



Modèles récursifs à double indice

Pietro Balestra, Hassan Ghassan

► To cite this version:

Pietro Balestra, Hassan Ghassan. Modèles récursifs à double indice. [Rapport de recherche] Laboratoire d'analyse et de techniques économiques(LATEC). 1994, 54 p., ref. bib. : 20 ref. <hal-01527155>

HAL Id: hal-01527155

<https://hal.science/hal-01527155v1>

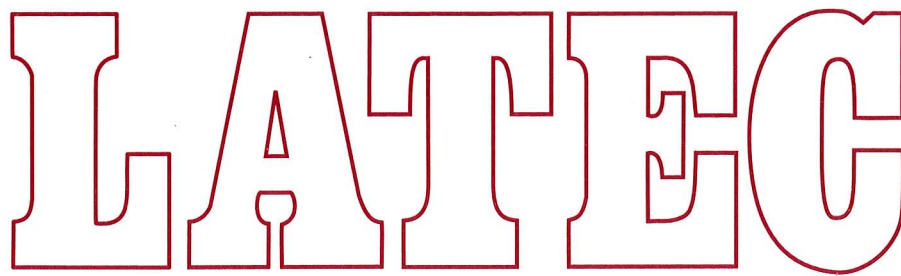
Submitted on 24 May 2017

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.



HAL Authorization



**LABORATOIRE D'ANALYSE
ET DE TECHNIQUES ÉCONOMIQUES**

UMR 5601 CNRS

DOCUMENT DE TRAVAIL



Pôle d'Economie et de Gestion

2, boulevard Gabriel - 21000 DIJON - Tél. 03 80 39 54 30 - Fax 03 80 39 54 43

ISSN : 1260-8556

n° 9406

**MODELES RECURSIFS
A DOUBLE INDICE**

Pietro BALESTRA et Hassan GHASSAN

septembre 1994

Résumé

L'objet de ce travail est d'étudier les résultats concernant l'estimation des modèles récurrents au cas où l'observation porte sur un ensemble d'unités statistiques (individus, secteurs, pays), chacune étant définie par un système d'équations de type récurrent. A la dimension temporelle, présente dans l'étude classique de ces systèmes, nous ajoutons la dimension transversale, d'où les "modèles récurrents à double indice". Le traitement de ces derniers modèles nous a conduit à regrouper les variables par individu puis par équation. Ce regroupement nous a permis d'assimiler notre modèle au système d'équations apparemment non-liées, avec la particularité que la variance relative aux individus i et j pour l'équation k n'est pas nécessairement nulle. L'estimation par la méthode du maximum de vraisemblance et par les moindres carrés généralisés purs, réalisables et itérés nous ont donné des estimateurs convergents. L'étude des propriétés asymptotiques de chacune des méthodes d'estimation a été effectuée en détail.

L'application économétrique a consisté à formaliser la structure productive de l'économie marocaine de cinq secteurs d'activités économiques. Après des tests statistiques d'exogénéité, il semble que les modèles sectoriels s'adaptent bien au schéma récurrent individuel et que les résultats s'avèrent très riches en contenus et se prêtent à des interprétations objectives. Par exemple, il apparaît nécessaire d'accroître l'interdépendance sectorielle sans concentrer les efforts sur l'augmentation des niveaux d'investissements sectoriels.

Mots clés : *Causalité, récurrent, modèle, estimation, asymptotique, intégration sectorielle.*

Abstract

The purpose of this document is to study the model belongs to the family of structural equation models with data varying both across individuals (sectors) and in time. A complete theoretical analysis is developed in this work for the case of a recursive structure. Maximum likelihood estimation and Zellner's "seemingly unrelated equations" estimators (iterated or not) are presented and their asymptotic properties are carefully derived.

The application of the estimation methods to the economy morocco produced statistically significant and economic meaningful results. Through a test of exogeneity, the recursivity of the system is confirmed. Policy recommendations are discussed and analysed.

Key words : *Causality, recursive system, model, estimation, asymptotic sectorial integration.*

Modèles récurrents à double indice

Pietro Balestra* et Hassan Ghassan†

Avril 1994

1 Introduction

Les phénomènes économiques sont par nature intrinséquement interdépendants. Pour les étudier, les systèmes d'équations simultanées sont un instrument d'analyse privilégié. Dans un système simultané, on distingue les variables exogènes (qui exercent une influence sur les variables endogènes sans rétroaction) et les variables endogènes (qui reçoivent l'influence des variables exogènes et s'influencent mutuellement). Parfois, la théorie économique suggère un schéma de causalité directe non seulement entre variables exogènes et endogènes, mais également entre les variables endogènes elles-mêmes [Wold et al.60]. On parle alors de système récurrent.

Le but de ce travail est d'étudier les résultats concernant l'estimation des modèles récurrents au cas où l'observation porte sur un ensemble d'unités statistiques (individus, secteurs, pays), chacune étant définie par un système d'équations de type récurrent. A la dimension temporelle, présente dans l'étude classique de tels systèmes, nous ajoutons la dimension transversale. D'où le nom "modèles récurrents à double indice".

Ce travail s'articule en plusieurs sections. Dans la section 1, on décrit formellement un système récurrent à double indice, suivie par la construction de la fonction de vraisemblance. Les sections suivantes sont consacrées à l'estimation du modèle par les techniques du maximum de vraisemblance et les moindres carrés généralisés purs, réalisables et itérés. Enfin, dans la dernière section, une application au cas de l'économie marocaine est donnée. Cette application conduite par l'un des deux auteurs, a inspiré le présent travail.

2 Spécification du modèle et hypothèses

Comme annoncé dans l'introduction, la formalisation que nous proposons généralise l'étude des systèmes récurrents au cas où l'observation porte sur un ensemble de N unités

*Université de Bourgogne

†ENESAD, Dijon

statistiques, appelées individus. Pour chaque individu, un système récursif à K équations est formulé. Enfin, chaque individu (et toutes les variables qui le concernent) est observé sur T périodes. Nous commençons par la spécification du modèle au niveau de l'individu (section 1.1), puis au niveau global (section 1.2).

2.1 Spécification individuelle

Au niveau de l'individu i ($i = 1, \dots, N$) et la période d'observation t ($t = 1, \dots, T$) la forme structurelle du modèle est définie par le système de K équations suivant :

$$\beta^{i'} y_t^i + \gamma^{i'} x_t^i + u_t^i = 0 \quad (1)$$

où :

- y_t^i est le vecteur $K \times 1$ des observations concernant les K variables endogènes (le système étant alors par définition complet)
- x_t^i est le vecteur $L \times 1$ des observations concernant les L variables exogènes
- $\beta^{i'}$ est la matrice $K \times K$ des coefficients des variables endogènes, supposée non singulière
- $\gamma^{i'}$ est la matrice $K \times L$ des coefficients des variables exogènes
- u_t^i est le vecteur $K \times 1$ d'erreurs.

Pour que la spécification soit complète, il est indispensable que le modèle 1 soit muni d'un corps d'hypothèses. Ces hypothèses ont la teneur suivante :

(i) H1 : Hypothèse de comportement

β^i et γ^i sont des matrices de constantes (inconnues), variables d'individu à individu. On admet par conséquent que chaque individu est caractérisé par un comportement qui lui est propre (mais qui reste invariant dans le temps).

Pour des raisons de commodité, on supposera également que les restrictions théoriques a-priori d'exclusion (les éléments nuls a-priori de β^i et γ^i) sont les mêmes pour tous les individus.

(ii) H2 : Récursivité

La matrice non-singulière $\beta^{i'}$ est triangulaire inférieure : ses éléments au-dessus de la diagonale principale sont tous nuls et ceux de la diagonale principale sont tous égaux à -1 (du fait de la normalisation).

(iii) H3 : Variables exogènes

Les variables exogènes sont fixes, non stochastiques, indépendantes des erreurs et telles que :

- (a) $\sum_t x_t' x_t'$ est une matrice de rang L
- (b) $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_t x_t' = x'$, fini
- (c) $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_t x_t' x_t' = M'$, définie positive

(iv) H1 : Hypothèses sur les erreurs

Les erreurs sont des variables aléatoires non-observables possédant des moments finis au moins jusqu'à l'ordre deux. En particulier, en désignant par $u_{kt}^i, k = 1, \dots, K$, l'élément typique du vecteur u_t^i , on postule que

(a) l'effet des perturbations est nul en moyenne :

$$E(u_{kt}^i) = 0 \quad \forall i, k, t$$

(b) la structure des variances-covariances est telle que :

$$E(u_{kt}^i u_{ls}^i) = \begin{cases} \sigma_k^{ij} & \text{si } k = l \text{ et } t = s \\ 0 & \text{sinon} \end{cases}$$

La structure des variances-covariances définie en (b) mérite d'être interprétée. La nullité de la covariance pour deux périodes différentes ($t \neq s$) exprime l'absence totale d'auto-corrélation (directe, entre individus et entre équations).

La nullité de la covariance pour deux équations différentes ($k \neq l$) définit (conjointement à H2) la récursivité du système. Le terme σ_k^{ii} désigne la variance de l'erreur de l'équation k de l'individu i et le terme σ_k^{ij} désigne la covariance contemporaine entre l'erreur de l'individu i et celle de l'individu j pour l'équation k . Elle n'est pas nécessairement nulle. C'est précisément cette possibilité d'une covariance non nulle qui donne à notre modèle toute sa spécificité.

De l'hypothèse H4, on tire les deux formulations suivantes :

$$E(u_t^i) = 0 \quad \text{et}$$

$$E(u_t^i u_s^j) = \begin{cases} \text{diag}(\sigma_1^{ii}, \dots, \sigma_K^{ii}) = D^{ii} & \text{si } i = j \text{ et } t = s \\ \text{diag}(\sigma_1^{ij}, \dots, \sigma_K^{ij}) = D^{ij} & \text{si } i \neq j \text{ et } t = s \\ 0 & \text{sinon} \end{cases}$$

Maintenant, si la seule information prise en compte est celle relative à l'individu i , le système est un système récursif traditionnel. En conséquence, l'estimation individuelle de chaque équation par la méthode des moindres carrés ordinaires fournit des estimateurs convergents et asymptotiquement efficaces, qui sont égaux (sous l'hypothèse de normalité) aux estimateurs du maximum de vraisemblance.

Il en va autrement lorsque toute l'information concernant l'ensemble des individus est prise en compte. Cet aspect du problème sera mis en évidence dans le prochain paragraphe.

2.2 Modèle global

Il s'agit de réunir dans un modèle global toute l'information concernant les N individus. Le regroupement des individus (pour la période t) peut s'opérer de deux manières différentes. Chacune d'elles présente un avantage particulier : soit en empilant les individus, soit en empilant les équations.

2.2.1 Regroupement par individus

En regroupant les N individus pour la période t , le modèle global s'écrit de la manière suivante :

$$\begin{bmatrix} \beta^{1'} & & \\ & \ddots & \\ & & \beta^{N'} \end{bmatrix} \begin{bmatrix} y_t^1 \\ \vdots \\ y_t^N \end{bmatrix} + \begin{bmatrix} \gamma^{1'} & & \\ & \ddots & \\ & & \gamma^{N'} \end{bmatrix} \begin{bmatrix} x_t^1 \\ \vdots \\ x_t^N \end{bmatrix} + \begin{bmatrix} u_t^1 \\ \vdots \\ u_t^N \end{bmatrix} = 0 \quad (2)$$

ou encore sous forme compacte :

$$\beta' y_t + \gamma' x_t + u_t = 0 \quad (3)$$

où β' est triangulaire et diagonale par bloc, d'ordre $NK \times NK$, γ' est d'ordre $NK \times NL$, et y_t, x_t et u_t sont des vecteurs respectivement de dimension NK , NL et NK . Pour le vecteur de perturbations u_t , à partir de H4 nous déduisons :

$$E(u_t) = 0, \text{ et}$$

$$E(u_t u_s') = \begin{cases} [D^{ij}] := \Omega & \text{si } t = s \\ 0 & \text{sinon} \end{cases}$$

On note que la matrice Ω n'est pas diagonale, mais bloc-diagonale. Elle serait diagonale dans un cas très particulier où σ_k^{ij} quel que soit $i \neq j$ et quel que soit k . Par conséquent, ses propriétés ne sont pas faciles à étudier. Néanmoins, l'écriture du modèle sous la forme de l'équation (3) présente l'avantage considérable permettant de déduire immédiatement la fonction de vraisemblance du vecteur y en partant de la fonction de densité de u_t , en raison de la triangularité de β' (comme nous le verrons à la section 2).

2.2.2 Regroupement par équations

La k -ème équation pour l'individu i (la k -ème ligne du modèle (1)) s'écrit :

$$(\beta_k^i)' y_t^i + (\gamma_k^i)' x_t^i + u_{kt}^i = 0 \quad (4)$$

où $(\beta_k^i)'$ et $(\gamma_k^i)'$ sont respectivement la k -ème ligne de $(\beta^i)'$ et $(\gamma^i)'$.

En raison de la triangularité de $(\beta^i)'$, de la normalisation (le k -ème coefficient de (β_k^i) étant égal à -1) et en tenant compte d'éventuelles autres restrictions nulles sur certains paramètres, l'équation (4) prend la forme suivante :

$$y_{kt}^i = (\tilde{\beta}_k^i)' y_t^{i/k} + (\tilde{\gamma}_k^i)' x_t^{i/k} + u_{kt}^i \quad (5)$$

où

y_{kt}^i est la variable endogène expliquée de la k -ème équation pour l'individu i

$y_t^{i/k}$ est le vecteur des variables endogènes explicatives dans cette équation (les variables endogènes y_{pt}^i d'indice $p \geq k$ ne figurent pas dans le vecteur $y_t^{i/k}$)

$x_t^{i/k}$ est le vecteur des variables exogènes figurant dans l'équation (si toutes les variables exogènes interviennent dans l'équation, alors $x_t^{i/k}$ est égal à x_t^i)

$(\tilde{\beta}_k^i)'$ et $(\tilde{\gamma}_k^i)'$ sont les parties correspondantes de $(\beta_k^i)'$ et $(\gamma_k^i)'$.

Le nombre total de variables explicatives de la k -ème équation est désigné par s_k avec $s_k \leq L + k - 1$.

Il est commode de réunir toutes les variables explicatives dans un seul vecteur ainsi que les paramètres correspondants.

En posant :

- $Z_t^{i/k} = \begin{bmatrix} y_t^{i/k} \\ x_t^{i/k} \end{bmatrix}$ le vecteur $s_k \times 1$ des variables explicatives
- $(\theta_k^i)' = [(\tilde{\beta}_k^i)' \ (\tilde{\gamma}_k^i)']$ le vecteur $1 \times s_k$ des paramètres

il vient alors :

$$y_{kt}^i = (\theta_k^i)' Z_t^{i/k} + u_{kt}^i = (Z_t^{i/k})' (\theta_k^i) + u_{kt}^i \quad (6)$$

En empilant tous les individus pour l'équation k , on obtient :

$$\begin{bmatrix} y_{kt}^1 \\ y_{kt}^2 \\ \vdots \\ y_{kt}^N \end{bmatrix} = \begin{bmatrix} (Z_t^{1/k})' & & & \\ & (Z_t^{2/k})' & & \\ & & \ddots & \\ & & & (Z_t^{N/k})' \end{bmatrix} \begin{bmatrix} \theta_k^1 \\ \theta_k^2 \\ \vdots \\ \theta_k^N \end{bmatrix} + \begin{bmatrix} u_{kt}^1 \\ u_{kt}^2 \\ \vdots \\ u_{kt}^N \end{bmatrix} \quad (7)$$

ou encore sous forme compacte :

$$y_{kt} = (Z_{kt} \theta_k + u_{kt} \quad (8)$$

Finalement, pour l'ensemble des K équations on arrive à l'expression suivante :

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Kt} \end{bmatrix} = \begin{bmatrix} Z_{1t} & & & \\ & Z_{2t} & & \\ & & \ddots & \\ & & & Z_{Kt} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_K \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{Kt} \end{bmatrix} \quad (9)$$

que l'on peut écrire :

$$\tilde{y}_t = Z_t \theta + \tilde{u}_t \quad (10)$$

Dans l'expression ci-dessus \tilde{y}_t et \tilde{u}_t contiennent les mêmes éléments que y_t et u_t de la formule (3), mais dans un ordre différent. L'intérêt de cette expression est d'explicitier le vecteur θ des paramètres inconnus incorporant toutes les restrictions théoriques a-priori.

L'hypothèse H4 nous conduit aux propriétés statistiques suivantes concernant les vecteurs de perturbation de (8) :

$$\begin{aligned} E(u_{kt}) &= 0, \text{ et} \\ E(u_{kt}u_{lt}') &= \begin{cases} [\sigma_k^{ij}] := \tilde{\Omega} & \text{si } k = l \\ 0 & \text{sinon} \end{cases} \end{aligned} \quad (11)$$

Il s'en suit que la matrice des variances-covariances du vecteur \tilde{u}_t est une matrice diagonale par bloc, à savoir :

$$E(\tilde{u}_t \tilde{u}_t') = \text{diag}(\tilde{\Omega}_1, \dots, \tilde{\Omega}_K) := \tilde{\Omega} \quad (12)$$

En l'absence de restrictions linéaires sur les éléments \tilde{u}_t (ce que nous admettons) la matrice $\tilde{\Omega}$ doit être définie positive. Cette propriété est satisfaite dès lors que chaque bloc $\tilde{\Omega}_k$ est une matrice définie positive.

Le déterminant et l'inverse de $\tilde{\Omega}$ se déduisent facilement. Ils sont donnés par :

- $|\tilde{\Omega}| = \prod_{k=1}^K |\tilde{\Omega}_k|$
- $\tilde{\Omega}^{-1} = \text{diag}(\tilde{\Omega}_1^{-1}, \dots, \tilde{\Omega}_K^{-1})$

2.2.3 Matrice de passage de u_t à \tilde{u}_t

Comme nous l'avons déjà signalé, les vecteurs u_t et \tilde{u}_t contiennent les mêmes éléments. Ils sont au total NK en nombre. Il est possible de les ordonner tous dans une matrice $K \times N$. A cet effet, nous posons :

$$U_t = [u_t^1, \dots, u_t^N]$$

Dans cette matrice, les N colonnes représentent les vecteurs u_t^i de l'équation (1), $i = 1, \dots, N$. En revanche, les K lignes représentent les vecteurs $(u_{kt})'$ de l'équation (8), avec

$k = 1, \dots, K$. On a donc l'écriture équivalente :

$$U_t = \begin{bmatrix} u'_{1t} \\ \vdots \\ u'_{Kt} \end{bmatrix}$$

Par les propriétés de l'opération vec (qui empile les colonnes d'une matrice) [Balestra76], nous avons :

$$u_t = \text{vec } U_t \quad \text{et} \quad \tilde{u}_t = \text{vec } U'_t$$

Comme, par ailleurs : $\text{vec } U'_t = P_{N,K} \text{vec } U_t$ où $P_{N,K}$ est la matrice unitaire permutée, on obtient :

$$\tilde{u}_t = P_{N,K} u_t \quad (13)$$

La matrice orthogonale $P_{N,K}$ est donc la matrice de passage de u_t à \tilde{u}_t . Elle va nous permettre de passer indifféremment, par une transformation linéaire non-singulière, de la formulation (3) à la formulation (10) et aussi de dériver les propriétés de la matrice Ω de celles de $\tilde{\Omega}$. Ce que nous verrons en détail dans ce qui suit.

En pré-multipliant l'équation (3) par $P_{N,K}$, il vient :

$$-P_{N,K}(\beta' y_t + \gamma' x_t) = P_{N,K} u_t = \tilde{u}_t$$

En comparant avec (10), on a la relation suivante :

$$\tilde{y}_t - Z_t \theta = -P_{N,K}(\beta' y_t + \gamma' x_t) \quad (14)$$

L'équivalence ci-dessus sera exploitée dans la dérivation de la fonction de vraisemblance du modèle.

