\ -\rho & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & -\rho & 1 \end{bmatrix} \quad (2.5)$$

Le modèle transformé s'écrit :

$$\begin{cases} Y_i^* = \delta Y_{i-1}^* + \beta X_i^* + \phi S_T^* \alpha_i + U_i^* \\ U_i^* = S_T^* \alpha_i + V_i^* \end{cases} \quad (2.6)$$

avec

$$\begin{aligned} Y_i^* &= \left[(1-\rho^2)^{\frac{1}{2}} y_{i1}, y_{i2} - \rho y_{i1}, \dots, y_{iT} - \rho y_{iT-1} \right] & Y_{i-1}^* &= \left[(1-\rho^2)^{\frac{1}{2}} y_{i0}, y_{i1} - \rho y_{i0}, \dots, y_{iT-1} - \rho y_{iT-2} \right] \\ X_i^* &= \left[(1-\rho^2)^{\frac{1}{2}} x_{i1}, x_{i2} - \rho x_{i1}, \dots, x_{iT} - \rho x_{iT-1} \right] & S_T^* &= \left[(1-\rho^2)^{\frac{1}{2}}, 1-\rho, \dots, 1-\rho \right] \\ V_i^* &= \left[(1-\rho^2)^{\frac{1}{2}} v_{i1}, \varepsilon_{i2}, \varepsilon_{i2}, \dots, \varepsilon_{iT} \right] \end{aligned}$$

Le vecteur S_T^* peut encore s'écrire :

$$S_T^* = (1-\rho)S_T^a \quad \text{où } S_T^a = (a, 1, \dots, 1) \quad \text{et } a = \left(\frac{1+\rho}{1-\rho} \right)^{\frac{1}{2}}$$

REMARQUE

La matrice C est aussi qualifiée par certains auteurs de matrice de transformation PRAIS-WINSTEN (PW transformation matrix)

Sachant que :

$$U_i^* = S_T^* \alpha_i + V_i^* = (1-\rho)S_T^a \alpha_i + V_i^* \quad (2.7)$$

$$E(U_i^* U_i^{*\prime}) = (1-\rho)^2 (S_T^a)' \sigma_\alpha^2 + \sigma_\varepsilon^2 I_T \quad (2.8)$$

Posons :

$$J_T^a = \frac{S_T^a S_T^{a\prime}}{d^2} = \frac{1}{d^2} \begin{bmatrix} a^2 & a & \dots & a \\ a & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a & 1 & \dots & 1 \end{bmatrix} \quad \text{où } d^2 = S_T^a S_T^a = a^2 + (T-1) \quad (2.9)$$

D'où

$$\Sigma_i^* = E(U_i^* U_i^{*\prime}) = d^2 (1-\rho)^2 J_T^a \sigma_\alpha^2 + \sigma_\varepsilon^2 I_T \quad (2.10)$$

Nous pouvons aussi écrire :

$$\Sigma_i^* = [d^2 (1-\rho)^2 \sigma_\alpha^2 + \sigma_\varepsilon^2] J_T^a + \sigma_\varepsilon^2 E_T^*$$

soit

$$\Sigma_i^* = \sigma_\alpha^2 J_T^a + \sigma_\varepsilon^2 E_T^a \quad (2.11)$$

avec

$$\sigma_a^2 = d^2(1-\rho)^2 \sigma_\alpha^2 + \sigma_\varepsilon^2$$

$$E_T^a = I_T - J_T^a$$

Pour l'ensemble des individus, nous avons :

$$\Sigma^* = \sigma_a^2 B_a + \sigma_\varepsilon^2 W_a \quad (2.12)$$

où

$$\begin{aligned} B_a &= I_N \otimes J_T^a \\ W_a &= I_N \otimes E_T^a \end{aligned} \quad (2.13)$$

B_a et W_a sont une autre manière de présenter les opérateurs between et within¹, ils ont donc les mêmes propriétés.

Nous les qualifierons d'opérateur a-between et a-within. Ils sont idempotentes d'ordre NT et de rang N et N(T-1) respectivement.

Utilisant ces propriétés, BALTAGI et LI (1991) montrent que :

$$\Sigma^{*p} = (\sigma_a^2)^p B_a + (\sigma_\varepsilon^2)^p W_a \quad (2.14)$$

p est un scalaire quelconque.

$$|\Sigma^*| = (\sigma_a^2)^N (\sigma_\varepsilon^2)^{N(T-1)} \quad (2.15)$$

$|\Sigma^*|$ désigne son déterminant.

Une autre décomposition de cette matrice peut être obtenue de la manière suivante :

$$\Sigma^* = \sigma_\varepsilon^2 [W_a + \theta^{-1} B_a] \quad (2.16)$$

où

$$\theta = \frac{\sigma_\varepsilon^2}{(1-\rho)^2 d^2 \sigma_\alpha^2 + \sigma_\varepsilon^2}$$

III ETUDE DU COMPORTEMENT DES ESTIMATEURS USUELS DU MODELE AR(1) A ERREURS COMPOSEES AVEC AUTOCORRELATION

Nous travaillerons avec le modèle transformé donné en (2.6).

Nous allons ici présenter les propriétés des estimateurs usuels (intra, inter, MCO, MCG, MCQG) d'abord lorsque N et T tendent vers l'infini (comportement asymptotique), ensuite lorsque seule N tend vers l'infini (comportement semi-asymptotique). Nous nous sommes servi de la loi des grands nombres pour calculer les limites en probabilité nécessaires à l'étude du comportement asymptotique des estimateurs. Ces limites sont présentées dans Matondzi (1997). Sous la forme générale, les estimateurs usuels s'écrivent en utilisant la paramétrisation proposée par MADDALA :

$$\hat{\varphi}(\lambda_\rho) - \varphi = \left[\Delta^* \Delta^* + (\lambda_\rho - 1) \Delta^* B_a \Delta^* \right]^{-1} \left[\Delta^* U^* + (\lambda_\rho - 1) \Delta^* B_a U^* \right] \quad (3.1)$$

avec

$$\Delta^* = (Y_{-1}^* \ X^* \ Z^*) \quad \varphi = (\delta \beta \phi)'$$

Dans cette paramétrisation, on retrouve les estimateurs usuels : $\hat{\varphi}(0)$ est l'estimateur intra-individuel ou within qui ne dépend que de la dimension intra des variables, $\hat{\varphi}(\infty)$ est

¹ DORMONT. B donne une bonne présentation des propriétés de ces opérateurs.

l'estimateur inter qui ne dépend que de la dimension inter des variables, $\hat{\phi}(1)$ est l'estimateur des MCO, $\hat{\phi}(\theta)$ est l'estimateur des MCG, $\hat{\phi}(\hat{\theta})$ est l'estimateur des MCQG. Ces trois derniers estimateurs dépendent de la dimension totale des variables.

3.1 COMPORTEMENT ASYMPTOTIQUE DES ESTIMATEURS

L'étude du comportement asymptotique des estimateurs sous la forme générale n'est pas possible pour l'ensemble des estimateurs [voir Matondzi (1997a)]. Nous allons donc partager les estimateurs en deux groupes contenant des estimateurs dont les comportements asymptotiques sont voisins.

