



UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES
SIDI BEL-ABBÈS

BP 89 SBA 22000 –ALGERIE–

TEL/FAX 048-54-43-44

THESE

Présentée par:
HAMLILI ABDELKADER

Pour obtenir le Diplôme de Doctorat

Spécialité : Mathématiques
Option : Probabilités-Statistiques

Intitulée

**TEST OF CAUSALITY BETWEEN THE ECONOMICS
INDEXES IN ALGERIA**

Thèse soutenue le
Devant le jury composé de :

Président :

GHERRIBALLAH Abdelkader *Professeur* à L'Université S.B.A

Directeur de thèse :

CHIKRELMEZOUAR ZOUAOUI *Maître de Conférence A* à L'Université de
Béchar

Co-Directeur de thèse :

Attouch Mohammed Kadi *Professeur* à L'Université de S.B.A

Examineurs :

MECHAB Boubaker *Maître de Conférence A* à L'Université de S.B.A

KANDOUCI Abdeljebbar *Professeur* à L'Université de Saida

GUENDOUDI Toufik *Professeur* à L'Université de Saida

Résumé

Dans cette thèse , nous sommes intéressés essentiellement à l'étude de la relation de causalité entre deux séries chronologiques (stationnarité ,test de racine unitaire ,cointégration et VECM)

Dans un premier temps , nous considérons deux séries $X(t), Y(t)$ et on les étudie étape par étape selon le plan de la causalité au sens de Granger. ie stat, test de racine, VECM ainsi que la causalité

Dans un second temps , nous apportons une contribution au sujet, en faisant introduire la simulation pour bien motiver et illustrer nos démarches et enfin nous aboutissons à la conclusion de cette thèse.

Remerciements

Je tiens à exprimer toute ma reconnaissance à mon directeur de thèse monsieur Chikr EL Mezouar. Je le remercie de m'avoir encadré orienté et conseillé.

Je tiens à remercier sincèrement monsieur le professeur Ghariballah Abdelkader qui, s'est toujours montré à l'écoute et très disponible tout au long de réalisation de thèse, ainsi pour l'inspiration , l'aide et le temps qu'il a bien voulu me consacrer.

Mes remerciements vont aussi à messieurs qui me font l'honneur d'être parmi les membres du jury.

Je remercie particulièrement ma femme et mes enfants, qui ont su croire en moi et qui m'ont apporté toute leur aide quand j'en ai eu besoin.

Enfin, Mes remerciements les plus sincères à toutes les personnes qui ont contribué soit de près ou de loin à l'élaboration de cette thèse.

Table des matières

1	Introduction	6
1.1	Présentation	6
1.2	Définitions et notations	6
1.2.1	Causalité à la croisée des chemins	6
1.2.2	Les théories probabilistes de la causalité	7
1.2.3	Causalité générique et causalité singulière	8
1.2.4	Contexte bibliographique	8
2	Tests de stationnarité	13
2.1	Introduction	13
2.2	Tests de stationnarité	14
2.2.1	Les processus stochastiques	14
2.2.2	Les processus aléatoires stationnaires	15
2.2.3	Stationnarité faible ou large d'un processus Stochastique (au second ordre)	15
2.2.4	Les processus aléatoires non stationnaires	16
2.2.5	Processus DS	17

2.3	Tests de racine unitaire	18
2.3.1	Test de Dickey Fuller simple (DFS)	18
2.3.2	Test de Dickey Fuller Augmenté (ADF)	25
2.3.3	Test de Phillips et Perron(PP)	35
2.3.4	Test Kwiatkowski, Phillips, Schmidt et Shin (KPSS)	41
2.4	Les modèles VAR	43
2.4.1	Représentation d'un modèle VAR	44
2.4.2	Stationnarité d'un modèle VAR	44
2.4.3	Détermination du nombre de décalage	46
2.4.4	Décomposition de la variance et fonction de réponse impulsionnelle	46
3	Tests de cointegration	48
3.1	Définition de la cointegration	49
3.1.1	Conditions de la cointegration	49
3.2	La méthode de Engle–Granger	50
3.2.1	Pré-test chaque variable	51
3.2.2	Estimation de modèle à correction d'erreur	54
3.2.3	Limites de la méthode Engle-Granger	60
3.3	Méthode de Phillips-Ouliaris	63
3.3.1	Le test du rapport de la variance	64
3.3.2	La trace statistique multivariée	67
3.3.3	Limitation des méthodes Phillips-Ouliaris	68
3.4	La procédure de Johansen	69

