UNE SYNTHÈSE DE L'EXOGÉNÉITÉ DANS LES MODÈLES VECTORIELS À CORRECTION D'ERREURS

C. RAULT*

RÉSUMÉ

Cet article propose une revue sur l'exogénéité dans les modèles Vectoriels à Correction d'Erreurs (VAR-ECM) à la Johansen, en insistant sur les points communs et les différences avec la littérature maintenant bien établie sur l'exogénéité dans les modèles vectoriels autorégressifs (VAR). L'étude de l'exogénéité a en effet été faite de manière détaillée dans le cadre stationnaire par Florens, Mouchart et Richard (1979), Engle *et alii* (1983), Florens et Mouchart (1985), ainsi que par Monfort et Rabemananjara (1990). Dans le cadre non stationnaire des modèles VAR-ECM, les travaux théoriques sur l'exogénéité se sont multipliés ces dernières années (*cf.* en particulier Johansen, 1992b ; Urbain, 1992, 1995 ; Hendry et Mizon, 1993 ; Ericsson *et al.*, 1998 ; Rault, 2000, 2005 ; Rault et Pradel, 2003), et des avancées récentes ont vu le jour. Il s'agit ici de reprendre, synthétiser et articuler ces avancées ayant trait à l'exogénéité, dans les modèles linéaires à variables non stationnaires intégrées d'ordre 1 et co-intégrées et de présenter certaines pistes de recherche de cette littérature.

Mots-clés : Modèles VAR, VAR-ECM, co-intégration, représentation canonique, exogénéité faible, exogénéité forte, non-causalité.

ABSTRACT

This paper proposes a survey on weak-exogeneity in Vector Error Correction Models (VAR-ECM), and makes the connection with the literature now well-established on weak-exogeneity in Vectorial Autoregressive Models (VAR). Indeed, in the stationary case, considerable interest has been shown in the issue of weak exogeneity testing by Florens, Mouchart and Richard (1979), Engle *et alii* (1983), Florens and Mouchart (1985), as well as by Monfort and Rabemananjara (1990). In the non stationary case theoretical and applied works devoted to exogeneity have increasead these past years (e.g. Johansen, 1992b ; Urbain, 1992, 1995 ; Hendry and Mizon, 1993 ; Ericsson *et al.*, 1998 ; Rault, 2000, 2005 ; Rault and Pradel, 2003), and some recent advances have recently emerged. The aim of this article is to unify, supplement, and synthesize previously recent results on exogeneity in linear VAR-ECM models with I (1) variables and to present some further possible developments of this literature.

Keywords : VAR models, VAR-ECM, cointegration, canonical representation, weak exogeneity, exogeneity, non causality.

---

* LEO, Université d'Orléans, CNRS UMR 6221, Rue de Blois, B.P.6739, 45067 Orléans Cedex 2.
E-mail : chrault@hotmail.com, web-site : http ://membres.lycos.fr/chrault/index.html.
Je remercie les éditeurs ainsi que deux rapporteurs anonymes pour leurs remarques constructives sur une première version de ce texte. Je demeure seul responsable des éventuelles insuffisances de ce travail.

1. Introduction

La modélisation économétrique a connu un profond renouvellement ces trente dernières années à la fois sur un plan empirique et sur un plan théorique. Alors que jusqu'au début des années soixante-dix les modèles structurels constituaient l'axe de recherche principal d'un point de vue empirique, et que leur utilisation semblait être solidement établie dans la pratique, la deuxième moitié des années soixante-dix est venue progressivement remettre en cause cet acquis. Trois étapes principales ont marqué cette évolution. Tout d'abord, la critique de Lucas (1976) a porté sur le traitement des anticipations des agents et l'invariance des paramètres des équations structurelles à la politique économique. Ensuite, la critique de Sims (1980) a remis en question la démarche suivie par la Cowles Commission, en particulier l'utilisation de contraintes *a priori* « non crédibles » pour identifier les paramètres des modèles structurels. Enfin, les développements récents sur la non stationnarité des séries économiques (Nelson et Plosser, 1982 ; Engle et Granger, 1987 ; Johansen, 1991) ont conduit à des modifications importantes dans la manière d'opérer la modélisation.

Le renouvellement de la modélisation économétrique s'est fait essentiellement dans quatre directions : l'amélioration des modèles macro-économétriques, le développement des modèles vectoriels autorégressifs (VAR), des modèles d'équilibre général dynamique stochastique (DSGE) et des modèles vectoriels à correction d'erreurs (VAR-ECM) [1].

La première direction s'est attachée à incorporer certaines innovations dans les modèles macro-économétriques utilisés à des fins de prévisions et d'évaluation de politiques économiques. On pense notamment pour la France, aux modèles *Amadeus* (Insee), *Banque de France (*Banque de France *et al.*, 1996, *Hermes* (Erasme-École Centrale Paris), *Metric* (Direction de la Prévision), *Mosaïque* (OFCE) ; pour le Royaume Uni, au modèle construit par la London Business School (LBS) ; pour les USA, au modèle MPS (Ando et Modigliani, 1969). Bien que ces modèles aient certes connu des améliorations significatives, leur évolution n'en est pas moins demeurée extrêmement lente et ils se situent encore, pour la plupart, dans la lignée de l'approche initiée par la Cowles Commission. En particulier, la distinction entre variables endogènes et exogènes est souvent encore opérée *a priori* et des restrictions souvent non testées sont imposées sur leurs paramètres afin de les identifier.

La seconde direction, pour une meilleure gestion de la dynamique, suit une approche descriptive et est moins ambitieuse sur l'analyse économique sous-jacente au modèle. Elle regroupe, à la suite de la méthodologie développée par Doan, Litterman et Sims (1984), Litterman (1986), les modèles VAR non-contraints, bayésiens et structurels. Ceux-ci sont fréquemment utilisés dans les études empiriques à des fins de prévisions, mais sont d'une utilité

1. On considère ici uniquement des modèles économétriques dynamiques à temps discret, ce qui nous conduit à ne pas parler des modèles micro-économétriques, ou des modèles de finance par exemple. On n'abordera pas non plus ici les modèles à seuils, ni les modèles à changement de régimes markoviens, qui ont bien sûr connu un développement considérable depuis le début des années quatre-vingt.

très limitée pour l'évaluation des politiques économiques [2]. L'approche VAR structurelle, par exemple, cherche à spécifier un cadre d'analyse interprétable économiquement en imposant des restrictions sur la matrice de variance-covariance des divers chocs [3]. Néanmoins, elle se limite souvent à la mise en œuvre d'une analyse impulsionnelle et à l'étude des contributions respectives de différentes impulsions à la dynamique d'un système de variables, sans s'attacher à modéliser de façon explicite la structure de l'économie sous la forme de relations de comportements spécifiques.

La troisième direction découle d'une sorte de renoncement raisonné aux méthodes d'estimation économétriques, auxquelles on préfère des méthodes de calibration directe des paramètres structurels d'un modèle théorique. Cette direction est étroitement liée à la méthodologie d'équilibre général dynamique stochastique (DSGE) utilisée dans la littérature ayant trait aux cycles réels (RBC). Plus précisément, elle s'est développée à la suite des travaux pionniers de Kydland et Prescott (1982), Long et Plosser (1983), et fournit un modèle d'équilibre général intertemporel de l'économie, qui repose sur les décisions d'optimisation des agents et des firmes. À l'origine, les modèles DSGE mettaient l'accent sur les facteurs réels (i.e. les chocs de productivité), mais plus récemment ils se sont développés dans plusieurs voies, pour tenir compte notamment des effets nominaux, des coûts d'ajustements, de l'hétérogénéité et du progrès technique endogène. En outre, bon nombre de modèles DGSE (*cf.* Kim et Pagan, 1995; Christiano, Eichenbaum et Evans, 1998) peuvent être approchés par des modèles VAR contraints, ce qui les rend plus conformes aux autres types de modélisation [4]. L'intérêt intrinsèque des modèles DGSE renforce cependant la nécessité de bien connaître la valeur des paramètres structurels, dans la mesure où ces valeurs conditionnent en grande partie la performance des modèles.

La quatrième direction, la plus récente, permet de gérer le problème de la non stationnarité des variables et concerne les modèles vectoriels à correction d'erreurs (VAR-ECM). Elle résulte d'une jonction entre l'approche économétrique dynamique selon la méthodologie de Hendry (1983), Hendry *et al.* (1988), Hendry (1995) et l'apport des méthodes statistiques d'analyse des séries temporelles de Box et Jenkins (1976). Les modèles VAR-ECM émergent en effet naturellement si l'on appréhende les propriétés statistiques des séries utilisées au travers des concepts d'intégration et de co-intégration, ce dernier fournissant alors un support statistique formel à l'utilisation de ces modèles. Depuis l'article fondamental de Engle et Granger (1987), l'étude de la co-intégration a donné lieu à une littérature importante, qu'il est impossible de présenter entièrement [5]. Deux grandes familles d'approches peuvent néanmoins être distinguées au sein de cette direction :

2. *cf.* par exemple Cooley et Leroy (1985), ou Pagan (1987).

3. Cette approche a été utilisée par Blanchard et Quah (1989), Clarida et Gali (1994), Astley et Garrat (1996) par exemple.

4. Pour une revue de la littérature sur les modèles DGSE, se reporter à Cooley (1995).

5. On consultera avec profit à ce sujet les tentatives de synthèse de Phillips et Loretan (1991) et Campbell et Perron (1991), ainsi que l'expérience de Monte Carlo de comparaison de plusieurs méthodes proposée par Gonzalo (1994) et Haug (1996).

– l'approche descriptive repose sur des modèles à correction d'erreurs dit «complets», c'est-à-dire dont toutes les variables sont traitées comme endogènes. L'efficacité de ce type d'approche a été soulignée, par exemple, par Phillips (1991) ou Johansen (1995b).
– l'approche structurelle s'appuie sur des modèles conditionnels, encore qualifiés de modèles «partiels». Ces modèles, qui décrivent seulement l'évolution de certaines variables, possèdent les mêmes propriétés d'efficacité que les VAR-ECM complets, lorsque l'hypothèse d'exogénéité des variables non modélisées est satisfaite.

Cette deuxième famille d'approche, basée sur des modèles de dimension moindre, est beaucoup plus facilement gérable d'un point de vue pratique. Elle permet également de tester des hypothèses structurelles découlant d'une théorie économique, ce qui favorise l'interprétation des paramètres de ces modèles. Par ailleurs, les estimateurs des vecteurs de co-intégration admettent, dans ce cas encore, des distributions limites standards (*cf.* Johansen, 1992a; Boswijk, 1992). Il faut cependant souligner qu'une hypothèse d'exogénéité erronée peut conduire à une inférence invalide et à des résultats biaisés. Tester cette hypothèse constitue par conséquent une étape cruciale et nécessaire dans tout exercice de modélisation de modèles partiels. L'économiste est ainsi confronté empiriquement à un choix parfois complexe entre d'une part, le risque de perte d'efficacité et l'obtention de résultats biaisés en utilisant un modèle partiel, et d'autre part, la perte de «gérabilité» et de précision en travaillant avec un modèle complet, composé d'un grand nombre de variables.

Cet article se place précisément dans le cadre de cette quatrième direction et de cette seconde famille d'approche, et a pour motivation de proposer une revue de la littérature récente de l'exogénéité dans les modèles VAR-ECM, tout en insistant sur les points communs et les différences avec la littérature maintenant bien établie sur l'exogénéité dans les modèles VAR. L'étude de l'exogénéité a en effet été faite de manière détaillée dans le cadre stationnaire par Florens, Mouchart et Richard (1979), Engle *et alii* (1983), Florens et Mouchart (1985), ainsi que par Monfort et Rabemananjara (1990). Dans le cadre non stationnaire des modèles VAR-ECM, les travaux théoriques sur l'exogénéité se sont multipliés ces dernières années (*cf.* en particulier Johansen, 1992b; Giannini et Mosconi, 1992; Urbain, 1992, 1995; Hendry et Mizon, 1993; Ericsson *et al.*, 1998; Hecq *et al.*, 2001; Rault, 2000, 2005; Rault et Pradel, 2003), et des avancées récentes ont vu le jour. Il s'agit ici de reprendre, synthétiser et articuler ces avancées ayant trait à l'exogénéité, dans les modèles linéaires à variables non stationnaires intégrées d'ordre 1 et co-intégrées [6].

6. La discussion suivante (suggérée par l'un des rapporteurs) permet d'éclairer le débat pour les non économètres. L'économétrie traite typiquement de séries chronologiques multivariées non stationnaires et cherche à mettre en évidence des relations « stables ». Pour y parvenir on peut recourir à deux méthodes : conditionner et marginaliser. Conditionner revient à rejeter dans les variables conditionnantes la non stationnarité et à utiliser des relations « conditionnellement stationnaires ». Marginaliser revient dans le cas linéaire à considérer des relations de co-intégration. L'intérêt de cet article est de combiner les deux. L'esprit du texte est de commencer par marginaliser (modèle non conditionnel avec co-intégration) puis d'introduire des hypothèses d'exogénéité.

La suite de ce travail s'organise comme suit. Dans la section 2, nous présentons le concept d'exogénéité dans le cadre des modèles VAR et faisons le lien entre les modèles descriptifs et les modèles structurels, en nous référant aux travaux de Monfort et Rabemanajara (1990). Dans la section 3, nous nous posons alors la question de savoir en quoi les conditions de d'exogénéité et de non-causalité données dans la section 2 se trouvent modifiées lorsqu'il existe des relations de co-intégration entre les variables étudiées. Ceci nous conduit alors tout naturellement à exposer les conditions proposées par Toda et Phillips (1991), Johansen (1992b), Hendry et Mizon (1993) [7] dans le cadre non stationnaire des modèles VAR-ECM et à discuter des principales différences avec les modèles VAR. Nous montrons par ailleurs que ces conditions usuelles de non-causalité et d'exogénéité peuvent s'avérer, dans certains cas, « trop contraignantes » théoriquement et également, pour certaines d'entre elles, difficiles à tester. Nous exposons alors en section 4 le cadre d'analyse développé par Rault (2000, 2005) et Rault et Pradel (2003), qui permet de clarifier statistiquement les hypothèses d'exogénéité faible, de non-causalité et d'exogénéité forte dans un modèle VAR-ECM. Pour ce faire, ces auteurs associent à chacune de ces notions une décomposition canonique de la matrice de co-intégration, qui conduit à la formulation de conditions nécessaires et suffisantes et à un test direct de chacune de ces hypothèses (*cf.* section 5). Dans la section 6 nous mettons en œuvre une expérience de Monte Carlo afin d'étudier les propriétés des procédures séquentielles de tests développées dans la section 5. Une dernière section présente certaines pistes de recherche de cette littérature.

2. Relations entre les modèles VAR et les modèles structurels

2.1 Enjeux et rappels

L'examen du concept d'exogénéité est particulièrement important en économétrie, car il est inhérent au développement d'une forme structurelle. Néanmoins, la propriété d'exogénéité est parfois posée *a priori* en référence à une théorie économique, sans avoir été testée au préalable. Les différentes formes d'exogénéité peuvent être classées en deux types : une exogénéité pour les perturbations et une exogénéité pour des paramètres. La première classe d'exogénéité englobe les concepts de prédétermination et d'exogénéité stricte. La deuxième regroupe les trois concepts d'exogénéité faible, d'exogénéité forte [8] et de super exogénéité [9] . Ces derniers constituent trois degrés d'exo-

7. *cf.* également Ericsson *et alii* (1998).

8. Il est important de bien noter une différence entre les notions d'exogénéité et de causalité (nécessaire pour définir l'exogénéité forte). L'exogénéité est toujours relative à un modèle, tandis que la causalité est une notion intrinsèque à la vraie loi (appelée quelquefois « Data Generating Process »).

9. Pour une définition formelle de chacun de ces concepts, nous renvoyons le lecteur aux travaux de Florens, Mouchart et Richard (1979), Engle *et alii* (1983), et de Florens et Mouchart (1985). Il est à noter que nous ne considérons pas ici le concept de super-exogénéité.

généité différents, suivant l'objectif recherché par le modélisateur (l'inférence statistique (estimation et tests), la prévision, la simulation de variantes de politiques économiques).

Une mauvaise spécification de l'exogénéité pour les perturbations peut biaiser sensiblement l'estimation des paramètres d'une forme structurelle et conduire à des résultats erronés. Un exemple simple est celui des Moindres Carrés Ordinaires qui reposent sur l'hypothèse d'une covariance nulle entre les régresseurs et la perturbation de l'équation considérée. Par contre, lorsqu'on s'intéresse à l'exogénéité pour des paramètres d'intérêt, il s'agit de savoir s'il est possible de ne considérer que les équations de comportement décrivant l'évolution des variables « endogènes » (modèle conditionnel), sans entraîner de perte d'information par rapport à la prise en compte simultanée du modèle conditionnel et du processus de génération des variables « exogènes » (modèle marginal). Comme le soulignent Florens, Marimoutou et Peguin-Feisolle (2004), « *le concept d'exogénéité formalise l'idée que le mécanisme générateur des variables exogènes ne contient pas d'information pertinente sur les paramètres d'intérêt qui n'apparaissent que dans le modèle conditionnel aux exogènes* ». L'idée intuitive est de savoir si les mécanismes engendrant les variables « exogènes » et ceux engendrant les variables « endogènes » sont séparables pour l'objectif poursuivi par le modélisateur.