En outre, en partant de (13), on a :

$$V(\tilde{u}_t) = P_{N,K} V(u_t) P_{K,N} \quad \text{ou encore} \quad \tilde{\Omega} = P_{N,K} \Omega P_{K,N}$$

où, par les propriétés de la matrice unitaire permutée on a :

$$P'_{N,K} = P_{K,N} = P_{N,K}^{-1}$$

On en déduit immédiatement :

$$\Omega = P_{K,N} \tilde{\Omega} P_{N,K}$$

par conséquent :

$$|\Omega| = |P_{K,N} \tilde{\Omega} P_{N,K}| = |\tilde{\Omega}| |P_{K,N} P_{N,K}| = |\tilde{\Omega}|$$

d'où :

$$|\Omega| = \prod_{k=1}^K |\tilde{\Omega}_k| \quad (15)$$

$$\Omega^{-1} = P_{K,N} \tilde{\Omega}^{-1} P_{N,K} \quad (16)$$

2.2.4 Prise en compte des observations pour toutes les périodes

Cette prise en compte peut se faire suivant les deux formes de regroupement des observations à l'instant t . Nous commençons le regroupement par individus. Transposons l'équation (3) et empilons pour obtenir :

$$\begin{bmatrix} y'_1 \\ \vdots \\ y'_t \\ \vdots \\ y'_T \end{bmatrix} \beta + \begin{bmatrix} x'_1 \\ \vdots \\ x'_t \\ \vdots \\ x'_T \end{bmatrix} \gamma + \begin{bmatrix} u'_1 \\ \vdots \\ u'_t \\ \vdots \\ u'_T \end{bmatrix} = 0$$

En forme compacte, cette expression s'écrit par :

$$Y\beta + X\gamma + U = 0 \quad (17)$$

où Y, X et U sont respectivement de dimension $T \times NK, T \times NL$ et $T \times NK$

La matrice d'erreurs U , en vertu de H4, possède les propriétés statistiques suivantes :

$$\begin{aligned} E(U) &= 0 \\ V(U) &= V(\text{vec } U)(\text{vec } U)' = I_T \otimes \Omega \end{aligned} \quad (18)$$

Le regroupement par équations conduit à deux écritures différentes. Premièrement, en empilant directement les vecteurs définis dans la formule (10) on obtient :

$$\begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_T \end{bmatrix} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_T \end{bmatrix} \theta + \begin{bmatrix} \tilde{u}_1 \\ \vdots \\ \tilde{u}_T \end{bmatrix} \quad (19)$$

Sous cette forme, le vecteur d'erreurs a une matrice de variances-covariances égale à $I_T \otimes \tilde{\Omega}$. Elle ne permet pas d'appréhender facilement la structure du modèle. La deuxième écriture est plus probante à cet égard.

De l'équation (6), en considérant l'ensemble des périodes, on peut écrire :

$$\begin{bmatrix} y_{k1}^i \\ \vdots \\ y_{kT}^i \end{bmatrix} = \begin{bmatrix} (Z_1^{i/k})' \\ \vdots \\ (Z_T^{i/k})' \end{bmatrix} \theta_k^i + \begin{bmatrix} u_{k1}^i \\ \vdots \\ u_{kT}^i \end{bmatrix}$$

ce qui peut s'écrire par :

$$y_k^i = Z^{i/k} \theta_k^i + u_k^i$$

En empilant les N individus, toujours pour l'équation k , il vient :

$$\begin{bmatrix} y_k^1 \\ \vdots \\ y_k^N \end{bmatrix} = \begin{bmatrix} Z^{1/k} & & \\ & \ddots & \\ & & Z^{N/k} \end{bmatrix} \begin{bmatrix} \theta_k^1 \\ \vdots \\ \theta_k^N \end{bmatrix} + \begin{bmatrix} u_k^1 \\ \vdots \\ u_k^N \end{bmatrix}$$

ce qui s'écrit comme suit :

$$y_k = Z_k \theta_k + u_k \quad (20)$$

Ce modèle représente un système d'équations apparemment non-liées de type de Zellner. Le vecteur d'erreur u_k possède la structure classique d'un tel système, à savoir :

$$\begin{aligned} E(u_k) &= 0 \\ E(u_k u_k') &= \tilde{\Omega}_k \otimes I_T \end{aligned}$$

La différence par rapport à un modèle Zellner classique est la présence de variables endogènes parmi les variables explicatives. L'écriture (20) jouera un rôle primordial lors de l'estimation.

Enfin, en empilant les K équations on aboutit au modèle complet suivant :

$$\begin{bmatrix} y_1 \\ \vdots \\ y_K \end{bmatrix} = \begin{bmatrix} Z_1 & & \\ & \ddots & \\ & & Z_K \end{bmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_K \end{bmatrix} + \begin{bmatrix} u_1 \\ \vdots \\ u_K \end{bmatrix}$$

ce qui s'écrit comme le système suivant :

$$\tilde{y} = Z\theta + \tilde{u} \quad (21)$$

avec

$$\begin{aligned} E(\tilde{u}) &= 0 \\ V(\tilde{u}) &= \tilde{\Omega} \otimes I_T \end{aligned} \quad (22)$$

3 Fonction de vraisemblance

Pour la préciser, il est nécessaire d'introduire une hypothèse supplémentaire concernant la forme de la fonction de densité des erreurs. Comme il est coutume en économétrie [Intriligator et al.84], nous adoptons l'hypothèse suivante :

(v) H5 : Les erreurs sont distribuées d'après la loi normale

$$u_t \sim N(0, \Omega)$$

Par conséquent, la fonction de densité du vecteur u_t prend la forme suivante :

$$f(u_t) = (2\pi)^{-\frac{NK}{2}} |\Omega|^{-\frac{1}{2}} \exp - \frac{1}{2} u_t' \Omega^{-1} u_t$$

Dans ce qui suit, nous dérivons la fonction de vraisemblance du vecteur y_t , puis nous montrerons qu'elle se factorise en un nombre K de vraisemblances individuelles, une pour chaque équation. Enfin, le passage à la fonction de vraisemblance pour l'ensemble des T périodes est obtenue immédiatement (par simple sommation des log-vraisemblances pour chaque période t), en raison de l'hypothèse fondamentale d'indépendance temporelle.



3.1 Vraisemblance pour y à l'observation t

Puisque la matrice β' est triangulaire avec des éléments égaux à -1 sur la diagonale principale, le Jacobien de la transformation de u_t à y_t (3) est unitaire. Il s'en suit que la fonction de densité de y_t est obtenue de celle de u_t par simple changement des variables. On a donc :

$$f(y_t) = (2\pi)^{-\frac{N_K}{2}} |\Omega|^{-\frac{1}{2}} \exp -\frac{1}{2}(\beta' y_t + \gamma' x_t)' \Omega^{-1} (\beta' y_t + \gamma' x_t)$$

et la fonction de vraisemblance logarithmique (à une constante additive près), notée \mathcal{L} prend la forme suivante :

$$\mathcal{L}(y_t|\beta, \gamma, \Omega) = -\frac{1}{2} \ln |\Omega| - \frac{1}{2}(\beta' y_t + \gamma' x_t)' \Omega^{-1} (\beta' y_t + \gamma' x_t) \quad (23)$$

3.2 Factorisation de la vraisemblance

Nous exploitons les résultats de la section (1.2) concernant la relation entre Ω et $\tilde{\Omega}$. De (15), nous déduisons immédiatement :

$$\frac{1}{2} \ln |\Omega| = \sum_{k=1}^K \frac{1}{2} \ln |\tilde{\Omega}_k|$$

En outre, de (16) et en utilisant (14) nous obtenons successivement :

$$\begin{aligned} (\beta' y_t + \gamma' x_t)' \Omega^{-1} (\beta' y_t + \gamma' x_t) &= (\beta' y_t + \gamma' x_t)' P'_{N,K} \tilde{\Omega}^{-1} P_{N,K} (\beta' y_t + \gamma' x_t) \\ &= (\tilde{y}_t - Z_t \theta)' \tilde{\Omega}^{-1} (\tilde{y}_t - Z_t \theta) \\ &= \sum_{k=1}^K (y_{kt} - Z_{kt} \theta_k)' \tilde{\Omega}_k^{-1} (y_{kt} - Z_{kt} \theta_k) \end{aligned}$$

L'introduction de ces deux résultats dans (23) donne :

$$\mathcal{L}(y_t|\beta, \gamma, \Omega) = \sum_{k=1}^K -\frac{1}{2} \ln |\tilde{\Omega}_k| - \frac{1}{2} (y_{kt} - Z_{kt} \theta_k)' \tilde{\Omega}_k^{-1} (y_{kt} - Z_{kt} \theta_k)$$

ce qui revient à écrire :

$$\mathcal{L}(y_t|\beta, \gamma, \Omega) = \sum_{k=1}^K (y_{kt}|\theta_k, \tilde{\Omega}_k) \quad (24)$$

La log-vraisemblance globale est donc la somme de K log-vraisemblances indépendantes, une pour chaque équation. L'avantage de cette décomposition est que la maximisation de la fonction de vraisemblance peut se faire séparément pour chaque équation (mais simultanément pour les N individus).

Dans le cas extrême, peu intéressant où $\sigma_k^{ij} = 0 (i \neq j)$ i.e. quand les individus eux-mêmes sont indépendants, la matrice $\tilde{\Omega}_k$ est diagonale et la décomposition peut se poursuivre au niveau de l'individu. En effet, la log-vraisemblance devient :

$$\begin{aligned}\mathcal{L}(y_{kt}|\theta_k, \tilde{\Omega}_k) &= \sum_{i=1}^n -\frac{1}{2} \log \sigma_k^{ii} - \frac{1}{2\sigma_k^{ii}} (y_{kt}^i - (Z_t^{i/k})' \theta_k^i)' (y_{kt}^i - (Z_t^{i/k})' \theta_k^i) \\ &= \sum_{i=1}^n \mathcal{L}(y_{kt}^i | \theta_k^i, \sigma_k^{ii})\end{aligned}$$

Dans ce cas extrême, l'estimation du maximum de vraisemblance se ramène à l'estimateur par les moindres carrés ordinaires équation par équation et individu par individu.

3.3 Vraisemblance pour l'échantillon complet

En raison de la factorisation, il suffit de donner la log-vraisemblance totale pour les T périodes pour une équation, la k -ème. Elle se trouve en sommant sur t les log-vraisemblances. On a alors :

$$\begin{aligned}\mathcal{L}(y_k|\theta_k, \tilde{\Omega}_k) &= \sum_{t=1}^T \mathcal{L}(y_{kt}|\theta_k, \tilde{\Omega}_k) \\ &= -\frac{T}{2} \ln |\tilde{\Omega}_k| - \frac{1}{2} \sum_{t=1}^T (y_{kt} - Z_{kt}\theta_k)' \tilde{\Omega}_k^{-1} (y_{kt} - Z_{kt}\theta_k)\end{aligned}\tag{25}$$

L'expression ci-dessus est identique à la vraisemblance du modèle (20). Pour montrer cette identité, il convient de représenter toutes les erreurs intervenant dans (25) (et aussi dans (20)) dans la matrice $T \times N$ suivante :

$$U_k = \begin{bmatrix} u_{k1}^1 \dots u_{k1}^N \\ \vdots \\ u_{kT}^1 \dots u_{kT}^N \end{bmatrix} = \begin{bmatrix} (u_{k1})' \\ \vdots \\ (u_{kT})' \end{bmatrix} = [u_k^1 \dots u_k^N]$$

dans laquelle la t -ème ligne est composée du vecteur $(u_{kt})'$ et la i -ème colonne est formée du vecteur u_k^i .

Analysons le terme sous le signe de sommation de (25) et développons :

$$\begin{aligned}\sum_t (y_{kt} - Z_{kt}\theta_k)' \tilde{\Omega}_k^{-1} (y_{kt} - Z_{kt}\theta_k) &= \sum_t u_{kt}' \tilde{\Omega}_k^{-1} u_k \\ &= \text{tr} \sum_t \tilde{\Omega}_k^{-1} u_{kt} u_{kt}' = \text{tr} \tilde{\Omega}_k^{-1} \sum_t u_{kt} u_{kt}'\end{aligned}$$

d'où :

$$\sum_t (y_{kt} - Z_{kt}\theta_k)' \tilde{\Omega}_k^{-1} (y_{kt} - Z_{kt}\theta_k) = \text{tr} \tilde{\Omega}_k^{-1} U_k' U_k \tag{26}$$

D'autre part, en partant de (20) sous l'hypothèse de normalité de u_k , la vraisemblance logarithmique de y_k s'écrit directement comme :

$$\mathcal{L}(y_k|\theta_k, \tilde{\Omega}_k) = -\frac{T}{2} \ln |\tilde{\Omega}_k| - \frac{1}{2} u_k' (\tilde{\Omega}_k^{-1} \otimes I) u_k \tag{27}$$

dans laquelle, par commodité on a laissé u_k au lieu de $y_k - Z_k\theta_k$.

Notons que $u_k = \text{vec } U_k$, de telle sorte que la forme quadratique dans (27) devient en utilisant les propriétés de l'opérateur vec [Balestra76] :

$$\begin{aligned} u_k'(\tilde{\Omega}_k^{-1} \otimes I)u_k &= (\text{vec } U_k)'(\tilde{\Omega}_k^{-1} \otimes I)\text{vec } U_k \\ &= \text{tr}\tilde{\Omega}_k^{-1}U_k'U_k \end{aligned} \quad (28)$$

Ainsi, l'identité entre les deux vraisemblances est prouvée. ■

4 Procédures d'estimation

Dans un modèle comme le nôtre, dans lequel les variables explicatives (exogènes et endogènes) sont indépendantes des erreurs en raison de la récursivité, deux procédures d'estimation sont envisageables : celle du maximum de vraisemblance et celle des moindres carrés généralisés (appelée également méthode Aitken).

Nous avons montré à la section précédente que la fonction de vraisemblance globale se factorise en K vraisemblances, une pour chaque équation. De l'autre côté, le problème de la minimisation de la somme généralisée des carrés des écarts se décompose lui aussi en la minimisation de K sommes généralisées des carrés, une pour chaque équation. Dès lors, le principe général d'estimation se dégage : l'estimation du modèle par l'une ou l'autre des deux méthodes se ramène à l'estimation séparée, équation par équation, pour l'ensemble des individus.

Nous développons en conséquence dans cette section la méthode du maximum de vraisemblance et la méthode des moindres carrés généralisés pour une équation générique, la k -ème par exemple.

4.1 La méthode du maximum de vraisemblance

4.1.1 Fonction de vraisemblance

La log-vraisemblance pour la k -ème équation (voir section 2) s'écrit comme suit :

$$\begin{aligned} \mathcal{L}(y_k|\theta_k, \tilde{\Omega}_k) &= -\frac{T}{2}\ln|\tilde{\Omega}_k| - \frac{1}{2}(y_k - Z_k\theta_k)'(\tilde{\Omega}_k^{-1} \otimes I)(y_k - Z_k\theta_k) \\ &= -\frac{T}{2}\ln|\tilde{\Omega}_k| - \frac{1}{2}\text{tr}\tilde{\Omega}_k^{-1}U_k'U_k \end{aligned} \quad (29)$$

avec $u_k = y_k - Z_k\theta_k$, $U_k = [u_k^1 \dots u_k^N]$ de dimension $T \times N$ et $u_k = \text{vec } U_k$.

Nous étudions d'abord la solution du problème de maximisation de la log-vraisemblance et dérivons ensuite la distribution limite des estimateurs du maximum de vraisemblance.

4.1.2 Les estimateurs du maximum de vraisemblance

Les conditions de premier ordre pour la maximisation de \mathcal{L} s'écrivent par :

$$(i) \quad \frac{\partial \mathcal{L}}{\partial \tilde{\Omega}_k} = -\frac{T}{2} \tilde{\Omega}_k^{-1} + \frac{1}{2} \tilde{\Omega}_k^{-1} U_k' U_k \tilde{\Omega}_k^{-1} = 0$$

$$(ii) \quad \frac{\partial \mathcal{L}}{\partial \theta_k} = Z_k' (\tilde{\Omega}_k^{-1} \otimes I) (y_k - Z_k \theta_k) = 0$$

ce qui donne les deux résultats suivants :

$$\tilde{\Omega}_k^* = \frac{1}{T} U_k' U_k \quad (30)$$

$$\theta_k^* = [Z_k' (\tilde{\Omega}_k^{-1} \otimes I) Z_k]^{-1} Z_k' (\tilde{\Omega}_k^{-1} \otimes I) y_k \quad (31)$$

Les estimateurs du maximum de vraisemblance de θ_k et $\tilde{\Omega}_k$ sont donnés par la solution simultanée des équations (30) et (31). Le système est non-linéaire et sans solution analytique, mais la structure en “zig-zag” des deux équations suggère une procédure numérique de solution simple et efficace. Elle comporte les pas suivants :

Pas 1 : condition initiale $\tilde{\Omega}_k = I$;

Pas 2 : on calcule θ_k par l'équation (31) et une nouvelle valeur de $\tilde{\Omega}_k$ par (30) ;

Pas 3 : on retourne au pas 2 jusqu'à convergence numérique.

4.1.3 La distribution asymptotique

Puisque chaque individu est caractérisé par son propre ensemble de coefficients de réaction, la seule analyse asymptotique pertinente concerne la situation où le nombre N d'individus est fixe et le nombre T d'observations croît indéfiniment. Dans ce contexte (et il en est de même pour l'étude des propriétés asymptotiques des estimateurs des moindres carrés généralisés de la section 3), nous évaluons et calculons les limites de la matrice Hessienne pour aboutir enfin à la distribution asymptotique.

4.1.3.1 Les dérivées deuxièmes et leurs limites

La dérivée deuxième directe par rapport à θ_k est :

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_k \partial \theta_k'} = -Z_k' (\tilde{\Omega}_k^{-1} \otimes I) Z_k$$

et sa limite est :

$$\lim_{T \rightarrow \infty} -\frac{1}{T} E\left(\frac{\partial^2 \mathcal{L}}{\partial \theta_k \partial \theta_k'}\right) = \lim_{T \rightarrow \infty} \frac{1}{T} Z_k' (\tilde{\Omega}_k^{-1} \otimes I) Z_k = Q_k^* \quad (32)$$

où Q_k^* est une matrice définie positive, dont l'expression analytique est fournie dans la section suivante. Pour calculer la dérivée croisée, on écrit d'abord la dérivée première par rapport à $\tilde{\Omega}_k$ en forme vectorielle :

$$\frac{\partial \mathcal{L}}{\partial \text{vec } \tilde{\Omega}_k} = -\frac{T}{2} \text{vec } \tilde{\Omega}_k^{-1} + \frac{1}{2} (\tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1}) \text{vec } U_k' U_k$$

on obtient alors :

$$\frac{\partial \mathcal{L}^2}{\partial \text{vec } \tilde{\Omega}_k \partial \theta_k'} = \frac{1}{2} (\tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1}) \frac{\partial \text{vec } U_k' U_k}{\partial \theta_k'}$$

or :

$$\begin{aligned} \frac{\partial \text{vec } U_k' U_k}{\partial \theta_k'} &= (I \otimes U_k') \frac{\partial \text{vec } U_k}{\partial \theta_k'} + (U_k' \otimes I) \frac{\partial \text{vec } U_k'}{\partial \theta_k'} \\ &= (I \otimes U_k') \frac{\partial \text{vec } U_k}{\partial \theta_k'} + (U_k' \otimes I) P \frac{\partial \text{vec } U_k}{\partial \theta_k'} \end{aligned} \quad (33)$$

où $\text{vec } U_k = u_k = y_k - Z_k \theta_k$ et P est la matrice unitaire permutée (ou matrice de commutation). Comme $\partial \text{vec } U_k / \partial \theta_k' = -Z_k$, on a :

$$\frac{\partial \mathcal{L}^2}{\partial \text{vec } \tilde{\Omega}_k \partial \theta_k'} = -\frac{1}{2} (\tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1}) [(I \otimes U_k') Z_k + (U_k' \otimes I) P Z_k]$$

et puisque $E(U_k') = 0$, la limite recherchée est :

$$-\frac{1}{T} \lim E \left(\frac{\partial \mathcal{L}^2}{\partial \text{vec } \tilde{\Omega}_k \partial \theta_k'} \right) = 0 \quad (34)$$

Pour la dérivée deuxième directe par rapport à $\text{vec } \tilde{\Omega}_k$ on obtient :

$$\begin{aligned} \frac{\partial \mathcal{L}^2}{\partial \text{vec } \tilde{\Omega}_k \partial (\text{vec } \tilde{\Omega}_k)'} &= \frac{\partial}{\partial (\text{vec } \tilde{\Omega}_k)'} \left[-\frac{T}{2} \text{vec } \tilde{\Omega}_k^{-1} + \frac{1}{2} (\tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1}) \text{vec } U_k' U_k \right] \\ &= \frac{T}{2} (\tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1}) - \frac{1}{2} [\tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1} U_k' U_k \tilde{\Omega}_k^{-1}] - \frac{1}{2} [\tilde{\Omega}_k^{-1} U_k' U_k \tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1}] \end{aligned}$$

En observant que $E(U_k' U_k) = \tilde{\Omega}_k$, la limite est donnée par :

$$-\frac{1}{T} \lim E \left(\frac{\partial \mathcal{L}^2}{\partial \text{vec } \tilde{\Omega}_k \partial (\text{vec } \tilde{\Omega}_k)'} \right) = \frac{1}{2} \tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1} \quad (35)$$

4.1.3.2 La matrice d'information et la distribution asymptotique

D'après ce qui précède et en l'absence de contraintes sur θ_k , la matrice d'information est bloc-diagonale. Le premier bloc, correspondant à l'estimateur de θ_k , est composé de la matrice définie positive $\tilde{\Omega}_k$. Il en résulte que :

$$\sqrt{T}(\tilde{\theta}_k - \theta_k) \rightarrow N(0, \tilde{Q}_k^{*-1}) \quad (36)$$

Quant au deuxième bloc, concernant $\vec{\hat{\Omega}}_k$ (l'estimateur du maximum de vraisemblance de $\tilde{\Omega}_k$), il faut tenir compte de la condition de symétrie de la matrice $\tilde{\Omega}_k$. Deux solutions sont envisageables : soit on élimine les éléments redondants de $\tilde{\Omega}_k$ (et on appellera $\vec{\hat{\Omega}}_{k*}$ le vecteur des éléments distincts de $\vec{\hat{\Omega}}_k$), soit on introduit la matrice d'information bordée. Nous choisissons ici la première.