3.4.1	Le test de la trace	72
3.4.2	La valeur propre maximale	73
3.4.3	Limitation de la procédure de Johansen	77
4	Causalité au sens de Granger(Cas indices économique Algeria)	79
4.1	Introduction	80
4.2	Methodolgy	80
4.3	Causality test and cointegration variables.	81
4.3.1	Unit root tests.	81
4.4	Cointegration.	82
4.5	An error correction model.	82
4.6	Causality test.	83
4.7	Empirical results and interpretation	83
4.7.1	Statistical data proprieties.	83
4.8	Unit root tests	84
4.9	Johansen cointegration tests.	87
4.10	An error correction model estimate	87
4.11	Granger causality test	87
4.12	Pairwise Granger causality test	87
4.13	Conclusion	88
4.14	Conclusion	89
	Bibliographie	91

Chapitre 1

Introduction

1.1 Présentation

Cette thèse est consacré à l'étude des tests de causalité appliqués à des indices économétriques. Plus précisément, nous proposons d'étudier le Problème de la causalité au sens de Granger a des indices d'économétrie en Algeria. Il s'agit d'un sujet eoutemporain noté par l'importance de l'utilisation de la notion de cointégration aux séries chronologiques comme outil de prévision permettant aux économistes d'analyser correctement l'impact de la causalité au secteur économique.

1.2 Définitions et notations

1.2.1 Causalité à la croisée des chemins

La notion de causalité existe dans plusieurs domaines de la science

- En philosophie ; la causalité est liée à la notion de cause et effet.

- En physique, la causalité est un postulat, ou principe physique fondamental.
- En biologie et en écologie, la causalité est aussi l'un des postulats de base, au sein de systèmes complexes.
- En sciences sociales, la causalité est la recherche des facteurs ayant amené tel ou tel phénomène social.
- En histoire, la causalité cherche à expliquer en profondeur les causes des événements historiques.
- En droit français, le lien de causalité analyse plus précisément la responsabilité extra contractuelle.
- Généralement, la causalité est un rapport d'une cause à son effet.

1.2.2 Les théories probabilistes de la causalité

Il est difficile de s'orienter au sein du champ des tentatives d'explication de ce qu'est la causalité. L'introduction de concepts probabilistes dans l'analyse de la causalité rompt l'attachement séculaire de la causalité à la necessitation. La thèse de David Hume selon laquelle une cause est invariablement suivie de son effet, n'est pas satisfaisante : la plupart des causes ne suffisent pas à produire leurs effets, mais ne les produisent qu'en présence de certains facteurs, qui sont appelées "causes" sous la plupart des analyses de la causalité.

Les analyses régularisées supposent qu'il n'y a pas d'effet sans un ensemble de facteurs qui suffit à la produire, et donc qu'il suit régulièrement. Corrélativement, il n'y a de cause qu'appartenant à un ensemble de facteurs qui suffit à produire son effet.

En revanche, les théories probabilistes de la causalité sont compatibles avec cette hypothèse.

En effet, elles reposent sur la proposition de caractériser une cause par ceci qu'elle rend son

effet plus probable. Plus précisément, une cause A augmente la probabilité de son effet B au sens où la probabilité conditionnelle $p(B|A)$ est plus élevée que la probabilité absolue $p(B)$. L'idée qu'on trouve au fondement des théories probabilistes de la causalité est donc qu'il peut exister des effets sans ensemble de facteurs qui suffit à les produire.

1.2.3 Causalité générique et causalité singulière

La causalité générique et la causalité singulière sont généralement désignées comme des niveaux de causalité. La causalité générique est une relation entre propriétés, alors que la causalité singulière est une relation entre événements singuliers qui sont effectivement advenus. Par exemple : les énoncés : "fumer cause le cancer" et "le tabagisme de mon amis a causé son cancer du poumon" ainsi, la causalité générique est une relation entre la propriété de fumer et la propriété de souffrir d'une maladie, de l'autre coté, la causalité singulière est une relation entre événements qui sont déjà advenus. Les théories probabilistes de la causalité se présentent en premier lieu comme des théories de la causalité générique. La causalité singulière se prête moins bien que la causalité générique à une analyse fondée sur la notion d'augmentation de probabilités conditionnelles.

1.2.4 Contexte bibliographique

Historiquement, dans la littérature concernant notre champ d'étude, on distingue trois générations de modèles (Guttormsen, 2004). Les études de la première génération sont basées sur le traditionnel modèle VAR de Sims et le test causalité classique de Granger ; elles

supposent que les séries sont stationnaires. La deuxième et la troisième génération d'études traitent des séries non stationnaires et la cointégration apparaît comme la technique la plus appropriée. Dans la deuxième, la cointégration entre deux variables une fois testée, on estime le modèle à correction d'erreurs et on calcule le test de causalité de Granger. Dans la troisième génération d'études, on utilise l'approche multivariée de la cointégration basée sur la méthode du maximum de vraisemblance (Johansen, 1988).