Une des principales raisons de la popularité des modèles conditionnels provient de la distinction opérée par la théorie économique entre variables « endogènes » et variables « exogènes », ces dernières ayant une évolution propre qui ne fait pas l'objet de la modélisation. Sous l'hypothèse d'exogénéité, le modèle conditionnel synthétise exactement la même information concernant les paramètres d'intérêt que le modèle complet, mais présente l'avantage d'être de dimension moindre. À la fois plus facilement gérable et moins coûteux en temps d'analyse, il offre également la possibilité au modélisateur d'introduire des variables exogènes supplémentaires sans affecter le nombre d'équations.

2.2 L'approche de Monfort et Rabemananjara

Dans l'introduction générale de cette revue, nous avons brièvement présenté les modèles structurels et les VAR en opposant ces deux types d'approches de modélisation. La première incorpore des hypothèses *a priori* découlant d'une réflexion économique sur le phénomène à modéliser. La seconde repose au contraire sur l'utilisation de formes réduites non-contraintes et préconise le recours systématique aux méthodes statistiques d'analyse des séries temporelles pour tester des hypothèses structurelles. Bien que ces deux approches puissent paraître pour le moment antinomiques, nous nous proposons d'expliciter à présent les relations existant entre ces deux types de modélisation. Ces relations ont été étudiées par Monfort et Rabemananjara (1990), qui ont proposé dans le cadre stationnaire, une méthode d'analyse des notions de causalité, d'exogénéité, de prédétermination.

Soit la représentation canonique suivante d'un modèle VAR :

$$\begin{cases} \Phi(L)X_t = \mu + \varepsilon_t, t = 1, ..., T \\ \text{avec } \Phi(L) = I_n - \sum_{i=1}^{p} \Phi_i L^i,\ \Phi(0) = I_n, \\ \text{et où } (\varepsilon_t) \sim iid\, N(0_n, \Sigma). \end{cases} \tag{2.1}$$

où les racines en L du polynôme $\Phi(L)$ sont à l'extérieur du disque unité (condition de stationnarité). La partition des variables du vecteur X_t en $\begin{pmatrix} Y_t \\ Z_t \end{pmatrix}$, $Y_t \in R^g$, $Z_t \in R^k$, avec $g + k = n$, conduit à réécrire sans perte de généralité l'équation (2.1) comme suit :

$$\begin{bmatrix} \Phi_{YY}(L) & \Phi_{YZ}(L) \\ \Phi_{ZY}(L) & \Phi_{ZZ}(L) \end{bmatrix} \begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \mu_Y \\ \mu_Z \end{bmatrix} + \begin{bmatrix} \varepsilon_{Yt} \\ \varepsilon_{Zt} \end{bmatrix} \tag{2.2}$$

avec

$$\begin{cases} \Phi_{YY}(L) = Id_g - \sum\limits_{i=1}^{p} \Phi_{YY,i} L^i \\ \Phi_{YZ}(L) = - \sum\limits_{i=1}^{p} \Phi_{YZ,i} L^i \\ \Phi_{ZY}(L) = - \sum\limits_{i=1}^{p} \Phi_{ZY,i} L^i \\ \Phi_{ZZ}(L) = Id_k - \sum\limits_{i=1}^{p} \Phi_{ZZ,i} L^i, \end{cases}$$

où l'innovation $\begin{pmatrix} \varepsilon_{Yt} \\ \varepsilon_{Zt} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix}\right]$.

Il faut remarquer que la normalisation $\Phi(0) = I_n$ est typique des modèles VAR et qu'elle les différencie des modèles structurels. Les modèles VAR apparaissent en effet comme des modélisations marginales en formes réduites, dans la mesure où les variables Z_t contemporaines n'apparaissent pas dans le premier bloc d'équations de l'écriture (2.2). En revanche, les modèles structurels qui s'intéressent à la modélisation des variables Y_t conditionnellement aux variables Z_t sont des modèles conditionnels avec variables exogènes.

Sous l'hypothèse de normalité des perturbations, la loi marginale de Z_t sachant son passé et celui de Y_t est directement donnée par le second bloc d'équations de l'écriture (2.2). La loi conditionnelle de Y_t sachant (Z_t, $\underline{X_{t-1}}$) est, elle, une loi normale dont l'espérance est la régression linéaire théorique de Y_t sur (Z_t, $\underline{X_{t-1}}$). Cette régression théorique est obtenue par combinaison linéaire des équations en multipliant l'équation (2.2) par la matrice $\begin{bmatrix} Id_g & -\Sigma_{YZ}\,\Sigma_{ZZ}^{-1} \\ 0 & Id_k \end{bmatrix}$, ce qui conduit à la forme dite « bloc-récursive » d'un modèle VAR, décomposition en modèle conditionnel et marginal :

$$\begin{cases} \underbrace{Y_t = \sum_{i=1}^{p} \Phi_{YY,i}^{+}\, Y_{t-i} + \sum_{i=0}^{p} \Phi_{YZ,i}^{+}\, Z_{t-i} + \mu_Y^{+} + \eta_{Y,t}}_{\text{modèle conditionnel}} \\ \underbrace{Z_t = \sum_{i=1}^{p} \Phi_{ZY,i}\, Y_{t-i} + \sum_{i=1}^{p} \Phi_{ZZ,i}\, Z_{t-i} + \mu_Z + \varepsilon_{Z,t}}_{\text{modèle marginal}} \end{cases} \tag{2.3}$$

$$\text{avec} \begin{cases} \sum_{i=1}^{p} \Phi_{YY,i}^{+} L^i = Id_g - \Phi_{YY}(L) + \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Phi_{ZY}(L) \\ \sum_{i=0}^{p} \Phi_{YZ,i}^{+} L^i = -\Phi_{YZ}(L) + \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Phi_{ZZ}(L) \\ \eta_{Yt} = \varepsilon_{Yt} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\varepsilon_{Zt} \\ \Sigma_{YY}^{+} = \Sigma_{YY} - \Sigma_{YZ}\Sigma_{ZZ}^{-1} \\ \varphi_Y^{+} = (\Phi_{YY,1}^{+}, ..., \Phi_{YY,p}^{+}; \Phi_{YY,0}^{+}, ..., \Phi_{YY,p}^{+}; \Sigma_{YY}^{+}) \\ \varphi_Z = (\Phi_{ZY,1}, ..., \Phi_{ZY,p}; \Phi_{ZZ,1}, ...\Phi_{ZZ,p-1}; \Sigma_{ZZ}) \end{cases}$$

$$\text{et} \begin{pmatrix} \eta_{Yt} \\ \varepsilon_{Zt} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{YY}^{+} & 0 \\ 0 & \Sigma_{ZZ} \end{pmatrix}\right]$$

La normalisation induite par cette transformation fait que le modèle conditionnel est un modèle VAR comportant g équations, qui engendrent Y_t conditionnellement à son passé et aux valeurs passées et présentes de Z_t. Il correspond à la forme réduite non contrainte associée à un modèle structurel. Par contre, le modèle marginal qui décrit le processus de génération des variables Z_t conditionnellement à leur passé et à celui de Y_t admet la même expression que dans l'écriture (2.2). C'est un modèle VAR pour les variables Z_t, augmenté du passé de Y_t. En effet, l'orthogonalisation des résidus n'affecte pas les équations décrivant l'évolution des variables « exogènes » Z_t.

Il faut noter que, par construction la variable Z_t est prédéterminée dans le modèle conditionnel. Cette variable est également faiblement exogène pour les paramètres φ_Y^{+} de ce modèle : les perturbations des modèles conditionnel et marginal sont en effet indépendantes du fait des propriétés de la loi normale et il n'existe pas de contrainte intra-blocs [10]. Il est donc toujours possible d'obtenir une estimation efficace des paramètres du modèle conditionnel sans tenir compte du modèle marginal. Le concept d'exogénéité faible de la variable

10. L'hypothèse d'exogénéité faible ne sera plus automatiquement satisfaite dans le cadre des modèles à variables co-intégrées, mais elle prendra au contraire la forme d'une restriction testable.

Z_t apparaît par conséquent comme une notion vide de sens en dehors d'une forme structurelle.

En revanche, pour réaliser des prévisions de Y_t à partir du seul modèle conditionnel, il faut s'assurer que l'hypothèse d'exogénéité forte de la variable Z_t est satisfaite [11]. L'hypothèse d'exogénéité faible n'est en effet pas suffisante dans ce cas, puisque le passé de Y_t contribue à prédire Z_t dans le modèle marginal. Il faut alors faire appel à l'hypothèse de non-causalité au sens de Granger de Y vers Z pour faire disparaître les valeurs retardées de Y ($\Phi_{ZY,i}$, i =1,... p) du modèle marginal. Cette hypothèse de non-causalité qui prend la forme de restrictions paramétriques sur le modèle marginal n'est pas automatiquement satisfaite par construction, mais constitue au contraire une restriction testable. Soulignons enfin qu'il est possible de montrer (*cf. supra*) que les hypothèses d'exogénéité forte et d'exogénéité stricte sont ici confondues.

Considérons à présent le modèle structurel écrit sous la forme :

$$\sum_{i=0}^{p}(\Gamma_{Yi}Y_{t-i} + \Gamma_{Z_i}Z_{t-i}) + \mu = u_t, t = 1, \ldots, T \tag{2.4}$$

où

- u_t est un bruit blanc g dimensionnel de matrice de variance-covariance Ω,
- Γ_{Yi}, Γ_{Zi}, i = 0,...,p sont respectivement des matrices de dimensions (g, g), (g, k), supposées constantes au cours du temps, avec Γ_{Y0} inversible et où μ est un vecteur de taille (g,1),
- les paramètres structurels $\Gamma = (\Gamma_{Y0}, ..., \Gamma_{Yp}; \Gamma_{Z0}, ..., \Gamma_{Zp}; \mu)$ satisfont q contraintes linéaires identifiantes R vec(Γ) = ω, où R est une matrice connue de taille (q, g (n (p + 1) + 1)) de rang q et ω est un vecteur connu de taille q [12].
- cov $(u_t, Y_{t-k}) = 0$, cov $(u_t, Z_{t-k}) = 0$ $\forall$ k $\geqslant$ 1.

Dans ce cas, étant donné que l'on dispose d'une forme structurelle satisfaisant des contraintes identifiantes, dans laquelle Y_t a été désigné comme endogène et Z_t comme exogène (*cf.* équation 2.4), il s'agit par la suite de savoir si la représentation VAR bloc-récursive est compatible avec cette forme structurelle. Monfort et Rabemananjara (1990) ont introduit deux degrés distincts de compatibilité : la compatibilité avec une forme structurelle faible et la compatibilité avec une forme structurelle forte. Plus précisément, le modèle VAR étant mis sous forme bloc récursive, il faut que le modèle

11. Il est à noter que la notion d'absence de causalité au sens de Sims est équivalente à celle de Granger, mais elle permet de bien faire comprendre l'utilité de la non-causalité pour, par exemple, simuler une trajectoire des variables endogènes, conditionnellement à des valeurs passées, présentes, et futures des exogènes (*i.e.* des scénarios), en utilisant de façon séquentielle les lois des endogènes, sachant leur propre passé et uniquement le présent et le passé des exogènes.

12. Rappelons que la forme structurelle ainsi définie est identifiable si et seulement si Rang R $(\Gamma' \otimes I_n) = g^2$. Par ailleurs, une condition suffisante est donnée par la condition d'ordre q $\geqslant g^2$.

conditionnel soit compatible avec les contraintes identifiantes et que les Z_t vérifient les propriétés de prédétermination ou d'exogénéité stricte. Ceci conduit aux deux théorèmes suivants :

THÉORÈME 1. — **Conditions de compatibilité d'un modèle VAR avec une forme structurelle faible (Monfort et Rabemananjara, 1990, théorème 4)**

Le modèle VAR orthogonalisé est compatible avec une forme structurelle faible si et seulement si, il existe des matrices $\Gamma = (\Gamma_{Y0}, ..., \Gamma_{Yp}, \Gamma_{Z0}, ..., \Gamma_{Zp}, \mu)$ *vérifiant :*

$$\Gamma_{Y0}\Phi^{+}_{YZ,i} + \Gamma_{Zi} = 0, i = 0, ..., p$$

$$\Gamma_{Y0}\Phi^{+}_{YY,i} + \Gamma_{Yi} = 0, i = 1, ..., p$$

$$\Gamma_{Y0}\ \mu^{+}_{Y} + \mu = 0$$

$$R\ vec(\Gamma) = \omega.$$

THÉORÈME 2. — **Conditions de compatibilité d'un modèle VAR avec une forme structurelle forte (Monfort et Rabemananjara, 1990, théorème 5)**

Pour que le modèle VAR orthogonalisé soit compatible avec une forme structurelle forte, il est nécessaire et suffisant qu'il soit compatible avec une forme structurelle faible et que $\Phi_{ZY}(L) = 0$, *c'est-à-dire que Y ne cause pas Z au sens de Granger.*

Monfort et Rabemananjara ont ainsi montré que les hypothèses de prédétermination de Z_t et d'exogénéité faible de Z_t sont équivalentes, lorsque les contraintes identifiantes sont satisfaites et qu'il en est de même des hypothèses d'exogénéité stricte de Z_t et d'exogénéité forte de Z_t [13].

3. Exogénéité et non-causalité dans les modèles VAR-ECM

Il s'agit, dans cette section, d'examiner formellement ce que deviennent les équivalences précédentes, dans le cadre des modèles à variables non stationnaires co-intégrées « à la Johansen ». Nous considérons dans un premier temps l'hypothèse d'exogénéité faible, énonçons les conditions suffisantes proposées par Johansen (1992b) [14], ainsi que celles formulées par Hendry et Mizon (1993). Nous nous intéressons dans un second temps aux concepts de non-causalité au sens de Granger et d'exogénéité forte et exposons les

13. Comme souligné par l'un des rapporteurs, le théorème 1 est en réalité une condition d'exogénéité faible et le théorème 2 d'exogénéité forte.

14. *cf.* également Urbain (1992), ainsi que Boswijk (1992).

conditions fournies par Toda et Phillips (1991) et Johansen (1992b). À chaque fois, un intérêt particulier est accordé à la mise en évidence des principales différences avec l'exogénéité dans le cadre stationnaire et à la présentation des principales limites des conditions d'exogénéité de la littérature dans les modèles VAR-ECM. Avant d'aborder plus en détail ces différents concepts, nous rappelons dans une première sous-section les différentes écritures d'un modèle VAR-ECM.

3.1. Les différentes écritures d'un modèle VAR-ECM

Soit un processus vectoriel $\{X_t\}$, de dimension n, de moyenne zéro (donc, sans composante déterministe [15]), d'ordre p fini, admettant la représentation VAR-ECM suivante :

$$\Delta X_t = \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \alpha\beta' X_{t-1} + \varepsilon_t, t = 1, \ldots, T, \tag{2.5}$$

où Γ_i, α, β désignent respectivement des matrices de dimensions (n, n), (n, r), (n, r), $0 < r < n$, telles que $\Pi = \alpha\beta'$; et où ε_t est un vecteur de perturbations identiquement et indépendamment distribuées suivant une loi normale de moyenne zéro et de matrice de variance-covariance Σ.

On suppose également que l'équation (2.5) satisfait (i) $\left|(I_n - \sum\limits_{i=1}^{p-1} \Gamma_i z^i)(1-z) + \alpha\beta' z\right| = 0$ implique soit $|z| > 1$ soit $z = 1$, et que (ii) la matrice $\alpha'_\perp(I_n - \sum\limits_{i=1}^{p} \Gamma_i)\beta_\perp$ est inversible, où $\beta_\perp$ et $\alpha_\perp$ sont deux matrices de plein rang $n \times n - r$ telles que $\alpha'_\perp \alpha = \beta'_\perp \beta = 0$, ce qui exclut la possibilité que un ou plusieurs éléments de X_t soit intégrés d'ordre 2. Ces deux hypothèses assurent que $\{X_t\}$ et $\{\beta' X_t\}$ sont respectivement I (1) et I (0) et que le théorème de représentation de Engle et Granger (1987) est vérifié.

La partition des variables du vecteur X_t en $Y_t \in R^g, Z_t \in R^k$, avec g + k = n, entraîne, avec des notations évidentes, les partitions des matrices Γ_i, α et β respectivement en

$$\begin{bmatrix} \Gamma_{YY,i} & \Gamma_{YZ,i} \\ \Gamma_{ZY,i} & \Gamma_{ZZ,i} \end{bmatrix}, \begin{bmatrix} \alpha_Y \\ \alpha_Z \end{bmatrix}, \begin{bmatrix} \beta_Y \\ \beta_Z \end{bmatrix}.$$

Elle conduit en outre aux deux écritures suivantes de l'équation (2.5) [16] :

15. Cette hypothèse simplificatrice n'entraîne aucune perte de généralité et a pour seul objet d'alléger les notations et faciliter ainsi la présentation.

16. voir à ce sujet Urbain (1992), ou Rault (1997).

Écriture 1 : canonique

$$\left\{\begin{array}{ll} \Delta Y_t = \sum_{i=1}^{p-1} \Gamma_{YY,i} \Delta Y_{t-i} & + \sum_{i=1}^{p-1} \Gamma_{YZ,i} \Delta Z_{t-i} \\ & + \alpha_Y \, [\beta'_Y \;\|\; \beta'_Z] \begin{bmatrix} Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \varepsilon_{Y,t} \\ \Delta Z_t = \sum_{i=1}^{p-1} \Gamma_{ZY,i,} \Delta Y_{t-i} & + \sum_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} \\ & + \alpha_Z \, [\beta'_Y \;\|\; \beta'_Z] \begin{bmatrix} Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \varepsilon_{Z,t} \end{array}\right. \tag{2.6}$$

avec $\begin{pmatrix} \varepsilon_{Yt} \\ \varepsilon_{Zt} \end{pmatrix} \sim \mathrm{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix}\right]$.