On montre alors que [Balestra76] la matrice d'information pour $\vec{\hat{\Omega}}_{k*}$ est égale à :

$$\frac{1}{2} Q \bar{P} (\tilde{\Omega}_k^{-1} \otimes \tilde{\Omega}_k^{-1}) \bar{P}' Q'$$

où

- $\bar{P} = P_{N,N} + I - G$, (P la matrice unitaire permutée)
- $G = \text{diag}(e_1 e_1' \dots e_N e_N')$, (e_i le i -ème vecteur élémentaire)
- $Q = \text{diag}(Q_1 \dots Q_N)$ $Q'_i = (e_1 e_2 \dots e_i)$.

On montre également que son inverse, en désignant par \bar{R} la matrice idempotente $\frac{1}{2}(I+P)$, est égal à :

$$2Q\bar{R}((\tilde{\Omega}_k \otimes \tilde{\Omega}_k)\bar{R}Q')$$

on en déduit finalement que :

$$\sqrt{T} \text{vec} (\hat{\Omega}_k - \tilde{\Omega}) \rightarrow N(0, 2Q\bar{R}((\tilde{\Omega}_k \otimes \tilde{\Omega}_k)\bar{R}Q')) \quad (37)$$

5 La méthode des moindres carrés généralisés

5.1 La méthode de l'Aitken pure : modèle et estimateurs

Le modèle individuel de base s'écrit comme suit :

$$y_k^i = Z_k^i \alpha_k^i + u_k^i$$

avec $Z_k^i = (y^{i/k} \ x^{i/k})$ et $\alpha_k^{i'} = (\beta^{i/k'} \ \gamma^{i/k'})$

où :

$y^{i/k}$ est l'ensemble des variables endogènes explicatives dans la k -ème équation, il sera noté par y_* pour alléger les notations ;

$x^{i/k}$ est l'ensemble des variables exogènes figurant dans l'équation, il sera noté par x_Δ ;

α_k^i est un vecteur colonne d'ordre s_k ou encore $K_* + L_\Delta$ avec ($K_* < k$).

L'absence d'auto-corrélation entre les équations individuelles, qui est due au caractère récursif du système individuel de base, s'écrit par :

$$E(u_k^{i'} y_l^i) = 0 \quad \text{pour } k \neq l$$

ce qui fait que les erreurs u_k^i ne sont pas corrélées avec y_\star^i .

Cette caractéristique importante justifie l'emploi de la méthode des moindres carrés ordinaires, ce qui donne :

$$\hat{\alpha}_k^i = (Z_k^{i'} Z_k^i)^{-1} Z_k^{i'} y_k^i$$

En considérant tous les individus le modèle précédent devient :

$$y_k = Z_k \alpha_k + u_k \quad (38)$$

en posant

$$u_k = \begin{bmatrix} u_k^1 \\ \vdots \\ u_k^N \end{bmatrix} \quad \alpha_k = \begin{bmatrix} \alpha_k^1 \\ \vdots \\ \alpha_k^N \end{bmatrix} \quad Z_k = \begin{bmatrix} Z_k^1 & & \\ & \ddots & \\ & & Z_k^N \end{bmatrix}$$

où u_k est un vecteur aléatoire d'ordre $NT \times 1$ avec :

$$E(u_k) = 0 \quad \text{et} \quad E(u_k u_k') = \Omega_k \otimes I_T$$

Soit y_\star^k une matrice d'ordre $NT \times NK_\star$, correspondant à la partie aléatoire de Z_k , elle s'écrit comme suit :

$$y_\star^k = \text{diag} (y_\star^1, \dots, y_\star^N)$$

Un calcul simple de la covariance entre y_\star^k et u_k montre la non-corrélation entre elles, car :

$$E(u_k' y_\star^k) = 0$$

ce qui permet de traiter le modèle avec interdépendance individuelle comme un système de Zellner, et du coup on applique la méthode de l'Aitken pure pour arriver à l'estimation simultanée suivante :

$$\hat{\alpha}_k = [Z_k' (\Omega_k^{-1} \otimes I_T) Z_k]^{-1} Z_k' (\Omega_k^{-1} \otimes I_T) y_k \quad (39)$$

5.1.1 Propriétés de base des estimateurs dans l'échantillon

Les estimateurs $\hat{\alpha}_k^i$ comptent dans l'estimation de la matrice des variances-covariances des erreurs. On montre facilement que ces estimateurs sont centrés ou sans biais. En effet, puisque la quantité

$$E[(Z_k^{i'} Z_k^i)^{-1} Z_k^{i'}]$$

est finie, alors $E(\hat{\alpha}_k^i) = \alpha_k^i$.

Quant à la variance $V(\hat{\alpha}_k^i)$, puisque le vecteur $(Z_k' Z_k)^{-1} Z_k' u_k^i$ et son transposé ne sont pas indépendants, on utilise le lemme suivant sur l'espérance mathématique conditionnelle.

Lemme 1 Soit $g(x_1, x_2)$ une fonction de variables aléatoires, on a :

$$E[g(x_1, x_2)] = E_{x_2}\{E_{x_1/x_2}[g(x_1, x_2)/x_2]\}$$

Ce qui permet d'obtenir finalement que :

$$V(\hat{\alpha}_k^i) = \sigma_k^2 E(Z_k' Z_k)^{-1}$$

On montre aussi qu'il est à variance minimale parmi les estimateurs centrés.

Considérons les paramètres $\hat{\alpha}_k$ du modèle 38, qui correspond à un système de régressions apparemment non-liées. Par l'absence de corrélation entre Z_k et u_k , et puisque

$$E[(Z_k' \Phi_k^{-1} Z_k)^{-1} Z_k' \Phi_k^{-1}]$$

est une quantité finie alors : $E(\hat{\alpha}_k) = \alpha_k$

A l'aide du lemme précédent, la matrice des variances-covariances de $\hat{\alpha}_k$ notée $\Phi(\hat{\alpha}_k)$ s'écrit comme suit :

$$\Phi(\hat{\alpha}_k) = E(\hat{\alpha}_k - \alpha_k)(\hat{\alpha}_k - \alpha_k)' = E(Z_k' \Phi_k^{-1} Z_k)^{-1} \quad (40)$$

Pour montrer que $\hat{\alpha}_k$ est à variance minimale, on suppose habituellement qu'il existe α_k^* un autre estimateur α_k^* tel que :

$$\alpha_k^* = C_k y_k \quad \text{avec} \quad C_k = A_k + B_k$$

où

$$A_k = (Z_k' \Phi_k^{-1} Z_k)^{-1} Z_k' \Phi_k^{-1}$$

B_k est une matrice arbitraire non nulle avec une composante aléatoire non corrélée avec u_k .

En vertu de l'absence de corrélation entre u_k et les matrices B_k et A_k , on obtient :

$$\begin{aligned} E(\alpha_k^*) &= \alpha_k \quad \text{si} \quad E(B_k Z_k) = 0 \\ \Phi(\alpha_k^*) &= E(Z_k' \Phi_k^{-1} Z_k)^{-1} + E(B_k \Phi_k B_k') \end{aligned}$$

puisque Φ_k est définie positive et B_k non nulle alors $B_k \Phi_k B_k'$ est définie non négative ; il en est de même pour son espérance mathématique. Par conséquent $\hat{\alpha}_k$ est plus efficace que α_k^* car :

$$\Phi(\hat{\alpha}_k) \leq \Phi(\alpha_k^*)$$

5.1.2 Convergence et distribution asymptotique

4.2.3.1 Convergence des estimateurs

Pour montrer la convergence de $\hat{\alpha}_k$, on a besoin de formulations appropriées des hypothèses H3 et H5 sur le comportement asymptotique des variables explicatives.

Soit $Z_k^i = (y_*^i \quad x_{\Delta}^i)$, une matrice composée d'une partie aléatoire y_*^i et d'une partie fixe x_{Δ}^i , on postule que :

H3b : Soit $x_{\Delta t}^{i'} = x_t^{i'} L$ où $L' = (I \quad O)$ le t -ème vecteur de x_{Δ}^i avec L_{Δ} composantes. $x_{\Delta t}^i$ est un vecteur de variables certaines uniformément bornées telle que :

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_t x_{\Delta t}^i x_{\Delta t}^{i'} = \lim_{T \rightarrow \infty} \frac{1}{T} x_{\Delta}^{i'} x_{\Delta}^i = M_{\Delta}^i \text{ pour tout } i$$

avec M_{Δ}^i définie non négative (dnn) et M_{Δ}^i bornée.

H4b : Soit y_{*t}^i le t -ème vecteur, avec K_* variables aléatoires, en nombre fini, indépendantes d'une équation à une autre et identiquement distribuées, de la matrice y_*^i d'ordre $T \times K_*$ définie par la forme réduite suivante : $y_*^i = x^i \Pi_*^i + v_*^i$ avec

$$\begin{aligned} E(v_{*t}^i) &= 0 \\ E(v_{*t}^i v_{*t}^{i'}) &= \Omega_*^i \text{ définie positive} \end{aligned}$$

(a) Détermination de $\text{plim} \frac{1}{T} Z_k' \Phi_k^{-1} Z_k$

On va tout d'abord calculer la limite en probabilité de $\hat{\alpha}_k$ en partant du modèle suivant, qui ressemble au modèle 8 : $y_{kt} = Z_{kt} \alpha_k + u_{kt}$ où Z_{kt} est d'ordre $N \times N_{sk}$, et avec

$$E(u_{kt}) = 0 \quad \text{et} \quad E(u_{kt} u_{kt}') = \Omega_k$$

on obtient :

$$\begin{aligned} \text{plim } \hat{\alpha}_k &= \text{plim} \left(\frac{1}{T} Z_k' \Phi_k^{-1} Z_k \right)^{-1} \frac{1}{T} Z_k' \Phi_k^{-1} y_k \\ &= \text{plim} \left(\frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} Z_{kt} \right)^{-1} \frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} y_{kt} \end{aligned} \quad (41)$$

Le calcul du premier facteur de cette dernière expression donne une matrice dont le terme générique est défini par :

$$\sum_t Z_{kt}' \Omega_k^{-1} Z_{kt} = \sigma_{ij}^k \sum_t Z_{kt}^i Z_{kt}^{j'} = \sigma_{ij}^k Z_k^{i'} Z_k^j = \left[\sigma_{ij}^k \begin{pmatrix} \sum_t y_{*t}^i y_{*t}^{j'} & \sum_t y_{*t}^i x_{\Delta t}^{j'} \\ \sum_t x_{\Delta t}^i y_{*t}^{j'} & \sum_t x_{\Delta t}^i x_{\Delta t}^{j'} \end{pmatrix} \right] \quad (42)$$

et dont on calcule alors les limites en probabilité des trois quantités distinctes.

(a.1) Calcul de $\text{plim } \frac{1}{T} \sum_t y_{*t}^i y_{*t}^{j'}$

$$\text{plim } \frac{1}{T} y_{*}^{i'} y_{*}^j = \text{plim } \frac{1}{T} (\Pi_{*}^{i'} x^{i'} x^j \Pi_{*}^j + \Pi_{*}^{i'} x^{i'} v_{*}^j + v_{*}^{i'} x^j \Pi_{*}^j + v_{*}^{i'} v_{*}^j) \quad (43)$$

on obtient par l'hypothèse H3 :

- $\text{plim } \frac{1}{T} (\Pi_{*}^{i'} x^{i'} x^j \Pi_{*}^j = \Pi_{*}^{i'} (\text{plim } \frac{1}{T} x^{i'} x^j) \Pi_{*}^j = \Pi_{*}^{i'} M^{ij} \Pi_{*}^j$

pour le second et le troisième termes du résultat 43, nous procédons sur le résultat suivant :

- $\text{plim } \frac{1}{T} \Pi_{*}^{i'} x^{i'} v_{*}^j = \Pi_{*}^{i'} \text{plim } \frac{1}{T} x^{i'} v_{*}^j$

par la propriété de la convergence quadratique, qui implique la convergence en probabilité :

(i) $E(\frac{1}{T} x^{i'} v_{*}^j) = \frac{1}{T} x^{i'} E(v_{*}^j) = 0$

(ii) $V(\frac{1}{T} x^{i'} v_{*}^j) = \frac{1}{T^2} x^{i'} E(v_{*}^j v_{*}^{j'}) x^i = \frac{1}{T^2} x^{i'} \sum_{k=1}^{K_{*}} \delta_k^{jj} x^i$

alors $\text{plim } V(\frac{1}{T} x^{i'} v_{*}^j) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_k \delta_k^{jj} \lim_{T \rightarrow \infty} \frac{1}{T} x^{i'} x^i = 0 \times M^i = 0,$

par conséquent : $\text{plim } \frac{1}{T} x^{i'} v_{*}^j = 0$

quant au dernier terme, on utilise le théorème suivant de Khintchine étendu au cas de vecteurs aléatoires.

Théorème Soit un échantillon de vecteurs aléatoires indépendants : $X_t (t = 1, 2, \dots, T)$ tel que $X_t \sim \text{iid } (U, \Omega)$ alors

$$\text{plim } \frac{1}{T} \sum_t X_t = U \text{ et } \text{plim } \frac{1}{T} \sum_t (X_t - \bar{X})(X_t - \bar{X})' = \Omega.$$

Ce qui nous donne :

- $\text{plim } \frac{1}{T} v_{*}^{i'} v_{*}^j = \text{plim } \frac{1}{T} \sum_t v_{*t}^i v_{*t}^{j'} = E(v_{*t}^i v_{*t}^{j'}) = \Omega_{*}^{ij}$

cette matrice des variances-covariances n'est pas forcément définie positive, sauf pour si i est égale à j .

Il en résulte que :

$$\text{plim } \frac{1}{T} y_{*}^{i'} y_{*}^j = \Pi_{*}^{i'} M^{ij} \Pi_{*}^j + \Omega_{*}^{ij} := Q_{*}^{ij}$$

(a.2) Calcul de $\text{plim } \frac{1}{T} \sum_t y_{\Delta t}^i x_{\Delta t}^{j'}$

$$\text{plim } \frac{1}{T} \sum_t (\Pi_{\star}^i x_t^i + v_{\star t}^i) x_{\Delta t}^{j'} = \text{plim } \frac{1}{T} \sum_t \Pi_{\star}^i x_t^i x_{\Delta t}^{j'} + \text{plim } \frac{1}{T} \sum_t v_{\star t}^i x_{\Delta t}^{j'}$$

pour le premier terme, en utilisant l'hypothèse H3c, nous avons :

$$\bullet \text{plim } \frac{1}{T} \sum_t \Pi_{\star}^i x_t^i x_{\Delta t}^{j'} = \Pi_{\star}^i \text{plim } \frac{1}{T} \sum_t x_t^i x_{\Delta t}^{j'} = \Pi_{\star}^i \text{plim } \frac{1}{T} \sum_t x_t^i x_t^{j'} L = \Pi_{\star}^i M^{ij} L$$

pour le second terme de l'expression ci-dessus, on a :

$$\bullet \text{plim } \frac{1}{T} \sum_t v_{\star t}^i x_{\Delta t}^{j'} = (\text{plim } \frac{1}{T} \sum_t v_{\star t}^i x_t^{j'}) L = (\text{plim } \frac{1}{T} v_{\star}^i x^j) L = 0 \times L = 0$$

il en résulte que :

$$\text{plim } \frac{1}{T} \sum_t y_{\Delta t}^i x_{\Delta t}^{j'} = \Pi_{\star}^i M^{ij} L \quad \text{et} \quad \text{plim } \frac{1}{T} \sum_t x_{\Delta t}^i y_{\star t}^j = L' M^{ij} \Pi_{\star}^j$$

Ainsi, le résultat cherché se détermine comme suit :

$$\text{plim } \frac{1}{T} \sum_t Z_{kt}^i Z_{kt}^{j'} = \begin{bmatrix} Q_{\star}^{ij} & \Pi_{\star}^i M^{ij} L \\ L' M^{ij} \Pi_{\star}^j & M_{\Delta}^{ij} \end{bmatrix} := Q_k^{ij} \quad (44)$$

Finalement, la limite en probabilité du premier facteur de la formulation 41 est la quantité suivante :

$$\text{plim } \frac{1}{T} Z_k' \Phi_k^{-1} Z_k = [(\sigma_{ij}^k Q_k^{ij})] := Q_k^* \quad (45)$$

(a.3) Définition de Q_k^*

Pour préciser la définition de Q_k^* , on manipule la matrice suivante :

$$Q_k^* = \begin{bmatrix} \sigma_{11}^k Q_k^{11} & \sigma_{12}^k Q_k^{12} & \dots & \sigma_{1N}^k Q_k^{1N} \\ \sigma_{21}^k Q_k^{21} & \sigma_{22}^k Q_k^{22} & \dots & \sigma_{2N}^k Q_k^{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1}^k Q_k^{N1} & \sigma_{N2}^k Q_k^{N2} & \dots & \sigma_{NN}^k Q_k^{NN} \end{bmatrix}$$

(a.3.1) Les conditions d'inversion de Q_k^i

Au préalable, l'étude de la définition de Q_k^* nécessite de chercher les conditions d'inversion des matrices Q_k^i ou ce qui revient au même, de déterminer leurs définitions.

Par les hypothèses H3 et H3b, M^i et Ω_\star^i sont définies positives ; soit Π_\star^i une matrice non nulle de coefficients, de rang complet et d'ordre $L \times K_\star$ avec $L > K_\star$, on a :

$$Q_\star^i = \Pi_\star^{i'} M^i \Pi_\star^i + \Omega_\star^i$$

Théorème[Balestra72] Soit A et B deux matrices. Si A d'ordre n est définie positive et B d'ordre (n, m) et de rang complet égale à m , alors $B'AB$ est définie positive.

Par ce théorème $\Pi_\star^{i'} M^i \Pi_\star^i$ est définie positive, et puisque la somme de deux matrices définies positives est définie positive, alors Q_\star^i est définie positive. Par conséquent, Q_\star^i est inversible.