Le travail pionnier de la première génération est celui de Kraft et Kraft (1978). Utilisant la technique de Sims, ces deux auteurs ont trouvé une causalité unidirectionnelle entre le PIB et la consommation d'énergie aux USA sur la période 1947-1974. Le papier de Kraft et Kraft a été critiqué par Akarca et Long (1980). Ces derniers ont noté que la période choisie était instable, car elle incluait le premier choc pétrolier. Ils ont souligné que les résultats obtenus ne seraient pas les mêmes si cette période était écourtée de deux ans. Ils ont donc repris l'analyse avec la même technique, sur une période plus homogène allant de 1950 à 1968. Le test a révélé le manque de causalité entre le PIB et la consommation d'énergie. Pratiquement, tous les articles qui ont suivi ont été consacrés aux séries américaines avec des résultats très variés (cf. par exemple Yu et Hwang (1984), Yu et Choi (1985)).

Le premier texte de la deuxième génération semble être celui de Nachane, Nadkani et Karnik (1988). Adoptant l'approche de Engle et Granger de la cointégration, ils ont trouvé une relation de long terme entre la consommation d'énergie et la croissance économique pour onze pays en développement et cinq pays développés. La méthodologie d'Engle et Granger a été ensuite appliquée par de nombreux auteurs à différents pays avec des résultats parfois ambigus. L'étude de H. Yang (2000) sur la province chinoise de Taiwan a conduit à une causalité bidirectionnelle entre croissance et consommation d'énergie sur la période 1954-1997. L'analyse

de A. A. Quell et M. Butt (2001) en ce qui concerne le Pakistan, en intégrant l'emploi comme variable additionnelle et en appliquant une version du test de Granger proposé par Hsiao, infère que la croissance économique cause la consommation totale d'énergie.

Masih et Masih (1996) sont parmi les premiers auteurs à utiliser la méthodologie de Johansen dans la recherche des relations entre énergie et croissance. Ainsi, dans une série d'articles concernant six pays asiatiques (L'Inde, le Pakistan, l'Indonésie, la Malaisie, Singapour et les Philippines), leurs études ont abouti aux conclusions suivantes :

qu'il existe une relation de long terme entre les deux variables dans le cas de l'Inde, le Pakistan et l'Indonésie ; - Que la consommation d'énergie n'cause pas le PIB en Inde ;

- Que le PIB n'cause pas la consommation d'énergie en Indonésie ;

- Qu'il existe une causalité bidirectionnelle entre la consommation d'énergie et le PIB au Pakistan ;

- Que pour les trois pays restants (Malaisie, Singapour et Philippines), l'utilisation d'un VAR ordinaire a révélé l'inexistence de relation causale entre le PIB et la consommation d'énergie.

Le cas de la Turquie a été testé sur la période 1960-1995 par U. Stoytas et alii (2001). Le résultat indique une causalité unidirectionnelle de la consommation d'énergie vers le PIB : la consommation d'énergie affecte positivement le PIB ; et cela suggère qu'il est possible, qu'à long terme, le programme d'économie de l'énergie puisse influencer la croissance économique.

La relation entre l'énergie et la croissance économique a été étudiée, toujours dans le cadre d'un modèle multivarié incluant le capital, le travail, l'énergie et le PIB, par W. Ok et K. Lee (2003) pour la Corée sur la période 1970- 1999. Le modèle à correction d'erreur indique une causalité bidirectionnelle de long terme entre l'énergie et la croissance économique, et unidirectionnelle de court terme de l'énergie vers la croissance économique. Depuis les travaux de ce