Écriture 2 : bloc-récursive

$$\left\{\begin{array}{ll} \Delta Y_t = \sum_{i=1}^{p-1} \Gamma^+_{YY,i} \Delta Y_{t-i} & + \sum_{i=1}^{p-1} \Gamma^+_{YZ,i} \Delta Z_{t-i} \\ & \underbrace{+ \alpha^+_Y \, [\beta'_Y \;\|\; \beta'_Z] \begin{bmatrix} Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \eta_{Y,t}}_{\text{modèle conditionnel}} \\ \Delta Z_t = \sum_{i=1}^{p-1} \Gamma_{ZY,i,} \Delta Y_{t-i} & + \sum_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} \\ & \underbrace{+ \alpha_Z \, [\beta'_Y \;\|\; \beta'_Z] \begin{bmatrix} Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \varepsilon_{Z,t}}_{\text{modèle marginal}} \end{array}\right. \tag{2.7}$$

$$\text{avec} \left\{\begin{array}{l} \Gamma^+_{YY}(L) = \Gamma_{YY}(L) - \Sigma_{YZ}\Sigma^{-1}_{ZZ}\Gamma_{ZY}(L) = Id_g - \sum_{i=1}^{p-1} \Gamma^+_{YY,i} L^i \\ \Gamma^+_{YZ}(L) = \Gamma_{YZ}(L) - \Sigma_{YZ}\Sigma^{-1}_{ZZ}\Gamma_{ZZ}(L) = - \sum_{i=0}^{p-1} \Gamma^+_{YZ,i} L^i \\ \alpha^+_Y = \alpha_Y - \Sigma_{YZ}\Sigma^{-1}_{ZZ} \, \alpha_Z \\ \eta_{Yt} = \varepsilon_{Yt} - \Sigma_{YZ}\Sigma^{-1}_{ZZ} \, \varepsilon_{Zt} \\ \Sigma^+_{YY} = \Sigma_{YY} - \Sigma_{YZ}\Sigma^{-1}_{ZZ} \end{array}\right.$$

et $\begin{pmatrix} \eta_{Yt} \\ \varepsilon_{Zt} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma^+_{YY} & 0 \\ 0 & \Sigma_{ZZ} \end{pmatrix}\right]$.

Il convient de noter que, bien que l'écriture 1 soit qualifiée de canonique (*cf.* Johansen, 1995b), il reste en réalité un degré de liberté pour définir les matrices α et β puisque $\Pi = \alpha\ \beta' = (\alpha\ \Psi\ ^{-1})\ (\Psi\ \beta')$ pour toute matrice Ψ non singulière, de dimensions (r, r).

3.2 Conditions suffisantes usuelles d'exogénéité et de non-causalité dans un modèle VAR-ECM

Comme dans la représentation bloc-récursive d'un modèle VAR [17] Z_t intervient à la date t dans le modèle régissant Y_t, c'est-à-dire dans le premier bloc d'équations du modèle VAR-ECM orthogonalisé (*cf.* équation 2.7). De ce fait, par construction et non en tant que propriété intrinsèque, la variable Z_t est, comme dans le cas du modèle VAR othogonalisé, automatiquement prédéterminée dans le modèle conditionnel.

Par contre, l'orthogonalisation des perturbations du modèle VAR-ECM n'assure plus la faible exogénéité de la variable Z_t pour les paramètres du modèle conditionnel. En effet, les paramètres des modèles conditionnel et marginal ne varient pas indépendamment car il existe des contraintes inter-équations, le paramètre $\beta' = [\beta'_Y\ \beta'_Z]$ figurant simultanément dans ces deux modèles. Ceci constitue une différence majeure avec le modèle VAR orthogonalisé, dans lequel la variable Z_t était faiblement exogène par construction pour les paramètres du modèle conditionnel. Une source de biais importante serait d'estimer les vecteurs de co-intégration à partir du seul modèle conditionnel. Nous avons donc un problème intéressant où une variable Z_t est prédéterminée, mais pas faiblement exogène pour les paramètres du modèle conditionnel. Toutefois, la faible exogénéité pourra être obtenue en imposant une condition supplémentaire. Johansen (1992b, théorème 1) a formulé un résultat simple, indiquant une condition suffisante permettant une analyse valide à partir du seul modèle conditionnel.

Proposition 1. — **Condition suffisante d'exogénéité faible pour tous les (α, β) (Johansen, 1992b)**

Dans le cadre du modèle VAR-ECM orthogonalisé (équation 2.7), pour que la variable Z_t soit faiblement exogène pour les paramètres de long terme (α_Y^+, β) du modèle conditionnel, il suffit que $\alpha_Z = 0$, c'est-à-dire que les vecteurs de co-intégration ne figurent pas dans le modèle marginal.

Ainsi, avant d'estimer les paramètres de long terme du modèle conditionnel $\{\Delta Y_t / \Delta Z_t\}$ uniquement à partir de ce dernier, il suffit de s'assurer au préalable que la variable Z_t est faiblement exogène pour ces paramètres, c'est-à-dire que les processus générant respectivement $\{\Delta Y_t / \Delta Z_t\}$ et $\{\Delta Z_t\}$ sont séparables du point de vue de l'inférence. Cette condition est satisfaite ici pour $\alpha_Z = 0$ et assure que les paramètres des modèles conditionnel et marginal sont libres en variation. Dès lors, l'estimateur du maximum de vraisemblance des paramètres de long terme peut être obtenu en maximisant la vraisemblance

17. *cf.* section 2.

du modèle conditionnel, en ignorant celle du modèle marginal. Autrement dit, on se place dans le modèle conditionnel de $\Delta \mathrm{Y}_t$ à $\Delta \mathrm{Z}_t$ fixé.

Remarque. — Si dans une étude empirique, les paramètres d'intérêt du modélisateur sont donnés par l'ensemble des paramètres du modèle conditionnel, alors $\alpha_Z = 0$ reste une condition suffisante assurant l'exogénéité faible de la variable Z_t. Par contre, dans le cas où les paramètres d'intérêt sont l'ensemble des paramètres des g premières équations du modèle VAR-ECM écrit sous forme réduite (équation 2.6) [18], l'exogénéité faible sera obtenue si la forme réduite est déjà orthogonalisée, c'est-à-dire si $\Sigma_{YZ} = 0$.

La condition d'exogénéité formulée dans la proposition 1 est très facilement testable, puisqu'elle n'implique que des contraintes linéaires sur les paramètres de la matrice α. Elle apparaît néanmoins comme assez contraignante et cela pour deux raisons.

D'une part, d'un point de vue théorique, elle interdit l'existence de relations de long terme propres aux équations décrivant l'évolution des variables exogènes. Ceci contraint le modèle marginal à être un modèle VAR en différences, alors qu'on aurait pu au contraire envisager la possibilité que ce modèle comporte des relations de long terme ne faisant intervenir par exemple que des variables exogènes, avec des paramètres qui ne feraient pas partie du modèle conditionnel (forme structurelle).

D'autre part, elle suppose que le modélisateur a un intérêt pour l'ensemble des relations de co-intégration existant entre les variables considérées. Pourtant, il faut souligner que très souvent dans les études empiriques, celui-ci n'est en réalité intéressé que par un sous ensemble de ces relations et plus particulièrement par celles figurant dans le modèle conditionnel et beaucoup plus rarement par les relations apparaissant dans le modèle marginal. Déjà en 1995, Ericsson avait remarqué que c'était une hypothèse « *trop restrictive* », puisque selon lui, « *toute étude empirique doit raisonnablement se limiter à un sous ensemble de vecteurs de co-intégration de l'économie* ». En outre, une difficulté typique apparaît au niveau empirique lorsque les tests de co-intégration révèlent l'existence de r relations d'équilibre, alors que, selon la théorie économique, il ne devrait en exister que m, avec $m < r$.

Une condition plus souple a été étudiée par Hendry et Mizon (1993) [19] qui ont proposé le résultat ci-dessous.

Considérons la partition des r vecteurs de co-intégration β en deux sous-ensembles $[\beta_1 \ \beta_2]$, avec $r_1 + r_2 = r$, $r_1 > 0$, $r_2 > 0$, qui conduit au découpage de la matrice α en $\begin{bmatrix} \alpha^+_{Y1} & \alpha^+_{Y2} \\ \alpha_{Z1} & \alpha_{Z2} \end{bmatrix}$, où α^+_{Y1}, α^+_{Y2}, α_{Z1}, α_{Z2} sont respectivement des matrices de dimensions $g \times r_1$, $g \times r_2$, $k \times r_1$, $k \times r_2$ et supposons que $\alpha^+_{Y1} \neq 0$.

18. Remarquons que ces paramètres ne présentent en réalité un intérêt pour l'économiste que dans le cas particulier de séparabilité (*cf.* Granger et Haldrup, 1997).

19. voir également Ericsson *et al.*(1998).

Dans ce cas, le modèle VAR-ECM orthogonalisé (équation 2.7) peut être réécrit sous la forme :

$$\left\{\begin{array}{l} \Delta Y_t = \sum_{i=1}^{p-1} \Gamma^+_{YY,i} \Delta Y_{t-i} + \sum_{i=0}^{p-1} \Gamma^+_{YZ,i} \Delta Z_{t-i} \\ \qquad\qquad \underbrace{+\alpha^+_{Y1}\beta'_1 X_{t-1} + \alpha^+_{Y2}\beta'_2 X_{t-1} + \eta_{Y,t}}_{\text{modèle conditionnel}} \\ \\ \Delta Z_t = \sum_{i=1}^{p-1} \Gamma_{ZY,i,} \Delta Y_{t-i} + \sum_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} \\ \qquad\qquad \underbrace{+\alpha_{Z1}\beta'_1 X_{t-1} + \alpha_{Z2}\beta'_2 X_{t-1} + \varepsilon_{Z,t}}_{\text{modèle marginal}} \end{array}\right. \tag{2.7a}$$

et l'on peut énoncer la proposition suivante :

Proposition 2. — **Condition suffisante d'exogénéité faible pour β_1(Hendry et Mizon, 1993)**

Dans le cadre du modèle VAR-ECM orthogonalisé (équation 2.7a), pour que la variable Z_t soit faiblement exogène pour les vecteurs de co-intégration $\beta_1 = \begin{bmatrix} \beta_{Y1} \\ \beta_{Z1} \end{bmatrix}$, il suffit que $\left\{\begin{array}{l} \alpha_{Z1} = 0 \\ \alpha^+_{Y2} = 0 \end{array}\right.$.

Cette condition d'exogénéité est moins restrictive que celle donnée par Johansen (1992b). Elle autorise en particulier l'existence de relations de co-intégration dans le modèle marginal. Néanmoins, elle peut être critiquée à au moins deux titres. D'une part, la partition des r vecteurs de co-intégration en r_1 etr_2 est donnée *a priori*, de manière à ce que β_1 et β_2 figurent respectivement dans les modèles conditionnel et marginal. D'autre part, seuls les paramètres de long terme du modèle conditionnel sont considérés comme possibles paramètres d'intérêt pour le modélisateur. Pourtant, dans certaines études appliquées, celui-ci peut également s'intéresser aux paramètres de court terme du modèle conditionnel et souhaiter modéliser par exemple la dynamique d'ajustement vers l'équilibre de long terme. Pour pallier ces critiques, Rault (2005) introduit une décomposition canonique de la matrice Π permettant de déterminer explicitement cette partition et proposent une condition nécessaire et suffisante d'exogénéité faible (*cf.* section 4).

Si l'on souhaite étudier les relations causales entre les variables, on peut énoncer la proposition suivante :

Proposition 3. — **Condition nécessaire et suffisante de non-causalité (Toda et Phillips, 1991)** [20]

Dans le cadre du modèle VAR-ECM orthogonalisé (équation 2.7), Y ne cause pas Z au sens de Granger (1969) si et seulement si :

$$\{\Gamma_{ZY,1}\,, = \Gamma_{ZY,2} = ... = \Gamma_{ZY,p-1} = \alpha_Z \beta'_Y = 0\}$$

$$\Longleftrightarrow \{\Gamma_{ZY}(L) = 0 \;\; et \;\; \alpha_Z \beta'_Y = 0\}\,.$$

20. voir également Yum (1994), Lemme 5.1, ainsi que Bruneau (1996).

Les définitions suivantes sont d'usage courant :

DEFINITION 1. — *Dans le cadre de l'équation (2.7a),Y ne cause pas Z à court terme si et seulement si* $\Gamma_{ZY}(L) = 0$.

DEFINITION 2. — *Dans le cadre de l'équation (2.7a), Y ne cause pas Z à long terme si et seulement si* $\alpha_Z \beta'_Y = 0$.

L'hypothèse de non-causalité de Y vers Z correspond donc toujours, dans le cadre d'un modèle VAR-ECM, à des restrictions paramétriques sur le modèle marginal, indépendantes des paramètres d'intérêt du modélisateur. La condition $\{\alpha_Z \beta'_Y = 0\}$ implique cependant des contraintes non-linéaires sur les paramètres de long terme du modèle, ce qui complique considérablement la procédure pour tester l'hypothèse de non-causalité par rapport à un modèle VAR. C'est la raison pour laquelle d'autres conditions paramétriques ont été proposées par Giannini et Mosconi (1992). Néanmoins, dans leurs théorèmes, ils ne distinguent pas clairement la nullité de certains blocs de paramètres à laquelle il est toujours possible de se ramener sans perte de généralité, de la nullité des blocs de paramètres résultant de la propriété de non causalité. Ces observations ont conduit Rault (2000) à rechercher une condition paramétrique plus précise et plus simple à tester que celle de Toda et Phillips (*cf.* section 4).

En revanche, si l'on cherche à tester l'hypothèse d'exogénéité forte de la variable Z_t cette expression non linéaire se simplifie et on peut énoncer la proposition suivante :

PROPOSITION 4. — **Condition suffisante d'exogénéité forte pour tous les** (α, β) **(Johansen, 1992b)**

Dans le cadre du modèle VAR-ECM orthogonalisé (équation 2.7), pour que la variable Z_t *soit fortement exogène pour les paramètres de long terme du modèle conditionnel, il suffit que* $\begin{cases} \alpha_Z = 0 \\ \Gamma_{ZY,i} = 0, \ i = 1, ..., p-1 \end{cases}$.

Les hypothèses d'exogénéité faible de la variable Z_t et de non-causalité de Y vers Z apparaissent ainsi beaucoup plus simples à tester simultanément que l'hypothèse de non-causalité de Y vers Z seule. En effet, l'hypothèse d'exogénéité faible, qui implique la restriction $\{\alpha_Z = 0\}$, conduit à une simplification de la condition assurant la non-causalité de long terme. L'hypothèse d'exogénéité forte peut donc être testée avec une statistique admettant une distribution limite du chi-deux. Il est à noter que les hypothèses d'exogénéité faible de la variable Z_t et de non causalité de Y vers Z ne sont plus ici séparées comme c'était le cas dans un modèle VAR.

4. Conditions nécessaires et suffisantes d'exogénéité et de non-causalité dans les modèles VAR-ECM

Les conditions d'exogénéité faibles étudiées jusqu'à présent sont suffisantes, la moins contraignante étant celle de Hendry et Mizon (1993). Cette dernière impose néanmoins une partition structurelle des r vecteurs de co-intégration en β_1 et β_2. On a certes parfois un *a priori* de la partition des variables du vecteur X_t en variables endogènes Y_t et exogènes Z_t, mais il est également important d'avoir dans un premier temps une condition intrinsèque d'exogénéité faible. C'est l'objet de la première sous-section, dans laquelle nous présentons une condition nécessaire et suffisante d'exogénéité faible dans le cadre d'une décomposition canonique de la matrice Π (*cf.* Rault, 2005). Cette décomposition permet de mettre en évidence le nombre minimum de relations de co-intégration figurant dans les équations des variables Y_t.

Par ailleurs, un autre point concerne la non-causalité dans les modèles VAR-ECM. En effet, dès que l'on examine cette notion, on s'intéresse à la présence ou non des variables endogènes dans les relations de long terme. On a ainsi besoin de déterminer le nombre minimum de relations de long terme qui contiennent obligatoirement des variables endogènes. Cela nécessite par conséquent une décomposition spécifique de la matrice Π pour faire apparaître dans l'espace de co-intégration les relations de co-intégration qui doivent contenir des variables endogènes. Ceci constitue l'objectif de la seconde sous-section, qui donne également une condition nécessaire et suffisante de non-causalité plus simple à tester que celle de Toda et Phillips (1991). Nous verrons en particulier que, dès lors que l'hypothèse de non-causalité intervient, les relations de long terme figurant dans le modèle marginal ne contiennent que des variables exogènes (*cf.* Rault, 2000).

Enfin, lorsqu'on étudie l'exogénéité forte, on considère simultanément les hypothèses d'exogénéité faible et de non-causalité. Il s'agit donc de s'intéresser à la fois aux relations de co-intégration qui doivent figurer obligatoirement dans le modèle conditionnel et à celles ne contenant pas de variables endogènes qui doivent apparaître nécessairement dans le modèle marginal. La troisième sous-section s'inscrit dans cette optique. Elle développe en effet un cadre d'analyse s'appuyant sur une troisième décomposition canonique de la matrice de long terme (*cf.* Pradel et Rault, 2003). Elle fournit aussi une condition nécessaire et suffisante d'exogénéité forte très simple à mettre en œuvre dans les travaux appliqués.