3.1.1 LES ESTIMATEURS WITHIN, MCG ET MCQG

Les estimateurs de ce groupe s'écrivent :

$$\begin{aligned} \begin{bmatrix} \hat{\delta}(\lambda_\rho) - \delta \\ \hat{\beta}(\lambda_\rho) - \beta \end{bmatrix} &= \begin{bmatrix} \sum_{i=1}^N \left\{ Y_{i-1}^* \prime Y_{i-1}^* + (\lambda_\rho - 1)d^2 b_{i-1}^{2*} \right\} & \sum_{i=1}^N \left\{ Y_{i-1}^* \prime X_i^* + (\lambda_\rho - 1)d^2 b_i g_i \right\} \\ \sum_{i=1}^N \left\{ X_i^* \prime Y_{i-1}^* + (\lambda_\rho - 1)d^2 g_i b_i \right\} & \sum_{i=1}^N \left\{ X_i^* \prime X_i^* + (\lambda_\rho - 1)d^2 g_i^2 \right\} \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} \sum_{i=1}^N \left\{ Y_{i-1}^* \prime U_i^* + (\lambda_\rho - 1)d^2 b_i c_i \right\} \\ \sum_{i=1}^N \left\{ X_i^* \prime U_i^* + (\lambda_\rho - 1)d^2 g_i c_i \right\} \end{bmatrix} \end{aligned} \quad (3.2)$$

$$\hat{\phi}(\lambda_\rho) = \frac{\sum_{i=1}^N (b_i - \hat{\delta}(\lambda_\rho) b_{i-1} - \hat{\beta}(\lambda_\rho) g_i) z_i}{(1-\rho) \sum_{i=1}^N z_i^2} \quad (3.3)$$

$$\hat{\sigma}_\varepsilon^2(\lambda_\rho) = \frac{\sum_{i=1}^N [\hat{U}_i^* \prime \hat{U}_i^* + (\lambda_\rho - 1) \hat{U}_i^* \prime J_T^* \hat{U}_i^*]}{N(T-1) - 3} \quad (3.4)$$

$$\hat{\sigma}_a^2(\lambda_\rho) = \frac{\sum_{i=1}^N [b_i - \hat{\delta}(\lambda_\rho) b_{i-1} - \hat{\beta}(\lambda_\rho) g_i - \hat{\phi}(\lambda_\rho) h_i]^2}{N-3} \quad (3.5)$$

$$\hat{\sigma}_\alpha^2(\lambda_\rho) = \frac{1}{(1-\rho)^2} \hat{\sigma}_a^2(\lambda_\rho) - \frac{1}{d^2} \hat{\sigma}_\varepsilon^2(\lambda_\rho) \quad (3.6)$$

avec

$$d^2 = a^2 + (T-1) \quad a = \left(\frac{1+\rho}{1-\rho} \right)^{\frac{1}{2}} \quad b_{i-1} = \frac{1}{d^2} \sum_{i=1}^N \left[a y_{i0}^* + \sum_{t=2}^T y_{it-1}^* \right]$$

$$g_i = \frac{1}{d^2} \sum_{i=1}^N \left[a x_{i1}^* + \sum_{t=2}^T x_{it}^* \right] \quad h_i = (1-\rho) z_i \quad y_{i0}^* = (1-\rho^2)^{\frac{1}{2}} y_{i0} \quad y_{it}^* = y_{it} - \rho y_{it-1} \quad \forall t \geq 2$$

$$x_{i1}^* = (1-\rho^2)^{\frac{1}{2}} x_{i1} \quad x_{it}^* = x_{it} - \rho x_{it-1} \quad \forall t \geq 2 \quad u_{i1}^* = (1-\rho^2)^{\frac{1}{2}} u_{i1} \quad u_{it}^* = u_{it} - \rho u_{it-1} \quad \forall t \geq 2$$

Il est clair d'après (3.2) et (3.3) que les estimateurs ce groupe sont asymptotiquement convergents si :

$$P \lim_{N,T \rightarrow \infty} \frac{1}{NT} \begin{bmatrix} \sum_{i=1}^N \left\{ Y_{i-1}^* U_i^* + (\lambda_p - 1) d^2 b_{i-1} c_i \right\} \\ \sum_{i=1}^N \left\{ X_i^* U_i^* + (\lambda_p - 1) d^2 g_i c_i \right\} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (3.7)$$

$$P \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N c_i h_i = 0 \quad (3.8)$$

Nous avons vu précédemment que les estimateurs within, MCG et MCQG correspondent aux valeurs suivantes du paramètre λ_p : 0, θ et $\hat{\theta}$.

Les estimateurs de ce groupe asymptotiquement équivalents puisque $P \lim_{N,T \rightarrow \infty} \theta = P \lim_{N,T \rightarrow \infty} \hat{\theta} = 0$.

D'autre part, dans Matondzi (1997b), nous démontrons que :

$$\begin{aligned} P \lim_{N,T \rightarrow \infty} \frac{1}{Nd^2} \sum_{i=1}^N Y_{i-1}^* U_i^* &= \frac{(1-\rho)^2}{1-\delta} \sigma_\alpha^2 & P \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N b_{i-1} c_i &= \frac{(1-\rho)^2}{1-\delta} \sigma_\alpha^2 \\ P \lim_{N,T \rightarrow \infty} \frac{1}{Nd^2} \sum_{i=1}^N U_i^* U_i^* &= (1-\rho)^2 \sigma_\alpha^2 + \sigma_\varepsilon^2 & P \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N c_i^2 &= (1-\rho)^2 \sigma_\alpha^2 \\ P \lim_{N,T \rightarrow \infty} \frac{1}{Nd^2} \sum_{i=1}^N X_i^* U_i^* &= P \lim_{N,T \rightarrow \infty} \frac{1}{Nd^2} \sum_{i=1}^N z_i S_T^* U_i^* = 0 & P \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N g_i c_i &= P \lim_{N,T \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N h_i c_i = 0 \end{aligned}$$

Les estimateurs de ce groupe asymptotiquement équivalents sont convergents, sauf pour l'estimateur de $\hat{\sigma}_\alpha^2(\lambda_p)$ pour lequel nous avons :

$$P \lim_{N,T \rightarrow \infty} \hat{\sigma}_\alpha^2(\lambda_p) = (1-\rho)^2 \sigma_\alpha^2 \quad (3.9)$$

REMARQUE

$\sigma_\alpha^2 = d^2 (1-\rho)^2 \sigma_\alpha^2 + \sigma_\varepsilon^2$ où $d^2 = a^2 + (T-1)$. Lorsque $T \rightarrow \infty$, $\sigma_\alpha^2 \rightarrow (1-\rho)^2 \sigma_\alpha^2$.

Nous pouvons donc considérer que $\hat{\sigma}_\alpha^2(\lambda_p)$ est un estimateur convergent de σ_α^2 .

3.1.2 LES ESTIMATEURS DES MCO ET BETWEEN

Ils sont tous les deux asymptotiquement biaisés. Ces biais sont dues d'une part à la corrélation de l'effet individuel avec la variable endogène et d'autre part à l'autocorrélation des erreurs. Une analyse numérique de ces biais en fonction des paramètres de base du modèle est faite dans Matondzi (1997b). Les résultats obtenus montrent que c'est l'estimateur des MCO qui a les biais les plus faibles.

En ce qui concerne l'estimateur des MCO, nous avons montré que :

- Le biais asymptotique de $\hat{\delta}_{MCO}$ est positif. Cet estimateur surestime la vraie valeur de ce paramètre. Toute chose égale par ailleurs, ce biais est une fonction croissante de σ_α^2 et décroissante de $\beta, \sigma_\varepsilon^2$ et σ_x^2 . Ce biais, en général faible surtout pour une valeur positivement élevée de ρ , est une fonction décroissante de ce paramètre;

- L'estimateur de β a un biais asymptotique dont le signe peut être positif ou négatif en fonction de β et ρ . Il est positif si β positif et ρ négatif ou l'inverse, et il est négatif lorsque β et ρ sont tous deux positifs ou négatifs. Ce biais est en général faible, mais nous observons cependant que cette faiblesse est plus marquée pour δ élevé en valeur absolue que pour des valeurs moyennes de ce paramètre. Ce biais en valeur absolue décroît avec ρ .

- Le biais de l'estimateur de ϕ , il est du signe contraire de ce paramètre. Il est faible et décroît en valeur absolue avec ρ .

Les biais des estimateurs $\hat{\delta}_B$, $\hat{\beta}_B$ et $\hat{\phi}_B$ sont identiques à ceux obtenus lorsque le modèle ne comporte pas d'autocorrélation. [cf. SEVESTRE.P (1983)].

Ainsi, le biais de $\hat{\delta}_B$ est positif et décroît avec δ . Quant aux biais de $\hat{\beta}_B$ et $\hat{\phi}_B$, ils sont du signe contraire de ces paramètres et croissent en valeur absolue avec ces paramètres.

Les estimateurs de σ_ε^2 et σ_α^2 sont également biaisés, ces biais sont la conséquence directe des biais des paramètres de base.