Une autre procédure de déterminer les conditions d'inversion de Q_k^i , consiste à calculer son déterminant, ce qui nous donne :

$$|Q_k^i| = |Q_\star^i| |P^i| = |M_\Delta^i| |R^i|$$

avec

$$\begin{aligned} P^i &= M_\Delta^i - L' M^i \Pi_\star^i (Q_\star^i)^{-1} \Pi_\star^{i'} M^i L := M_\Delta^i - B_\star^i \\ R^i &= Q_\star^i - \Pi_\star^{i'} M^i L (M_\Delta^i)^{-1} L' M^i \Pi_\star^i \end{aligned}$$

En choisissant de partir de P^i , l'expression $L' M^i \Pi_\star^i$ peut s'écrire comme suit :

$$L' M^i \Pi_\star^i = M_\Delta^i \Pi_{\star_\Delta}^i + M_{\Delta,\Delta\Delta}^i \Pi_{\star_{\Delta\Delta}}^i \text{ et } \Pi_\star^{i'} M^i L = \Pi_{\star_\Delta}^{i'} M_\Delta^i + \Pi_{\star_{\Delta\Delta}}^{i'} M_{\Delta\Delta,\Delta}^i$$

nous avons :

$$B_\star^i = M_\Delta^i \Pi_{\star_\Delta}^i (Q_\star^i)^{-1} \Pi_{\star_\Delta}^{i'} M_\Delta^i + M_{\Delta,\Delta\Delta}^i \Pi_{\star_{\Delta\Delta}}^i (Q_\star^i)^{-1} \Pi_{\star_{\Delta\Delta}}^{i'} M_{\Delta\Delta,\Delta}^i + 2 M_\Delta^i \Pi_{\star_\Delta}^i (Q_\star^i)^{-1} \Pi_{\star_{\Delta\Delta}}^{i'} M_{\Delta\Delta,\Delta}^i$$

Par conséquence, la matrice P^i peut s'écrire par :

$$P^i = M_\Delta^i (I_\Delta - G_{\star_\Delta}^i) \quad \text{avec } G_{\star_\Delta}^i = (M_\Delta^i)^{-1} B_\star^i \text{ et } G_{\star_\Delta}^i \neq I_\Delta$$

et dont le déterminant est développé comme suit :

$$|P^i| = |M_\Delta^i| |I_\Delta - G_{\star_\Delta}^i| = (-1)^{L_\Delta} |M_\Delta^i| |G_{\star_\Delta}^i - I_\Delta|$$

Il s'en suit que :

$$\begin{aligned} |Q_k^i| &= (-1)^{L_\Delta} |Q_\star^i| |M_\Delta^i| |G_{\star_\Delta}^i - I_\Delta| \\ |Q_k^i| &\neq 0 \quad \text{si} \quad |G_{\star_\Delta}^i - I_\Delta| \neq 0 \end{aligned} \tag{46}$$

Si $G_{\star_\Delta}^i$ admet L_Δ valeurs propres différentes alors $G_{\star_\Delta}^i$ est diagonalisable et,

$$|I_\Delta - G_{\star_\Delta}^i| = \prod_{l=1}^{L_\Delta} (1 - \lambda_l)$$

par conséquent, la condition pour que Q_k^i soit inversible est la suivante :

$$|G_{\Delta}^i - I_{\Delta}| \neq 0 \quad \text{si } \lambda_l \neq 1 \quad \text{pour tout } l = 1, 2, \dots, l_{\Delta}$$

Pour commencer, calculons le déterminant de la matrice suivante en prenant le cas où N est égale à 2, ce qui nous donne :

$$|Q_k^*| = |\sigma_{11}^k Q_k^1| |P_k|$$

$$\text{où } P_k = \sigma_{22}^k Q_k^2 - \sigma_{21}^k Q_k^{21} (\sigma_{11}^k Q_k^1)^{-1} \sigma_{12}^k Q_k^{12} = a[bQ_k^2 - Q_k^{21}(Q_k^1)^{-1}Q_k^{12}]$$

$$\text{avec } a = \frac{(\sigma_{12}^k)^2}{\sigma_{11}^k} \quad \text{et} \quad b = \frac{\sigma_{11}^k \sigma_{22}^k}{(\sigma_{12}^k)^2}$$

puisque, d'après ce qui précède, la matrice Q_k^i est régulière pour tout i , on a donc :

$$P_k = aQ_k^2[bI - (Q_k^2)^{-1}Q_k^{21}(Q_k^1)^{-1}Q_k^{12}]$$

d'où :

$$\begin{aligned} |Q_k^*| &= |\sigma_{11}^k Q_k^1| |aQ_k^2| |bI - (Q_k^2)^{-1}Q_k^{21}(Q_k^1)^{-1}Q_k^{12}| \\ &= (-1)^{s_k} (\sigma_{12}^k)^{2s_k} |Q_k^1| |Q_k^2| |(Q_k^2)^{-1}Q_k^{21}(Q_k^1)^{-1}Q_k^{12} - bI| \end{aligned}$$

ainsi, la condition d'inversion est donnée par :

$$|Q_k^*| \neq 0 \text{ si les valeurs propres de } (Q_k^2)^{-1}Q_k^{21}(Q_k^1)^{-1}Q_k^{12} \text{ sont différentes de } b.$$

Par ailleurs, la forme quadratique associée à Q_k^* pour $X' = (X^1' \ X^2')$ non nul est définie par : $q(X) = X'Q_k^*X$, ce qui donne en utilisant la symétrie de la matrice :

$$\begin{aligned} q(X) &= \sigma_{11}^k X^1' Q_k^1 X^1 + \sigma_{21}^k X^2' Q_k^{21} X^1 + \sigma_{12}^k X^1' Q_k^{12} X^2 + \sigma_{22}^k X^2' Q_k^2 X^2 \\ &= \sigma_{11}^k X^1' Q_k^1 X^1 + 2\sigma_{21}^k X^2' Q_k^{21} X^1 + \sigma_{22}^k X^2' Q_k^2 X^2 \end{aligned}$$

Théorème [Acher et al.70] Toute matrice A symétrique associée à une forme définie positive ou définie négative est non-singulière.

Puisque Q_k^i est définie positive alors pour tout i $X^{i'} Q_k^i X^i > 0$. Si Q_k^{ij} est définie non négative alors $q(X) > 0$ pour tout X non nul, par conséquent : Q_k^* est définie positive.

Dans le cas général, la forme quadratique associée à Q_k^* , pour tout X^i non nul, s'écrit comme suit :

$$q(X) = \sum_i \sigma_{ii}^k X^{i'} Q_k^i X^i + \sum_i \sum_{j \neq i} \sigma_{ij}^k X^{i'} Q_k^{ij} X^j \quad (47)$$

Si on suppose que Q_k^{ij} est définie non négative, alors $X^{i'} Q_k^{ij} X^j \geq 0$. Il s'en suit que : $q(X) > 0$ pour tout $X \neq 0$. Finalement, d'après l'hypothèse précédente et le dernier théorème, on obtient le résultat suivant :

$$|Q_k^*| \neq 0 \text{ si } X^{i'} Q_k^{ij} X^j \geq 0$$

(a.3.2) Condition d'inversion de Q_k^*

Une autre façon de traiter la matrice Q_k^* est de chercher les conditions de son inversion en utilisant l'algorithme de Gauss généralisé [Gantmacher66]. En effet, au i -ème bloc ligne de Q_k^* , ajoutons le premier bloc ligne multiplié à gauche par :

$$-\sigma_{i1}^k Q_k^{i1} (\sigma_{11}^k Q_k^{11})^{-1} \quad \text{pour } i = 2, 3, \dots, N$$

nous obtenons la matrice suivante que l'on note $Q_k^*(1)$:

$$Q_k^*(1) = \begin{bmatrix} \sigma_{11}^k Q_k^{11} & \sigma_{12}^k Q_k^{12} & \dots & \sigma_{1N}^k Q_k^{1N} \\ 0 & P_k^{22}(1) & \dots & P_k^{2N}(1) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & P_k^{N2}(1) & \dots & P_k^{NN}(1) \end{bmatrix} \quad (48)$$

où $P_k^{ij}(1) = \sigma_{ij}^k Q_k^{ij} - \sigma_{i1}^k Q_k^{i1} (\sigma_{11}^k Q_k^{11})^{-1} \sigma_{1j}^k Q_k^{1j}$ pour $i, j = 2, 3, \dots, N$

Si la matrice $P_k^{22}(1)$ est régulière au même titre que Q_k^1 , le processus itératif peut être continué. En effet, on a :

$$|P_k^{22}(1)| = (-1)^{s_k} a^{s_k} |Q_k^2| |((Q_k^2)^{-1} Q_k^{21} (Q_k^1)^{-1} Q_k^{12} - bI|$$

par conséquent, $|P_k^{22}(1)| \neq 0$ si les valeurs propres de $(Q_k^2)^{-1} Q_k^{21} (Q_k^1)^{-1} Q_k^{12}$ sont différentes de b .

C'est ainsi, que nous parvenons à la formulation suivante, où $(P_k^{ij}(1))$ est un bloc générique :

$$|Q_k^*| = |Q_k^*(1)| = |\sigma_{11}^k Q_k^{11}| |(P_k^{ij}(1))| \quad i, j = 2, 3, \dots, N \quad (49)$$

En posant $Q_k^*(2) = [(P_k^{ij}(1))]$, et à son i -ème bloc, ajoutons le premier bloc ligne multiplié à gauche par :

$$-P_k^{i2}(1) (P_k^{22}(1))^{-1} \quad \text{pour } i = 3, 4, \dots, N$$

nous obtenons la matrice suivante que l'on note $Q_k^*(3)$:

$$Q_k^*(3) = \begin{bmatrix} P_k^{22}(1) & P_k^{23}(1) & \dots & P_k^{2N}(1) \\ 0 & P_k^{33}(2) & \dots & P_k^{3N}(2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & P_k^{N3}(2) & \dots & P_k^{NN}(2) \end{bmatrix} \quad (50)$$

où

$$P_k^{ij}(2) = P_k^{ij}(1) - P_k^{i2}(1) (P_k^{22}(1))^{-1} P_k^{2j}(1) \quad \text{pour } i, j = 3, 4, \dots, N$$

Si la matrice $P_k^{33}(2)$ est régulière, le processus itératif continue. En effet :

$$|P_k^{33}(2)| = (-1)^{s_k} |P_k^{33}(1)| |(P_k^{33}(1))^{-1} P_k^{32}(1) (P_k^{22}(1))^{-1} P_k^{23}(1) - I|$$

par conséquent,

$|P_k^{33}(1)| \neq 0$ si les valeurs propres de $(Q_k^3)^{-1} Q_k^{31} (Q_k^1)^{-1} Q_k^{13}$ sont différentes de $\sigma_{11}^k \sigma_{33}^k / (\sigma_{13}^k)^2$.

En posant $\bar{P}_k(1) = (P_k^{33}(1))^{-1} P_k^{32}(1) (P_k^{22}(1))^{-1} P_k^{23}(1)$ avec $\bar{P}_k(1) \neq I$, et en faisant le calcul, on trouve que :

$$\begin{aligned} \bar{P}_k(1) &= b_0 [Q_k^3 - b_1 Q_k^{21} (Q_k^1)^{-1} Q_k^{12}]^{-1} [Q_k^{32} - b_2 Q_k^{31} (Q_k^1)^{-1} Q_k^{12}] \\ &\quad [Q_k^2 - b_3 Q_k^{21} (Q_k^1)^{-1} Q_k^{12}]^{-1} [Q_k^{23} - b_4 Q_k^{21} (Q_k^1)^{-1} Q_k^{13}] \end{aligned}$$

avec

$$b_0 = \frac{\sigma_{32}^k \sigma_{23}^k}{\sigma_{22}^k \sigma_{33}^k} \quad b_1 = \frac{(\sigma_{13}^k)^2}{\sigma_{11}^k \sigma_{33}^k} \quad b_2 = \frac{\sigma_{11}^k \sigma_{32}^k}{\sigma_{12}^k \sigma_{31}^k} \quad b_3 = \frac{(\sigma_{21}^k)^2}{\sigma_{11}^k \sigma_{22}^k} \quad \text{et} \quad b_4 = \frac{\sigma_{11}^k \sigma_{23}^k}{\sigma_{13}^k \sigma_{21}^k}$$

En posant, $\bar{P}_k(1) = b_0 A_k$, on peut écrire que $|\bar{P}_k(1) - I| = b_0^{2k} |A_k - b_0^{-1} I|$, ce qui nous permet de déduire que :

$$|\bar{P}_k(1) - I| \neq 0 \quad \text{si les valeurs propres de } A_k \text{ sont différentes de } b_0^{-1}.$$

Par conséquent, on obtient que :

$$|P_k^{33}(2)| \neq 0 \quad \text{si } |P_k^{33}(1)| \neq 0 \quad \text{et } |\bar{P}_k(1) - I| \neq 0$$

Comme $P_k^{33}(2)$ est régulière sous les conditions précisées ci-dessus, le processus itératif continue.

La formule donnant le déterminant de Q_k^* à partir de $Q_k^*(1)$ réduit le calcul de $|Q_k^*|$ constitué par N^2 blocs, au calcul d'un déterminant d'ordre moins élevé constitué par $(N-1)^2$ blocs. Ce dernier déterminant peut être soumis au processus itératif décrit ci-dessus.

Lorsqu'on arrive à un déterminant partitionné en $[N - (N-2)]^2$ blocs nous aurons ce qui suit :

$$|Q_k^*| = |\sigma_{11}^k Q_k^1| |P_k^{22}(1)| \dots |P_k^{jj}(j-1)| \dots \begin{vmatrix} P_k^{N-1, N-1}(N-2) & P_k^{N-1, N}(N-2) \\ P_k^{N, N-1}(N-2) & P_k^{N, N}(N-2) \end{vmatrix}$$

comme toutes les matrices $P_k^{jj}(j-1)$ sont régulières selon des conditions précises, nous considérons le dernier déterminant comme suit :

$$|Q_k^*(N-2)| = \begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} A & B \\ 0 & D - CA^{-1}B \end{vmatrix} = \begin{vmatrix} A - BD^{-1}C & 0 \\ C & D \end{vmatrix}$$

ce déterminant est obtenu de la façon suivante : à la deuxième ligne, on a ajouté la première ligne multipliée à gauche par $-CA^{-1}$ ($|A| \neq 0$). Le membre de droite est obtenu de la même façon avec $|D| \neq 0$.

Dans notre cas, les quatre matrices sont carrées et d'ordre s_k . On peut déduire de la précédente formulation les formules de Schur, lesquelles réduisent les calculs d'un déterminant d'ordre $2s_k$ au calcul d'un déterminant d'ordre s_k , ce qui nous donne :

$$|Q_k^*(N-2)| = |AD - AC'A^{-1}B| = |AD - BD^{-1}CD|$$

si A et C sont permutables i.e $AC = CA$ alors :

$$|Q_k^*(N-2)| = |AD - CB|$$

de même, si $DC = CD$ on a :

$$|Q_k^*(N-2)| = |AD - BC|$$

par conséquent, $|Q_k^*(N-2)| \neq 0$ si les valeurs propres de $D^{-1}A^{-1}CB$ sont différentes de 1.

Finalement, on obtient la formulation suivante :

$$|Q_k^*| = (\sigma_{11}^k)^{s_k} |Q_k^1| |P_k^{22}(1)| \dots |P_k^{jj}(j-1)| \dots |Q_k^*(N-2)| \neq 0 \quad \text{si} \quad |Q_k^*(N-2)| \neq 0 \quad (51)$$

Ainsi, cette condition de l'inversion montre qu'il s'agit d'une quantité finie, et le résultat final est donc le suivant :

$$Q_k^* = \text{plim} \frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} Z_{kt} \text{ est définie positive} \quad (52)$$

(b) Détermination de $\text{plim} \frac{1}{T} Z_k' \Phi_k^{-1} u_k$

Pour achever le calcul de $\text{plim} \hat{\alpha}_k$, on doit déterminer maintenant la limite en probabilité suivante, qui permet d'obtenir :

$$\text{plim} \frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} u_{kt} = \text{plim} \frac{1}{T} \sum_t \begin{bmatrix} \sum_j \sigma_{1j}^k \begin{pmatrix} y_{*t}^1 \\ x_{\Delta t}^1 \end{pmatrix} u_{kt}^j \\ \vdots \\ \sum_j \sigma_{Nj}^k \begin{pmatrix} y_{*t}^N \\ x_{\Delta t}^N \end{pmatrix} u_{kt}^j \end{bmatrix} \quad (53)$$

il s'agit de calculer pour tout i les quantités suivantes :

$$\text{plim} \frac{1}{T} \sum_t y_{*t}^i u_{kt}^j \quad \text{et} \quad \text{plim} \frac{1}{T} \sum_t x_{\Delta t}^i u_{kt}^j$$

- $\text{plim} \frac{1}{T} \sum_t x_{\Delta t}^i u_{kt}^j$

par la propriété de la convergence quadratique, pour tout $i, j = 1, 2, \dots, N$, on obtient que :

$$(i) \quad E\left(\frac{1}{T} \sum_t x_{\Delta t}^i u_{kt}^j\right) = \frac{1}{T} \sum_t x_{\Delta t}^i E(u_{kt}^j) = 0$$

$$(ii) \quad V\left(\frac{1}{T} \sum_t x_{\Delta t}^i u_{kt}^j\right) = \frac{1}{T^2} V(x_{\Delta}^i u_k^j) = \frac{1}{T^2} x_{\Delta}^i V(u_k^j) x_{\Delta}^i = \frac{1}{T^2} \sigma_k^{jj} x_{\Delta}^i x_{\Delta}^i$$

on en déduit que :

$$\text{plim } V\left(\frac{1}{T} \sum_t x_{\Delta t}^i u_{kt}^j\right) = \text{plim } \frac{1}{T} \sigma_k^{jj} \text{plim } \frac{1}{T} x_{\Delta}^i x_{\Delta}^i = 0 \times Q'_{\Delta} = 0$$

par conséquent, avec la convergence en probabilité, on obtient :

$$\text{plim } \frac{1}{T} \sum_t x_{\Delta t}^i u_{kt}^j = 0 \quad \text{pour tout } i, j = 1, 2, \dots, N$$

D'après ce qui précède, puisque $v_{*t}^i = (B_{*+}^i)^{-1} u_{*t}^i$ et en utilisant le théorème de Khintchine, on arrive au résultat suivant :

$$\begin{aligned} \bullet \text{plim } \frac{1}{T} \sum_t y_{*t}^i u_{kt}^j &= \Pi_{*+}^{i'} \text{plim } \frac{1}{T} \sum_t x_t^i u_{kt}^j + \text{plim } \frac{1}{T} \sum_t v_{*t}^i u_{kt}^j \\ &= \text{plim } \frac{1}{T} \sum_t v_{*t}^i u_{kt}^j = (B_{*+}^i)^{-1} \text{plim } \frac{1}{T} \sum_t u_{kt}^j u_{*t}^i \\ &= \text{Cov}(u_{kt}^j, u_{*t}^i) \end{aligned}$$

or,

$$\text{Cov}(u_{kt}^j, u_{*t}^i) = \begin{bmatrix} E(u_{kt}^j u_{1t}^i) \\ \vdots \\ E(u_{kt}^j u_{lt}^i) \\ \vdots \\ E(u_{kt}^j u_{k-1,t}^i) \end{bmatrix} = [E(u_{kt}^j u_{lt}^i)] = 0$$

par conséquent, en raison de la récursivité du système initial de base i.e. $E(u_{kt}^i u_{lt}^j) = 0$ pour tout $k \neq l$, nous avons pour tout $i, j = 1, 2, \dots, N$, le résultat suivant :

$$\text{plim } \frac{1}{T} \sum_t v_{*t}^i u_{kt}^j = 0$$

$$\text{et donc, } \text{plim } \frac{1}{T} \sum_t y_{*t}^i u_{kt}^j = 0$$

Finalement on obtient le résultat suivant :

$$\text{plim } \frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} u_{kt} = 0 \tag{54}$$

Nous concluons donc par la convergence de $\hat{\alpha}_k$ car :

$$\begin{aligned} \text{plim } \hat{\alpha}_k &= \alpha_k + \text{plim } \left(\frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} Z_{kt} \right)^{-1} \text{plim } \frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} u_{kt} \\ &= \alpha_k + (Q_k^*)^{-1} \times 0 = \alpha_k \end{aligned} \quad (55)$$

4.2.3.2 La distribution asymptotique

On s'intéresse maintenant à la propriété de l'efficacité asymptotique de $\hat{\alpha}_k$, qui nous permettra de caractériser sa distribution limite.

L'analyse asymptotique consiste à préciser la distribution asymptotique de l'expression suivante :

$$\sqrt{T}(\hat{\alpha}_k - \alpha_k) = \left(\frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} Z_{kt} \right)^{-1} \frac{1}{\sqrt{T}} \sum_t Z_{kt}' \Omega_k^{-1} u_{kt}$$

Le premier facteur est déjà analysé, il suffit alors d'étudier le second facteur.

Théorème centrale généralisé Soit $\{y_t\}$ une suite de vecteurs aléatoires indépendants et identiquement distribués où y_t est d'ordre $m \times 1$ tel que $y_t \sim \text{iid}(0, \Omega)$ avec Ω définie positive; et soit la matrice Z_T définie par :

$$Z_T = \frac{1}{\sqrt{T}} \sum_t X_t y_t$$

où X_t est une matrice d'ordre $k \times m$ de constantes certaines et uniformément bornées i.e. $|X_t| < R$ pour tout t avec R finie et telle que :

$$\text{plim } \frac{1}{T} \sum_t X_t \Omega X_t' = M$$

avec M finie et définie et positive,

alors Z_T tend en distribution vers le vecteur aléatoire Z tel que :

$$Z_T \xrightarrow{(d)} Z \sim N(0, M).$$

Ce théorème ne peut être directement appliqué dans notre cas, car Z_{kt}' , en contenant une partie aléatoire, n'est pas totalement fixe.

Seulement une écriture pertinente du vecteur $Z_{kt}' \Omega_k^{-1} u_{kt}$ nous permet d'obtenir ce qui suit :

$$Z_{kt}' \Omega_k^{-1} u_{kt} = \left[\left(\sum_{j=1}^N \sigma_{ij}^k Z_{kt}^i u_{kt}^j \right) \right] \quad \text{pour } i = 1, 2, \dots, N \quad (56)$$

où le terme générique est :

$$Z_{kt}^i u_{kt}^j = \begin{pmatrix} \Pi_{*}^{i'} x_t^i + v_{*t}^i \\ x_{\Delta t}^i \end{pmatrix} u_{kt}^j$$

pour séparer les variables certaines de celles aléatoires, il suffit d'écrire le vecteur Z_{kt}^i sous la forme d'un produit matriciel, ce qui donne :

$$\begin{pmatrix} \Pi_{\star}^{i'} x_t^i + v_{\star_t}^i \\ x_{\Delta_t}^i \end{pmatrix} = \begin{pmatrix} \Pi_{\star}^{i'} x_t^i & I_{\star} \\ x_{\Delta_t}^i & 0_{\Delta^{\star}} \end{pmatrix} \begin{pmatrix} 1 \\ v_{\star_t}^i \end{pmatrix} = A_t^i \begin{pmatrix} 1 \\ v_{\star_t}^i \end{pmatrix} \quad (57)$$

où A_t^i est une matrice d'ordre $s_k \times (1 + K_{\star})$, non aléatoire et uniformément bornée.