Soulignons dès maintenant que les conditions nécessaires et suffisantes de non-causalité et d'exogénéité proposées s'expriment comme des conditions minimales sur les paramètres d'une forme canonique, dans laquelle la nullité de certains blocs de paramètres n'entraîne aucune perte de généralité. En outre, les trois décompositions canoniques suggérées exploitent le fait que les matrices α et β ne sont définies qu'à une transformation linéaire près de matrice Ψ non singulière, de dimensions (r, r).

4.1. L'exogénéité faible

La présence ou l'absence d'exogénéité faible dépend de manière cruciale des paramètres d'intérêt du modélisateur, mais contrairement à ce qu'il est souvent supposé dans le cadre des modèles à variables co-intégrées, ces paramètres d'intérêt ne se composent pas nécessairement de l'ensemble des vecteurs de co-intégration. En effet, lorsqu'on travaille avec un modèle VAR-ECM ou un modèle conditionnel, les paramètres d'intérêt peuvent être, par exemple, tous les paramètres du modèle conditionnel, ou simplement les vecteurs de co-intégration faisant partie de ce modèle. Pour tenir compte de cette possibilité, nous proposons le résultat ci-dessous qui fournit une décomposition canonique envisageable de la matrice Π. Cette décomposition constitue un cadre propice dans lequel une condition nécessaire et suffisante d'exogénéité faible moins restrictive peut ensuite être formulée.

THÉORÈME 3. — **(Rault, 2005)**
Soit $\Pi = \alpha\beta'$ *une matrice de dimensions* (n, n)*, singulière, de rang* $r (0 < r < n)$ *et partitionnons* α *en* $\begin{bmatrix} \alpha_Y \\ \alpha_Z \end{bmatrix}$.

(i) Si on définit $m_1 =$ *rang* (α_Y)*, avec* $m_1 > 0$ *et* $r - m_1 > 0$[21]*, alors il est toujours possible de reparamétrer les matrices* α *et* β *de la façon suivante :*

$$\beta = [\beta_1\ \beta_2] = \begin{bmatrix} \beta_{Y1} & \beta_{Y2} \\ \beta_{Z1} & \beta_{Z2} \end{bmatrix}$$

$$\alpha = [\alpha_1\ \alpha_2] = \begin{bmatrix} \alpha_{Y1} & 0_{(g, r-m_1)} \\ \alpha_{Z1} & \alpha_{Z2} \end{bmatrix},$$

où β_{Y1}, α_{Y1}, β_{Z1}, α_{Z1}, β_{Y2}, β_{Z2}, α_{Z2} *sont des sous-matrices respectivement de dimensions* $g \times m_1$, $g \times m_1$, $k \times m_1$, $k \times m_1$, $g \times r - m_1$, $k \times r - m_1$, $k \times r - m_1$ *avec rang* $(\alpha_{Y1}) = m_1$ *et rang* $(\alpha_{Z2}) = r - m_1$. [22].

(ii) m_1 *est défini de manière unique et est invariant par rapport à la reparamétrisation choisie. Il vérifie max* $(0, r - k) \leqslant m_1 \leqslant$ *min* (g, r)[23].

En tenant compte de ces nouvelles écritures, l'équation (2.7) devient :

$$\begin{cases} \text{modèle conditionnel} \\ \Delta Y_t = \sum\limits_{i=1}^{p-1} \Gamma^+_{YY,i} \Delta Y_{t-i} + \sum\limits_{i=0}^{p-1} \Gamma^+_{YZ,i} \Delta Z_{t-i} + \alpha^+_{Y1} \beta'_1 X_{t-1} + \eta_{Y,t} \\ \text{modèle marginal} \\ \Delta Z_t = \sum\limits_{i=1}^{p-1} \Gamma_{ZY,i} \Delta Y_{t-i} + \sum\limits_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} + \alpha_{Z1} \beta'_1 X_{t-1} \\ \qquad + \alpha_{Z2} \beta'_2 X_{t-1} + \varepsilon_{Z,t} \end{cases} \tag{2.7b}$$

21. Nous supposons que β_1 et β_2 contiennent chacun au moins un vecteur de co-intégration afin d'exclure le cas où $\beta_1 = \beta$.

22. Rappelons que $Y_t \in R^g$ et $Z_t \in R^k$.

23. Cette condition assure que la matrice α est de rang r.

Le théorème 3 fournit une indication sur les m_1 relations de long terme faisant nécessairement partie du modèle conditionnel et sur les (r-m_1) n'en faisant pas obligatoirement partie. Cette écriture canonique est obtenue par un changement de base dans l'espace d'ajustement. À cette partition correspond un nouveau β découpé en $[\beta_1\ \beta_2]$. Ce théorème n'entraîne aucune perte de généralité et nécessite uniquement la détermination du rang m_1.

L'écriture canonique donnée ci-dessus est très commode pour tester des restrictions paramétriques découlant de l'hypothèse d'exogénéité faible. Ceci nous conduit à introduire le résultat suivant :

Proposition 5. — **Condition nécessaire et suffisante d'exogénéité faible pour les paramètres du modèle conditionnel (Rault, 2005)**
Dans le cadre du modèle VAR-ECM bloc-récursif (équation 2.7b), si $m_1 > 0$, Z_t est faiblement exogène pour l'ensemble des paramètres du modèle conditionnel, i.e.$\Psi = (\Gamma^+_{YY,i}, i = 1, ..., p-1\,; \Gamma^+_{YZ,i}, i = 0, \ldots, p-1\,; \alpha^+_{Y1}\,;\beta_1')$, si et seulement si $\alpha_{Z1} = 0$ dans l'écriture canonique donnée par le théorème 3.

Remarque. — Si $\mathrm{m}_1 = 0$, la question ne se pose pas.

À la différence de la condition d'exogénéité faible donnée par Hendry et Mizon (1993) (*cf.* section 3.2, proposition 2), où la partition de β en $[\beta_1\ \beta_2]$ est donnée *a priori* de manière à ce que β_1 et β_2 figurent respectivement dans les modèles conditionnel et marginal, Rault (2005) détermine ici explicitement cette partition en exploitant le fait que les matrices α et β ne sont pas uniques (*cf. infra*). Cette condition demeure la même si on ne s'intéresse qu'aux paramètres de long terme du modèle conditionnel.

4.2. La non-causalité

Nous avons vu dans la section 3.2 qu'il pouvait s'avérer complexe de tester l'hypothèse de non-causalité de Y vers Z indépendamment de l'hypothèse d'exogénéité faible de la variable Z_t. En effet, la condition usuelle de non-causalité de long terme de Y vers Z dans les modèles VAR-ECM $\{\alpha_Z\beta_Y' = 0\}$ implique des contraintes non linéaires sur les paramètres de long terme. De ce fait, les statistiques usuelles de tests de non-causalité nécessitent en général de vérifier des conditions particulières pour suivre asymptotiquement une loi du chi-deux (*cf.* par exemple Toda et Phillips, 1991, théorème 2). Afin de remédier à cette difficulté Rault (2000) a développé un cadre d'analyse permettant de formuler une condition nécessaire et suffisante de non-causalité pouvant être facilement testée avec des statistiques admettant, dans tous les cas de figure, une distribution limite du chi-deux. Dans cette optique, on peut formuler le théorème suivant :

Théorème 4. — **(Rault, 2000)**
Soit $\Pi = \alpha\beta'$ une matrice de dimensions (n, n), singulière, de rang $r (0 < r < n)$ et partitionnons β en $\begin{bmatrix} \beta_Y \\ \beta_Z \end{bmatrix}$.

(i) *Si on définit* r_1= *rang* (β_Y), *avec* $r_1 > 0$ *et* $r - r_1 > 0$ [24], *il est toujours possible de reparamétrer les matrices* α *et* β *de la façon suivante :*

$$\beta = [\beta_1 \ \beta_2] = \begin{bmatrix} \beta_{YY} & 0 \\ \beta_{ZY} & \beta_{ZZ} \end{bmatrix}$$

$$\alpha = [\alpha_1 \ \alpha_2] = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ} \\ \alpha_{ZY} & \alpha_{ZZ} \end{bmatrix},$$

où β_{YY}, α_{YY}, β_{ZY}, α_{ZY}, α_{YZ}, β_{ZZ}, α_{ZZ} *sont des sous-matrices respectivement de dimensions* $g \times r_1$, $g \times r_1$, $k \times r_1$, $k \times r_1$, $g \times r\text{-}r_1$, $k \times r\text{-}r_1$, $k \times r\text{-}r_1$, *avec rang* $(\beta_{YY}) = r_1$ *et rang* $(\beta_{ZZ}) = r - r_1$.

(ii) r_1 *est défini de manière unique et est invariant par rapport à la reparamétrisation choisie. Il vérifie max* $(0, r-k) \leqslant r_1 \leqslant$ *min* (g, r) [25].

En tenant compte de ces nouvelles écritures, l'équation (2.7) devient :

$$\begin{cases} \text{modèle conditionnel} \\ \Delta Y_t = \sum\limits_{i=1}^{p-1} \Gamma^+_{YY,i} \Delta Y_{t-i} + \sum\limits_{i=0}^{p-1} \Gamma^+_{YZ,i} \Delta Z_{t-i} + \alpha^+_{YY} \beta'_1 X_{t-1} \\ \qquad\qquad + \alpha^+_{YZ} \beta'_{ZZ} Z_{t-1} + \eta_{Y,t} \\ \text{modèle marginal} \\ \Delta Z_t = \sum\limits_{i=1}^{p-1} \Gamma_{ZY,i} \Delta Y_{t-i} + \sum\limits_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} + \alpha_{ZY} \beta'_1 X_{t-1} \\ \qquad\qquad + \alpha_{ZZ} \beta'_{ZZ} Z_{t-1} + \varepsilon_{Z,t} \end{cases} \tag{2.7c}$$

Le théorème 4 fournit une indication sur le nombre minimum de relations de long terme contenant les g premières variables Y. Il est obtenu par un changement de base dans l'espace co-intégrant. Ce changement de base permet de scinder l'ensemble des relations de co-intégration en deux sous-ensembles : un premier de dimension r_1, mélangeant variables endogènes et exogènes et un second de dimension $(r - r_1)$ ne comportant que des variables exogènes. Par la suite, nous désignerons par « relation de long terme purement exogène » un vecteur de co-intégration ne contenant que des variables exogènes. A cette partition correspond une nouvelle matrice α découpée en $\alpha = [\alpha_1 \ \alpha_2]$. À l'instar du théorème 3, ce théorème n'entraîne aucune perte de généralité et requiert uniquement la détermination du rang r_1 de la sous-matrice β_{YY} [26].

Étant donnée cette décomposition canonique, on peut énoncer la proposition ci-dessous :

24. Nous supposons que β_1 et β_2 contiennent chacun au moins un vecteur de cointégration afin d'exclure le cas où β_1=β.

25. Cette condition assure que la matrice β est de rang r.

26. Nous présentons dans la section 5 une procédure séquentielle permettant de déterminer ce rang r_1.

PROPOSITION 6. — **Condition nécessaire et suffisante de non-causalité au sens de Granger (Rault, 2000)**
Dans le cadre du modèle VAR-ECM bloc-récursif (équation 2.7c), Y ne cause pas Z au sens de Granger (1969), si et seulement si $\begin{cases} \Gamma_{ZY,i} = 0, \ i = 1, ..., p-1 \\ \alpha_{ZY} = 0 \end{cases}$
dans l'écriture canonique donnée par le théorème 4.

La condition de non-causalité donnée ci-dessus est très simple à utiliser en pratique, puisque la non-causalité de long terme est équivalente à $\alpha_{ZY} = 0$ dans le cadre du théorème 4. Elle peut par conséquent être testée avec des statistiques admettant une loi limite du chi-deux. Ce n'est pas en effet une condition non linéaire comme celle de Toda et Phillips (1991). Elle n'implique au contraire que des contraintes linéaires sur les paramètres de la matrice α. Il est à noter que cette condition demeure inchangée si on considère les paramètres des g premières équations du modèle VAR-ECM écrit sous forme canonique.

4.3. Exogénéité forte

Nous nous intéressons à présent à l'hypothèse d'exogénéité forte et exposons une troisième décomposition de la matrice de co-intégration Π introduite par Rault et Pradel (2003). Cette décomposition constitue un cadre approprié dans lequel peut être formulée une condition nécessaire et suffisante d'exogénéité forte. Cette condition autorise en particulier les modèles conditionnel et marginal à contenir des relations de long terme purement exogènes. Pour ce faire, il nous faut au préalable formuler le résultat suivant :

THÉORÈME 5. — **(Pradel et Rault, 2003)**
Soit $\Pi = \alpha\beta'$ *une matrice de dimensions (n,n), singulière, de rang* $r(0 < r < n)$ *et considérons la reparamétrisation*

$$\beta = \begin{bmatrix} \beta_{YY} & 0 \\ \beta_{ZY} & \beta_{ZZ} \end{bmatrix}, \alpha = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ} \\ \alpha_{ZY} & \alpha_{ZZ} \end{bmatrix}$$

donnée dans le théorème 4. Dans ce cas :

(i) il existe un entier r_2 *tel que les matrices* α *et* β *puissent toujours être reparamétrées de la façon suivante :*

$$\alpha = [\alpha_1 \ \alpha_{21} \ \alpha_{22}] = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ_1} & \alpha_{YZ_2} \\ \alpha_{ZY} & 0 & \alpha_{ZZ_2} \end{bmatrix}$$

$$\beta = [\beta_1 \ \beta_{21} \ \beta_{22}] = \begin{bmatrix} \beta_{YY} & 0 & 0 \\ \beta_{ZY} & \beta_{ZZ_1} & \beta_{ZZ_2} \end{bmatrix},$$

où α_{YY}, β_{YY}, α_{ZY}, β_{ZY}, α_{YZ_1}, β_{ZZ_1}, α_{YZ_2}, α_{ZZ_2} , β_{ZZ_2} *sont des sous-matrices respectivement de dimensions* $g \times r_1$, $g \times r_1$, $k \times r_1$, $k \times r_1$, $g \times r^*$, $k \times r^*$, $g \times r_2$, $k \times r_2$, $k \times r_2$, *avec* $r_1 + r_2 + r^* = r$ *et rang* $(\alpha_{ZZ_2}) = r_2 \geqslant 0$.

(ii) si en outre $\alpha_{ZY} = 0$,*(ou* $r_1 = 0$*) alors* r_2 *est défini de manière unique et est invariant à la reparamétrisation choisie. Il vérifie max (0, r-k)*$\leqslant r_1 + r_2 \leqslant$*min (g,r)* [27]

Compte tenu de la reparamétrisation donnée dans le point (i) du théorème 5, l'équation (2.7) se réécrit comme suit :

$$\begin{cases} \text{modèle conditionnel} \\ \Delta Y_t = \sum\limits_{i=1}^{p-1} \Gamma^+_{YY,i} \Delta Y_{t-i} + \sum\limits_{i=0}^{p-1} \Gamma^+_{YZ,i} \Delta Z_{t-i} + \alpha^+_{YY} \beta'_1 X_{t-1} \\ \qquad + \left(\alpha^+_{YZ_1} \beta'_{ZZ_1} + \alpha^+_{YZ_2} \beta'_{ZZ_2} \right) Z_{t-1} + \eta_{Y,t} \\ \text{modèle marginal} \\ \Delta Z_t = \sum\limits_{i=1}^{p-1} \Gamma_{ZY,i} \Delta Y_{t-i} + \sum\limits_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} + \alpha_{ZY} \beta'_1 X_{t-1} \\ \qquad + \alpha_{ZZ_2} \beta'_{ZZ_2} Z_{t-1} + \varepsilon_{Z,t} \end{cases} \tag{2.7d}$$

Le théorème 5 est démontré à l'aide d'un deuxième changement de base dans l'espace dual de l'espace des vecteurs de co-intégration (c'est-à-dire dans l'espace d'ajustement engendré par les colonnes de la matrice α). Ce changement de base conduit à partitionner les (r-r_1) relations de long terme ne comportant que des variables exogènes en deux sous-ensembles de dimensions respectives r^* et r_2 ; un premier sous-ensemble qui ne figure que dans le modèle conditionnel et un second sous-ensemble qui apparaît à la fois dans les modèles conditionnel et marginal. À cette deuxième partition sur l'espace d'ajustement α correspond un nouveau β dont l'écriture est donnée dans le théorème 5. Le fait que les matrices α et β soient de rang r entraîne que $rang\left(\alpha^+_{YZ_1}\right) = rang\left(\beta_{ZZ_1}\right) = r^* \geqslant 0$ et que $rang\left(\beta_{ZZ_2}\right) = r_2 \geqslant 0$. Ceci nous conduit à formuler la proposition ci-dessous.