Nous pouvons au terme de cette analyse asymptotique des estimateurs usuels établir le classement suivant (les chiffres dans le tableau désignent le rang de l'estimateur) :

TABLEAU 3-1 CLASSEMENT ASYMPTOTIQUE DES ESTIMATEURS

METHODE	δ	β	ϕ
WITHIN	1	1	1
MCG	1	1	1
MCQG	1	1	1
MCO	4	4	4
BETWEEN	5	5	5

Ce classement est identique à celui obtenu par SEVESTRE et TROGNON (1983 et 1985) et montre que l'estimateur between est celui qui a le biais asymptotique le plus élevé dans cette classe d'estimateurs.

3.2 COMPORTEMENT SEMI-ASYMPTOTIQUE

Sous le semi-asymptotique, tous les estimateurs sont asymptotiquement biaisés et ceci quelle que soit l'hypothèse faite sur les observations initiales. Ce résultat est différent de celui obtenu lorsque l'autocorrélation est absente du modèle. En effet, dans ce dernier cas, l'estimateur des MCG est convergent sous les hypothèses où les observations initiales sont fixes ou indépendantes des erreurs du modèle.

La complexité des biais asymptotiques ne permet pas une étude analytique du comportement de ces biais en fonction des divers paramètres. Nous avons procédé dans Matondzi(1997b) à une étude numérique de ces biais et nous avons remarqué que les valeurs des biais sont fortement influencées par les hypothèses sur les observations initiales et la structure de la variable exogène x_{it} . Les biais sont plus faibles lorsque cette dernière variable est AR(1) que lorsqu'elle est fixe. Le classement établi sous les différentes hypothèses montre que c'est l'estimateur des MCG qui le meilleur estimateur à faible biais.

IV LA METHODE DU MAXIMUM DE VRAISEMBLANCE

La première méthode qui vient à l'esprit pour trouver des estimateurs convergents même sous le semi-asymptotique est celle du maximum de vraisemblance. Ici, nous allons travailler sur le modèle sans transformation.

Quand les erreurs sont normalement distribuées, la fonction de densité associée au modèle (2.1), conditionnellement à la situation initiale y_{i0} , s'écrit pour un individu :

$$f_i(y_{i1}, y_{i2}, \dots, y_{iT} / y_{i0}) = (2\pi)^{-\frac{T}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(Y_i - M_i\phi)' \Sigma_i^{-1} (Y_i - M_i\phi)\right\} \quad (4.1)$$

où Σ_i est la matrice des variances-covariances pour un individu et

$$M_i = (Y_{i1}, X_i, S_T z_i) \quad \varphi = (\delta \quad \beta \quad \phi)'$$

Les individus étant indépendants, la log-vraisemblance conditionnelle associée au modèle pour l'ensemble des individus s'écrit :

$$\log L_c = \sum_{i=1}^N \log f_i(y_{i1}, y_{i2}, \dots, y_{iT} / y_{i0}) \quad (4.2)$$

La log-vraisemblance non conditionnelle (la vraie log-vraisemblance du modèle) doit tenir compte de la densité marginale des observations initiales y_{i0} , soit $f_i(y_{i0})$. Cette log-vraisemblance s'écrit :

$$\log L = \sum_{i=1}^N \log [f_i(y_{i1}, y_{i2}, \dots, y_{iT} / y_{i0}) f_i(y_{i0})] \quad (4.3)$$

On voit bien que cette log-vraisemblance dépend fortement des hypothèses faites sur les observations initiales. La connaissance du statut de ces observations est donc capitale. Comme nous l'avons vu précédemment, plusieurs hypothèses peuvent être faites sur les observations initiales. Chaque hypothèse sur les observations initiales présente une difficulté particulière dans l'élaboration de la fonction de densité jointe.

Nous allons ici donner une illustration de cette difficulté en considérant deux cas simple : le premier tiré de la première famille postule l'indépendance de y_{i0} par rapport au modèle, et le deuxième, réfute cette indépendance et constitue une généralisation de plusieurs hypothèses de la deuxième famille.

L'estimation des paramètres par la méthode du maximum de vraisemblance nécessite des calculs extrêmement complexes. La lourdeur des expressions en cause ne nous permet pas de les présenter ici ceci pour rendre le texte plus clair à la lecture. Toutefois le lecteur intéressé peut se référer à Matondzi (1997b) pour avoir une présentation exhaustive de cette méthode.

4.1 Maximum de vraisemblance : cas 1

Ce cas suppose que les observations initiales sont des constantes fixes. La première observation n'influence donc pas la fonction de densité.

La log-vraisemblance doit être une fonction des paramètres. Ces paramètres doivent être fixes, indépendants de N et T.

Pour le modèle (2.1) les paramètres sont :

- les coefficients δ , β et ϕ ;
- les deux variances σ_α^2 et σ_ϵ^2 ;
- le coefficient d'autocorrélation ρ .

La log-vraisemblance s'écrit alors :

$$\begin{aligned} L = \text{cste} - \frac{N(T-1)}{2} \log \sigma_\epsilon^2 - \frac{N}{2} \log [(1-\rho)^2 d^2 \sigma_\alpha^2 + \sigma_\epsilon^2] - \frac{1}{2\sigma_\epsilon^2} \left\{ (Y^* - \tilde{\Delta}^* \psi - \phi Z^*)' W_\alpha \right. \\ \left. \times (Y^* - \tilde{\Delta}^* \psi - \phi Z^*) \right\} - \frac{1}{2[(1-\rho)^2 d^2 \sigma_\alpha^2 + \sigma_\epsilon^2]} \left\{ (Y^* - \tilde{\Delta}^* \psi - \phi Z^*)' B_\alpha (Y^* - \tilde{\Delta}^* \psi - \phi Z^*) \right\} \end{aligned} \quad (4.4)$$

Posons :

$$\begin{aligned} U^* &= Y^* - \tilde{\Delta}^* \psi - \phi Z^* \\ \lambda_1 &= \sigma_\epsilon^2 \\ \lambda_2 &= \sigma_\epsilon^2 + (1-\rho)^2 d^2 \sigma_\alpha^2 \end{aligned}$$

La log-vraisemblance (4.4) peut aussi s'écrire :

$$l(\psi, \phi, \sigma_\alpha^2, \sigma_\varepsilon^2, \rho) = \text{cste} - \frac{N(T-1)}{2} \log \lambda_1 - \frac{N}{2} \log \lambda_2 - \frac{1}{2\lambda_1} U'^* W_a U^* - \frac{1}{2\lambda_2} U'^* B_a U^* \quad (4.5)$$

Cette log-vraisemblance est toujours fonction des vrais paramètres.

4.1.1 Les conditions de premier ordre

$$\frac{\partial L}{\partial \psi} = \frac{1}{\lambda_1} \left\{ \tilde{\Delta}' W_a (Y^* - \tilde{\Delta} \psi - \phi Z^*) \right\} + \frac{1}{\lambda_2} \left\{ \tilde{\Delta}' B_a (Y^* - \tilde{\Delta} \psi - \phi Z^*) \right\} \quad (4.6)$$

$$\frac{\partial L}{\partial \phi} = \frac{1}{\lambda_1} \left\{ Z'^* W_a (Y^* - \tilde{\Delta} \psi - \phi Z^*) \right\} + \frac{1}{\lambda_2} \left\{ Z'^* B_a (Y^* - \tilde{\Delta} \psi - \phi Z^*) \right\} \quad (4.7)$$

Pour le calcul des dérivées par rapport aux variances, notons :

$$\frac{\partial \lambda_1}{\partial \sigma_\varepsilon^2} = 1 \quad \frac{\partial \lambda_2}{\partial \sigma_\varepsilon^2} = 1 \quad \frac{\partial \lambda_1}{\partial \sigma_\alpha^2} = 1 \quad \frac{\partial \lambda_2}{\partial \sigma_\alpha^2} = (1-\rho)^2 d^2$$

Nous avons donc :

$$\frac{\partial L}{\partial \sigma_\varepsilon^2} = -\frac{N(T-1)}{2} \frac{1}{\lambda_1} - \frac{N}{2} \frac{1}{\lambda_2} + \frac{1}{2\lambda_1^2} U'^* W_a U^* + \frac{1}{2\lambda_2^2} U'^* B_a U^* \quad (4.8)$$

$$\frac{\partial L}{\partial \sigma_\alpha^2} = -\frac{N(1-\rho)^2 d^2}{2\lambda_2} + \frac{(1-\rho)^2 d^2}{2\lambda_2^2} U'^* B_a U^* \quad (4.9)$$