En posant la composante aléatoire suivante

$$g_{kt}^{ij} = \begin{pmatrix} u_{kt}^j \\ v_{\star_t}^i u_{kt}^j \end{pmatrix}$$

nous avons donc :

$$Z_{kt}^i u_{kt}^j = A_t^i g_{kt}^{ij} \quad \text{pour } i, j = 1, 2, \dots, N \quad (58)$$

Les g_{kt}^{ij} sont des vecteurs indépendants et identiquement distribués tels que :

$$\begin{aligned} E(g_{kt}^{ij}) &= E \begin{pmatrix} u_{kt}^j \\ v_{\star_t}^i u_{kt}^j \end{pmatrix} = 0 \\ V(g_{kt}^{ij}) &= E(u_{kt}^j)^2 E \begin{pmatrix} 1 & v_{\star_t}^{i'} \\ v_{\star_t}^i & v_{\star_t}^i v_{\star_t}^{i'} \end{pmatrix} = \sigma_k^{jj} \begin{pmatrix} 1 & 0' \\ 0 & \Omega_{\star}^i \end{pmatrix} \end{aligned}$$

Par l'hypothèse H4b, Ω_{\star}^i est définie positive, et donc $V(g_{kt}^{ij})$ est définie positive.

On montre facilement l'expression suivante, qui est, d'après le résultat 46, définie positive :

$$\text{plim } \frac{1}{T} \sum_t A_t^i V(g_{kt}^{ij}) A_t^{i'} = \sigma_k^{jj} Q_k^i \quad (59)$$

Finalement, pour l'ensemble des individus, on obtient la formulation suivante :

$$Z_{kt}' \Omega_k^{-1} u_{kt} = [(\sum_j \sigma_{ij}^k A_t^i g_{kt}^{ij})] = A_t [(\sum_j \sigma_{ij}^k g_{kt}^{ij})] := A_t g_{kt} \quad (60)$$

où $A_t = \text{diag} (A_t^1, \dots, A_t^i, \dots, A_t^N)$ est une matrice de constantes certaines et uniformément bornée et g_{kt} est un vecteur de variables aléatoires indépendantes et identiquement distribuées tel que :

$$(i) \ E(g_{kt}) = 0 \quad \text{car} \quad E(\sum_j \sigma_{ij}^k g_{kt}^{ij}) = \sum_j \sigma_{ij}^k E(g_{kt}^{ij}) = 0 \quad \text{pour tout } i,$$

$$(ii) \ V(g_{kt}) = E \begin{bmatrix} \sum_j \sigma_{1j}^k g_{kt}^{1j} \\ \vdots \\ \sum_j \sigma_{Nj}^k g_{kt}^{Nj} \end{bmatrix} \left[\sum_j \sigma_{1j}^k g_{kt}^{1j'} \dots \sum_j \sigma_{Nj}^k g_{kt}^{Nj'} \right] = \left[\left\{ E(\sum_j \sigma_{ij}^k g_{kt}^{ij}) (\sum_i \sigma_{ji}^k g_{kt}^{ji'}) \right\} \right]$$

En posant $\bar{v}_{*t}' = (1 \ v_{*t}^{'})$ et en vertu de l'absence de corrélation entre u_{kt}^i et v_{*t}^j pour tout i et j , nous obtenons ce qui suit :

$$\begin{aligned} E(\sum_j \sigma_{ij}^k g_{kt}^{jj'}) (\sum_i \sigma_{ji}^k g_{kt}^{ii'}) &= E(\sum_j \sigma_{ij}^k u_{kt}^j \bar{v}_{*t}^i) (\sum_i \sigma_{ji}^k u_{kt}^i \bar{v}_{*t}^j) \\ &= E[(\sum_j \sigma_{ij}^k u_{kt}^j) (\sum_i \sigma_{ji}^k u_{kt}^i) \bar{v}_{*t}^i \bar{v}_{*t}^j] \\ &= \{E(\sum_j \sigma_{ij}^k u_{kt}^j) (\sum_i \sigma_{ji}^k u_{kt}^i)\} E(\bar{v}_{*t}^i \bar{v}_{*t}^j) \end{aligned} \quad (61)$$

or, au niveau de la dernière égalité, on a :

$$\begin{aligned} E(\sum_j \sigma_{ij}^k u_{kt}^j) (\sum_i \sigma_{ji}^k u_{kt}^i) &= E[\sum_p \sigma_{ip}^k \sigma_{jp}^k (u_{kt}^p)^2 + \sum_p \sum_{q \neq p} \sigma_{ip}^k \sigma_{jq}^k u_{kt}^p u_{kt}^q] \\ &= \sum_p \sigma_{ip}^k \sigma_{jp}^k \sigma_k^{pp} + \sum_p \sum_{q \neq p} \sigma_{ip}^k \sigma_{jq}^k \sigma_k^{pq} \end{aligned} \quad (62)$$

En prenant à titre d'exemple le cas de deux individus i.e. $p = 1, 2$ et sachant que :

$$I_2 = \Omega_k \Omega_k^{-1} = \begin{bmatrix} \sigma_k^{11} \sigma_{11}^k + \sigma_k^{12} \sigma_{21}^k & \sigma_k^{11} \sigma_{12}^k + \sigma_k^{12} \sigma_{22}^k \\ \sigma_k^{21} \sigma_{11}^k + \sigma_k^{22} \sigma_{21}^k & \sigma_k^{21} \sigma_{12}^k + \sigma_k^{22} \sigma_{22}^k \end{bmatrix}$$

on en déduit alors que :

$$\begin{aligned} \sum_p \sigma_{1p}^k \sigma_{2p}^k \sigma_k^{pp} + \sum_p \sum_{q \neq p} \sigma_{1p}^k \sigma_{2q}^k \sigma_k^{pq} &= \sigma_{11}^k \sigma_{21}^k \sigma_k^{11} + \sigma_{12}^k \sigma_{22}^k \sigma_k^{22} + \sigma_{11}^k \sigma_{22}^k \sigma_k^{12} + \sigma_{12}^k \sigma_{21}^k \sigma_k^{21} \\ &= \sigma_{11}^k (\sigma_{21}^k \sigma_k^{11} + \sigma_{22}^k \sigma_k^{12}) + \sigma_{12}^k (\sigma_{21}^k \sigma_k^{21} + \sigma_{22}^k \sigma_k^{22}) = \sigma_{12}^k \end{aligned}$$

Par conséquent, on a :

$$E(\sum_j \sigma_{ij}^k g_{kt}^{jj'}) (\sum_i \sigma_{ji}^k g_{kt}^{ii'}) = \sigma_{ij}^k E(\bar{v}_{*t}^i \bar{v}_{*t}^j)$$

et finalement, on a le résultat suivant :

$$V(g_{kt}) = \left[\left\{ \sigma_{ij}^k \begin{pmatrix} 1 & 0' \\ 0 & \Omega_{*}^{ij} \end{pmatrix} \right\} \right] \quad \text{pour } i, j = 1, 2, \dots, N \quad (63)$$

En vertu de l'avant dernier théorème, on ne développera pas les conditions d'inversion de $V(g_{kt})$, mais on s'intéressera à la forme quadratique qui lui est associée.

Soit X un vecteur colonne d'ordre $N(1 + K_*)$, qui s'écrit par :

$$X' = (X^{1'}, \dots, X^{i'} \dots, X^{N'}) \quad \text{où } X^{i'} = (x_1^i \ X_*^i)$$

la forme quadratique définie par $q(X) = X' V(g_{kt}) X$ s'écrit comme suit :

$$\begin{aligned} q(X) &= \sum_i X^{i'} \Omega_{g_k}^i X^i + \sum_i \sum_{j \neq i} X^{i'} \Omega_{g_k}^{ij} X^j \\ &= \sum_i \sigma_{ii}^k ((x_1^i)^2 + X_*^{i'} \Omega_*^i X_*^i) + \sum_i \sum_{j \neq i} \sigma_{ij}^k (x_1^i x_1^j + X_*^{i'} \Omega_*^{ij} X_*^j) \end{aligned} \quad (64)$$

Comme Ω_*^i est définie positive et si σ_{ii}^k est positive, alors $\sigma_{ii}^k ((x_1^i)^2 + X_*^{i'} \Omega_*^i X_*^i)$ est positive; et pour que $q(X)$ soit définie positive, il faut que le second terme du membre de droite de la formulation 64 soit définie non négative. Cette condition est réalisée si l'une des deux conditions suivantes est vérifié :

- $\sigma_{ij}^k \geq 0$, $x_1^i x_1^j \geq 0$ et $\Omega_{\mathbf{x}}^{ij}$ définie non négative
- $\sigma_{ij}^k \leq 0$, $x_1^i x_1^j \leq 0$ et $\Omega_{\mathbf{x}}^{ij}$ définie négative

En supposant la première condition satisfaite, il en résulte que $V(g_{kt})$ est définie positive.

Il nous reste maintenant d'étudier la définiture de l'expression suivante, pour réunir les conditions d'application du théorème centrale limite généralisé :

$$\text{plim } \frac{1}{T} \sum_t A_t V(g_{kt}) A_t' = \text{plim } \frac{1}{T} \sum_t \left[\left\{ A_t^i \sigma_{ij}^k \begin{pmatrix} 1 & 0' \\ 0 & \Omega_{\mathbf{x}}^{ij} \end{pmatrix} A_t^{j'} \right\} \right] \quad \text{pour } i, j = 1, 2, \dots, N$$

or, d'après la formulation 44, on arrive donc pour tout $i, j = 1, 2, \dots, N$ à :

$$\text{plim } \frac{1}{T} \sum_t A_t V(g_{kt}) A_t' = [(\sigma_{ij}^k Q_k^{ij})] := Q_k^* \quad (65)$$

par conséquent, d'après le résultat précédent 52, la quantité Q_k^* est définie positive. Ainsi, nous venons de construire le vecteur aléatoire g_{kt} tel que :

$$g_{kt} \sim \text{iid } (0, V(g_{kt})) \quad \text{avec } V(g_{kt}) \text{ définie positive}$$

On considère maintenant le deuxième facteur W_T définie par :

$$W_T = \frac{1}{\sqrt{T}} \sum_t A_t g_{kt}$$

avec

$$\begin{aligned} E(W_T) &= \frac{1}{\sqrt{T}} \sum_t A_t E(g_{kt}) = 0 \\ V(W_T) &= \frac{1}{T} E(\sum_t A_t g_{kt})(\sum_t A_t g_{kt})' \\ &= \frac{1}{T} \sum_t E(A_t g_{kt} g_{kt}' A_t') + \frac{1}{T} \sum_t \sum_{s \neq t} E(A_t g_{kt} g_{ks}' A_s') \end{aligned} \quad (66)$$

Vu l'indépendance temporelle des observations pour t différent de s , on obtient donc :

$$V(W_T) = \frac{1}{T} \sum_t A_t E(g_{kt} g_{kt}') A_t' = \frac{1}{T} \sum_t A_t V(g_{kt}) A_t'$$

en utilisant la convergence quadratique qui entraîne la convergence en probabilité, on a :

$$\text{plim } V(W_T) = Q_k^*$$

et, par conséquent :

$$W_T \xrightarrow{(d)} W$$

En vertu du théorème centrale limite généralisé, alors W_T tend en distribution vers le vecteur aléatoire W , tel que :

$$W_T \xrightarrow{(d)} W \sim N(0, Q_k^*)$$

Pour terminer l'étude de la limite en distribution de $\sqrt{T}(\hat{\alpha}_k - \alpha_k)$, on utilise le théorème suivant de Dhrymes [Dhrymes73] sur la limite d'un produit de deux suites.

Théorème Soit la suite $\{y_t, w_t\}$ où $\{y_t\}$ est une suite de vecteurs aléatoires qui converge en distribution vers y , $\{w_t\}$ est une suite de matrices qui converge en probabilité vers la matrice des constantes C finie et de rang complet pour les lignes, telle que : $\text{plim } \{w_t\} = C$ alors la suite $\{w_t y_t\}$ converge en distribution limite vers Cy .

A titre d'exemple, soit $\{w_t\}$ une suite de matrices qui convergent en probabilité vers C d'ordre (m, n) et de rang m , on a :

$$\text{si } \{y_t\} \sim N(0, \Omega) \quad \text{alors} \quad \{w_t y_t\} \xrightarrow{(d)} N(0, C\Omega C')$$

Par l'application de ce théorème et puisque Q_k^* est symétrique, on arrive alors au résultat final suivant :

$$\sqrt{T}(\hat{\alpha}_k - \alpha_k) \xrightarrow{(d)} N[0, (Q_k^*)^{-1} Q_k^* (Q_k^*)^{-1}] = N[0, (Q_k^*)^{-1}] \quad (67)$$

L'intérêt de la variance asymptotique est de permettre le calcul d'un écart-type asymptotique servant à préciser la convergence asymptotique, à réaliser des tests chaque fois que le calcul dans l'échantillon n'est pas facile à faire, et elle indique la vitesse à laquelle on se rapproche de la convergence. Intuitivement, parmi deux estimateurs convergents, on choisit celui qui possède la convergence asymptotique la plus rapide i.e. la variance asymptotique la plus petite.

Par définition, on appelle variance asymptotique d'un estimateur convergent θ_T de θ , si elle existe, la variance associée à la fonction de répartition divisée par T . Il ressort donc de la distribution asymptotique du résultat 67 que :

$$\text{Var-As } (\hat{\alpha}_k) = \frac{1}{T} Q_k^{*-1}.$$

6 La méthode des moindres carrés généralisés réalisables

6.1 Principe de la méthode

Pour passer à la méthode de l'Aitken réalisable, on estime les différentes composantes σ_k^{ij} par la méthode des moindres carrés individuelles pour obtenir dans une première étape $\hat{\Omega}_k$, puis dans une seconde étape on détermine l'estimateur réalisable $\hat{\alpha}_k(r)$.

Pour réaliser l'estimation de σ_k^{ii} , on procède comme suit, d'abord on sait que :

$$\begin{aligned}\hat{u}_k^i &= y_k^i - \hat{y}_k^i = y_k^i - Z_k^i \hat{\alpha}_k^i \\ &= y_k^i - Z_k^i (Z_k^{i'Z_k^i})^{-1} Z_k^{i'Z_k^i} y_k^i = M_k^i y_k^i\end{aligned}$$

avec $M_k^i = I_T - Z_k^i (Z_k^{i'Z_k^i})^{-1} Z_k^{i'Z_k^i}$ et $M_k^i Z_k^i = 0$, on en déduit donc que :

$$\hat{u}_k^i \hat{u}_k^i = y_k^{i'Z_k^i} M_k^i y_k^i = u_k^{i'Z_k^i} M_k^i u_k^i := SS^i$$

Pour déterminer un estimateur centré de σ_k^{ii} , on calcule $E(SS^i)$:

$$\begin{aligned}E(SS^i) &= E[u_k^{i'Z_k^i} u_k^i - u_k^{i'Z_k^i} Z_k^i (Z_k^{i'Z_k^i})^{-1} Z_k^{i'Z_k^i} u_k^i] = E(\text{tr } u_k^i u_k^{i'}) - E[\text{tr } Z_k^i (Z_k^{i'Z_k^i})^{-1} Z_k^{i'Z_k^i} u_k^i u_k^{i'}] \\ &= \sigma_k^{ii} \text{tr } I_T - \sigma_k^{ii} E[\text{tr } (Z_k^{i'Z_k^i})^{-1} Z_k^{i'Z_k^i}] = \sigma_k^{ii}(T - s_k)\end{aligned}$$

Ce qui permet de proposer un estimateur centré pour σ_k^{ii} , car $E(\hat{\sigma}_k^{ii}) = \sigma_k^{ii}$:

$$\hat{\sigma}_k^{ii} = \frac{SS^i}{T - s_k}$$

On montre par ailleurs, que $\hat{\sigma}_k^{ii}$ est convergent quant T tend vers l'infini et s_k est fini :

$$\begin{aligned}\text{plim } \hat{\sigma}_k^{ii} &= \text{plim } \frac{1}{T - s_k} u_k^{i'Z_k^i} u_k^i - \text{plim } \frac{1}{T - s_k} u_k^{i'Z_k^i} Z_k^i (Z_k^{i'Z_k^i})^{-1} Z_k^{i'Z_k^i} u_k^i \\ &= \text{plim } \frac{1}{T} u_k^{i'Z_k^i} u_k^i - \text{plim } \frac{1}{T} u_k^{i'Z_k^i} Z_k^i \text{plim } (Z_k^{i'Z_k^i})^{-1} \text{plim } \frac{1}{T} Z_k^{i'Z_k^i} u_k^i\end{aligned}$$

à partir du résultat 44 et en utilisant le théorème de Khintchine, on obtient alors que :

$$\begin{aligned}\text{plim } \hat{\sigma}_k^{ii} &= \text{plim } \frac{1}{T} u_k^{i'Z_k^i} u_k^i - 0'(Q_k^i)0 = \text{plim } \frac{1}{T} u_k^{i'Z_k^i} u_k^i \\ &= \text{plim } \frac{1}{T} u_k^{i'Z_k^i} u_k^i = \text{plim } \frac{1}{T} \sum_t (u_{kt}^i)^2 = E(u_{kt}^i)^2 = \sigma_k^{ii}\end{aligned}$$

La matrice Ω_k est symétrique, elle contient donc $(N/2)(N + 1)$ éléments à estimer. On présentera pour les covariances des estimateurs seulement convergents, car la valeur exacte de $E(\text{tr } M^i M^j)$ n'est pas connue, elle est définie par :

$$E(\text{tr } M^i M^j) = T - 2s_k + E[\text{tr } (Z_k^{i'Z_k^i})^{-1} Z_k^{i'Z_k^i} Z_k^j (Z_k^{j'Z_k^j})^{-1} Z_k^{j'Z_k^j}]$$

nous optons alors pour l'estimateur convergent de la covariance suivant :

$$\hat{\sigma}_k^{ij} = \frac{1}{T} SS^{ij}$$

En développant la limite en probabilité de la somme des carrés des écarts pour i différent de j , on obtient :

$$\begin{aligned}\text{plim } \frac{1}{T} SS^{ij} &= \text{plim } \frac{1}{T} u_k^{i'Z_k^i} u_k^j - \text{plim } \frac{1}{T} u_k^{i'Z_k^i} Z_k^i (\frac{1}{T} Z_k^{i'Z_k^i})^{-1} \frac{1}{T} Z_k^{i'Z_k^i} u_k^j \\ &\quad + \text{plim } \frac{1}{T} u_k^{i'Z_k^i} Z_k^i (\frac{1}{T} Z_k^{i'Z_k^i})^{-1} \frac{1}{T} Z_k^{i'Z_k^i} Z_k^j (\frac{1}{T} Z_k^{j'Z_k^j})^{-1} \frac{1}{T} Z_k^{j'Z_k^j} u_k^j \\ &\quad - \text{plim } \frac{1}{T} u_k^{i'Z_k^i} Z_k^j (\frac{1}{T} Z_k^{j'Z_k^j})^{-1} \frac{1}{T} Z_k^{j'Z_k^j} u_k^j\end{aligned}$$

d'après le résultat 44 et en vertu du théorème de Khintchine, nous obtenons :

$$\begin{aligned} \text{plim } \frac{1}{T} S S' &= \text{plim } \frac{1}{T} u_k' u_k - 0'(Q_k^i)^{-1} 0 - 0'(Q_k^j)^{-1} 0 + 0'(Q_k^i)^{-1} Q_k^{ij} (Q_k^j)^{-1} 0 \\ &= \text{plim } \frac{1}{T} \sum_t u_{kt}' u_{kt} = E(u_{kt}' u_{kt}) = \sigma_k^{ij} \end{aligned}$$

Ainsi, à partir des deux derniers résultats habituels, on arrive à une quantité aléatoire estimée finie :

$$\hat{\Omega}_k = (\hat{\sigma}_k^{ij}) \quad \text{pour } i, j = 1, 2, \dots, N$$

ce qui permet de déterminer l'estimateur réalisable suivant :

$$\hat{\alpha}_k(r) = [Z_k'(\hat{\Omega}_k^{-1} \otimes I)Z_k]^{-1} Z_k'(\hat{\Omega}_k^{-1} \otimes I)y_k \quad (68)$$