Proposition 7. — **Condition nécessaire et suffisante d'exogénéité forte pour les paramètres du modèle conditionnel (Pradel et Rault, 2003)**

Dans le cadre du modèle VAR-ECM bloc-récursif (équation 2.7d), si $r_2 < k$, *alors la variable* Z_t *est fortement exogène pour l'ensemble des paramètres du modèle conditionnel, si et seulement si :* $\begin{cases} \alpha_{ZY} = 0 \\ \alpha^+_{YZ_2} = 0 \\ \Gamma_{ZY,i} = 0, \ i = 1, ..., p-1 \end{cases}$ *dans l'écriture canonique donnée par le théorème 5.*

Si $r_2 = k$, cela implique qu'il existe autant de relations de long terme purement exogènes que de variables Z_t, ce qui entraîne que les Z_t sont stationnaires [28]. Il n'y a dans ce cas aucun paramètre β_{ZZ_2} à estimer puisque cette sous-matrice

27. Cette condition garantit que la matrice α est de rang r.

28. Il est à noter que la condition $rank\left(\beta_{ZZ}\right) = r - r_1$ implique que $k \geqslant r - r_1 = r_2 + r^*$. Dans le cas où $r_2 = k$ alors $r^* = 0$.

est alors égale à I_k : ces variables peuvent par conséquent figurer aussi bien dans le modèle conditionnel que dans le modèle marginal.

La condition d'exogénéité forte donnée ci-dessus demeure la même si le modélisateur n'est intéressé que par les paramètres de long terme du modèle conditionnel (c'est-à-dire par les r_1 vecteurs de co-intégration mélangeant variables endogènes et exogènes, par les r^* relations de long terme purement exogènes, ainsi qu'éventuellement par leurs poids associés). Par ailleurs, cette condition n'interdit pas, à l'instar de la condition nécessaire et suffisante de non-causalité donnée dans la sous-section précédente, que le modèle marginal contienne des relations de long terme purement exogènes. Elle ne le contraint donc pas non plus à être un modèle VAR en différences premières, puisque outre la nullité des paramètres $\Gamma_{ZY,i} = 0$, $i = 1, \ldots, p-1$ associés aux variables en différences, c'est uniquement la partie correspondante de α et non des lignes complètes qui doit disparaître pour assurer cette hypothèse. Par ailleurs, les modèles conditionnel et marginal contiennent deux ensembles distincts de relations de long terme purement exogènes. Le modélisateur a en effet maintenant la possibilité d'être également intéressé par un sous-ensemble des relations de long terme purement exogènes figurant dans le modèle conditionnel, en plus des relations d'équilibre mélangeant variables exogènes et endogènes.

L'exogénéité forte de la variable Z_t pour les paramètres du modèle conditionnel implique que :

(i) aucune des r_1 relations de long terme contenant à la fois les exogènes et les endogènes ne figure dans le modèle marginal,

(ii) aucune des r_2 relations de long terme ne comportant que des variables exogènes et faisant partie du modèle marginal, ne figure dans le modèle conditionnel,

(iii) Y ne cause pas Z au sens de Granger.

Il est ainsi possible d'obtenir sous cette hypothèse une estimation efficace des paramètres du modèle conditionnel et d'effectuer des prévisions de Y_t à partir uniquement de ce modèle (système partiel), en négligeant, sans aucune perte d'information, le modèle marginal.

5. Tests d'exogénéité dans les modèles VAR-ECM

Les conditions nécessaires et suffisantes d'exogénéité et de non-causalité introduites dans la section 4 nécessitent d'écrire au préalable la matrice Π sous l'une des trois décompositions canoniques données par les théorèmes 3, 4 et 5. Par la suite, ces conditions ont toutes été exprimées, dans ce cadre, en termes de nullité de coefficients des matrices α, β et Γ : cela permet d'utiliser les statistiques de chi-deux classiques grâce aux résultats de Johansen. Or, comme nous avons pu le voir, ces trois représentations requièrent la détermination du rang d'une ou de deux sous-matrices, selon l'hypothèse que l'on cherche à caractériser.

Dans cette optique, cette section présente une procédure séquentielle de tests de rangs, adaptée à chacun de ces cas. Une première sous-section est consacrée à l'hypothèse d'exogénéité faible, une seconde, à l'hypothèse de non-causalité et une troisième à l'hypothèse d'exogénéité forte. L'ensemble de cette section s'appuie sur les travaux de Rault (2000, 2005) et Rault et Pradel (2003).

5.1 L'hypothèse d'exogénéité faible

Nous présentons ici la procédure séquentielle de test introduite par Rault (2005) qui permet de déterminer le rang m_1 de la sous-matrice α_Y, de manière à pouvoir réécrire la matrice Π sous la forme canonique donnée par le théorème 3. Puis, nous explicitons comment tester dans ce cadre la condition nécessaire et suffisante d'exogénéité faible établie dans la proposition 5.

La procédure de tests développée par Rault (2005) s'appuie sur un test d'hypothèses proposé par Johansen et Juselius (1992), ainsi que par Hunter (1992), qui permet de tester des restrictions linéaires sur les paramètres de la matrice α. Le modèle est défini par les restrictions

$$\alpha = (H_1\ \vartheta_1, \kappa_1),$$

où H_1 est une matrice connue de dimensions (n, k) et où ϑ_1 et κ_1 sont respectivement des matrices de dimensions (k, j), $(n, r-j)$, dont les paramètres doivent être estimés, avec $j \leqslant k \leqslant n$. En d'autres termes, il s'agit de tester $n-k$ restrictions sur un sous-ensemble de j vecteurs de l'espace d'ajustement, les $r-j$ autres vecteurs variant librement. Ce type d'hypothèse peut être testé avec une statistique du rapport de vraisemblance admettant une loi limite du chi-deux [29]. Il faut noter que nous ne considérons ici que des contraintes de nullité de coefficients, puisqu'il s'agit de déterminer le rang m de la matrice α_Y. Plus précisément, posons $m_a = \min(g, r)$, $m_b = \max(0, r-k)$ et considérons la séquence suivante d'hypothèses nulles :

$$\left\{\begin{array}{l}
H_{0,1}\left\{\begin{array}{l}\text{Il existe une base de vecteurs de l'espace d'ajustement telle que}\\ \alpha = (H_1\theta_{r-m_a+1}, \kappa_{r-m_a+1})\\ \text{avec } H_1 = \begin{pmatrix} 0_{(g,k)} \\ I_k \end{pmatrix},\ \text{c'est-à-dire } m_b \leqslant rang\,(\alpha_Y) \leqslant m_a - 1.\end{array}\right\} \\
\vdots \\
\text{pour } j = 2, \ldots, m_a - m_b,\ \text{tant que } H_{0,j-1} \text{ n'est pas rejetée,} \\
H_{0,j}\left\{\begin{array}{l}\text{Il existe une base de vecteurs de l'espace d'ajustement telle que}\\ \alpha = (H_1\theta_{r-m_a+j}, \kappa_{r-m_a+j})\\ \text{avec } H_2 = \begin{pmatrix} 0_{(g,k)} \\ I_k \end{pmatrix}, \text{c'est-à-dire } m_b \leqslant rang\,(\alpha_Y) \leqslant m_a - j,\end{array}\right\}
\end{array}\right.$$

29. Pour plus de détails, nous renvoyons le lecteur à Johansen et Juselius (1992), ainsi qu'à Hunter (1992), qui illustre la mise en oeuvre de ce test sur des données de taux de change.

Pour tester ces différentes hypothèses, on peut utiliser la procédure séquentielle suivante [30] :

$$\begin{cases} \text{- Étape 1 : tester } H_{0,1} \text{ avec la statistique } \xi_1 \text{ au niveau } \alpha_1 \\ \text{et rejeter } H_{0,1} \, (\Longrightarrow rang\,(\alpha_Y) = m_a) \quad \text{si } \xi_1 \geqslant \chi^2_{1-\alpha_1}(v_1), \\ \vdots \\ \text{pour } j = 2, \ldots, m_a - m_b, \text{ tant que } H_{0,j-1} \text{ n'est pas rejetée} \\ \text{- Étape j : tester } H_{0,j} \text{ avec la statistique } \xi_j \text{ au niveau } \alpha_j \\ \text{et rejeter } H_{0,j} \; (\Longrightarrow rang\,(\alpha_Y) = m_a - j + 1) \text{ si } \xi_j \geqslant \chi^2_{1-\alpha_j}(v_j), \\ \text{sinon accepter } H_{0,j} \; (\Longrightarrow rang\,(\alpha_Y) = m_a - j) \text{ si } \xi_j < \chi^2_{1-\alpha_j}(v_j). \end{cases}$$
avec $\nu_j = (g - r + j)j$

Chaque hypothèse nulle est testée avec une statistique du rapport de vraisemblance

$$\xi_j = -2\ln Q(H_j/H_1) = T\left[\sum_{i=1}^{j} \ln(1-\widehat{\rho}_i) + \sum_{i=1}^{r-j} \ln(1-\widehat{\lambda}_i) - \sum_{i=1}^{r} \ln(1-\widetilde{\lambda}_i)\right] \quad (2.8)$$

qui suit asymptotiquement sous $H_{0,j}$ une loi du chi-deux, avec v_j = (n – k – (r – j))j degrés de liberté. H_1 correspond à l'hypothèse de co-intégration $\Pi = \alpha\beta'$ et $\widetilde{\lambda}_i$, $\widehat{\rho}_i$, $\widetilde{\lambda}_i$ désignent respectivement les valeurs propres du modèle VAR-ECM non contraint, des j vecteurs contraints de l'espace d'ajustement et des r-j vecteurs non contraints de l'espace d'ajustement.

Il faut souligner que, comme cette statistique de test est obtenue sous des arguments asymptotiques, elle constitue une mauvaise approximation de la distribution limite en petits échantillons et nécessite donc des corrections. Le vrai seuil est supérieur au seuil nominal du test, c'est la raison pour laquelle il est préférable de considérer ici des statistiques du rapport de vraisemblance corrigées de la taille de l'échantillon (*cf.* Psaradakis, 1994). Ces statistiques sont obtenues en remplaçant T par T – (s / n) + 0.5 (n – r (n – k) / (n + 1)) dans l'équation (2.8), où s désigne le nombre de paramètres à estimer dans le modèle VAR-ECM non-contraint [31]. Cette correction en petits échantillons conduit à une puissance et à un seuil empiriques plus satisfaisants, lorsqu'on teste des restrictions linéaires dans un modèle multivarié gaussien (voir par exemple à ce sujet Anderson, 1984, ch 8)).

Ayant déterminé le rang m_1 de la sous-matrice α_Y, la matrice Π peut être réécrite sous la forme canonique donnée par le théorème 3 et l'hypothèse d'exogénéité faible implique alors les restrictions paramétriques suivantes :

$$H_{0,we} : \alpha_{Z1} = 0.$$

Comme ces restrictions ne correspondent qu'à des nullités de coefficients dans le modèle marginal, plusieurs types de tests usuels peuvent être sollicités (test

30. Il faut noter que cette procédure détermine correctement le rang de la matrice α_Y, même si certaines colonnes de α_Y sont combinaisons linéaires des autres (voir à ce sujet l'expérience de Monte Carlo fournie en section 6).

31. c'est-à-dire dans l'équation $\Delta X_t = \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \alpha\beta' X_{t-1} + \varepsilon_t, t = 1, ..., T$.

du rapport de vraisemblance, test du multiplicateur de Lagrange, test de Wald). Ces tests sont tous relativement simples à mettre en œuvre dans les travaux appliqués, les statistiques de tests étant directement données par les logiciels standards comme PC-FIML (1994,1997), CATS in RATS, ou MALCOLM 2.4 (Mosconi, 1998) par exemple. Il est à souligner que le test du rapport de vraisemblance est ici en général préférable aux tests de Wald ou du multiplicateur de Lagrange puisque les restrictions sont non linéaires sur Π, même si elles sont linéaires sur α. Le test du rapport de vraisemblance est en effet invariant par rapport à la manière dont ces restrictions s'expriment sur Π.

5.2 L'hypothèse de non-causalité

Nous détaillons ici la procédure séquentielle de tests fournie par Rault (2000) pour déterminer le rang r_1 de la matrice β_Y, qui est requis par la décomposition canonique de la matrice Π du théorème 4. Nous indiquons également comment tester la condition nécessaire et suffisante de non-causalité établie dans la proposition 6.

La procédure de tests repose, comme précédemment, sur l'utilisation d'une classe particulière de tests d'hypothèses structurelles proposée par Johansen et Juselius (1992). Il s'agit néanmoins à présent de tester des restrictions linéaires sur les vecteurs de co-intégration β. Le modèle est caractérisé par les restrictions

$$\beta = (H_2\ \varphi, \Psi),$$

où H_2 est une matrice connue de dimensions (n, k) et où φ et Ψ sont respectivement des matrices de dimensions (k, j), $(n, r-j)$, dont les paramètres doivent être estimés, avec $j \leqslant k \leqslant n$. Il s'agit en fait de tester $n-k$ restrictions sur un sous-ensemble de j vecteurs de co-intégration, les r-j autres vecteurs variant librement. Comme l'ont montré Johansen et Juselius (1992), une telle hypothèse peut être testée avec une statistique du rapport de vraisemblance admettant une loi asymptotique de type chi-deux. Pour déterminer le rang r_1 de la matrice β_Y, nous ne considérons ici que des contraintes d'exclusion de variables. Pour ce faire, posons $r_a = min(g, r)$, $r_b = max(0, r-k)$ et considérons maintenant la séquence suivante d'hypothèses nulles :

$$\left\{\begin{array}{l}
H_{0,1}\left\{\begin{array}{l}\text{Il existe une base de vecteurs de co-intégration telle que}\\ \beta=(H_2\ \varphi_{r-r_a+1},\Psi_{r-r_a+1})\\ \text{avec } H_2=\begin{pmatrix}0_{(g,k)}\\ I_k\end{pmatrix}\text{, c'est-à-dire } r_b\leqslant rang\,(\beta_Y)\leqslant r_a-1.\end{array}\right\}\\
\vdots\\
\text{pour } j=2,\ldots,m_a-m_b\text{, tant que } H_{0,j-1}\text{ n'est pas rejetée,}\\
H_{0,j}\left\{\begin{array}{l}\text{Il existe une base de vecteurs de co-intégration telle que}\\ \beta=(H_2\ \varphi_{r-r_a+j},\Psi_{r-r_a+j})\\ \text{avec } H_2=\begin{pmatrix}0_{(g,k)}\\ I_k\end{pmatrix}\text{, c'est-à-dire } r_b\leqslant rang\,(\beta_Y)\leqslant r_a-j,\end{array}\right\}
\end{array}\right.$$

Pour tester ces différentes hypothèses, on peut mettre en œuvre la procédure séquentielle suivante [32] :

$$\begin{cases} \text{- Étape 1 : tester } H_{0,1} \text{ avec la statistique } \xi_1 \text{ au niveau } \alpha_1 \\ \text{et rejeter } H_{0,1} \ (\Longrightarrow rang\,(\beta_Y) = r_a) \quad \text{si } \xi_1 \geqslant \chi^2_{1-\alpha_1}(v_1), \\ \vdots \\ \text{pour } j = 2, ..., m_a - m_b, \text{ tant que } H_{0,j-1} \ \text{ n'est pas rejetée} \\ \text{- Étape } j : \text{ tester } H_{0,j} \text{ avec la statistique } \xi_j \text{ au niveau } \alpha_j \\ \text{et rejeter } H_{0,j} \ (\Longrightarrow rang\,(\beta_Y) = r_a - j + 1) \ \text{ si } \xi_j \geqslant \chi^2_{1-\alpha_j}(v_j), \\ \text{sinon accepter } H_{0,j} \ (\Longrightarrow rang\,(\beta_Y) = r_a - j) \ \text{ si } \xi_j < \chi^2_{1-\alpha_j}(v_j). \end{cases}$$

avec $\nu_j = (g - r + j)j$

Chaque hypothèse nulle est testée avec une statistique du rapport de vraisemblance

$$\begin{aligned} \xi_j &= -2 \ln Q \ (H_j/H_1) \\ &= T \left[\sum_{i=1}^{j} \ln\,(1-\widehat{\rho}_i) + \sum_{i=1}^{1-j} \ \ln\,(1-\widehat{\lambda}_i) - \sum_{i=1}^{r} \ \ln\,(1-\widetilde{\lambda}_i) \right] \end{aligned}$$

qui suit asymptotiquement sous $H_{0,j}$ une loi du chi-deux, avec $v_j = (n - k - (r-j))j$ degrés de liberté. H_1 correspond à l'hypothèse de co-intégration $\Pi = \alpha\beta'$ et $\widetilde{\lambda}_i$, $\widehat{\rho}_i$, $\widetilde{\lambda}_i$ désignent respectivement les valeurs propres du modèle VAR-ECM non contraint, des j relations de co-intégration contraintes et des $r - j$ relations de co-intégration non contraintes.

Dans le cadre de l'écriture canonique donnée par le théorème 4, l'hypothèse de non-causalité correspond alors aux restrictions paramétriques suivantes :

$$H_{0,nc} : \begin{cases} \Gamma_{ZY,i} = 0, \ i = 1, ..., p-1 \\ \alpha_{Z1} = 0 \end{cases}$$

Comme pour l'exogénéité faible (*cf.* section 4.1), ces restrictions peuvent également ici être testées avec les statistiques usuelles de tests (test du rapport de vraisemblance, test du multiplicateur de Lagrange, test de Wald), qui admettent sous l'hypothèse nulle une distribution limite du chi-deux et qui sont facilement calculées dans les travaux appliqués avec les logiciels économétriques usuels.

5.3. L'hypothèse d'exogénéité forte

Nous détaillons ici la procédure séquentielle de tests de rangs développée par Rault et Pradel (2003) permettant de déterminer les rangs r_1 et r_2 des matrices β_Y et α_{ZZ_2} [33]. Ces rangs permettent de réécrire la matrice Π sous la forme

32. Soulignons que cette procédure détermine correctement le rang de la matrice β_Y, même si certaines colonnes de β_Y sont combinaisons linéaires des autres.