La résolution des équations normales obtenues à partir des conditions de premier ordre nous donne les estimateurs suivants :

$$\hat{\psi} = \left\{ \frac{1}{\hat{\lambda}_1} \tilde{\Delta}' W_a \tilde{\Delta}_a + \frac{1}{\hat{\lambda}_2} \tilde{\Delta}' B_a \tilde{\Delta}_a \right\}^{-1} \left\{ \tilde{\Delta}' \left(\frac{1}{\hat{\lambda}_1} W_a + \frac{1}{\hat{\lambda}_2} B_a \right) (Y^* - \hat{\phi} Z^*) \right\} \quad (4.10)$$

$$\hat{\phi} = \left\{ \frac{1}{\hat{\lambda}_1} Z'^* W_a Z^* + \frac{1}{\hat{\lambda}_2} Z'^* B_a Z^* \right\}^{-1} \left\{ Z'^* \left(\frac{1}{\hat{\lambda}_1} W_a + \frac{1}{\hat{\lambda}_2} B_a \right) (Y^* - \tilde{\Delta} \hat{\psi}) \right\} \quad (4.11)$$

Pour avoir les estimateurs de λ_1 et λ_2 , il suffit de partir de (4.9) en remarquant que :

$$\frac{\partial L}{\partial \sigma_\alpha^2} = 0 \quad \Rightarrow \quad \frac{N}{2\lambda_2^2} = \frac{1}{2\lambda_2^2} U'^* B_a U^* \quad (4.12)$$

et en utilisant ce résultat dans (4.8), il vient :

$$\hat{\lambda}_1 = \frac{1}{N(T-1)} (Y^* - \tilde{\Delta} \hat{\psi} - \hat{\phi} Z^*)' W_a (Y^* - \tilde{\Delta} \hat{\psi} - \hat{\phi} Z^*) \quad (4.13)$$

et de (4.9) nous avons :

$$\hat{\lambda}_2 = \frac{1}{N} (Y^* - \tilde{\Delta} \hat{\psi} - \hat{\phi} Z^*)' B_a (Y^* - \tilde{\Delta} \hat{\psi} - \hat{\phi} Z^*) \quad (4.14)$$

REMARQUES

- 1°) Les estimateurs de σ_α^2 et σ_ε^2 s'obtiennent facilement par identification de (4.13) et (4.14).
 2°) Les estimateurs obtenus sont interdépendants. Le système formé par les équations de vraisemblance n'admet donc pas de solution analytique explicite. Seul le recours à une procédure itérative permet d'étudier leurs comportements asymptotiques.
 3°) Pour trouver la solution numérique, puisqu'elle est calculée pour N et T fixes (donnés dans l'échantillon), il est plus commode de travailler avec λ_1 et λ_2 .
 3°) L'estimateur du maximum de vraisemblance est solution de (4.10), (4.11), (4.13) et (4.14). Il s'agit d'un système non linéaire.
 Un algorithme de solution en zigzag peut être obtenu comme suit :

- Première étape : on commence par une estimation convergente des paramètres δ , β et ϕ qui ne demande pas la connaissance de λ_1 et λ_2 . On pourrait par exemple utiliser les estimateurs within ou les estimateurs des variables instrumentales simples qui seront à la section V. La précision des « estimateurs de début » dans cette procédure est importante, c'est pour cette raison que nous conseillons de commencer la procédure en utilisant les estimateurs des variables instrumentales.

Soient $\tilde{\delta}$, $\tilde{\beta}$ et $\tilde{\phi}$ les estimateurs obtenus par cette méthode.

- Deuxième étape : on calcule $U^* = Y^* - \tilde{\Delta}^* \tilde{\psi} - Z^* \tilde{\phi}$ et on estime λ_1 et λ_2 en utilisant (4.13) et (4.14). Dans ces expressions, les paramètres inconnus sont remplacés par leurs estimations obtenues dans la première étape.

- Troisième étape : retour à la première, on calcule de nouveau les paramètres δ , β et ϕ en utilisant cette fois les expressions (4.11) et (4.12), puis on retourne à la deuxième étape.

- On itère le processus jusqu'à la convergence numérique.

Dans les développements nous avons implicitement supposé que le coefficient d'autocorrélation ρ était connu ce qui n'est toujours pas le cas. Malheureusement, la dérivée de la log-vraisemblance par rapport à ρ nous donne un polynôme de degré 9. La complexité de l'expression du polynôme ne nous a pas permis de faire une étude comparable à celle de KOOPMANS.T(1942).

Nous suggérons donc lorsque ρ est inconnu de faire un balayage sur l'intervalle $]-1,1[$ pour achever l'estimation du modèle par la méthode du maximum de vraisemblance.

4.1.2 Les conditions de deuxième ordre

$$\frac{\partial^2 L}{\partial \psi \partial \psi'} = -\frac{1}{\lambda_1} \tilde{\Delta}^*{}' W_\alpha \tilde{\Delta}^* - \frac{1}{\lambda_2} \tilde{\Delta}^*{}' B_\alpha \tilde{\Delta}^* \quad (4.15)$$

$$E \left[-\frac{\partial^2 L}{\partial \psi \partial \psi'} \right] = \frac{1}{\lambda_1} \tilde{\Delta}^*{}' W_\alpha \tilde{\Delta}^* + \frac{1}{\lambda_2} \tilde{\Delta}^*{}' B_\alpha \tilde{\Delta}^* \quad (4.16)$$

$$\frac{\partial^2 L}{\partial \psi \partial \phi} = -\frac{1}{\lambda_1} \tilde{\Delta}^*{}' W_\alpha Z^* - \frac{1}{\lambda_2} \tilde{\Delta}^*{}' B_\alpha Z^* = -\frac{1}{\lambda_2} \tilde{\Delta}^*{}' B_\alpha Z^* \quad (4.17)$$

car $W_\alpha Z^* = 0$

$$E \left[-\frac{\partial^2 L}{\partial \psi \partial \phi} \right] = \frac{1}{\lambda_2} \tilde{\Delta}^*{}' B_\alpha Z^* \quad (4.18)$$

$$\frac{\partial^2 L}{\partial \psi \partial \sigma_\epsilon^2} = -\frac{1}{\lambda_1^2} \tilde{\Delta}' W_a U^* - \frac{1}{\lambda_2^2} \tilde{\Delta}' B_a U^* \quad (4.19)$$

$$E \left[-\frac{\partial^2 L}{\partial \psi \partial \sigma_\epsilon^2} \right] = \frac{1}{\lambda_1^2} \tilde{\Delta}' W_a U^* + \frac{1}{\lambda_2^2} \tilde{\Delta}' B_a U^* \quad (4.20)$$

$$\frac{\partial^2 L}{\partial \psi \partial \sigma_\alpha^2} = -\frac{(1-\rho)^2 d^2}{\lambda_2^2} \tilde{\Delta}' B_a U^* \quad (4.21)$$

$$E \left[-\frac{\partial^2 L}{\partial \psi \partial \sigma_\alpha^2} \right] = \frac{(1-\rho)^2 d^2}{\lambda_2^2} \tilde{\Delta}' B_a U^* \quad (4.22)$$

$$\frac{\partial^2 L}{\partial \phi \partial \phi} = -\frac{1}{\lambda_1} Z' W_a Z^* - \frac{Z' B_a Z^*}{\lambda_2} = -\frac{Z' B_a Z^*}{\lambda_2} \quad (4.23)$$

$$E \left[-\frac{\partial^2 L}{\partial \phi \partial \phi} \right] = \frac{Z' B_a Z^*}{\lambda_2} \quad (4.24)$$

$$\frac{\partial^2 L}{\partial \phi \partial \sigma_\epsilon^2} = -\frac{1}{\lambda_1^2} Z' W_a U^* - \frac{1}{\lambda_2^2} Z' B_a U^* \quad (4.25)$$

$$E \left[-\frac{\partial^2 L}{\partial \phi \partial \sigma_\epsilon^2} \right] = 0 \quad (4.26)$$

$$\frac{\partial^2 L}{\partial \phi \partial \sigma_\alpha^2} = -\frac{(1-\rho)^2 d^2}{\lambda_2^2} Z' B_a U^* \quad (4.27)$$

$$E \left[-\frac{\partial^2 L}{\partial \phi \partial \sigma_\alpha^2} \right] = 0 \quad (4.28)$$