6.2 Propriétés de base des estimateurs

En partant de ce dernier estimateur, on étudie certaines habituelles propriétés dans l'échantillon. D'abord, il s'agit d'un estimateur sans biais, car par l'absence de corrélation entre u_k et Z_k et en utilisant le lemme de l'espérance mathématique conditionnelle du fait de l'indépendance entre $\hat{\Omega}_k$ et u_k , on obtient :

$$\begin{aligned} E(\hat{\alpha}_k) &= \alpha_k + E[(Z_k' \hat{\Phi}_k^{-1} Z_k)^{-1} Z_k' \hat{\Phi}_k^{-1} u_k] \\ &= \alpha_k + E[(Z_k' \hat{\Phi}_k^{-1} Z_k)^{-1} Z_k' \hat{\Phi}_k^{-1} E(u_k) / \hat{\Omega}_k] = \alpha_k \end{aligned} \quad (69)$$

Quant à la propriété de la convergence, on utilise le lemme suivant :

Lemme 2 Soit $\hat{\Omega}_k^{-1}$ une matrice aléatoire telle que :

$$\text{plim } \hat{\Omega}_k^{-1} = \Omega_k^{-1}$$

alors :

$$\text{plim } \frac{1}{T} \sum_t Z_{kt}' \Omega_k^{-1} Z_{kt} = \text{plim } \frac{1}{T} \sum_t Z_{kt}' \hat{\Omega}_k^{-1} Z_{kt} \quad (70)$$

Preuve Une réécriture de cette dernière expression avec l'opérateur vec et le fait que la matrice $\hat{\Omega}_k^{-1}$ soit convergente, nous permet d'obtenir :

$$\begin{aligned} \text{plim vec } \frac{1}{T} \sum_t Z_{kt}' (\Omega_k^{-1} - \hat{\Omega}_k^{-1}) Z_{kt} &= \text{plim } \frac{1}{T} \sum_t (Z_{kt}' \otimes Z_{kt}') (\text{vec } \Omega_k^{-1} - \text{plim vec } \hat{\Omega}_k^{-1}) \\ &= \text{plim } \frac{1}{T} \sum_t (Z_{kt}' \otimes Z_{kt}') \times 0 \end{aligned}$$

Il nous reste d'étudier le premier facteur pour conclure par la nullité ou non du dernier membre de droite. Soit une matrice $\text{diag}(Q_k^1, \dots, Q_k^N)$ finie et définie positive, telle que :

$$\text{plim } \frac{1}{T} \sum_t Z_{kt}' Z_{kt} = \text{diag}(Q_k^1, \dots, Q_k^N)$$

en utilisant l'opérateur vec , on obtient :

$$\begin{aligned}\text{vec} [\text{diag} (Q_k^1, \dots, Q_k^N)] &= \text{plim} \text{vec} \frac{1}{T} \sum_t Z_{kt}' I_N Z_{kt} \\ &= \text{plim} \frac{1}{T} \sum_t (Z_{kt}' \oplus Z_{kt}') \text{vec} I_N\end{aligned}$$

et puisque $\text{vec} I_N$ est une quantité finie alors :

$$\text{plim} \frac{1}{T} \sum_t (Z_{kt}' \oplus Z_{kt}') \text{est finie}$$

Il en résulte finalement que :

$$\text{vec} \text{plim} \frac{1}{T} \sum_t Z_{kt} (\Omega_k^{-1} - \hat{\Omega}_k^{-1}) Z_{kt}' = 0. \quad \blacksquare$$

Il nous reste donc à analyser l'expression suivante que l'on peut déterminer facilement à partir de la formule 56 :

$$\text{plim} \frac{1}{T} Z_k' \hat{\Phi}_k^{-1} u_k = [(\text{plim} \sum_j \hat{\sigma}_{ij}^k \text{plim} \frac{1}{T} \sum_t Z_{kt}^i u_{kt}^j)] \quad \text{pour } i = 1, 2, \dots, N$$

or, par la convergence des estimateurs $\hat{\sigma}_k^{ij}$ et par le résultat 54, on a :

$$\text{plim} \sum_j \hat{\sigma}_{ij}^k \text{plim} \frac{1}{T} \sum_t Z_{kt}^i u_{kt}^j = \sum_j \text{plim} \hat{\sigma}_{ij}^k \text{plim} \frac{1}{T} \sum_t Z_{kt}^i u_{kt}^j = \sum_j \sigma_{ij}^k 0 = 0$$

il en résulte que :

$$\text{plim} \frac{1}{T} Z_k' \hat{\Phi}_k^{-1} u_k = 0.$$

Finalement, on obtient le résultat :

$$\text{plim} \hat{\alpha}_k(r) = \alpha_k \quad (71)$$

6.3 La distribution asymptotique ou l'équivalence asymptotique

Pour étudier l'équivalence asymptotique entre l'estimateur pur des moindres carrés généralisés et l'estimateur réalisable, nous utilisons le lemme suivant :

Lemme 3 Soit un vecteur aléatoire u_{kt} à N composantes où Ω_k est définie positive, tel que :

$$u_{kt} \sim \text{iid} (0, \Omega_k)$$

et soit $\hat{\Omega}_k$ une matrice aléatoire telle que : $\text{plim} \hat{\Omega}_k = \Omega_k$ alors :

$$\frac{1}{\sqrt{T}} \sum_t Z_{kt}' (\Omega_k^{-1} - \hat{\Omega}_k^{-1}) u_{kt} \xrightarrow{(p)} 0$$

Preuve Il s'agit de prouver que les deux termes de l'expression ci-dessus ont la même distribution limite i.e. $N(0, \Omega_k^*)$ avec Q_k^* définie positive.

Une écriture intéressante nous permet d'avoir :

$$Z_{kt}'(\Omega_k^{-1} - \hat{\Omega}_k^{-1})u_{kt} = \text{vec} [Z_{kt}'(\Omega_k^{-1} - \hat{\Omega}_k^{-1})u_{kt}] = (u_{kt}' \otimes Z_{kt}') \text{vec} (\Omega_k^{-1} - \hat{\Omega}_k^{-1})$$

ce qui donne, en utilisant le lemme précédent, que :

$$\begin{aligned} \text{plim } \frac{1}{\sqrt{T}} \sum_t Z_{kt}'(\Omega_k^{-1} - \hat{\Omega}_k^{-1})u_{kt} &= \text{plim } \frac{1}{\sqrt{T}} \sum_t (u_{kt}' \otimes Z_{kt}') \text{plim } \text{vec} (\Omega_k^{-1} - \hat{\Omega}_k^{-1}) \\ &= \text{plim } \frac{1}{\sqrt{T}} \sum_t (u_{kt}' \otimes Z_{kt}') \times 0 := \bar{W}_T \times 0 \end{aligned}$$

Il reste alors à analyser le premier facteur, pour savoir s'il s'agit d'une quantité finie. L'expression $u_{kt} \otimes Z_{kt}'$ s'écrit de façon intéressante comme suit :

$$u_{kt}' \otimes Z_{kt} = [\text{diag } u_{kt}^1(Z_{kt}^1 \dots Z_{kt}^N) \dots \text{diag } u_{kt}^N(Z_{kt}^1 \dots Z_{kt}^N)]$$

or, $Z_{kt}^i u_{kt}^j = A_t^i g_{kt}^{ij}$ pour tout $i, j = 1, 2, \dots, N$ et en posant w_T^{ij} comme suit :

$$w_T^{ij} = \frac{1}{\sqrt{T}} \sum_t A_t^i g_{kt}^{ij}$$

où, d'après la formule 59 A_t^i est telle que :

$$\text{plim } \frac{1}{T} \sum_t A_t^i V(g_{kt}^{ij}) A_t^{i'} = \sigma_k^{jj} Q_k^i$$

est définie positive.

Par le théorème centrale limite généralisé, on arrive à :

$$w_T^{ij} \xrightarrow{(d)} w^{ij} \sim N(0, \sigma_k^{jj} Q_k^i)$$

Pour l'ensemble des individus, en posant \bar{W}_T comme suit :

$$\bar{W}_T = \frac{1}{\sqrt{T}} \sum_t (u_{kt}' \otimes Z_{kt}')$$

on obtient le résultat suivant :

$$\bar{W}_T = [\text{diag } (w_T^{11} \dots w_T^{1N}) \dots \text{diag } (w_T^{N1} \dots w_T^{NN})]$$

avec

$$E(\bar{W}_T) = 0$$

$$V(\bar{W}_T) = E\{ \text{diag } (\sum_i w_T^{i1} w_T^{i1'} \dots \sum_i w_T^{iN} w_T^{iN'}) \} \quad \text{pour } i = 1, 2, \dots, N$$

Le calcul de $E(w_T^{ij} w_T^{ij'})$ nous donne, en vertu de l'indépendance temporelle, ce qui suit :

$$\begin{aligned} E(w_T^{ij} w_T^{ij'}) &= \frac{1}{T} E(\sum_t A_t^i g_{kt}^{ij} (\sum_t g_{kt}^{ij'} A_t^{i'})) \\ &= \frac{1}{T} E[\sum_t (A_t^i g_{kt}^{ij} g_{kt}^{ij'} A_t^{i'}) + \sum_t \sum_{s \neq t} (A_t^i g_{kt}^{ij} g_{ks}^{ij'} A_s^{i'})] \\ &= \frac{1}{T} E(\sum_t A_t^i g_{kt}^{ij} g_{kt}^{ij'} A_t^{i'}) := \frac{1}{T} \sum_t A_t^i V(g_{kt}^{ij}) A_t^{i'} \end{aligned} \quad (72)$$

et en utilisant la propriété de la convergence quadratique, qui entraîne la convergence en probabilité, alors :

$$\text{plim } V(w_T^{ij}) = \text{plim } E(w_T^{ij} w_T^{ij'}) = \sigma_k^{jj} Q_k^i$$

par conséquent :

$$\begin{aligned} \text{plim } V(\overline{W}_T) &= \text{diag} (\sum_i \sigma_k^{11} Q_k^i \dots \sum_i \sigma_k^{NN} Q_k^i) \\ &= \text{diag} (\sigma_k^{11} \dots \sigma_k^{NN}) \otimes \sum_i Q_k^i := V_k \end{aligned} \quad (73)$$

avec

$$V_k = \text{diag} (\sigma_k^{11} \dots \sigma_k^{NN}) \otimes Q_k^1 + \dots + \text{diag} (\sigma_k^{11} \dots \sigma_k^{NN}) \otimes Q_k^N$$

or, $\text{diag} (\sigma_k^{11} \dots \sigma_k^{NN}) \otimes Q_k^i$ est définie positive pour tout i , donc V_k est définie positive.

Par le théorème centrale limite généralisé, on a :

$$\overline{W}_T \xrightarrow{(d)} \overline{W} \sim N(0, V_k)$$

où \overline{W} est une matrice aléatoire finie, ce qui prouve que \overline{W}_T est finie.

Corollaire [Dhrymes74] Si la suite $\{w_t\}$ tend vers 0 et si $\{y_t\}$ suit la distribution $N(0, Q)$ avec Q définie positive, alors :

$$\{w_t y_t\} \rightarrow 0 \text{ et } \text{plim } \{w_t y_t\} = 0$$

Finalement, d'après ce corollaire du théorème sur la limite d'un produit de deux suites, on déduit le résultat suivant :

$$\frac{1}{\sqrt{T}} \sum_t Z_{kt}' (\Omega_k^{-1} - \hat{\Omega}_k^{-1}) u_{kt} \xrightarrow{(p)} 0 \quad (74)$$

qui exprime le fait que les deux termes de ce dernier résultat ont la même distribution limite.

En conséquence, par les deux derniers lemmes, on aboutit à l'équivalence asymptotique entre l'estimateur pur et réalisable, car :

$$\sqrt{T}[\hat{\alpha}_k(r) - \alpha_k] \quad \text{et} \quad \sqrt{T}[\hat{\alpha}_k(u) - \alpha_k]$$

ont la même distribution limite i.e. $N[0, (Q_k^*)^{-1}]$. ■

Il ressort de cette distribution asymptotique que la matrice des variances-covariances et son estimation sont déterminés respectivement comme suit :

$$\begin{aligned} \text{Var-As } [\hat{\alpha}_k(r)] &= \frac{1}{T} Q_k^{*-1} \\ \text{Var-as } (\hat{\alpha}_k(r)) &= (\sum_t Z_{kt}' \hat{\Omega}_k^{-1} Z_{kt})^{-1} \end{aligned} \quad (75)$$

7 La méthode du Zellner itéré

7.1 Principe de la méthode

Comme il a été indiqué par plusieurs économètres dont Kmenta, Gilbert et Zellner, l'estimateur réalisable $\hat{\alpha}'_k(r)$ n'est pas toujours efficace dans les petits échantillons et qu'il est possible de trouver d'autres estimateurs pour lesquels la variance est susceptible d'être plus petite.

Le point de départ concernant la procédure itérative est de considérer l'estimateur réalisable suivant :

$$\alpha_k^*(r) = (Z_k' \Phi^{*-1} Z_k)^{-1} Z_k' \Phi^{*-1} y_k$$

Si la matrice $\Phi_k^* = \sigma_k I_{NT}$, on obtient alors, en adoptant la notation suivante $\alpha_k^* = \alpha_k(1)$:

$$\alpha_k^*(1) = (Z_k' Z_k)^{-1} Z_k' y_k \Rightarrow \alpha_k^{i*} = (Z_k^{i'} Z_k^i)^{-1} Z_k^{i'} y_k^i \quad \text{pour tout } i$$

A partir du modèle $y_k^i = Z_k^i \alpha_k^i + u_k^i$, on arrive aux estimateurs convergents pour les différents paramètres σ_k^{ij} , ce qui permet d'obtenir pour $i \neq j$:

$$\Omega_k^* = \Omega_k(1) \quad \text{avec} \quad \sigma_k^{ii*} = \frac{1}{T - s_k} u_k^{i*'} u_k^{i*} \quad \text{et} \quad \sigma_k^{ij*} = \frac{1}{T} u_k^{i*'} u_k^{j*}$$

La première itération s'exprime donc comme suit, avec $u_k(1) = y_k - Z_k \alpha_k(1)$:

$$\alpha_k(2) = [Z_k' \Phi_k^{-1}(1) Z_k]^{-1} Z_k' \Phi_k^{-1}(1) y_k$$

La seconde itération consiste tout d'abord à calculer $u_k(2)$ telle que :

$$u_k(2) = y_k - Z_k \alpha_k(2)$$

A partir des estimateurs convergents $\sigma_k^{ij}(2)$, on obtient $\Omega_k(2)$, et on déduit alors l'estimateur $\alpha_k(3)$ défini par :

$$\alpha_k(3) = [Z_k' \Phi_k^{-1}(2) Z_k]^{-1} Z_k' \Phi_k^{-1}(2) y_k$$

Ainsi, les différentes itérations appliquent les équations suivantes :

$$\alpha_k(p+1) = [Z_k' \Phi_k^{-1}(p) Z_k]^{-1} Z_k' \Phi_k^{-1}(p) y_k \tag{76}$$

avec $\Omega_k(0) = \sigma_k I_N$ et $u_k(p) = y_k - Z_k \alpha_k(p)$ pour $p = 1, 2, 3, \dots$

Il reste à préciser les formules donnant $\Omega_k(p)$ à partir de $u_k(p)$ en utilisant des estimateurs simultanés, par exemple ceux du maximum de vraisemblance :

$$\Omega_k(p) = [\sigma_k^{ij}(p)] \quad \text{avec} \quad \begin{cases} \sigma_k^{ij}(p) = \frac{1}{T} u_k^{i'}(p) u_k^j(p) \\ \sigma_k^{ii}(p) = \frac{1}{T - s_k} u_k^{i'}(p) u_k^i(p) \end{cases} \tag{77}$$

Si u_k est distribuée suivant une loi normale, on a :

$$\mathcal{L}_k = \text{cste} - \frac{T}{2} \ln |\Omega_k| - \frac{1}{2} \text{tr } \Phi_k^{-1} u_k u_k'$$

les estimateurs du maximum de vraisemblance $(\tilde{\alpha}_k, \tilde{\Omega}_k)$ sont solution du système suivant :

$$\begin{cases} Z_k' \Phi_k^{-1} Z_k \alpha_k - Z_k' \Phi_k^{-1} y_k = 0 & (1) \\ \Phi_k - u_k u_k' = 0 & (2) \end{cases} \quad (78)$$

nous déduisons de l'expression (2) que :

$$\Omega_k^*(\alpha_k) = \frac{1}{T} \sum_t u_{kt} u_{kt}' := \frac{1}{T} \epsilon_k' \epsilon_k$$

avec $\epsilon_k' = (u_{k1}, \dots, u_{kT})$ et de dimension $N \times T$. Pour Ω_k^* régulière cela suppose que $T \geq N$, i.e. ϵ_k soit une matrice rectangulaire haute.

Ainsi, on aura à minimiser la vraisemblance suivante :

$$\mathcal{L}_k^* = \text{cste} - \frac{T}{2} \ln |\Omega_k^*| - \frac{1}{2} \text{tr } \Phi_k^{*-1} u_k u_k' = \text{cste} - \frac{T}{2} \ln |\Omega_k^*| \quad (79)$$

il en résulte que $\tilde{\alpha}_k$ est solution du système qui suit :

$$\frac{\partial \mathcal{L}_k^*}{\partial \alpha_k} = 0.$$

Finalement, l'estimateur du maximum de vraisemblance est défini par :

$$\tilde{\alpha}_k = (Z_k' \tilde{\Phi}_k^{-1} Z_k)^{-1} Z_k' \tilde{\Phi}_k^{-1} y_k$$

avec $\tilde{\Omega}_k = \Omega_k^*(\tilde{\alpha}_k)$.

Par ailleurs, le problème de choisir entre l'estimateur itéré $\alpha_k(p)$ et l'estimateur réalisable $\hat{\alpha}_k$ se pose de façon pertinente travers le nombre d'itérations suffisant. Dans la pratique selon Malinvaud [Malinvaud81], un petit nombre d'itérations peut suffire à la convergence.

Le calcul de $\alpha_k(p)$ n'exige pas automatiquement celui de $\alpha_k(p+1)$. En retenant un critère de convergence sur un échantillon donné, on itère jusqu'à ce que $|\alpha_k(p) - \alpha_k(p-1)| < \epsilon$ où ϵ est une valeur infiniment petite.

Comme, nous le verrons dans ce qui suit, si u_k suit une loi normale, la convergence aura pour conséquence que les estimateurs $\alpha_k(p)$ et $\Omega_k(p)$ pour $p = 1, 2, \dots$ résolvent précisément les équations auxquelles l'estimateur du maximum de vraisemblance est soumis. Autrement, il s'agirait d'un estimateur du quasi-maximum de vraisemblance [Philipps76].