33. Il ne s'agit pas bien évidemment ici de « tests de rang » (type Wilcoxon) au sens que l'on entend dans les procédures non paramétriques.

canonique donnée par le théorème 5. Nous précisons ensuite comment tester dans ce cadre la condition nécessaire et suffisante d'exogénéité forte donnée dans la proposition 7.

Les rangs r_1 et r_2 peuvent être déterminés à l'aide de la procédure ci-dessous, composée de deux étapes :

1) Tout d'abord, déterminer le rang r_1 de la matrice β_Y qui correspond à la reparamétrisation des matrices α et β donnée dans le théorème 4. Pour ce faire, mettre en œuvre la procédure séquentielle de tests présentée dans la sous-section précédente.

2) Ensuite, pour la valeur du rang r_1 déterminée à l'étape 1, tester H_0 : $\alpha_{ZY} = 0$ en utilisant une statistique de test admettant asymptotiquement une distribution chi-deux (*cf.* par exemple Toda and Phillips, 1991,1993).

Si l'hypothèse nulle n'est pas rejetée, passer à l'étape (3) sinon arrêter la procédure de tests.

Le second cas implique que l'hypothèse d'exogénéité forte est également rejetée : cette hypothèse nécessite en effet que $\alpha_{ZY} = 0$.

3) Enfin, pour r_1 déterminé à l'étape 1) et $\alpha_{ZY} = 0$, considérant le modèle donné par la décomposition canonique du théorème 4 :

$$\beta = [\beta_1 \ \beta_2] = \begin{bmatrix} \beta_{YY} & 0 \\ \beta_{ZY} & \beta_{ZZ} \end{bmatrix}, \alpha = [\alpha_1 \ \alpha_2] = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ} \\ \alpha_{ZY} & \alpha_{ZZ} \end{bmatrix}$$

et pour $j = 1, \ldots, \min(g, r - r_1)$ tester les restrictions :

$$\alpha_{ZZ} = \begin{bmatrix} 0_{(k,\ j)} & \alpha_{ZZ_2} \end{bmatrix}$$

(où α_{ZZ} et α_{ZZ_2} sont des sous-matrices respectivement de dimensions $k\times$ r-r_1 and $k\times\ r-r_1-j$) sous lesquelles les paramètres de long-terme peuvent s'écrire comme suit :

$$\beta = [\beta_1 \ \beta_{21} \ \beta_{22}] = \begin{bmatrix} \beta_{YY} & 0 & 0 \\ \beta_{ZY} & \beta_{ZZ_1} & \beta_{ZZ_2} \end{bmatrix},$$

$$\alpha = [\alpha_1 \ \alpha_{21} \ \alpha_{22}] = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ_1} & \alpha_{YZ_2} \\ 0 & 0 & \alpha_{ZZ_2} \end{bmatrix}$$

où β_{YY}, α_{YY}, β_{ZY}, β_{ZZ_1},α_{YZ_1},β_{ZZ_2}, α_{YZ_2}, α_{ZZ_2} sont des sous-matrices respectivement de dimensions $g\times r_1$, $g\times r_1$, $k\times r_1$, $k\times j$, $g\times j$, $k\times r-r_1-j$, $g\times r-r_1-j$, $k\times r-r_1-j$.

Les écritures ci-dessus correspondent à un modèle défini par :

$$H_{0,j} \left\{ \begin{array}{l} \text{Il existe une base des espaces co-intégrant et d'ajustement} \\ \text{telle que } \beta = (\beta_1, H_1\beta_{ZZ_1}, H_1\beta_{ZZ_2}),\ \alpha = (H_2\alpha_{YY}, H_2\alpha_{YZ_1}, \alpha_{22}) \\ \text{avec } rang\,(\beta_{YY}) = r_1,\ rang\,(\beta) = rang\,(\alpha) = r \end{array} \right\}$$

où $H_1 = \begin{pmatrix} 0_{(g,k)} \\ I_k \end{pmatrix}$ et $H_2 = \begin{pmatrix} I_g \\ 0_{(g,k)} \end{pmatrix}$

avec β_1, β_{ZZ_1}, β_{ZZ_2}, α_{YY}, α_{YZ_1}, α_{22}, respectivement de dimensions $n \times r_1$, $k \times j$, $k \times r - r_1 - j$, $g \times r_1$, $g \times j$, $n \times r - r_1 - j$.

Les paramètres à estimer sont $\beta_1, \beta_{ZZ_1}, \beta_{ZZ_2}, \alpha_{YY}, \alpha_{YZ_1}, \alpha_{22}$, étant donné que certaines contraintes apparemment imposées sur les paramètres de long terme, en particulier sur la matrice β, ne sont que des contraintes identifiantes dans notre modèle. L'intervalle de variation pour j est calculé à partir des conditions de rang suivantes :

$$\begin{array}{rcl} 1 & \leqslant & j \leqslant k \\ j & \leqslant & g \\ 0 & \leqslant & r - r_1 - j \end{array}$$

Pour tester les hypothèses $H_{0,j}$, Rault et Pradel (2003) procèdent comme dans Johansen et Juselius (1992), Johansen (1995a), Giannini and Mosconi (1992), Konishi and Granger (1993) et utilisent la version appropriée du « switching algorithm » développé par ces auteurs (*cf.* Hecq *et al.*, 2001) afin de tester les restrictions linéaires sur les paramètres de long terme, les statistiques de tests suivant asymptotiquement dans ce cas une loi du chi-deux.

Plus précisément, comme il s'agit de déterminer ici le rang r_2 de la sous-matrice α_{ZZ} considérons maintenant la séquence suivante d'hypothèses nulles :

$$\begin{cases} H_{0,1} \text{ i.e. } rang\,(\alpha_{ZZ}) \leqslant \min\,(k, r - r_1 - 1)\,, \\ \vdots \\ \text{pour } j = 2, ..., , \min\,(g, r - r_1)\,, \text{ tant que } H_{0,j-1} \text{ n'est pas rejetée,} \\ H_{0,j} \text{ i.e. } rang\,(\alpha_{ZZ}) \leqslant \min\,(k, r - r_1 - j) \end{cases}$$

Pour tester ces différentes hypothèses, on peut mettre en œuvre la procédure séquentielle suivante :

$$\begin{cases} \text{Étape 1 : tester } H_{0,1} \text{ avec la statistique } \xi_1 \text{ au niveau } \alpha_1 \\ \text{et rejeter } H_{0,1} \text{ si } \xi_1 > \chi^2_{1-\alpha_1}(v_1) \Longrightarrow rang\,(\alpha_{ZZ}) = r - r_1, \\ \vdots \\ \text{pour } j = 2, ..., \min\,(g, r - r_1)\,, \text{ tant que } H_{0,j-1} \text{ n'est pas rejetée,} \\ \text{Étape j : tester } H_{0,j} \text{ avec la statistique } \xi_j \text{ au niveau } \alpha_j \\ \text{et rejeter } H_{0,j} \text{ si } \xi_j > \chi^2_{1-\alpha_j}(v_j) \Longrightarrow rang\,(\alpha_{Z2}) = r - r_1 - j + 1 \\ \text{Sinon accepter } H_{0,j} \text{ si } \xi_j < \chi^2_{1-\alpha_j}(v_j) \Longrightarrow rang\,(\alpha_{Z2}) = r - r_1 - j \end{cases}$$

Chaque hypothèse nulle est testée avec une statistique du rapport de vraisemblance :

$$\begin{aligned} \xi_j = -2\ln Q(L_{0,j}/L_1) = T \Bigg[\sum_{i=1}^{r_1} \ln(1 - \widehat{\rho}_{1i}) + \sum_{i=1}^{r^*} \ln(1 - \widehat{\rho}_{2i}) \\ + \sum_{i=1}^{r_2} \ln(1 - \widehat{\rho}_{3i}) - \sum_{i=1}^{r} \ln(1 - \widetilde{\lambda}_i) \Bigg] \end{aligned} \tag{1}$$

qui est asymptotiquement distribuée sous l'hypothèse nulle comme un $\chi^2_{1-\alpha_j}(v_j)$, où ν_j désigne le nombre de degrés de liberté, calculé comme le nombre de coefficients égaux à zéro sous l'hypothèse H_{oj}, qui vont au delà de ceux résultant de la normalisation. H_1 correspond à l'hypothèse de co-intégration $\Pi = \alpha\beta'$ et $\widetilde{\lambda}_j$ désignent les valeurs propres associées au modèle VAR-ECM non contraint, et les $\widehat{\rho}_{1j}, \widehat{\rho}_{2j}, \widehat{\rho}_{3j}$ correspondent respectivement aux $r_1, j, (r-r_1-j)$ relations de co-intégration contraintes et leurs coefficients associés.

Une fois ces rangs r_1 et r_2 déterminés, la matrice Π peut être réécrite sous l'expression canonique donnée dans le théorème 5 et l'hypothèse d'exogénéité forte conduit alors aux restrictions paramétriques suivantes (*cf.* propositions 7) :

$$H_{0,se} : \begin{cases} \Gamma_{ZY,i} = 0, \ i = 1, ..., p-1 \\ \\ \alpha^+_{YZ_2} = 0 \end{cases}$$

Ces restrictions linéaires peuvent ici encore être testées avec les statistiques usuelles.

6. Expérience de Monte Carlo [34]

Afin d'étudier les propriétés des procédures séquentielles de tests de rangs proposées dans la section 5, nous mettons en œuvre une expérience de Monte Carlo.

6.1. Conditions expérimentales des simulations

Les simulations sont menées sur un processus vectoriel $(\mathrm{X}_t)_{\ t=1,...,T}$, de dimension n, de moyenne zéro (donc, sans composante déterministe), d'ordre $p = 1$, admettant la représentation autorégressive suivante :

$$\begin{cases} X_t = A \ X_{t-1} + \varepsilon_t, t = 1, \ldots, T \\ \\ X_0 \end{cases} \tag{2.9}$$

où

- $(\varepsilon_t) \sim \text{iid}\, N(0_n, \Sigma)$,
- Σ est une matrice de variance-covariance régulière, définie positive,
- A est une matrice carrée de dimensions (n, n), supposée constante au cours du temps.

34. Cette section est une version étendue de Rault (2000).

6.1.1. Conditions pour générer un processus $I(1)$, co-intégré d'ordre r

Pour que le processus X_t soit intégré d'ordre 1 et admette r relations de co-intégration (avec $0 < r < n$), l'équation (2.9) doit vérifier les trois hypothèses ci-dessous [35] :

(H_1) Les racines de l'équation $\det(I_n - Az) = 0$ sont soit égales à un, soit de module strictement supérieur à un,

(H_2) $\text{Rang}(\Pi) = \text{Rang}(A - I_n) = r$, avec $\Pi = \alpha\beta'$,

(H_3) La matrice $\alpha'_\perp A\,\beta_\perp$ est *r*égulière.

En pratique, pour effectuer des simulations, le problème qui se pose est de générer des processus vérifiant ces trois conditions. Une condition nécessaire et suffisante a été donnée par Yum (1994) qui a démontré le théorème suivant :

THÉORÈME 6. — **(Yum, 1994, Lemme 5.3, p 361)**
Soit deux matrices α et β de dimensions (n, r) de rang r, telles que $|\lambda_i(\beta'\alpha) + 1| < 1$, $i = 1, \ldots, r$. Si on pose $A = \alpha\beta' + I_n$, le processus X_t donné par l'équation (2.9) est $I(1)$ et admet r comme rang de co-intégration.

En particulier, une condition suffisante pour que X_t soit $I(1)$ avec une seule relation de co-intégration est que les vecteurs α et β (non nuls) satisfassent la condition $-2 < \beta'\alpha < 0$.

6.1.2. Description des processus générés

Il s'agit de simuler des processus admettant la représentation donnée par l'équation (2.9), qui peut être réécrite de manière équivalente sous la forme VAR-ECM ci-dessous :

$$\begin{cases} \Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t, t = 1, \ldots, T \\ X_0 \end{cases} \tag{3.0}$$

Dans tout ce qui suit, nous considérons comme précédemment la partition du vecteur X_t en deux sous-vecteurs $\binom{Y_t}{Z_t}$, $Y_t \in R^g, Z_t \in R^k$, avec $g+k = n$ et nous choisissons les valeurs des paramètres des matrices α et β de manière à ce que les conditions du théorème 6 soient satisfaites. Les différentes configurations envisagées sont les suivantes : nous considérons pour chacun des deux cas à étudier cinq processus de génération des données (DGP) X_t de dimensions 11 ($g = 5$, $k = 6$), intégrés d'ordre 1, avec 4 relations de co-intégration, exprimés sous forme VAR-ECM, sans dynamique de court terme ($p = 1$). Ces DGP se distinguent soit par le rang m_1 de la matrice α_Y (première décomposition canonique), soit par le rang r_1 de la matrice β_Y (deuxième décomposition canonique) [36], qui varient de 0 à 4. Les tableaux 1 et 4 (*cf.* annexe) résument les dix DGP utilisés dans cette expérience.

35. *cf.* Johansen (1995b).

36. *cf.* théorèmes 3 et 4.

Dans chaque cas, nous considérons 10000 échantillons de taille T. Pour que les conditions initiales varient arbitrairement, nous ajoutons 100 observations à la taille de la série retenue, c'est-à-dire que nous générons $T+100+p$ points et ignorons les 100 premiers, où p désigne le nombre de retards du modèle VAR-ECM estimé. Le vecteur des innovations ε_t suit un processus gaussien de dimension 11, de moyenne zéro et de matrice de variance-covariance I_{11}. Les valeurs initiales ($t=0$) sont mises à zéro pour toutes les variables du modèle, c'est-à-dire que $X_0 = 0_{11}$ et $X_1 = \varepsilon_t \sim N(0_{11}, I_{11})$. Toutes les simulations sont menées en utilisant le langage de programmation matriciel GAUSS, les ε_t sont générés avec la fonction « RNDN » et chaque test est effectué au seuil nominal de 5 %.

Pour chaque DGP, nous incluons cinq tailles d'échantillons ($T = 50$, 100, 200, 500, 1000) et nous utilisons les statistiques de tests du rapport de vraisemblance corrigées de la taille de l'échantillon pour $T \leqslant 100$. Lors de chaque répétition, le nombre de retards « p » et la dimension de l'espace co-intégrant « r » sont supposés, dans un premier temps, avoir été correctement déterminés au préalable à l'aide de tests du rapport de vraisemblance (par exemple), de manière à ce que l'on puisse se concentrer exclusivement sur les performances de nos deux procédures séquentielles de tests de rangs. Les résultats tabulés de ces exercices de simulation sont donnés dans les tableaux 2, 3, 5 et 6 (*cf.* annexe). Les nombres dans le corps des tableaux 2, 5 et 3, 6 donnent respectivement les seuils et puissances des tests intermédiaires de l'hypothèse nulle $H_{0,j}$ ($j = 1, \ldots, 4$), de seuil nominal 5 % et le seuil global de la procédure séquentielle considérée.

6.2. Résultats des simulations

Les résultats de ces deux procédures de tests de rangs sont très similaires et nous allons donc les commenter en les exposant dans le cadre d'une décomposition canonique particulière. Nous avons choisi la deuxième (qui permet d'étudier, rappelons-le, l'hypothèse de non-causalité).

Tous les tests de l'hypothèse nulle $H_{0,j}$ ($j = 1, \ldots, 4$) souffrent de distorsion de seuil en petit échantillon ($T = 50, 100$), puisque les seuils réels sont différents des seuils nominaux. Au fur et à mesure que la taille de l'échantillon s'accroît, les seuils réels approximent assez bien le seuil nominal. Il faut souligner que l'asymptotique est atteint d'autant plus tard que le nombre de restrictions testées est élevé. Pour l'hypothèse nulle la moins contrainte ($H_{0,1}$), le seuil réel est proche du seuil théorique des tests pour des échantillons comportant 500 observations (5.03 % pour $T = 500$), tandis que pour l'hypothèse nulle la plus contrainte ($\mathrm{H}_{0,4}$) le seuil réel est encore de 5.15 % pour des échantillons comportant 1000 points. On constate en outre pour chaque test une puissance égale à 1 pour les DGP considérés.

En ce qui concerne la procédure séquentielle, la multiplicité des tests peut conduire à un problème de seuil global. Nos simulations mettent effectivement en évidence ce phénomène, qui est cependant plus marqué pour des échantillons de petite taille ($T = 50, 100$). On constate en effet que, pour des tests intermédiaires de seuil nominal 5 %, le seuil global de la procédure séquentielle dépend du nombre de tests nécessaires pour conclure (pour $T = 100$ par exemple, il est respectivement de 6.16 %, 7.05 %, 9.43 %, 10.6 % pour $r_1 = 3, \ldots 1, 0$). Au contraire, pour des échantillons de plus grandes tailles ($T = 200$, 500, ou 1 000), le seuil global semble ne pas varier énormément, ce qui traduit le fait que la procédure de test ne souffre pas de distorsion de seuil en grand échantillon. Pour n'importe quel rang r_1, le seuil global estimé est toujours trè s proche de 5 % (il est respectivement de 5.03 %, 5.19 %, 5.47 %, 5.62 % pour $r_1 = 3, \ldots, 0$ et $T = 500$). Ce résultat est dû au fait que les différentes hypothèses nulles $H_{0,j}$ ($j = 1, \ldots, 4$) sont suffisamment séparées pour que le seuil global de la procédure soit pratiquementégal au seuil du dernier test que l'on met en œuvre. Par conséquent, si nous devons effectuer j tests pour déterminer le rang r_1 de la matrice β_Y, le seuil global de la procédure séquentielle est donné en grand échantillon par $\alpha = 1 - (1 - \alpha_j)$. Il faut souligner que ce résultat n'est pas vrai à distance finie.