$$\frac{\partial^2 L}{\partial \sigma_\epsilon^2 \partial \sigma_\epsilon^2} = \frac{N(T-1)}{2\lambda_1^2} + \frac{N}{2\lambda_2^2} - \frac{1}{\lambda_1^3} U' W_a U^* - \frac{1}{\lambda_2^3} U' B_a U^* \quad (4.29)$$

$$E \left[-\frac{\partial^2 L}{\partial \sigma_\epsilon^2 \partial \sigma_\epsilon^2} \right] = -\frac{N(T-1)}{2\lambda_1^2} - \frac{N}{2\lambda_2^2} + \frac{1}{\lambda_1^3} U' W_a U^* + \frac{1}{\lambda_2^3} U' B_a U^* \quad (4.30)$$

$$\frac{\partial^2 L}{\partial \sigma_\epsilon^2 \partial \sigma_\alpha^2} = \frac{N(1-\rho)^2 d^2}{2\lambda_2^2} - \frac{(1-\rho)^2 d^2}{\lambda_2^3} U' B_a U^* \quad (4.31)$$

$$E \left[-\frac{\partial^2 L}{\partial \sigma_\epsilon^2 \partial \sigma_\alpha^2} \right] = -\frac{N(1-\rho)^2 d^2}{2\lambda_2^2} + \frac{(1-\rho)^2 d^2}{\lambda_2^3} U' B_a U^* \quad (4.32)$$

$$\frac{\partial^2 L}{\partial \sigma_\alpha^2 \partial \sigma_\alpha^2} = \frac{N(1-\rho)^4 d^4}{2\lambda_2^2} - \frac{(1-\rho)^4 d^4}{\lambda_2^3} U' B_a U^* \quad (4.33)$$

$$E\left[-\frac{\partial^2 L}{\partial \sigma_\alpha^2 \partial \sigma_\alpha^2}\right] = -\frac{N(1-\rho)^4 d^4}{2\lambda_2^2} + \frac{(1-\rho)^4 d^4}{\lambda_2^3} U^*{}' B_\alpha U^* \quad (4.44)$$

Pour étudier les distributions asymptotiques des estimateurs, nous devons tenir compte des variabilités de chaque paramètre.

Posons :

$$\Phi = \begin{bmatrix} \psi \\ \phi \\ \sigma_\varepsilon^2 \\ \sigma_\alpha^2 \end{bmatrix} \quad D = \begin{bmatrix} \sqrt{Nd^2} I_2 & 0 & 0 & 0 \\ 0 & \sqrt{N} & 0 & 0 \\ 0 & 0 & \sqrt{Nd^2} & 0 \\ 0 & 0 & 0 & \sqrt{N} \end{bmatrix} \quad \Gamma = E\left[-\frac{\partial^2 L}{\partial \Phi \partial \Phi'}\right]$$

La limite en probabilité suivante $\text{Plim}_{N,T \rightarrow \infty} D^{-1} \Gamma D^{-1}$ nous permet d'avoir la matrice des dérivées secondes ou marice d'information suivante :

$$\Theta = \begin{bmatrix} a_1 & a_2 & 0 & 0 & 0 \\ a_2 & a_3 & 0 & 0 & 0 \\ 0 & 0 & a_4 & 0 & 0 \\ 0 & 0 & 0 & a_5 & 0 \\ 0 & 0 & 0 & 0 & a_6 \end{bmatrix} \quad (4.45)$$

avec

$$\frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} \frac{\sigma_\varepsilon^2}{1-\delta^2} + \frac{\beta^2}{1-\delta^2} (1-2\delta\rho + \rho^2) \sigma_x^2 & \rho\beta\sigma_x^2 \\ \rho\beta\sigma_x^2 & (1+\rho^2)\sigma_x^2 \end{bmatrix} = \frac{1}{\sigma_\varepsilon^2} H = \begin{bmatrix} a_1 & a_2 \\ a_2 & a_3 \end{bmatrix}$$

$$\frac{\sigma_z^2}{\sigma_\alpha^2} = a_4 \quad \frac{1}{2(\sigma_\varepsilon^2)^2} = a_5 \quad \frac{1}{2(\sigma_\alpha^2)^2} = a_6$$

La matrice d'information étant bloc diagonale, nous avons donc les distributions asymptotiques suivantes :

$$\sqrt{Nd^2} (\hat{\psi} - \psi) \rightarrow N(0, \sigma_\varepsilon^2 H^{-1}) \quad (4.46)$$

$$\sqrt{N} (\hat{\phi} - \phi) \rightarrow N\left(0, \frac{\sigma_\alpha^2}{\sigma_z^2}\right) \quad (4.47)$$

Nous constatons à ce niveau que les estimateurs within, MCG et MCQG de ces paramètres asymptotiquement équivalents, sont efficaces car ils ont la même distribution asymptotiques que les estimateurs du maximum de vraisemblance.

$$\sqrt{Nd^2} (\hat{\sigma}_\varepsilon^2 - \sigma_\varepsilon^2) \rightarrow N(0, 2(\sigma_\varepsilon^2)^2) \quad (4.48)$$

$$\sqrt{N} (\hat{\sigma}_\alpha^2 - \sigma_\alpha^2) \rightarrow N(0, 2(\sigma_\alpha^2)^2) \quad (4.49)$$

Pour l'ensemble des paramètres, nous avons :

$$D(\hat{\Phi}_{MV} - \Phi) \rightarrow N[0; \Theta^{-1}] \quad (4.50)$$

La borne de Cramer Rao est l'inverse de la matrice d'information donc les estimateurs obtenus sont efficaces.

Sous le semi-asymptotique, les vitesses de convergence sont les mêmes pour tous les estimateurs. Rechercher cette distribution revient à calculer la limite en probabilité suivante :

$\text{Plim}_{N \rightarrow \infty} \frac{1}{N} E \left[-\frac{\partial^2 L}{\partial \Phi \partial \Phi'} \right]$. Cette matrice des limites en probabilité des dérivées secondes n'est pas bloc diagonale.

Ainsi, nous avons la distribution semi-asymptotique suivante :

$$\sqrt{N}(\hat{\Phi}_{MV} - \Phi) \rightarrow N \left\{ 0; \left[\text{Plim}_{N \rightarrow \infty} \frac{1}{N} E \left(-\frac{\partial^2 L}{\partial \Phi \partial \Phi'} \right) \right]^{-1} \right\} \quad (4.51)$$

4.2 Maximum de vraisemblance : cas 2

Il s'agit ici d'utiliser l'hypothèse suivante sur les observations initiales :

$$y_{i0} = k_1 \alpha_i + v_{i0} \quad (4.52)$$

Le modèle s'écrit pour un individu :

$$\begin{cases} y_{i0} = k_1 \alpha_i + k_2 v_{i0} \\ y_{i1} = \delta y_{i0} + \beta x_{i1} + \phi z_i + \alpha_i + v_{i1} \\ \vdots \\ y_{iT} = \delta y_{iT-1} + \beta x_{iT} + \phi z_i + \alpha_i + v_{iT} \end{cases} \Leftrightarrow \begin{cases} y_{i0} = \kappa_i \\ y_i = \delta y_{i-1} + \beta x_i + S_T z_i + S_T \alpha_i + v_i \end{cases} \quad (4.53)$$

La matrice des variances-covariances pour un individu s'écrit :

$$\Omega_{(T+1, T+1)} = E(U_i U_i') = \begin{bmatrix} k_1^2 \sigma_\alpha^2 + k_2^2 \frac{\sigma_\varepsilon^2}{1-\rho^2} & k_1 S_T' \sigma_\alpha^2 + k_2 \frac{\sigma_\varepsilon^2}{1-\rho^2} [\rho \ \rho^2 \ \dots \ \rho^T] \\ k_1 S_T \sigma_\alpha^2 + k_2 \frac{\sigma_\varepsilon^2}{1-\rho^2} \begin{bmatrix} \rho \\ \rho^2 \\ \vdots \\ \rho^T \end{bmatrix} & \sigma_\alpha^2 S_T S_T' + \sigma_\varepsilon^2 \frac{1}{1-\rho^2} \tilde{v}_i \end{bmatrix} \quad (4.54)$$

avec

$$S_T' = [1 \ \dots \ 1] \quad \tilde{v}_i = \begin{bmatrix} 1 & \rho & \dots & \rho^{T-1} \\ \rho & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho^{T-1} & \dots & \rho & 1 \end{bmatrix}$$