7.2 Propriétés

En ce qui concerne la convergence des estimateurs par l'Aitken itéré, en posant que : $\Omega_k^* = \Omega_k(1) := \hat{\Omega}_k$, en utilisant les formulations 76 et 77 et la convergence de l'estimateur réalisable et sachant que $\text{plim } \hat{\Phi}_k = \Phi_k$, alors

$$\text{plim } \alpha_k(2) = \alpha_k$$

En travaillant avec $u_k(2) = y_k - Z_k \alpha_k(2)$, on a donc :

$$\begin{aligned} u_k(2) &= y_k - Z_k [(Z_k' \hat{\Phi}_k^{-1} Z_k)^{-1} Z_k' \hat{\Phi}_k^{-1} y_k] = [I - Z_k (Z_k' \hat{\Phi}_k^{-1} Z_k)^{-1} Z_k' \hat{\Phi}_k^{-1}] y_k \\ &= [I - Z_k (Z_k' \hat{\Phi}_k^{-1} Z_k)^{-1} Z_k' \hat{\Phi}_k^{-1}] u_k(1) := M_{k(1)} u_k(1) \end{aligned}$$

ce qui nous permet d'obtenir :

$$\alpha_k(3) = [Z_k' \Phi_k^{-1}(2) Z_k]^{-1} Z_k' \Phi_k^{-1}(2) y_k \quad \text{avec} \quad y_k = Z_k \alpha_k(2) + u_k(2)$$

par conséquent pour la deuxième itération on a :

$$\begin{aligned} \text{plim } \alpha_k(3) &= \text{plim } \alpha_k(2) + \text{plim } [Z_k' \Phi_k^{-1}(2) Z_k]^{-1} Z_k' \Phi_k^{-1}(2) u_k(2) \\ &= \alpha_k + \text{plim } [\frac{1}{T} Z_k' \Phi_k^{-1}(2) Z_k]^{-1} \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(2) u_k(2) \end{aligned}$$

Pour déterminer le deuxième terme du second membre, on utilisera le lemme suivant :

Lemme 4 Soit $\Omega_k^{-1}(p)$ une matrice aléatoire telle que :

$$\text{plim } \Omega_k^{-1}(p) = \text{plim } \Omega_k^{-1}(p-1) = \dots = \text{plim } \Omega_k^{-1}(1) = \Omega_k^{-1}$$

alors

$$\text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p) Z_k = \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p-1) Z_k = \dots = \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1} Z_k \quad (80)$$

Preuve Nous avons pour tout $p \geq 2$, l'égalité suivante :

$$\text{vec } \frac{1}{T} \sum_t Z_{kt}' [\Omega_k^{-1}(p-1) - \Omega_k^{-1}(p)] Z_{kt} = \frac{1}{T} \sum_t (Z_{kt}' \otimes Z_{kt}') [\text{vec } \Omega_k^{-1}(p-1) - \text{vec } \Omega_k^{-1}(p)]$$

donc, la limite en probabilité est déterminée comme suit :

$$\text{plim } \text{vec } \frac{1}{T} \sum_t Z_{kt}' [\Omega_k^{-1}(p-1) - \Omega_k^{-1}(p)] Z_{kt} = \text{plim } \frac{1}{T} \sum_t (Z_{kt}' \otimes Z_{kt}') [\text{vec } \Omega_k^{-1}(p-1) - \text{vec } \Omega_k^{-1}(p)]$$

or, d'après notre hypothèse et puisque la quantité

$$\text{plim } \frac{1}{T} \sum_t (Z_{kt}' \otimes Z_{kt}')$$

est finie, on obtient alors le résultat qui achève la preuve :

$$\text{plim vec } \frac{1}{T} \sum_t Z_{kt}' [\Omega_k^{-1}(p-1) - \Omega_k^{-1}(p)] Z_{kt} = \text{plim } \frac{1}{T} \sum_t (Z_{kt}' \otimes Z_{kt}') 0 = 0. \quad \blacksquare$$

Il nous reste à déterminer l'autre facteur du second membre que nous l'on peut définir de façon générale, à partir de la formulation 56, comme suit :

$$\text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p) u_k(p) = [(\text{plim } \sum_j \sigma_{ij}^k(p) \text{plim } \frac{1}{T} Z_k^{i'} u_k^j(p))] \quad \text{pour } i = 1, 2, \dots, N$$

on montre facilement que $u_k(p) = M_k(p-1)u_k$, par conséquent :

$$\begin{aligned} \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p) u_k(p) &= \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p) M_k(p-1) u_k \\ &= \text{plim } [\frac{1}{T} Z_k' \Phi_k^{-1}(p) u_k - \frac{1}{T} Z_k' \Phi_k^{-1}(p) Z_k (\frac{1}{T} Z_k' \Phi_k^{-1}(p-1) Z_k)^{-1} \frac{1}{T} Z_k' \Phi_k^{-1}(p-1) u_k] \end{aligned}$$

or, pour tout $p \geq 2$ et pour $i = 1, 2, \dots, N$, d'après le précédent lemme et par le résultat 54, on obtient :

$$\begin{aligned} \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p) u_k(p) &= \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p) u_k - Q_k^* (Q_k^*)^{-1} \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p-1) u_k \\ &= \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p) u_k - \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p-1) u_k \\ &= [((\sigma_{ij}^k - \sigma_{ij}^k) 0)] = 0 \end{aligned} \quad (81)$$

on en déduit finalement que :

$$\text{plim } \alpha_k(3) = \alpha_k + (Q_k^*)^{-1} 0 = \alpha_k$$

et plus généralement, on a ce qui suit :

$$\text{plim } \alpha_k(p) = \text{plim } \{ \alpha_k[p - (p-1)] + \sum_{\tau=1}^{p-1} [\frac{1}{T} Z_k' \Phi_k^{-1}(p-\tau) Z_k]^{-1} \frac{1}{T} Z_k' \Phi_k^{-1}(p-\tau) u_k(p-\tau) \}$$

or,

$$\begin{aligned} \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p-\tau) u_k(p-\tau) &= \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p-\tau) M_k(p-\tau-1) u_k \\ &= \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p-\tau) u_k - \text{plim } \frac{1}{T} Z_k' \Phi_k^{-1}(p-\tau-1) u_k \end{aligned}$$

d'après le résultat 80 et le lemme précédent, on a finalement démontré la convergence des estimateurs itérés :

$$\text{plim } \alpha_k(p) = \text{plim } \alpha_k(1) + (p-1)(Q_k^*)^{-1} \times 0 = \text{plim } \alpha_k(1) = \alpha_k \quad (82)$$

En ce qui concerne la distribution asymptôtique des estimateurs itérés, selon l'idée de Zellner [Zellner62] ils ont les mêmes propriétés asymptôtiques que l'estimateur réalisable, et il s'agira de prouver pour tout $p \geq 2$ que l'expression suivante, converge en probabilité vers zéro :

$$\frac{1}{\sqrt{T}} \sum_t Z_{kt}' [\Omega_k^{-1}(p-1) - \Omega_k^{-1}(p)] u_{kt}(p-1).$$

8 Comparaison des méthodes

8.1 L'équivalence asymptotique

La comparaison entre $\hat{\alpha}_k(r)$ et $\tilde{\alpha}_k$, dans des échantillons de taille finie et petite, a nécessité de faire des essais empiriques et donc d'utiliser la méthode de Monte-Carlo pour en déterminer les distributions. Certaines expériences réalisées ne font pas apparaître une supériorité asymptotique de l'estimateur $\tilde{\alpha}_k$ par rapport à l'estimateur réalisable $\hat{\alpha}_k(r)$. On peut même s'interroger sur leurs efficacités relatives.

Par ailleurs, dans le cas où les équations du système 78 admettent une solution unique [Kmenta et al.74], alors les estimateurs $\alpha_k(p)$ et $\Omega_k(p)$ coïncideront en cas de convergence respectivement avec $\tilde{\alpha}_k$ et $\tilde{\Omega}_k$.

Les essais de Kmenta et Gilbert [Kmenta et al.68] ont précisé, qu'en cas de normalité des erreurs, la procédure du Zellner itéré [Zellner et al.62] converge vers les estimateurs du maximum de vraisemblance. Ils ont trouvé dans leurs expériences que l'estimateur de l'Aitken itéré et celui du maximum de vraisemblance sont numériquement équivalents. Cette équivalence numérique a été aussi montrée par Ruble. [Ruble68]

L'efficacité de $\tilde{\alpha}_k$ est nécessairement meilleure dans les échantillons de grande taille. Certains travaux [Bement et al.69] [Taylor77] ont montré dans des cas simples que la dimension de y_k doit être supérieure à 20 pour avoir l'efficacité de $\tilde{\alpha}_k$ sur $\alpha_k(r)$.

Les propriétés des estimateurs $\tilde{\alpha}_k$ et $\alpha_k(r)$ sont très voisines. De même, les conditions qui assurent l'équivalence asymptotique entre $\hat{\alpha}_k$ et $\alpha_k(r)$ garantissent généralement l'équivalence asymptotique entre $\tilde{\alpha}_k$ et $\hat{\alpha}_k$.

On montre que l'efficacité asymptotique de $\tilde{\alpha}_k$ est maximale dans la classe de tous les estimateurs convergents, car un tel estimateur atteint asymptotiquement la limite de l'inégalité de Cramer-Rao ; en effet, la matrice inverse d'information donnée par $(\tilde{Q}_k^*)^{-1}$ pour α_k correspond, en cas de normalité des erreurs, exactement à la matrice des variances-covariances de la distribution asymptotique. On a donc :

$$\text{Var-as}(\tilde{\alpha}_k) = \frac{1}{T}(\tilde{Q}_k^*)^{-1}$$

$$\text{Var-as}(\hat{\alpha}_k) = \text{Var-as}(\hat{\alpha}_k(r)) = \frac{1}{T}(Q_k^*)^{-1}$$

La matrice des variances-covariances asymptotiques peut être estimée, à partir de ce qui précède, comme suit :

$$\widehat{\text{Var-as}}(\tilde{\alpha}_k) = (\sum_t Z_{kt}' \hat{\Omega}_k^{-1} Z_{kt})^{-1}$$

$$\widehat{\text{Var-as}}(\hat{\alpha}_k) = (\sum_t Z_{kt}' \Omega_k^{-1} Z_{kt})^{-1}$$

Au niveau de l'estimateur du maximum de vraisemblance et de l'estimateur pur, on s'aperçoit que :

$$(\tilde{Q}_k^*)^{-1} - (Q_k^*)^{-1} \simeq 0$$

Théorème [Balestra72] Soit A et B deux matrices du même ordre et définies positives. Si $A - B$ est définie non négative, alors $B^{-1} - A^{-1}$ est définie non négative.

En utilisant ce théorème, on arrive à montrer l'équivalence asymptotique des estimateurs considérés :

$$Q_k^* - \tilde{Q}_k^* \simeq 0 \quad (83)$$

8.2 Convergence numérique du Zellner itéré

On s'intéressera à la convergence de l'estimateur du Zellner itéré vers celui du maximum de vraisemblance. En supposant que :

$$\tilde{u}_k = y_k - Z_k \tilde{\alpha}_k$$

la vraisemblance logarithmique à maximiser s'écrit comme suit :

$$\mathcal{L}_k^{**} = \text{cste} + \frac{1}{2} \ln |\Phi_k^{-1}| - \frac{1}{2} \text{tr} \Phi_k^{-1} \tilde{u}_k \tilde{u}_k'$$

Cette vraisemblance ne dépend que de Ω_k , qui est symétrique. Pour exploiter cette symétrie, on considère le vecteur w défini par :

$$w = (w^1, w^2, \dots, w_*^N) \text{ où pour } i > j \quad w^l = \sigma_k^{ij} \text{ et } N_* = \frac{N}{2}(N+1)$$

Considérons maintenant le différentiel de \mathcal{L}_k^{**} , qui n'est fonction que de w , défini par :

$$d\mathcal{L}_k^{**} = \frac{1}{2} \text{tr} (\Phi_k - \tilde{u}_k \tilde{u}_k') d\Phi_k^{-1}$$

L'estimation cherchée de Ω_k doit satisfaire à l'équation suivante :

$$\text{tr} [(\tilde{\Phi}_k - \tilde{u}_k \tilde{u}_k') \tilde{\Phi}_k^l] = 0 \quad \text{pour } l = 1, 2, \dots, N_* \quad (84)$$

où Φ_k^l est la matrice des dérivées partielles de Φ_k^{-1} par rapport à w^l .

Ainsi, les équations suivantes définissent comment les estimateurs $\tilde{\alpha}_k$ et $\tilde{\Phi}_k$ sont déterminés simultanément, et que leur calcul simultané exige une procédure itérative :

$$\begin{cases} \tilde{\alpha}_k &= (Z_k' \tilde{\Phi}_k^{-1} Z_k)^{-1} Z_k' \tilde{\Phi}_k^{-1} y_k & (1) \\ \tilde{u}_k &= y_k - Z_k \tilde{\alpha}_k & (2) \\ 0 &= \text{tr} [(\tilde{\Phi}_k - \tilde{u}_k \tilde{u}_k') \tilde{\Phi}_k^l] & (3) \end{cases}$$

Un tel système peut avoir une multiplicité de solutions, qui maximise \mathcal{L}_k . L'équation (3) permet de calculer l'estimation de $\tilde{\Phi}_k$ à partir des \tilde{u}_k . La matrice $\tilde{u}_k \tilde{u}_k'$ joue le rôle de la matrice des variances-covariances empiriques des résidus, car :

$$E(u_k^i u_k^{j'})$$

est remplacée par $\tilde{u}_k^i \tilde{u}_k^j$

Cette même équation montre que les \tilde{w}^l vont être définies de façon à annuler N_* combinaisons linéaires des éléments de la matrice des différences entre covariances théoriques et covariances empiriques correspondantes.

Par exemple pour $N = 2$, cherchons à déterminer \tilde{w}^l . Le nombre de paramètres à estimer est $N_* = 3$ i.e. \tilde{w}^1 , \tilde{w}^2 , et \tilde{w}^3 .

Le calcul de Ω_k^1 , nous donne ce qui suit, en prenant $w^1 = \sigma_k^{22}$:

$$\Omega_k^1 = \frac{\partial \Omega_k^{-1}}{\partial \sigma_k^{22}} = \frac{1}{d_k^2} \begin{pmatrix} -\sigma_k^{12} \sigma_k^{21} & \sigma_k^{11} \sigma_k^{12} \\ \sigma_k^{11} \sigma_k^{12} & -\sigma_k^{11} \sigma_k^{11} \end{pmatrix}$$

où $d_k = \sigma_k^{11} \sigma_k^{22} - (\sigma_k^{12})^2$; on déduit facilement, à partir de la formulation 84 que :

$$(\sigma_k^{11})^2 [\sigma_k^{22} T - \sum_t (\tilde{u}_{kt}^2)^2] - 2\sigma_k^{11} \sigma_k^{12} [\sigma_k^{12} T - \sum_t \tilde{u}_{kt}^1 \tilde{u}_{kt}^2] + (\sigma_k^{12})^2 [\sigma_k^{11} T - \sum_t (\tilde{u}_{kt}^1)^2] = 0$$

ce qui permet de déterminer par le maximum de vraisemblance des estimateurs convergents pour les w^l , on a donc :

$$\sigma_k^{ij} T - \sum_t \tilde{u}_{kt}^i \tilde{u}_{kt}^j = 0$$

ainsi, l'équation (3) s'écrit par :

$$\tilde{\sigma}_k^{ij} = \frac{1}{T} \sum_t \tilde{u}_{kt}^i \tilde{u}_{kt}^j.$$

On se trouve dans un cas où le système ((1), (2), (3)) admet une solution unique qui maximise \mathcal{L}_k . Dans le cas où la solution analytique est difficile à obtenir, une procédure itérative de calcul est nécessaire pour avoir une solution numérique qui maximise la vraisemblance.

Avec l'hypothèse de normalité des erreurs, l'estimateur itéré converge numériquement vers celui du maximum de vraisemblance, car pour ce dernier on itère jusqu'à ce que \mathcal{L}_k atteigne son maximum. Dans la pratique, un critère de convergence doit être choisi de façon précise.

La convergence entre les deux estimateurs tient au fait qu'avec la méthode du Zellner itéré on incrémente jusqu'à ce que l'on ne peut plus améliorer l'objectif d'une variance minimale i.e. jusqu'à ce que $\Omega_k(p)$ détermine un estimateur $\alpha_k(p+1)$ identique à $\alpha_k(p)$. Par conséquent, on itère jusqu'à ce que les estimateurs $[\alpha_k(p), \Omega_k(p)]$ coïncideront avec $(\tilde{\alpha}_k, \tilde{\Omega}_k)$ qui sont d'ailleurs déterminés de façon simultanée.

Ainsi, si les équations de l'expression suivante :

$$\text{tr} [(\Phi_k(p) - u_k(p)u_k'(p))\Phi_k^l(p)] = 0 \quad \text{avec } l = 1, 2, \dots, N_* \quad \text{et } p = 1, 2, \dots \quad (85)$$

ont une solution unique alors, par une procédure numérique, les estimateurs itérés s'identifierait avec le calcul des estimateurs du maximum de vraisemblance :

$$\alpha_k(p) = \tilde{\alpha}_k \quad \text{et} \quad \Omega_k(p) = \tilde{\Omega}_k.$$

9 Application

9.1 Introduction

Puisqu'on n'a pas une connaissance parfaite ou du moins des informations en terme de mesure sur le fonctionnement de certaines variables économiques et sur les effets qu'elles causent, la théorie économétrique propose des perspectives de compréhension en établissant, à la lumière des analyses statistiques et économiques, des modèles sous forme d'équations de comportements.

L'analyse économique et statistique préalable nous a permis de distinguer, jusqu'à un certain niveau dépendant de la problématique de recherche, les variables d'effets des variables causes. Ce qui nous a suggéré de considérer, puis de tester des modèles récurifs [Ghassan90] qui font apparaître une chaîne causale entre les variables endogènes.

Parmi les principales caractéristiques de l'économie du Maroc, on a retenu : le déficit du commerce extérieur de la majorité des secteurs d'activités, le déséquilibre sectoriel des investissements, et la faiblesse de l'auto-financement des secteurs. L'hypothèse soutenue dans ce travail est que la combinaison des deux premières caractéristiques conduit à la troisième.

Ces traits fondamentaux de l'économie du Maroc nous ont incité d'approfondir de façon originale le phénomène récent de l'endettement extérieur pour une structure productive dépendante en déterminant un système d'équations de comportements.

Les effets de l'endettement extérieur tiennent à leurs interactions avec les forces de concentration des revenus, et en matière de rentabilité la dette extérieure dépend en principe de son emploi dans différents secteurs et de la maîtrise du processus d'investissement dans ses derniers.

L'effort à effectuer pour une substitution progressive et accrue à l'importation des biens intermédiaires et finaux est considérable, étant donné la prédominance d'inputs industriels et d'outputs importés. Sur ce point l'unanimité est presque acquise, il en va de même de la nécessité de valoriser les matières premières et les produits intermédiaires d'exportation.

L'effort d'investissement conséquent à ces derniers objectifs exige, d'une manière qui peut paraître à plusieurs reprises comme étant paradoxale, des moyens financiers élevés. La réussite de cet effort nécessite d'élever les niveaux de revenus locaux, en maîtrisant la hausse probable des prix, pour accroître la demande effective et potentielle en produits industriels et autres.

Il est retenu donc un système récurif sectoriel comportant trois relations et dont le choix n'est pas dû au hasard. Puisque le déficit extérieur sectoriel correspond à un endettement extérieur chaque fois que les moyens propres font défaut, il est devenu donc un moyen puissant pouvant orienter l'investissement sectoriel et arbitrer sur le choix des projets de développement. Une telle influence laisse des traces sur la structure de la répartition des revenus et déterminera le niveau d'autonomie financière du secteur.

On postule donc que les mécanismes de l'endettement extérieur s'adaptent bien à une expérimentation économique réursive. Le modèle sectoriel comporte alors un déterminisme

restreint de type récursif stipulant des relations de causalité généralement non réversible entre les variables explicatives et celles expliquées.

La considération de plusieurs secteurs nécessite de préciser la nature et la forme de l'interdépendance entre secteurs différents. Cette interdépendance passe bien par les variables de consommation, d'investissement et de financement, on suggère qu'au lieu de la "variabiliser", il est intéressant de la repérer par les éléments aléatoires, que composent les erreurs notées u_{ik} où k est l'indice de l'équation et i est l'indice de l'individu i.e. le secteur d'activité. Il s'agit essentiellement de postuler que la covariance entre u_{ik} et u_{jk} n'est pas forcément nulle.

L'intérêt théorique de ce modèle nous a encouragé de procéder à une étude générale des modèles récursifs à double indice et analyser les différents aspects de l'estimation, que nous avons développé dans les premières sections.

9.2 Modèle et définitions

Le modèle individuel relatif au secteur d'activité i est défini comme suit :

$$\begin{cases} k &= f(z_1) + u_1 \\ m &= g(z_2, k) + u_2 \\ n &= h(z_3, k, m) + u_3 \end{cases}$$

où k , m et n sont des variables endogènes désignant respectivement le taux du déficit extérieur sectoriel, le taux d'investissement sectoriel et un indice d'auto-financement sectoriel, qui est une proxy-variable du taux d'auto-financement défini par le rapport entre la capacité d'auto-financement sectorielle et l'investissement sectoriel. La variable z correspond à la composante exogène du modèle, et u désigne une variable d'erreur aléatoire non observable.

Certaines variables exogènes méritent qu'on les explicite leur signification pour éviter des contre-sens. La variable v , qui désigne le taux de valeur ajoutée, exprime le revenu déduit par unité produite d'un secteur. La variable g , qui correspond à l'indice d'intégration, exprime le degré d'intégration par les achats à la structure productive.

La variable r , qui est un indicateur d'inégalité dans la répartition des revenus sectoriels, est plus difficile à interpréter. Sa formulation a demandé plus d'efforts. On soutient que la concentration des revenus commence au niveau sectoriel i.e. chacun des facteurs productifs essaiera d'avoir plus de revenus que les autres. On a supposé que le niveau de partage de la valeur ajoutée entre le profit et le salaire se détermine, dans l'activité de production et d'échange, selon des mécanismes de frottement socio-économique entre le groupe du capital et le groupe du travail. Ensuite, j'ai quantifié ces différents éléments de détermination à l'aide de règles statistique et comptable pour exprimer r en fonction du revenu du capital. Le résultat obtenu finalement est que l'indicateur r traduit le fait que plus le revenu du capital dans le secteur augmente plus r est élevé.

9.3 Résultats économétriques et interprétation

Il est impératif au préalable de préciser que les résultats sont obtenus par application de la méthode des moindres carrés ordinaires, puis par la méthode de l'Aitken itéré, pour être ensuite soumis à une série de tests étudiant la qualité statistique et économique des estimateurs. On a utilisé les tests habituels de Fisher-Student et de Durbin-Watson, mais le test clé de mon modèle est le test d'exogénéité.

Ce test a été réalisé pour la deuxième équation du modèle sur la variable k_i , ce qui revient à tester, à partir de la statistique t , l'hypothèse de nullité des paramètres des variables résiduelles u_i estimées. Les cinq paramètres de mon application me donne $|t| < t_\alpha$ pour $\alpha = 0.05$ ou $\alpha = 0.01$. Ce qui autorise d'adopter l'hypothèse nulle pour tous les paramètres associés aux variables u_i estimées, j'en déduis que les variables k_i peuvent être exogènes dans le modèle où elles sont explicatives.