Nous avons également effectué des simulations complémentaires (non reportées ici) afin d'étudier la robustesse de cette procé dure de tests en cas d'erreur de spécification sur le nombre (r) de relations de co-intégration. Il en ressort les deux résultats suivants :

(i) Si r est surestimé, c'est-à-dire si on retient un système générateur de vecteurs de l'espace co-intégrant au lieu d'une base, les résultats reportés dans les tableaux 3 à 6 ne sont pas modifiés : en effet, si on retient $r = 5$ au lieu de $r = 4$, on peut alors faire apparaître par combinaison linéaire une colonne de zéros de dimension $(n, 1)$ dans la matrice β, ce qui rajoute simplement une étape dans la procédure séquentielle de tests du rang considérée, mais n'altère pas sa performance.

(ii) Par contre, si r est sous-estimé, c'est-à-dire si on retient moins de vecteurs linéairement indépendants qu'il n'en faut pour engendrer l'espace co-intégrant, la procédure séquentielle de tests est biaisée (sa puissance est inférieure à son seuil).

7. Conclusion et extensions

Dans ce travail, nous nous sommes attachés à dresser une revue de la littérature des principaux résultats sur l'exogénéité dans les modèles vectoriels à correction d'erreurs, en faisant le parallèle avec la littérature maintenant bien connue sur l'exogénéité dans les modèles VAR. Monfort et Rabemananjara (1990) on en effet proposé une méthode permettant de tester soigneusement chaque aspect d'une forme structurelle dans le cadre d'une modélisation VAR non-contrainte. En particulier, il est possible de construire un modèle à équations simultanées à partir de la représentation VAR canonique en orthogonalisant les résidus. Ce modèle conditionnel de nature plus « structurelle » est analysable séparément pourvu qu'une hypothèse d'exogénéité faible soit satisfaite. Dans le cas d'un modèle VAR, elle l'est toujours. Les choses sont en revanche beaucoup moins claires dans les modèles à variables non stationnaires et l'introduction de la possibilité de co-intégration entre les variables complique grandement l'analyse. En effet, la co-intégration en impliquant une perte de rang de la matrice de long terme impose des restrictions qui font que l'hypothèse d'exogénéité n'est plus automatiquement vérifiée. Il devient alors nécessaire de formuler explicitement les conditions assurant l'exogénéité d'une variable pour les paramètres d'intérêt de l'étude que l'on souhaite entreprendre. Nous avons pour ce faire soigneusement exposé les diverses conditions d'exogénéité de la littérature et discuté les forces et faiblesses de chacune d'entre elle, puis détaillé la manière de les tester dans un travail appliqué. Trois procédures séquentielles de tests de rangs ont ainsi été proposées, une pour chaque décomposition canonique. En outre, une expérience de Monte Carlo a été mise en œuvre pour étudier en petits en grands échantillons les propriétés des deux premières procédures.

Deux grands types extensions permettraient, à notre sens, d'approfondir les résultats de la littérature sur l'exogénéité dans les modèles vectoriels à corrections d'erreurs. La première est un prolongement des travaux de Rault (2000, 2005), Rault et Pradel (2003) dans lesquels ont été proposées trois décompositions canoniques de la matrice Π, selon l'hypothèse à laquelle on s'intéresse (l'exogénéité faible, la non-causalité et l'exogénéité forte). Ces auteurs ne cherchaient alors qu'à donner des conditions nécessaires et suffisantes d'exogénéité, qui permettent de distinguer clairement la nullité des blocs de paramètres que l'on peut toujours obtenir sans perte de généralité à l'aide de changement de bases, de la nullité des blocs de paramètres découlant de la propriété testée. Á ce titre, la question de l'identifiabilité des vecteurs de co-intégration n'a pas été discutée. Pourtant, il est fréquent que des restrictions identifiantes fournies par la théorie économique soient utilisées pour identifier les différents vecteurs de co-intégration. Il serait alors intéressant d'étudier comment ces hypothèses structurelles pourraient être incorporées dans ce cadre d'analyse. Étendre ce travail aux modèles sur-identifiés paraît également être une voie de recherche prometteuse.

Un autre prolongement possible serait de mettre au point une stratégie séquentielle permettant de tester chaque aspect d'une forme structurelle à correction d'erreurs (MSCE) dans le cadre d'une modélisation VAR-ECM non contrainte. Cela constituerait une extension de la méthode d'analyse des notions de non-causalité et d'exogénéité proposée par Monfort et Rabemananjara (1990) au cadre non stationnaire des modèles à variables co-intégrées.

On peut par ailleurs se poser la question plus générale de savoir si un modèle VAR-ECM, même reparamétré, est assimilable à une forme structurelle. Il est certes indéniable qu'un modèle VAR-ECM est interprétable économiquement (à la différence d'un modèle VAR) et qu'il offre une description séparée entre, d'une part, les relations de long terme, représentées par les vecteurs de co-intégration qui sont considérés comme des équilibres de longue période, et d'autre part, des liens de court terme traduisant les ajustements dynamiques vers ces équilibres de long terme. Cependant, dans sa discussion sur le travail de Clement et Mizon (1991), Sims (1991) critique cette assimilation, car les modèles VAR-ECM, mêmes reparamétrés, découlant d'une analyse des relations de co-intégration, ne sont interprétables qu'a *postériori*. On ne cherche en effet d'explications économiques que sous contrainte des vecteurs de co-intégration (ou, du moins de leur nombre et de leurs composantes). Qui plus est, les paramètres estimés, ainsi que les perturbations de leurs équations, ne sont pas toujours interprétés en terme de comportement des agents économiques, d'implications sur la forme des fonctions de production, d'utilité ou de politiques économiques. La démarche est tout à fait différente de celle suivie par Monfort et Rabemananjara (1990), qui cherchent à vérifier l'adéquation d'un modèle VAR orthogonalisé et d'une forme structurelle. La théorie économique est alors antérieure à la démarche statistique, ne serait-ce que par la réflexion préalable sur le phénomène économique à modéliser et sur les *a priori* concernant le sens de l'action de certaines variables. On peut formuler la différence en ces termes : ce n'est pas l'interprétation qui doit justifier les relations empiriques, mais les relations empiriques qui doivent confirmer une théorie économique.

Il serait donc intéressant de poursuivre ces travaux dans le cadre non stationnaire et de mettre au point des tests en présence de séries co-intégrées, qui reprendraient la tentative de réconciliation d'une économétrie traditionnelle basée sur l'existence de formes structurelles et de l'approche fondée sur les séries temporelles. Il serait également utile d'examiner les développements récents de l'exogénéité dans le cadre des modèles non linéaires, non paramétriques et de panel. Ceci fera l'objet de travaux ultérieurs.

TABLEAU 1. – Processus de génération des données (DGP) ($n = 11$, $g = 5$, $k = 6$)[1] associés à la procédure de test de la section 5.1

DGP (1) : $m_l = 4$				DGP (2) : $m_l = 3$				DGP (3) : $m_l = 2$				DGP (4) : $m_l = 1$				DGP (5) : $m_l = 0$			
Beta				Beta				Beta				Beta				Beta			
0.10	1.70	0.60	-0.10	0.10	1.70	0.60	-0.10	0.10	1.70	0.60	-0.10	0.10	1.70	0.60	-0.10	1.00	-1.00	0.20	-1.00
-0.10	-2.00	-0.40	-0.40	-0.10	-2.00	-0.40	-0.40	-0.10	-2.00	-0.40	-0.40	-0.10	-2.00	-0.40	-0.40	-1.00	4.00	-4.00	-0.40
-0.30	0.50	-0.20	0.50	-0.30	0.50	-0.20	0.50	-0.30	0.50	-0.20	0.50	-0.30	0.50	-0.20	0.50	-3.00	-0.90	2.00	0.50
-1.00	0.10	0.20	0.20	-1.00	0.10	0.20	0.20	-1.00	0.10	0.20	0.20	-1.00	0.10	0.20	0.20	-1.00	1.00	0.80	0.50
0.10	-1.00	-0.80	0.30	0.10	-1.00	-0.80	0.30	0.10	-1.00	-0.80	0.30	0.10	-1.00	-0.80	0.30	-0.90	-1.00	-0.20	0.30
0.20	-0.10	-0.20	0.10	0.20	-0.10	-0.20	0.10	0.20	-0.10	-0.20	0.10	0.20	-0.10	-0.20	0.10	0.20	-0.10	-0.20	0.10
0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.80	0.20	0.10	0.30
0.10	-0.20	-0.30	-0.20	0.10	-0.20	-0.30	-0.20	0.10	-0.20	-0.30	-0.20	0.10	-0.20	-0.30	-0.20	0.10	0.30	-0.30	0.60
0.20	-0.10	0.20	-0.10	0.20	-0.10	0.20	-0.10	0.20	-0.10	0.20	-0.10	0.20	-0.10	0.20	-0.10	0.20	-0.50	0.20	-0.10
0.60	0.50	0.30	0.00	0.60	0.50	0.30	0.00	0.60	0.50	0.30	0.00	0.60	0.50	0.30	0.00	0.40	-0.20	0.30	0.00
0.10	-0.30	-0.30	0.00	0.10	-0.30	-0.30	0.00	0.10	-0.30	-0.30	0.00	0.10	-0.30	-0.30	0.00	0.10	-0.30	-0.30	0.20
alpha				alpha				alpha				alpha				alpha			
-0.50	-0.30	-0.40	0.00	-0.50	-0.30	-0.40	0.00	-0.50	-0.30	-0.60	0.00	-0.90	-0.30	-0.60	0.00	0.00	0.00	0.00	0.00
0.30	0.20	0.60	0.00	0.30	0.20	0.60	0.00	0.30	0.20	0.40	0.00	0.60	0.20	0.40	0.00	0.00	0.00	0.00	0.00
-0.20	-0.20	-0.20	0.00	-0.20	-0.20	-0.20	0.00	-0.20	-0.20	-0.40	0.00	-0.60	-0.20	-0.40	0.00	0.00	0.00	0.00	0.00
0.70	0.10	0.50	0.00	0.70	0.10	0.50	0.00	0.70	0.10	0.20	0.00	0.30	0.10	0.20	0.00	0.00	0.00	0.00	0.00
-0.90	-0.50	-1.10	0.00	-0.90	-0.50	-1.10	0.00	-0.90	-0.50	-1.00	0.00	-1.50	-0.50	-1.00	0.00	0.00	0.00	0.00	0.00
-1.00	0.00	-0.20	0.10	-1.00	0.00	-0.20	0.10	-1.00	0.00	-0.20	0.10	-1.00	0.00	-0.20	0.10	-1.00	0.00	-0.80	0.10
-0.20	0.40	0.00	0.10	-0.20	0.40	0.00	0.10	-0.20	0.40	0.00	0.10	-0.20	0.40	0.00	0.10	-0.20	0.40	0.00	0.10
-0.50	-0.50	0.60	0.00	-0.50	-0.50	0.60	0.00	-0.50	-0.50	0.60	0.00	-0.50	-0.50	0.60	0.00	-0.50	0.50	0.60	0.00
0.10	0.50	0.10	-0.20	0.10	0.50	0.10	-0.20	0.10	0.50	0.10	-0.20	0.10	0.50	0.10	-0.20	0.10	0.50	0.10	-0.20
0.10	0.30	-0.30	0.20	0.10	0.30	-0.30	0.20	0.10	0.30	-0.30	0.20	0.10	0.30	-0.30	0.20	0.10	-0.80	-4.00	0.20
-0.90	0.40	0.30	-0.50	-0.90	0.40	0.30	-0.50	-0.90	0.40	0.30	-0.50	-0.90	0.40	0.30	-0.50	-0.20	0.50	0.30	0.30

[1] On peut voir facilement que les DGP (1), (2) et (5) sont respectivement de rang $m_l = 4, 3, 0$. En revanche, il est beaucoup moins évident à première vue de noter que les DGP (3) et (4) sont de rang $m_l = 2$ et 1 : cela nécessite de remarquer que les colonnes α_Y de ces deux DGPs ne sont pas linéairement indépendantes, puisqu'elles sont respectivement liées par la relation $C_3 = 2\,C_2$, pour le DGP (3) et par $C_3 = 2\,C_2$, $C_1 = C_2 + C_3$ pour le DGP (4).

TABLEAU 2. – Seuil réel et puissance des tests intermédiaires de l'hypothèse nulle Ho,j (j = 1,..,4) (pourcentage de rejet), avec 10000 réplications au niveau nominal de signification de 5 % [2]

DGPS	DGP (1) : m_1= 4					DGP (2) : m_1 = 3					DGP (3) : m_1 = 2					DGP (4) : m_1 = 1					DGP (5) : m_1 = 0					Hypothèse testée
Taille de l'échantillon	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	
α_1 $W_1 = \xi_1 > A_1$ [3]	100	100	100	100	100	**7.87**	**6.35**	**5.23**	**5.09**	**5.03**	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	$H_{0,1}$: {rang (α_Y) ≤ 3} contre {rang (α_Y) = 4}
α_2 $W_2 = \xi_2 > A_2$	100	100	100	100	100	100	100	100	100	100	**12.4**	**7.22**	**6.12**	**5.27**	**5.11**	0.21	0.00	0.00	0.00	0.00	0.42	0.00	0.00	0.00	0.00	$H_{0,2}$: {rang (α_Y) ≤ 2} contre {rang (α_Y) ≥ 3 }
α_3 $W_3 = \xi_3 > A_3$	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	**17.5**	**9.57**	**7.02**	**5.58**	**5.11**	1.30	1.05	0.40	0.00	0.00	$H_{0,3}$: {rang (α_Y) ≤ 1} contre {rang (α_Y) ≥2 }
α_4 $W_4 = \xi_4 > A_4$	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	**20.4**	**9.86**	**7.12**	**5.72**	**5.19**	$H_{0,4}$: {rang (α_Y) = 0} contre {rang (α_Y) ≥1 }

TABLEAU 3. – Seuil global de la procédure séquentielle de tests (pourcentage de rejet), avec 10000 réplications au niveau nominal de de signification de 5 %

DGPS	DGP (2) : m_1 = 3 P (W_1)					DGP (3) : m_1 = 2 P ($\overline{W}_1$ W_2) [4]					DGP (4) : m_1 = 1 P ($\overline{W}_1$ $\overline{W}_2$ W_3)					DGP (5) : m_1 = 0 P ($\overline{W}_1$ $\overline{W}_2$ $\overline{W}_3$ W_4)				
Taille de l'échantillon T	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000
m_1 estimé = m_1	7.44	6.23	5.19	5.05	5.01	12.8	7.14	5.99	5.20	5.05	18.6	9.74	6.91	5.53	5.13	20.9	11.2	7.09	5.66	5.18

[2] La version corrigée de ces tests a été utilisée pour T = 50, 100.

[3] A_i, i = 1,..,4 désignent respectivement la valeur critique d'une loi du chi-deux au niveau de signification de 5%.

[4] P ($\overline{W}_1$ W_2) indique la probabilité d'être à la fois dans la région d'acceptation $\overline{W}_1$ du test 1 et dans la région critique W_2 du test 2.