Où $\Lambda = \left\{ [(a+T) + (T-2)\rho^2 - 2(T-1)\rho] k_2^2 + [k_1 - \rho k_2]^2 \right\}$

$$G' = \left[\frac{k_1}{k_2} - \rho \frac{1}{k_2}, 1 - \rho \left(1 + \frac{k_1}{k_2} \right) + \rho^2, (1-\rho)^2, \dots, (1-\rho)^2, 1 - \rho \right]$$

Sous ce cas, la log-vraisemblance s'écrit :

$$\begin{aligned}
L = & c - \frac{N(T-1)}{2} \text{Log} \sigma_\alpha^2 + \frac{N}{2} \text{Log}(1-\rho^2) - \frac{N}{2} \text{Log} \left\{ [(a+T) + (T-2)\rho^2 - 2(T-1)\rho] k_2^2 \right. \\
& + [k_1 - \rho k_2]^2 \left. \right\} - \frac{1}{2a\sigma_\alpha^2} \sum_{i=1}^N \left\{ [y_{i1} - \delta y_{i0} - \beta x_{i1} - \phi z_i]^2 + \sum_{t=2}^T [(y_{it} - \delta y_{it-1} - \beta x_{it} - \phi z_i) \right. \\
& - (y_{it-1} - \delta y_{it-2} - \beta x_{it-1} - \phi z_i)]^2 - \frac{1}{\Lambda} \left(k_2^2 \left[(1-\rho) \sum_{t=1}^T (y_{it} - \delta y_{it-1} - \beta x_{it} - \phi z_i) + \rho \left(\rho - \frac{1}{k_2} \right) (y_{i1} \right. \right. \\
& - \delta y_{i0} - \beta x_{i1} - \phi z_i) - \rho(1-\rho) \sum_{t=2}^T (y_{it} - \delta y_{it-1} - \beta x_{it} - \phi z_i) \left. \right]^2 + 2(k_1 - \rho k_2) y_{i0} \left[(1-\rho) \sum_{t=1}^T (y_{it} \right. \\
& - \delta y_{it-1} - \beta x_{it} - \phi z_i) - \rho(1-\rho) \sum_{t=2}^T (y_{it} - \delta y_{it-1} - \beta x_{it} - \phi z_i) \left. \right] + 2 \frac{\rho}{k_2} \left[(a+T) + (T-2)\rho^2 \right. \\
& - 2(T-1)\rho] k_2^2 + (k_1 - \rho k_2)(k_1 - 1) \left. \right\} (y_{i1} - \delta y_{i0} - \beta x_{i1} - \phi z_i) y_{i0} - [(a+T) - 2(T-1)\rho \\
& + (T-2)\rho^2] y_{i0}^2 \left. \right\} \quad (4.55)
\end{aligned}$$

Les expressions des estimateurs obtenus en résolvant les équations normales sont complexes. En égalant à zéro les autres équations de vraisemblance, nous avons des polynômes de degré supérieur à deux.

Ainsi :

$$\frac{\partial L}{\partial k_1} = B_0 + B_1 k_1 + B_2 k_1^2 + B_3 k_1^3 \quad (4.56)$$

$$\frac{\partial L}{\partial k_2} = A_0 + A_1 k_2 + A_2 k_2^2 + A_3 k_2^3 \quad (4.57)$$

De même $\frac{\partial L}{\partial \rho} = 0$ nous donne un polynôme de degré 7 .

Les estimateurs obtenus sont tous interdépendants. Il est donc souhaitable de résoudre le système formé par les équations de vraisemblance de façon itérative.

La complexité des expressions des estimateurs ne nous a pas permis de procéder à une étude comparable à celle présentée au cas 1. Toutefois, les conditions de régularité nous garantissent de leur existence et de leurs distributions asymptotique et semi-asymptotique.

La méthode du maximum de vraisemblance que nous venons de présenter est d'une application assez difficile surtout lorsque les hypothèses de la deuxième famille sont utilisées. Dans ce qui va suivre, nous allons présenter la méthode des variables qui est d'une application plus simple et permet d'obtenir des estimateurs convergents aussi bien lorsque N et T tendent vers l'infini que lorsque N seul tend vers l'infini et ceci quel que soit le statut des observations initiales.

V LA METHODE DES VARIABLES INSTRUMENTALES

Supposons à présent que nous disposons d'une matrice R d'ordre (NTxH) pouvant éventuellement contenir des variables figurant dans le modèle de départ. La matrice R est dite matrice d'instruments si :

$$1^\circ) \quad E(R'U) = 0 \quad \text{ou} \quad P \lim \frac{1}{NT} R'U = 0 \quad (5.1)$$

$$2^\circ) \quad P \lim \frac{1}{NT} R'R = \Omega_{RR} \quad \text{matrice non singulière définie positive} \quad (5.2)$$

$$3^\circ) \quad P \lim \frac{1}{NT} R'M = \Omega_{RM} \quad \text{matrice de rang complet} \quad (5.3)$$

Si $H < 3$, le nombre d'instruments est inférieur au nombre de variables explicatives, le système comporte trop peu d'équations et la démarche n'aboutit pas.

Si $H = 3$, la procédure conduit à une solution unique en prenant la limite en probabilité du modèle transformé.

Le modèle transformé par la matrice d'instruments s'écrit :

$$R'Y = R'M\phi + R'U \quad (5.4)$$

et l'estimateur des variables instrumentales simple est donné :

$$\hat{\phi}_{VI} = (R'M)^{-1} R'Y \quad (5.5)$$

Cet estimateur est convergent et suit asymptotiquement une loi normale.

$$\sqrt{N}(\hat{\phi}_{VI} - \phi) \xrightarrow{N, T \rightarrow \infty} N \left[0, P \lim \left(\frac{R'M}{NT} \right)^{-1} \frac{R'\Omega R}{NT} \left(\frac{M'R}{NT} \right)^{-1} \right] \quad (5.6)$$

Si $H > 3$, la procédure conduit à un système comportant plus d'équations que d'inconnues.

Dans cette situation, deux estimateurs convergents peuvent être obtenus.

Le premier qui est obtenu en appliquant les MCG sur le modèle transformé est défini par :

$$\hat{\phi}_{GIV} = [M'R(R'\Omega R)^{-1}R'M]^{-1} M'R(R'\Omega R)^{-1} R'Y \quad (5.7)$$

et sa variance asymptotique est donné par, lorsque Net T tendent vers l'infini :

$$Vas(\hat{\phi}_{GIV}) = \left[P \lim \frac{1}{NT} M'R(R'\Omega R)^{-1} R'M \right]^{-1} \quad (5.8)$$

Le deuxième est celui proposé par WHITE.H (1984) qui est obtenu en transformant le modèle

par $\Omega^{-\frac{1}{2}}$ pour le rendre sphérique et en utilisant $R^* = \Omega^{-\frac{1}{2}}R$ comme matrice d'instruments.

Les composantes de R^* sont des instruments puisqu'elles sont des combinaisons linéaires des instruments. Ainsi, si la condition d'orthogonalité suivante est vérifiée :

$$P \lim \frac{1}{NT} R^* U = P \lim \frac{1}{NT} R' \Omega^{-\frac{1}{2}} U = 0 \quad (5.9)$$

l'estimateur proposé par WHITE est défini par :

$$\hat{\phi}_{\text{WGIV}} = \left[M' \Omega^{-1} R (R' \Omega^{-1} R)^{-1} R' \Omega^{-1} M \right]^{-1} M' \Omega^{-1} R (R' \Omega^{-1} R)^{-1} R' \Omega^{-1} Y \quad (5.10)$$

soit

$$\hat{\phi}_{\text{WGMM}} = \left[\sum_{i=1}^N M_i' \Sigma^{-1} R_i \left(\sum_{i=1}^N R_i' \Sigma^{-1} R_i \right)^{-1} \sum_{i=1}^N R_i' \Sigma^{-1} M_i \right]^{-1} \sum_{i=1}^N M_i' \Sigma^{-1} R_i \left(\sum_{i=1}^N R_i' \Sigma^{-1} R_i \right)^{-1} \sum_{i=1}^N R_i' \Sigma^{-1} Y_i \quad (5.11)$$

Les deux estimateurs sont équivalents si [BALESTRA.P (1970)] :

$$\Omega R = R B \quad (5.12)$$

où B est une matrice non singulière.