Au niveau de la troisième équation du modèle individuel de base, on n'a pas introduit la régression étendue telle qu'elle est définie par Holly [Holly82] [Holly84], car au lieu de tester l'exogénéité d'un sous-ensemble de variables endogènes explicatives, on l'a testé sur l'ensemble. On a introduit donc dans la troisième équation les variables de résidus associées aux variables endogènes explicatives, et à partir de la statistique F de la loi de Fisher-Snedecor que l'on a testé simultanément la nullité des variables résiduelles u_{ij} estimées. Dans mon application, pour les cinq secteurs j'ai $F_* < F(\alpha)$ pour $\alpha = 0.01$; ce qui nous autorise d'accepter l'hypothèse de nullité pour les coefficients associées aux variables u_{ij} estimées, et donc on adopte l'exogénéité des variables k et m dans l'équation de n .

9.3.1 Le secteur de l'agriculture-forêt et pêche

L'estimation du modèle agricole a donné les résultats suivants :

$$\begin{cases} k_1 &= 0.17^\circ - 0.42v_1^* - 0.05h_1^{**} + 0.26w_1^{**} - 0.06p_1^* \\ m_1 &= -1.6^{**} - 0.03k_1^\circ + 2.35g_2^{**} + 0.16r_2^{**} + 0.29i^{**} \\ n_1 &= 4.78^{**} - 48.9k_1^{**} - 4.69v_1^{**} - 24.69r_1^{**} - 0.15h_1^{**} - 0.53m^* \end{cases}$$

le secteur agricole est particulièrement différent des autres secteurs par son importance stratégique, qui consiste à atteindre voire dépasser l'auto-suffisance alimentaire. Il est aussi différent par sa structure productive et financière.

La première équation montre que si le taux de valeur ajoutée augmente de 1 point alors le taux du déficit extérieur baisserait de 0.42 point; de manière simultanée, l'impact conjugué des variables k_1 et v_1 fait que l'indicateur d'autonomie financière est susceptible de s'accroître de 15.848 points. D'où la preuve de l'importance capitale de la valeur ajoutée agricole, dont l'augmentation apparaît diminuer le poids de la domination économique et commerciale par l'échange agricole avec l'extérieur, tout en améliorant son indépendance financière.

Cette dernière pourrait s'améliorer encore davantage si le degré d'inégalité dans la répartition des revenus agricoles diminue, comme le montre la troisième équation.

Le secteur agricole semble réagir de manière surprenante aux variations du taux d'intérêt. La deuxième équation montre que si le taux d'intérêt empirique noté i s'accroît de 1 point alors m_1 augmenterait de 0.29 point. Les exploitants agricoles pourraient vraisemblablement accepter le capital financier, quoi que leurs charges seraient augmentées, soit dans l'espoir d'avoir des récoltes et des écoulements satisfaisants, soit pour rembourser partiellement leurs dettes précédentes; ce dernier point est moins évident à justifier à partir de la deuxième équation.

L'absence de réactions du taux d'investissement par rapport au taux du déficit extérieur correspond probablement à ce que le secteur agricole n'est pas suffisamment modernisé et qu'il ne bénéficie pas de crédits agricoles appropriés aux réalités sociales de ces agents. D'autant plus, qu'il y a lieu de leur assurer des prix rémunérateurs.

9.3.2 Le secteur des industries agro-alimentaires

Le système récursif relatif à mon deuxième secteur se présente comme suit :

$$\begin{cases} k_2 &= 0.26^{**} - 0.92v_2^{**} + 0.01x_2^* - 0.26r_2^{**} - 0.01p_2^{**} - 0.04s^* \\ m_2 &= 47.8^{**} + 88.9k_2^{**} + 2.26g_2^{**} - 49.4w_2^{**} + 0.05h_2^* - 24.8x_2^{**} + 0.12p_2^{**} \\ n_2 &= -15.98^{**} - 60.15k_2^{**} + 6.45m_2^{**} + 66.61v_2^{**} - 2.81h_2^{**} + 0.99x_2^* \end{cases}$$

la première équation montre que si v_2 s'accroît de 0.01 point alors le déficit extérieur du secteur décroît de presque 0.01 point. De manière simultanée le taux d'investissement diminue de presque 0.818 point, et l'indicateur d'autonomie financière baisserait de presque 4.056 points.

Symétriquement, si v_2 décroît de 0.01 point alors k_2 augmente de presque 0.01 point. De façon simultanée m_2 s'accroît de presque 0.818 point, et n augmenterait de presque 4.056 points.

Ces deux propositions symétriques montrent l'importance de l'impact économique d'une petite variation du taux de la valeur ajoutée dans le secteur des IAA, qui est loin d'être homogène. L'accroissement de v paraît avoir exigé dans le passé de lourd fond financier, ce qui se traduit par la suite et à terme par une diminution de l'autonomie financière contemporaine et de l'investissement contemporain.

D'un autre côté si p_2 varie à la baisse de 1 point alors le taux d'investissement augmenterait de 0.769 point, et l'indicateur de l'autonomie financière augmenterait de presque 4.358 points. Il ressort de ces impacts que l'avantage concurrentiel par les prix quoi qu'il induit un accroissement du déficit avec l'extérieur, il améliore nettement l'autonomie financière du secteur et dans une moindre mesure le taux d'investissement.

9.3.3 Le secteur de l'énergie

Les résultats du modèle analysant le secteur de l'énergie sont les suivants :

$$\begin{cases} k_3 &= 0.18^{**} - 0.5h_3^{**} + 0.2x_3^{**} - 0.05r_3^{**} - 0.09p_3^{**} \\ m_3 &= 0.5^{**} - 0.4k_3^{**} - 0.45v_3^{**} - 0.26r_3^{**} - 0.06c_3^{**} \\ n_3 &= -9.49^{**} - 33.59k_3^{**} + 8.22m_3^{**} + 34.24g_3^{**} - 4.35r_3^{**} \end{cases}$$

si r_3 décroît de 1 point alors m_3 augmenterait de 0.24 point, et de manière simultanée n_3 augmenterait aussi de presque 0.293 point. Ce résultat montre que l'amélioration dans la répartition des revenus du secteur a des effets bénéfiques sur son investissement et sur sa capacité d'autonomie financière.

Il ressort de ce modèle que si l'indice des prix à la production p_3 s'accroît de 1 point alors k_3 diminuerait de 0.09 point, surtout si les prix à la consommation sont déterminés par l'administration d'un gouvernement. L'effet de cette variation unitaire du prix est double, car m_3 s'accroît de 0.036 point et n_3 de presque 2.727 points. Ce qui me semble cohérent, vu que l'accroissement des prix augmente la capacité financière du secteur et du coup son effort d'investissement.

9.3.4 Le secteur des mines

Les résultats de ce secteur sont les suivants :

$$\begin{cases} k_4 &= -0.49^{**} - 0.95v_4^{**} - 0.001h_4^{**} + 1.2w_4^{**} + 0.02p_4^{**} + 0.48x_4^{**} \\ m_4 &= -0.14^{**} + 1.09k_4^{**} + 0.82v_4^{**} + 0.39r_4^{**} + 0.3m^{**} \\ n_4 &= 10.4^{**} - 22.38k_4^{**} + 14.96m_4^{**} - 19.56v_4^{**} - 9.03r_4^{**} + 0.04h_4^{**} \end{cases}$$

il s'agit d'un secteur central pour l'économie du Maroc par ses ressources naturelles et financières et par ses activités. Son poids apparaît par exemple au niveau de la seconde équation, car en moyenne, dans 100 Dh d'investissement il y aurait autour de 30 Dh consacré au secteur minier.

Si on suppose une variation unitaire négative de k_4 i.e. un accroissement unitaire du taux de l'excédent de l'échange avec l'extérieur du secteur minier, alors l'indicateur de l'autonomie financière s'accroîtrait de presque 24.394 points, ce qui prouve le niveau élevé de fonds financiers que procurent les minerais pour l'économie du Maroc.

Quant à la transformation locale des matières premières, elle permet d'entraîner l'accroissement de l'excédent commercial avec l'extérieur, car si v_4 s'accroît de 1 point alors k_4 baisserait de 0.95 point. De manière simultanée, m_4 diminuerait de presque 0.215 point et n_4 diminuerait de presque 1.522 points, ce qui montre l'accroissement de la valeur ajoutée du secteur est loin de ne pas être coûteux.

L'inégalité dans la répartition des revenus du secteur exerce un impact non négligeable sur m_4 et n_4 . Si je suppose une variation unitaire positive de r_4 alors m_4 s'accroîtrait de

0.39 point et n_4 augmenterait de presque 3.195 points; ce qui montre qu'apparemment le partage du revenu du secteur en faveur du capital accroît son autonomie financière, mais décroît le niveau auquel pourrait aspirer le facteur travail. Ce résultat mérite une analyse plus fine en terme par exemple de la théorie des jeux concernant le partage des gains.

9.3.5 Le secteur autres-industries

Son système récursif estimé se présente comme suit :

$$\begin{cases} k_5 &= 0.32^{**} + 1.66k_5^{**} + 0.75x_5^{**} - 0.55p_5^{**} \\ m_5 &= -0.1^\circ - 0.45k_5^{**} - 0.12w_5^* - 0.26r_5^{**} + 0.33g_5^* \\ n_5 &= 16.99^{**} + 17.74k_5^{**} + 53.29m_5^{**} - 42.66g_5^{**} + 9.92r_5^{**} + 5.6h_5^* \end{cases}$$

la structure que fait ressortir ce modèle est que l'intégration industrielle paraît très coûteuse. La variation unitaire de g_5 a un effet double, m_5 s'accroîtrait de 0.33 point et n_5 diminuerait de presque 25.074 points, ce qui montre que le secteur industriel aurait besoin de fonds financiers élevés pour réaliser son intégration. Symétriquement, une baisse unitaire du degré d'intégration de ce secteur se traduit par un fort accroissement de l'indicateur de l'autonomie financière du capital étranger, qui est dominant dans ce secteur et à comportement fortement stratégique.

D'un autre côté, moins le secteur est déficitaire i.e. si on a une variation unitaire négative de k_5 alors m_5 s'accoîtrait de 0.45 point et n_5 augmenterait de presque 6.24 points. Ce qui montre qu'il est toujours opportun de chercher un avantage comparatif à partir de l'échange avec l'extérieur, et si cette condition est réalisée alors il est vraisemblable d'améliorer les résultats économiques du secteur.

On obtiendrait les mêmes effets, mais dans une moindre mesure si on réduit l'inégalité dans la répartition du secteur de 1 point.

9.4 Quelques résultats de synthèse

Le traitement statistique de l'échange extérieur au niveau global ou macro-économique montre que la structure productive s'oppose par la rigidité de l'offre globale, car en moyenne $\Delta Y/\Delta I < 1$ et $\Delta Y/\Delta C < 1$. Cependant, la demande continue de s'accroître dans un contexte de dépendance économique et de choix de projets à productivité immédiate et à rentabilité monétaire maximale.

D'où un accroissement élevé du déficit extérieur global. L'intérêt du modèle a consisté aussi à découvrir dans quelle mesure chaque secteur est responsable du déficit extérieur global.

L'équation de l'échange avec l'extérieur me suggère les relations suivantes prises de façon isolées. Pour le secteur agricole, on a la relation $k_1 = -1.82v_1$, qui signifie que variation positive de la valeur ajoutée agricole permet au secteur de se désendetter; pour le secteur

industriel, on a la relation $k_5 = 3.62v_5$, qui montre que : pour un produit industriel de 1 Dh, il y aurait une dette extérieure de 3 Dh 62.

Il est donc nécessaire de modifier la structure de l'offre dans le plus court terme et d'améliorer la contribution locale optimale aux besoins de l'économie en accroissant la valeur ajoutée de chaque secteur. L'exportation devrait être principalement le complémentaire à la production pour le marché local, sauf pour les branches où l'avantage comparatif est net sans se désintégrer de la structure productive.

L'analyse statistique a montré que la spécialisation dans le secteur minier et quelques branches des IAA n'a pas résulté d'une autonomie de développement de l'économie du Maroc. La difficulté qu'il y a à financer les investissements s'explique par l'orientation du capital vers certains secteurs pour des raisons de maximisation du profit.

En effet, l'étude des liens de causalité, entre m_i et les variables g_i et r_i , montre qu'il ne suffit pas d'agir sur le niveau des investissements, mais qu'il est nécessaire d'agir sur leur orientation dans le sens d'une intégration entre les secteurs. Pour éviter une évolution en dents de scie dans presque tous les secteurs productifs, il est important d'établir un ordre de priorité de l'investissement sectoriel basé sur les coefficients de g_i ; on a obtenu l'ordre suivant s_1, s_2, s_3, s_5 , qui ont respectivement les valeurs des coefficients 2.35, 2.26, 0.45, 0.33, et vérifient une variation induite de m_i positive.

Ce résultat montre que les priorités sectorielles ne sont pas celles formulées par les plans économiques, qui s'attendraient à ce que l'ordre soit s_5, s_1, s_4, s_2, s_3 . L'intégration proposée par le modèle se réaliserait dans le sens d'une planification diversifiant le tissu industriel, accélérant les échanges inter-sectoriels, recherchant des productivités élevées et améliorant le niveau de vie de chacun et de tout futur actif.

La faible intégration économique conduit à l'étroitesse du marché local et réciproquement. La résolution de ce cycle passe par un processus de déconcentration dans la répartition des revenus. Une telle stratégie n'est réalisable que si elle est clairement définie tout en engageant un maximum d'agents actifs et rationnels.

L'analyse statistique macro-économique m'a montré que l'économie du Maroc s'est engagé dans un processus d'effort à la consommation plus qu'à l'investissement, d'où l'urgence d'une intégration économique organisée. Sinon, nous serons amené à exporter de plus en plus d'agrumes pour importer des produits agricoles et alimentaires. Le capital national doit faire l'effort nécessaire pour que le secteur industriel et commercial évoluent en parallèle.

En général, il est admis que la concentration des revenus permet au secteur d'avoir une autonomie financière, qui facilite son intégration à la structure productive. Or, certains modèles contredisent cette proposition, car les coefficients associés à r_i sont négatifs, comme pour les secteurs 1, 2 et 4.

Les difficultés financières et de gestion économique des pouvoirs publics les ont conduit à privatiser certaines branches de l'industrie extractive, du raffinage de pétrole, de l'industrie de manufacture, du transport et communication, du commerce et des institutions financières. L'ampleur des programmes d'infrastructure hydraulique, agricole et de communication, de leurs coûts élevés et des délais pour leurs réalisations font que la marge de choix pour les

investissements publics serait réduite. Or, l'Etat doit identifier les causes structurelles du déséquilibre pour corriger les tendances de la consommation, de l'investissement et de la répartition des revenus.

Les industries de croissance tels l'électricité, la chimie, l'automobile, le pétrole et la mécanique doivent avoir leur importance dans les programmes de développement pour constituer un groupe cohérent d'activités liées par de nombreuses interdépendances. Le problème de l'utilisation optimale des ressources s'impose, en plus d'une grande discipline dans l'économie des moyens et dans la compression des coûts sans sacrifier la qualité.

Références

- [Acher et al.70] Acher (J.) et Gardelle (J.). – *Algèbre linéaire et programmation linéaire*. – Paris, Dunod, 1970, 2-ème édition, volume 5.
- [Balestra72] Balestra (P.). – *Calcul matriciel pour économistes*. – Albeuve (Suisse), Castella, 1972.
- [Balestra76] Balestra (P.). – *La dérivation matricielle (Techniques et résultats pour économistes)*. – Rapport technique, I.M.E., Dijon, 1976.
- [Bement et al.69] Bement (T.R.) et Williams (J.S.). – Variance of wieghted regression estimators when sampling errors are independent and heteroscedastic. *J.A.S.A.*, dec. 1969.
- [Dhrymes73] Dhrymes (P.J.). – Restricted and unrestricted reduced forms, asymptotic distribution and relative efficiency. *Econometrica*, janv. 1973.
- [Dhrymes74] Dhrymes (P.J.). – *Econometrics, statistical foundation and applications*. – New York, Harper and Row, 1974.
- [Gantmacher66] Gantmacher (F.R.). – *Théorie des matrices*. – Paris, Dunod, 1966.
- [Ghassan90] Ghassan (H.). – *Modèles récurifs à double indice, concentration des revenus et analyse d'une structure sectorielle*. – Université de Bourgogne, Dijon, Doctorat "Economie mathématique et conométrie", 1990.
- [Holly82] Holly (A.). – Testing recursiveness in a triangular simultaneous equations model. *Cahiers et séminaires d'économétrie*, vol. 15, 1982.
- [Holly84] Holly (A.). – *Tests d'exogénéité*. – Rapport technique, I.M.E., Dijon, 1984.
- [Intriligator et al.84] Intriligator (M.D.) et Griffiches (Zui). – *Handbook of econometrics*. – North Holland, 1984.

- [Kmenta et al.68] Kmenta (J.) et Gilbert (Roy F.). – Small sample properties of alternative estimators of seeming by unrelated regressions. *J.A.S.A.*, vol. 63, 1968.
- [Kmenta et al.74] Kmenta (J.) et Oberhofer (W.). – A general procedure for obtaining maximum likelihood estimates in generalized regression models. *Econometrica*, mai 1974.
- [Malinvaud81] Malinvaud (E.). – *Méthodes statistiques de l'économétrie*. Paris, Dunod, 1981 volume 3-ème édition.
- [Philipps76] Philipps (P.C.). – The iterated minimum-distance estimator and the quasi-maximum likelihood estimator. *Econometrica*, May 1976.
- [Ruble68] Ruble (W.L.). – *Improving the computation of simultaneous stochastic linear equations estimates*. – Michigan state University, Dept. of Agricultural Economics, oct. 1968.
- [Taylor77] Taylor (W.E.). – Small sample properties of a class of two stage aitken estimators. *Econometrica*, March 1977.
- [Wold et al.60] Wold (H.O.) et Strotz (R.H.). – Recursive and non recursive systems. *Econometrica*, vol. 28, n° 2, 1960.
- [Zellner et al.62] Zellner (A.) et Theil (H.). – Three-stage least squares, simultaneous estimation of simultaneous equations. *Econometrica*, vol. 30, 1962.
- [Zellner62] Zellner (A.). – An efficient method of estimating seemingly unrelated regression and tests for aggregation bias. *J.A.S.A.*, vol. 57, 1962.

Documents statistiques et données

Direction de la Statistique

Comptes nationaux rétrospectifs 1969–1979 (Documents de la comptabilité nationale).

Notes méthodologiques sur l'élaboration du TFI 1969, Rapport n14. 1972.

Etudes économiques et statistiques n1 (TFI 1975) 1982.

TES de l'économie du Maroc pour l'année 1978.

Annuaire de l'économie du Maroc pour les années 69, 75, 77 et 79.

Direction de l'industrie

Notes sur l'évolution de l'investissement industriel.

Division des études et de la balance des paiements (Office des changes)

Statistiques des échanges extérieurs du Maroc : de 71 à 83.

Ministère du plan :

Le Maroc en chiffres : de 69 à 83.

Liste des variables

k_i	Taux du Déficit Extérieur	$(M_i - X_i)/P_i$
m_i	Taux d'Investissement	$FBCF_i/VA_i$
n_i	Indicateur d'Auto-financement	$1 - (M_i - X_i)/FBCF_i$
v_i	Taux de Valeur Ajoutée	VA_i/P_i
h_i	Taux de Couverture Commerciale	X_i/M_i
g_i	Indice d'Intégration	CI_i/P_i
w_i	Taux d'Intraversion	$1 - (X_i/VA_i)$
p_i	Indice des Prix à la Production	Base 100 en 1969
r_i	Indicateur d'inégalité des Revenus	$(X_i - M_i)/(VA_i - CI_i), V_i \neq CI_i$
x_i	Taux d'Extraversion	$(M_i + X_i)/VA_i$
c_i	Taux d'Efficacité du Capital	$(VA_{it} - VA_{it_0})/\sum_{i=t_0+1}^t I_i$
i	Taux d'Intérêt Empirique	J_t/E_{t-1}
m	Taux d'Investissement global	$FBCF/Y$
s	Taux d'Epargne	S/Y

où :

X_i	Exportation du secteur i
M_i	Importation du secteur i
P_i	Production du secteur i
$FBCF_i$	Formation Brute du Capital Fixe du secteur i
VA_i	Valeur Ajoutée du secteur i
CI_i	Consommation Intermédiaire du secteur i
Y	Production macro-économique
$FBCF$	Investissement macro-économique
S	Epargne nationale
J_t	Intérêt de la dette de l'année t
E_{t-1}	Dette Extérieure annuelle de l'année $t - 1$.