TABLEAU 4. – Processus de génération des données (DGP) ($n = 11$, $g = 5$, $k = 6$)[1] associés à la procédure de test de la section 5.2

DGP (1) : $r_1 = 4$				DGP (2) : $r_1 = 3$				DGP (3) : $r_1 = 2$				DGP (4) : $r_1 = 1$				DGP (5) : $r_1 = 0$			
Beta				Beta				Beta				Beta				Beta			
1.00	– 1.00	0.20	– 1.00	1.00	– 1.00	0.20	0.00	1.00	– 1.00	– 1.00	0.00	– 1.50	– 0.50	– 1.00	0.00	0.00	0.00	0.00	0.00
– 2.00	4.00	– 4.00	– 0.60	– 2.00	4.00	– 4.00	0.00	– 2.00	4.00	4.00	0.00	6.00	2.00	4.00	0.00	0.00	0.00	0.00	0.00
– 3.00	– 0.90	2.00	0.90	– 3.00	– 0.90	2.00	0.00	– 3.00	– 0.90	– 0.90	0.00	– 1.35	– 0.45	– 0.90	0.00	0.00	0.00	0.00	0.00
– 1.00	1.00	0.80	0.50	– 1.00	1.00	0.80	0.00	– 1.00	1.00	1.00	0.00	1.50	0.50	1.00	0.00	0.00	0.00	0.00	0.00
– 0.90	– 1.00	– 0.20	0.30	– 0.90	– 1.00	– 0.20	0.00	– 0.90	– 1.00	– 1.00	0.00	– 1.50	– 0.50	– 1.00	0.00	0.00	0.00	0.00	0.00
0.20	– 0.10	– 0.20	0.10	0.20	– 0.10	0.20	0.40	0.20	– 0.10	0.20	0.40	0.20	0.10	– 0.20	– 0.40	– 0.20	– 0.10	0.20	– 0.40
0.80	0.20	0.10	0.30	0.80	0.20	0.10	0.30	0.80	0.20	0.10	0.30	0.80	0.20	0.10	0.30	3.00	– 0.90	0.10	0.50
0.10	0.30	– 0.30	0.60	0.10	0.30	– 0.30	0.60	0.10	0.30	– 0.30	0.60	0.10	0.30	– 0.30	0.60	0.10	0.70	– 0.30	0.60
0.20	– 0.50	0.70	– 0.10	0.20	– 0.50	0.70	– 0.10	0.20	– 0.50	0.70	– 0.10	0.20	– 0.50	0.70	– 0.10	0.20	– 0.50	0.70	1.20
0.40	– 0.20	0.50	0.00	0.40	– 0.20	0.50	0.00	0.40	– 0.20	0.50	0.00	0.40	– 0.20	0.50	0.00	0.40	– 0.20	0.50	0.00
0.10	– 0.30	0.40	0.20	0.10	– 0.30	0.40	0.20	0.10	– 0.30	0.40	0.20	0.10	– 0.30	0.40	0.20	0.10	– 0.30	– 2.00	0.20
alpha				alpha				alpha				alpha				alpha			
0.00	0.50	0.00	0.60	0.90	0.50	0.00	0.60	– 0.1	0.50	– 0.10	0.60	0.00	0.50	– 0.10	0.60	0.00	0.50	0.00	0.60
0.10	– 0.30	– 0.10	– 0.20	0.10	– 0.30	– 0.10	– 0.20	0.10	– 0.30	0.00	– 0.20	0.10	– 0.30	0.00	– 0.20	0.10	– 0.30	– 0.10	– 0.20
– 0.10	0.00	0.00	– 0.20	– 0.10	0.00	0.00	– 0.20	– 0.10	0.00	0.00	– 0.20	– 0.10	0.00	0.00	– 0.20	– 0.10	0.00	0.00	– 0.20
0.20	– 0.20	– 0.50	0.60	0.20	– 0.20	– 0.50	0.60	0.20	– 0.20	– 0.50	0.60	0.20	– 0.20	– 0.50	0.60	0.20	– 0.20	– 0.50	0.60
0.50	– 0.40	0.20	0.00	0.50	– 0.40	0.20	0.00	0.50	– 0.40	0.20	0.00	0.50	– 0.40	0.20	0.00	0.50	– 0.40	0.20	0.00
– 1.00	0.00	– 0.80	0.10	– 1.00	0.00	– 0.80	0.10	– 1.00	0.00	– 0.80	0.10	– 1.00	0.00	– 0.80	0.10	– 1.00	0.00	– 0.80	0.10
– 0.20	0.40	0.00	0.10	– 0.20	0.40	0.00	0.10	– 0.20	0.40	0.00	0.10	– 0.20	0.40	0.00	0.10	– 0.20	0.40	0.00	0.10
– 0.50	0.50	0.60	0.00	– 0.50	0.50	0.60	0.00	– 0.50	0.20	0.60	0.00	– 0.50	0.50	0.60	0.00	– 0.50	0.50	0.60	0.00
0.10	0.50	0.10	– 0.20	0.10	0.50	0.10	– 0.20	0.10	0.50	0.10	– 0.20	0.10	0.50	0.10	– 0.20	0.10	0.50	0.10	– 0.20
0.10	– 0.80	– 4.00	0.20	0.10	– 0.80	– 4.00	0.20	0.10	– 0.80	– 4.00	0.20	0.10	– 0.80	– 4.00	0.20	0.10	– 0.80	– 4.00	0.20
– 0.20	0.50	0.30	0.60	– 0.20	0.50	0.30	0.30	– 0.20	0.50	0.30	0.30	– 0.20	0.50	0.30	0.30	– 0.20	0.50	0.30	0.30

[1] On peut voir facilement que les DGP (1), (2) et (5) sont respectivement de rang $r_1 = 4, 3, 0$. En revanche, il est beaucoup moins évident à première vue de noter que les DGP (3) et (4) sont de rang $r_1 = 2$ et 1 : cela nécessite de remarquer que les colonnes β_Y de ces deux DGPs ne sont pas linéairement indépendantes, puisqu'elles sont respectivement liées par la relation $C_2 = C_3$, pour le DGP (3) et par $C_3 = 2\,C_2$, $C_1 = C_2 + C_3$ pour le DGP (4).

TABLEAU 5. – Seuil réel et puissance des tests intermédiaires de l'hypothèse nulle Ho,j (j = 1,..,4) (pourcentage de rejet), avec 10000 réplications au niveau nominal de signification de 5 %[2]

DGPS	DGP (1) : $r_1 = 4$					DGP (2) : $r_1 = 3$					DGP (3) : $r_1 = 2$					DGP (4) : $r_1 = 1$					DGP (5) : $r_1 = 0$					Hypothèse testée
Taille de l'échantillon	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	
α_1 $W_1 = \xi_1 > A_1$ [3]	100	100	100	100	100	**7.37**	**6.16**	**5.14**	**5.03**	**4.99**	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	$H_{0,1}$: {rang}(β_Y) ≤ 3} contre {rang (β_Y) = 4}
α_2 $W_2 = \xi_2 > A_2$	100	100	100	100	100	100	100	100	100	100	**12.3**	**7.04**	**5.95**	**5.19**	**5.03**	0.21	0.00	0.00	0.00	0.00	0.42	0.00	0.00	0.00	0.00	$H_{0,2}$: {rang}(β_Y) ≤ 2} contre {rang (β_Y) ≥ 3 }
α_3 $W_3 = \xi_3 > A_3$	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	**17.3**	**9.42**	**6.82**	**5.47**	**5.08**	1.30	1.05	0.40	0.00	0.00	$H_{0,3}$: {rang}(β_Y) ≤ 1 } contre {rang (β_Y) ≥ 2 }
α_4 $W_4 = \xi_4 > A_4$	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	**19.5**	**9.68**	**6.96**	**5.62**	**5.15**	$H_{0,4}$: {rang}(β_Y) = 0 } contre {rang (β_Y) ≥ 1 }

TABLEAU 6. – Seuil global de la procédure séquentielle de tests (pourcentage de rejet), avec 10000 réplications au niveau nominal de signification de 5 %

DGPS	DGP (2) : $r_1 = 3$ P (W_1)					DGP (3) : $r_1 = 2$ P ($\overline{W}_1$ W_2) [4]					DGP (4) : $r_1 = 1$ P ($\overline{W}_1$ $\overline{W}_2$ W_3)					DGP (5) : $r_1 = 0$ P ($\overline{W}_1$ $\overline{W}_2$ $\overline{W}_3$ W_4)				
Taille de l'échantillon T	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000
r_1 estimé = r_1	7.37	6.16	5.14	5.03	4.99	12.5	7.05	5.95	5.19	5.03	17.9	9.43	6.82	5.47	5.08	20.4	10.6	6.99	5.62	5.15

[2] La version corrigée de ces tests a été utilisée pour T = 50, 100.

[3] A_i, i = 1,..,4 désignent respectivement la valeur critique d'une loi du chi-deux au niveau de signification de 5%.

[4] P ($\overline{W}_1$ W_2) indique la probabilité d'être à la fois dans la région d'acceptation $\overline{W}_1$ du test 1 et dans la région critique W_2 du test 2.

Références

[1] ANDERSON T.W. (1984), « *An introduction to multivariate statistical analysis* », 2nd edn. (John Wiley & Sons, New York).

[2] ANDO A. et MODIGLIANI F. (1969), « Econometric evaluation of stabilisation policies », *American Economic Review*, n°59, pp 296-314.

[3] ASTLEY M. et GARRAT A. (1996), « Interpreting Sterling Exchange Rate movements », *Bank of England Quarterly Bulletin*, n°36, pp 394-404.

[4] Banque de France, Cepremap, Erasme, Insee, Ofce (1996), « Structures et propriétés de cinq modèles macroéconométriques français », *document de travail*, n° *G96601 de l'INSEE*, juin.

[5] BLANCHARD O.J. et QUAH D. (1989), « The dynamic effects of aggregate demand and supply disturbances », *American Economic Review*, n°79, pp 655-673.

[6] BOSWIJK P. (1992), « Cointegration, identification and exogeneity », *Thesis* (Tinbergen Institute, Amsterdam).

[7] BOX G.E.P. et JENKINS G.M. (1976), « *Time series analysis : forecasting and control* », second edition, Holden-Day, San Francisco.

[8] BRUNEAU C. (1996), « Analyse économétrique de la causalité », *thèse de doctorat*, université Paris IX.

[9] CAMPBELL J. et PERRON P. (1991), « Pitfalls and opportunities : what macroeconomists should know about unit roots », *NBER technical Working Paper 100*, Cambridge.

[10] CHRISTIANO L.J., EICHENBAUM M. et EVANS C.L. (1998), « Modeling Money ? » NBER Working Paper n° 3916, 98/1.

[11] CLARIDA R. et GALI J. (1994), « Sources of real exchange rate fluctuations : how important are nominal shocks ? », Carneige-Rochester Series on Public Policy, n°41, pp 1-56.

[12] CLEMENTS M.P. et MIZON G.E. (1991), « Empirical analysis of macroeconomic time series : VAR and structural models », *European Economic Review*, vol 35, pp 887-932 (with comments).

[13] COOLEY T. et LEROY S. (1985), « A theoretical macroeconometrics : a critique », *Journal of Monetary Economics*, vol 16, pp 283-308.

[14] COOLEY T. (1995), « *Frontiers of Business Cycle Research*. Princeton » (ed.) : Princeton University Press.

[15] DOAN T., LITTERMAN R. et SIMS C. (1984), « Forecasting and conditional projections using realistic priori distributions », *Econometric Reviews*, vol 3, pp 1-100.

[16] ENGLE R.F., HENDRY D.F. et RICHARD J.F. (1983), « Exogeneity », *Econometrica*, vol 51, pp 277-304.

[17] ENGLE R.F. et GRANGER C.W.J. (1987), « Cointegration and Error-Correction : Representation, Estimation and Testing », *Econometrica*, vol 55, pp 251-76.

[18] ERICSSON N.R., HENDRY D.F., MIZON G.E. (1998), « Exogenity, cointegration, and economic policy analysis : an overview », *Journal of Business and Economic Statistics*, vol 6, n°4, pp 370-387.

[19] FLORENS J.P., MARIMOUTOU V. et PEGUIN-FEISOLLE A. (2004), « *Econométrie, modélisation et inférence* », A. Colin, Paris.

[20] FLORENS J.P et MOUCHART M. (1985), « Conditioning in Dynamic Models », *Journal of Time Series Analysis*, vol 6, n° 1, pp 15-34.

[21] Florens J.P., Mouchart M., Richard J.F. (1979), « Specification and inference in linear models », *CORE Discussion Paper* n ° 7943, Université Catholique de Louvain.

[22] Giannini C. et Mosconi R. (1992), « Non-causality in cointegrated systems : representation estimation and testing », *Oxford Bulletin of Economics and Statistics*, vol 54, n°3, pp 399-417.

[23] Gonzalo J. (1994), « Five alternative methods of estimating long run equilibrium relationships », *Journal of Econometrics*, vol 60, pp 203-233.

[24] Granger C.W.J. (1969), « Investigating Causal Relations by Econometric Model and Cross-Spectral Methods », *Econometrica*, n°37, pp 424-438.

[25] Granger C.W.J. et Haldrup N. (1997), « Separation in cointegrated systems and P-T decompositions », *Oxford Bulletin of Economics and Statistics*, vol 59, n° 2, pp 449-463.

[26] Haug A.A. (1996), « Tests for cointegration : A monte-carlo comparison », *Journal of Econometrics,* vol 71, pp 89-115.

[27] Hecq A., Palm F.C., Urbain J.P. (2001), « Separation, Weak Exogeneity and P-T Decomposition in Cointegrated VAR Systems with Common Features », *Econometric Reviews,* vol 21, pp 273-307.

[28] Hendry D.F., Neale A.J. et Sbra F. (1988), « Econometric Analysis of Small Linear Systems using PC-FIML », *Journal of Econometrics*, vol 38, pp 203-226.

[29] Hendry D.F. et Mizon G.E. (1993), « Evaluating dynamic models by encompassing the VAR », in P.C.B. Phillips, ed., *Models, methods, and applications of econometrics* (Basil, Blackwell, Oxford), pp 272-300.

[30] Hendry D.F. (1983), « Econometric Modelling : The consumption function in retrospect », *Scottish Journal of Political Economy*, vol 30, pp 193-220.

[31] Hendry D.F. (1995), « *Dynamic Econometrics* », Oxford University Press.

[32] Hunter J. (1992), « Tests of cointegrating exogeneity for ppp and uncovered interest rate parity in the United Kingdon », *Journal of Policy Modeling*, vol 4, n° 14, pp 453-463.

[33] Johansen S. (1991), « Estimation and hypothesis testing of co-integration vectors in gaussian vectors autoregressive models », *Econometrica*, vol 6, pp 1551-1580.

[34] Johansen S. (1992a), « Determination of cointegration rank in the presence of a linear trend », *Oxford Bulletin of Economics and Statistics,* vol 54, pp 383-397.

[35] Johansen S. (1992b), « Cointegration in partial systems and the efficiency of single-equation analysis », *Journal of Econometrics,* vol 52, pp 389-402.

[36] Johansen S. (1995a), « Identifying restrictions of linear equations with applications to simultaneous equations and cointegration », *Journal of Econometrics*, vol 69, pp 111-132.

[37] Johansen S. (1995b), « *Likelihood-based inference in co-integrated vector autoregressive models* », Oxford University Press.

[38] Johansen S. et Juselius K. (1992), « Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for UK », *Journal of Econometrics*, vol 53, pp 211-244.

[39] Kim K. et Pagan A.R. (1995), « *The Econometric analysis of calibrated macroeconornic models* », Chapter 7 in Handbook of Applied Econometrics : Macroeconomics, ed. by M.H. Pesaran and M. Wickens. Oxford : Basil Blackwell.

[40] Konishi T., Granger C.W.J. (1993), « Separation in cointegrated system », *manuscript UCSD.*

[41] Kydland F. et Prescott E. (1982), « Time to build and aggregate fluctuations », *Econometrica*, vol 50, pp 1345-1370.

[42] Litterman R. (1986), « Forecasting with bayesian vector autoregressions – five years of experience », *Journal of Business and Economics Statistics,* vol 4, pp 25-38.

[43] Long J.B. et Plosser C. (1983), « Real Business Cycles », *Journal of Political Economy*, vol 91, pp 39-69.

[44] Lucas R. (1976), « *Econometric policy evaluation : a critique* », in : K. Brunner and A.H. Metzler (eds) The Phillips Curve and Labour Markets, Carnegie- Rochester Conference Series on Public Policy, vol 2, pp 19-46 (conference supplement), North Holland.

[45] Monfort A. et Rabemananjara R. (1990), « From a VAR model to a structural model, with an application to the wage-price spiral », *Journal of Econometrics*, vol 5, pp 203-227.

[46] Mosconi R.(1998), « MALCOLM : the Theory and Practice of Cointegration Analysis in RATS ». Venice : Ca Foscarina. *Available on-line at http//www.greta.it/malcolm*,

[47] Nelson C.R. et Plosser C.I. (1982), « Trends and random walks in macroeconomic time series », *Journal of Monetary Economics*, vol 10, pp 139-162.

[48] Pagan A. (1987), « Three econometric methodologies : a critical appraisal », *Journal of Economic Surveys*, vol 1, pp 3-24.

[49] Phillips P.C.B. (1991), « Optimal Inference in Cointegrated systems », *Econometrica*, vol 59, pp 283-306.

[50] Phillips P.C.B. et Loretan M. (1991), « Estimating long run equilibria », *Review of Economic Studies*, vol 58, pp 407-436.

[51] Pradel J. et Rault C. (2003), « Exogeneity in VAR-ECM models with purely exogenous long-run paths », *Oxford Bulletin of Economics and Statistics*, vol 65, n°5, pp 629-653.

[52] Psaradakis Z. (1994), « A comparison of tests of linear hypotheses in cointegrated vector autoregressive models », *Economic Letters* , vol 45, pp 137-144.

[53] Rault C. (1997), « Prédétermination, causalité, exogénéité dans un modèle vectoriel à correction d'erreurs : identifiabilité d'une forme structurelle », *Cahiers Eco-maths*, Université Paris I Panthéon Sorbonne, n°97. 60.

[54] Rault C. (2000), « Non-causality in VAR-ECM models with purely exogenous long run paths », *Economics Letters*, n°67-2.

[55] Rault C. (2005), « Further results on weak-exogeneity in vector error correction models », *Brazilian Review of Econometrics*, vol 25, n°2, pp 159-172.

[56] Sims C. (1980), « Macroeconomics and Reality », *Econometrica*, vol 48, pp 1-48.

[57] Sims C. (1991), « Comments on Clements, M.P. et Mizon G.E. [1991] paper », *European Economic Review*, vol 35, pp 922-932.

[58] Toda H.Y., Phillips P.C.B (1991), « Vector autoregressions and causality », *Cowles Foundation Discussion Paper*, n°9 77.

[59] Toda H.Y., Phillips P.C.B (1993), « Vector autoregressions and causality », *Econometrica*, vol 61, n°6, pp 1367-1393.

[60] Urbain J.P. (1992), « On weak exogeneity in error correction models », *Oxford Bulletin of Economics and Statistics*, vol 54, n° 2, pp 187-207.

[61] URBAIN J.P. (1995), « Partial versus full system modelling of cointegrated systems : an empirical illustration », *Journal of Econometrics*, vol 69, pp 177-210.

[62] YUM Y.S. (1994), « Tests de causalité dans les modèles macro-économétriques dynamiques », *thèse de doctorat*, université Paris I Panthéon Sorbonne.