Si par contre nous pouvons écrire :

$$M = R \pi + \mu \quad (5.13)$$

l'estimateur de WHITE sera plus efficient que celui obtenu en appliquant les MCG sur le modèle transformé.

5.1 LES CONDITIONS D'ORTHOGONALITE

Une fois exposée la méthode des variables variables, le problème qui se pose est de savoir comment trouver les conditions d'orthogonalité qui permettent d'utiliser cette méthode. Toutes les conditions d'orthogonalité proposées dans le cadre du modèle sans autocorrélation ne sont pas valides ici [voir Matondzi(1996)]. Les conditions d'orthogonalité suivantes peuvent être utilisées dans le cadre des modèles dynamiques à erreurs composées avec autocorrélation.

5.1.1 LA CONDITION DE LIVIATAN

Cet auteur propose d'utiliser une variable exogène retardée comme instrument pour y_{it-1} . Prenons pour simplifier x_{it-1} . Les autres variables étant leurs propres instruments.

$$E[u_{it} x_{it-1}] = 0 \quad (5.14)$$

La matrice d'instruments de liviatan s'écrit alors :

$$R'_{\text{IL}} = (X_{i-1i} \quad X_i \quad S_T z_i) \quad (5.15)$$

5.1.2 LA CONDITION DE BALESTRA-NERLOVE

Ces auteurs dans leur article de 1966 ont développé une méthode d'estimation inspirée de celle des doubles moindres carrés. Cette méthode se présente comme suit:

Régresser y_{it} sur les variables explicatives en écrivant le modèle comme suit :

$$y_{it} = \frac{1}{1 - \delta L} (\beta x_{it} + \phi z_i + \alpha_i + v_{it}) \quad (5.16)$$

ou encore

$$y_{it} = \beta_1 x_{it} + \phi_1 z_i + \alpha_i^* + v_{it}^* \quad (5.17)$$

et estimer les paramètres par les moindres carrés ordinaires.

Soient $\tilde{\beta}_1$ et $\tilde{\phi}_1$ les estimateurs obtenus. Nous avons :

$$\hat{y}_{it} = \tilde{\beta}_1 x_{it} + \tilde{\phi}_1 z_i \quad (5.19)$$

$$\hat{y}_{it-1} = \tilde{\beta}_1 x_{it-1} + \tilde{\phi}_1 z_i \quad (5.20)$$

Lorsque les variables exogènes sont indépendantes des erreurs, il est clair que :

$$P \lim \frac{1}{NT} \sum_{i=1}^N \sum_{t=2}^T \hat{y}_{it-1} (\alpha_i + v_{it}) = 0 \quad (5.21)$$

Ces auteurs proposent d'utiliser \hat{y}_{it-1} comme instrument pour la variable endogène retardée et la matrice d'instruments de Balestra-Nerlove s'écrit :

$$R'_{iBN} = (\hat{Y}_{i-1i} \ X_i \ S_T z_i) \quad (5.22)$$

Lorsque la variable z_i est corrélée avec l'effet individuel, elle doit être éliminée de (5.19) ; cette méthode est alors équivalente à la méthode de LIVIATAN.

5.1.3 LA CONDITION DE LIVIATAN 2

Nous qualifions cette méthode de Liviatan 2 car sa construction est inspirée de la méthode présentée par cet auteur et s'applique lorsque la variable z_i est corrélée avec l'effet individuel. Soit H la matrice ((T-1)XT) des différences premières.

$$H = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 1 & 0 \\ 0 & \dots & 0 & -1 & 1 \end{bmatrix} \quad (5.23)$$

Posons :

$$\bar{H} = I_N \otimes H \quad (5.24)$$

Transformons le modèle en le multipliant par \bar{H} .

$$\bar{H}Y = \bar{H}M\phi + \bar{H}U \quad (5.25)$$

L'élément typique de cette expression s'écrit :

$$(y_{it} - y_{it-1}) = \delta(y_{it-1} - y_{it-2}) + \beta(x_{it} - x_{it-1}) + (v_{it} - v_{it-1}) \quad (5.26)$$

soit

$$y^*_{it} = \delta y^*_{it-1} + \beta x^*_{it} + v^*_{it} \quad (5.27)$$

où $v^*_{it} = v_{it} - v_{it-1}$ est un MA(1) à racine unitaire.

Le modèle en différences premières supprime z_i et α_i .

La matrice des variances-covariances de ce nouveau modèle s'écrit :

Estimer les paramètres δ et β par la méthode des variables instrumentales en utilisant les instruments de LIVIATAN c'est-à-dire, prendre x^*_{it-1} comme instrument pour y^*_{it-1} .

Le vecteur d'instruments s'écrit alors :

$$R^*_{it} = (x^*_{it-1} \ x^*_{it}) \quad (5.28)$$

5.1.4 LA CONDITION D'ORTHOGONALITE DE AHN ET SCHMIDT (1995)

Elle est définie par :

$$E[(u_{it} - u_{it-2})u_{it-1}] = 0 \quad (5.29)$$

5.1.5 AUTRE CONDITION D'ORTHOGONALITE

Partant du modèle transformé (5.26), nous pouvons utiliser x^*_{it-1} comme instrument pour y^*_{it-1} car :

$$E[x^*_{it-1} y^*_{it-1}] = 0 \quad (5.30)$$

Nous avons aussi :

$$E[\tilde{y}_{it} \tilde{v}_{it}] = 0 \quad s < t-1 \quad (5.31)$$

avec

$$\tilde{y}_{it} = y^*_{it} - \hat{\rho} y^*_{it-1} \quad \text{et} \quad \tilde{v}_{it} = v^*_{it} - \hat{\rho} v^*_{it-1}$$

Cet estimateur est convergent et suit asymptotiquement une loi normale.

5.2 ESTIMATION DU COEFFICIENT D'AUTOCORRELATION

La formule d'estimation du coefficient d'autocorrélation dépend du modèle utilisé.

5.2.1 LE MODELE DE BASE

Soit

$$\hat{u}_{it} = \alpha_i + v_{it} - (\hat{\delta} - \delta)y_{it-1} - (\hat{\beta} - \beta)x_{it} + (\hat{\phi} - \phi)z_i \quad (5.32)$$

le résidu des variables instrumentales des modèles sans transformation [Liviatan I, Balestra-Nerlove et Ahn et Schmidt].

L'estimateur du coefficient d'autocorrélation suivant proposé dans le cadre des séries temporelles n'est malheureusement pas un estimateur convergent de ce paramètre.

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it} \hat{u}_{it-1}}{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it-1}^2} \quad (5.33)$$

La méthode qui permet d'obtenir une estimation convergente de ρ est celle proposée par BALTAGI et LI (1994).

A partir des résidus des variables instrumentales (5.32), calculer les quantités suivantes :

$$Q_0 = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it}^2}{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it}^2} = 1, \quad Q_1 = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it} \hat{u}_{it-1}}{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it-1}^2}, \quad Q_2 = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it} \hat{u}_{it-2}}{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it-2}^2}$$

Estimer ensuite ρ par :

$$\hat{\rho} = \frac{Q_1 - Q_2}{Q_0 - Q_1} \quad (5.34)$$

Nous proposons d'écrire cet estimateur plus simplement comme suit :

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=3}^T (\hat{u}_{it} - \hat{u}_{it-1}) \hat{u}_{it-2}}{\sum_{i=1}^N \sum_{t=3}^T (\hat{u}_{it-2} - \hat{u}_{it-1}) \hat{u}_{it-1}} \quad (5.35)$$

De façon générale, cet estimateur convergent de ρ s'écrit :

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=3}^T (\hat{u}_{it-1} - \hat{u}_{it-p}) \hat{u}_{it-q}}{\sum_{i=1}^N \sum_{t=3}^T (\hat{u}_{it-q} - \hat{u}_{it-p}) \hat{u}_{it-p}} \quad (5.36)$$

avec

$$q=p-1 \text{ et } s=p+1$$

5.2.2 LE MODELE TRANSFORME

Nous proposons les deux méthodes suivantes pour estimer le coefficient d'autocorrélation sur le modèle transformé :

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=3}^T \hat{v}_{it}^* \hat{v}_{it-2}^*}{\sum_{i=1}^N \sum_{t=3}^T \hat{v}_{it}^* \hat{v}_{it-1}^*} \quad (5.37)$$

avec

$$\hat{\rho} = 2r + 1 \quad (5.38)$$

$$r = \frac{\sum_{i=1}^N \sum_{t=3}^T \hat{v}_{it}^* \hat{v}_{it-1}^*}{\sum_{i=1}^N \sum_{t=3}^T \hat{v}_{it}^{*2}} \quad (5.39)$$

De façon générale nous avons :

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=3}^T \hat{v}_{it-q}^* \hat{v}_{it-s}^*}{\sum_{i=1}^N \sum_{t=3}^T \hat{v}_{it-q}^* \hat{v}_{it-p}^*} \quad (5.40)$$

$$r = \frac{\sum_{i=1}^N \sum_{t=3}^T \hat{v}_{it-s}^* \hat{v}_{it-p}^*}{\sum_{i=1}^N \sum_{t=3}^T \hat{v}_{it-s}^{*2}} \quad (5.40)$$

avec

$$q=p-1 \text{ et } s=p+1$$

CONCLUSION

L'étude des modèles dynamiques à erreurs à erreurs composées avec autocorrélation est assez récente. Nous avons voulu dans les sections précédentes présenter de façon unifiée un bref aperçu des résultats obtenus dans l'étude de ces modèles. Il ressort de notre analyse que :

- les estimateurs usuels n'ont pas tous le même comportement. Si, asymptotiquement, les estimateurs within, MCG et MCQG équivalents sont convergents, il n'en est pas de même pour les estimateurs between et MCO. En revanche, tous les estimateurs sont semi-asymptotiquement biaisés.
- l'estimateur du maximum de vraisemblance est d'une étude assez difficile surtout lorsque l'on tient compte de la loi des observations initiales.
- la méthode des variables instrumentales d'une application relativement simple est la méthode adaptée pour estimer de façon convergente (asymptotiquement et semi-asymptotiquement) les paramètres du modèle. Cependant, les conditions d'orthogonalité adaptées à ce modèle sont peu nombreuses. Il faut donc que les économètres se penchent sur cette question comme ils l'ont fait pour le modèle dynamiques sans autocorrélation.
- la formule classique d'estimation du coefficient d'autocorrélation n'est pas valide ici. Nous avons présenté quelques méthodes d'estimation convergente de ce coefficient et montré cette méthode varie en fonction du modèle considéré (modèle avec ou sans transformation).

BIBLIOGRAPHIE

- AHN, S.C. and SCHMIDT, P., 1995**, Efficient Estimation of models for dynamics Panel data, *Journal of Econometrics* 68, 5-27
- ANDERSON, T.W. and HSIAO, C., 1981**, Estimation of Dynamic Models with Error Components Models, *Journal of the American Statistical Association* 76, 598-606
- ANDERSON, T.W. and HSIAO, C., 1982**, Formulation and Estimation of Dynamic Models using Panel data, *Journal of Econometrics* 18, 47-82
- ARELLANO, M. and BOVER, O., 1995**, Another look at the Instrumental Variable Estimation of Error Component models, *Journal of econometrics* 68, 29-51
- BALESTRA, P. and NERLOVE, M., 1966**, Pooling Cross-Section and Time-series data in the Estimation of a dynamic model, *Econometrica* 34, 585-612
- BALTAGI, B.H., 1995**, Econometrics Analysis of Panel data, John WILEY and SONS
- BALTAGI, B.H. and LI, Q., 1991**, A transformation that will Circumvent the Problem of Autocorrelation in an Error Component Model, *Journal of Econometrics* 48, 385-393
- BALTAGI, B.H. and LI, Q., 1994**, Monte Carlo Results on Pure and Pretest Estimators Of an Error Component Model with Autocorrelated Disturbances, Unpublished paper
- BEACH, C.M. and MacKINNON, J.G., 1978**, A Maximum Likelihood Procedure for Regression with Autocorrelation Errors, *Econometrica* 46, 51-58
- BLUNDELL, R. and SMITH, R.S., 1991**, Conditions Initiales et Estimation Efficaces dans les Modèles Dynamiques sur Données de Panel, *Annales d'Economie et de Statistique* 20-21, 109-123
- DORMONT, B., 1989**, Introduction à l'Econometrie des Données de Panel: Théorie et Application à des Echantillons d'Entreprise, Editions du CNRS, Paris
- HATANAKA, M., 1974**, An Efficient Two-Step Estimation for the Dynamic Adjustment Model with Autoregressive Errors, *Journal of Econometrics* 2, 199-220
- HOLTZ-EAKIN, D., NEWEY, W and ROSEN, H.S., 1988**, Estimating Vector Autoregressions with Panel Data, *Econometrica* 56, 1371-1395
- HSIAO, C., 1986**, Analysis of Panel Data, Cambridge University Press
- KEANE, M.P and RUNKLE, D.E., 1992**, On the Estimation of Panel Data Models with Serial Correlation when the Instruments are not Strictly exogenous, *Journal of Business and Economics Statistics* 10, 1-9
- KIVIET, J.K., 1995**, On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamics Panel Data Models, *Journal of Econometric* 68, 53-78
- KOOPMANS, T., 1942**, Serial Correlation and Quadratic Forms in Normal Variable, *The Annals of Mathematical Statistics* 13, 14-33
- LIVIATAN, N., 1963**, Consistent Estimation of Distributed Lag, *International Economics Review* 4, 44-52

- MAGNUS, J.R and WOOLAND, A.D ., 1988**, On the Maximum Likelihood Estimation of Multivariate Regression Models Containing Serially Correlated Error Components, *International Economics Review* 29, 707-725
- MATONDZI NGOUMA, S.A.,1991**, Estimation des Modèles Dynamiques à Erreurs Composées, Mémoire de D.E.A., Université de Dijon
- MATONDZI NGOUMA, S.A.,1996**, Estimation des Modèles Dynamiques à Erreurs Composées avec Autocorrélation par la Méthode des Variables Instrumentales *LATEC, Document de Travail n° 9608 Université de Dijon*
- MATONDZI NGOUMA, S.A.,1997a**, Comportement des Estimateurs Usuels dans les Modèles Dynamiques à Erreurs Composées avec Autocorrélation , *LATEC, Document de Travail , Université de Dijon*
- MATONDZI NGOUMA, S.A.,1997b**, Modèles Dynamiques à Erreurs Composées avec Autocorrélation : Etude et Application à la Demande de Travail, Thèse de doctorat, Université de Dijon
- MATYAS, L. and SEVESTRE, P.,1992**, The Econometrics of Panel Data Kluwer Academic publishers
- OBERHOFER, W and KMENTA, J., 1974**, A General Procedure for Obtaining Maximum Likelihood Estimates in Generalized Regression Models, *Econometrica* 42, 579-590
- PRAIS, S.J and WINSTEN, C.B., 1954**, Trend Estimation and Serial Correlation Unpublished Cowles Commission Discussion paper Stat. No 383, University of Chicago
- REVANKAR,N.S., 1979**, Error Component Models with Serially Correlated Time Effects, *Journal of the Indian Statistical Association* 17, 137-160
- SEVESTRE,P., 1983**, Modèles Dynamiques à Erreurs Composées, Thèse de Doctorat 3^{ème} Cycle, Paris I
- SEVESTRE,P. et TROGNON,A., 1983**, Propriétés des Grands Echantillons d'une Classe d'Estimateurs des modèles Autoregressifs à Erreurs Composées, *Annales de l'INSEE* 50, 25-47
- SEVESTRE,P. et TROGNON,A., 1985**, A Note on Autoregressive Error Components Models, *Journal of Econometrics* 28, 231-245
- TROGNON, A., 1978**, Miscellaneous Asymptotic Properties of Ordinary Least Square and Maximum Likelihood Estimators in Dynamic Error Components Models, *Annales de l'INSEE* 30-31, 632-657
- WHITE,H., 1982**, Instrumental Variables Regression with Independent Observations, *Econometrica* 50, 483-499
- WHITE,H., 1984**, Asymptotic Theory for Econometricians, Academic Press, New-York