type abondent combinant cointégration, modèle à correction d'erreurs et causalité. On pourra consulter A. Guttormsen (op. cité) et P. Mozumber (2005) pour une revue assez complète sur ces travaux. Enfin, signalons deux études concernant l'Afrique subsaharienne. La première est celle de O. Ebohon (1996) sur la Tanzanie et le Nigeria. Utilisant le test classique de Granger, cet auteur trouve une causalité bidirectionnelle entre la croissance économique et la consommation d'énergie pour ces deux pays. La deuxième étude plus récente concerne le Malawi et a été réalisée par C.Jumbe (2004). S'appuyant sur la méthodologie de Engle et Granger de la cointégration et la causalité au sens de Granger, son analyse a abouti à la conclusion selon laquelle, d'une part, qu'il y a une causalité bidirectionnelle entre les consommations d'électricité et le PIB et d'autre part, qu'il existe une causalité unidirectionnelle du PIB non agricole vers les consommations d'électricité. Cela étant dit, dans la suite de ce texte, nous présentons dans la première partie, la méthodologie adoptée. On précise dans un premier temps, la notion de causalité utilisée. Elle repose sur la définition de Granger qui considère qu'une variable est causée par une autre dès lors qu'il existe des informations dans le passé de l'une qui soient utiles dans la prévision de l'autre, et qui ne sont pas déjà contenues dans son passé. Loin d'être exhaustive, cette définition est donc une étape essentielle d'une étude statistique. Elle s'inscrit dans le cadre théorique complet qui donne corps à la définition de la causalité au sens de Feigl (Bruneau et Nicolai, 1989) : *confirmed predictability according to law or set of laws*. Dans un deuxième temps, la causalité est étudiée dans le cadre des variables cointégrées en optant pour une approche en trois étapes (Stoytas, Sari et Ozdemir, 2001). Dans la première étape, on vérifie la stationnarité des séries, ainsi que leur ordre d'intégration à l'aide des tests de racine unitaire de Dickey-Fuller (1979) et Phillips-Perron (1988). Cela est nécessaire parce que d'une part, les tests de causalité sont très sensibles à la stationnarité des séries (Stock et

Watson, 1989) et d'autre part, il a été constaté que la plupart des séries macro économiques ne sont pas stationnaires (Nelson et Plosser, 1982). Dans l'étape suivante, on introduit la théorie de la cointégration qui est en fait la version multivariée du concept de racine unitaire. Celle-ci permet de spécifier les relations stables à long terme tout en analysant conjointement la dynamique de court terme des variables considérées. Dans la troisième et dernière étape, on décrit très brièvement le modèle à correction d'erreur qui, selon Engle et Granger, permet de représenter les séries cointégrées : c'est un mécanisme qui force la déviation de court terme par rapport à l'équilibre à une période donnée à revenir à la période suivante.

Enfin, on boucle cet exposé méthodologique par la présentation du test de Granger dans le cadre d'un modèle à correction d'erreur. L'objectif essentiel visé est celui de savoir si les deux séries étudiées sont dynamiquement interdépendantes ou si au contraire la liaison dynamique est unidirectionnelle. La deuxième partie du texte est consacrée aux résultats empiriques obtenus dans le cas l'Algérie suivant la méthodologie adoptée ci-dessus.

Chapitre 2

Tests de stationnarité

2.1 Introduction

Le développement des techniques économétriques permettant de traiter des séries non stationnaires a apporté des solutions rigoureuses à ce problème. Nous présenterons dans une première partie de ce chapitre les tests de non-stationnarité qui sont maintenant très largement utilisés dans la littérature empirique en distinguant les procédures qui postulent que la composante déterministe de la série considérée suit une tendance linéaire.

Cette distinction est primordiale car la mauvaise spécification de la composante déterministe peut nous conduire à ne pas rejeter à tort l'hypothèse de la racine unitaire. Nous prendrons soin de bien préciser les différents processus engendrant les données et les différents modèles de régression.

Nous expliciterons en particulier les relations qu'entretiennent les paramètres des processus engendrant les données et les paramètres des modèles de régression utilisés pour effectuer les tests ainsi que les implications de l'hypothèse de la racine unitaire sur ces paramètres.

Ces différents points sont d'une importance capitale dans l'élaboration d'une stratégie de test. Nous évoquerons également le problème de la sélection optimale du paramètre de troncature dans la procédure de test de Dickey et Fuller augmentée (1981) aussi bien que dans la procédure de test de Phillips et Perron (1988).

Dans une deuxième partie nous présenterons une discussion à ces stratégies et quelle est le test qui'il contient le meilleur caractéristiques .

2.2 Tests de stationarité

2.2.1 Les processus stochastiques

Définition 2.2.1. *Un processus aléatoire est une suite de variables aléatoires indexées dans le temps et définies sur un espace des états de la nature. Ainsi, pour chaque instant du temps, la valeur de la quantité étudiée X , est appelée variable aléatoire et l'ensemble des valeurs X , quand t varie est appelé processus aléatoire.*

Remarque 2.2.1. *La classe des processus aléatoires est très large, l'analyse des séries temporelles a donc été centrée sur une classe particulière de processus :*

-Les processus aléatoires stationnaires : ces processus sont caractérisés par le fait que leurs propriétés statiques sont stables dans le temps.

-Les processus aléatoires non stationnaires : ces processus sont caractérisés par l'instabilité de leurs propriétés statiques dans le temps.

2.2.2 Les processus aléatoires stationnaires

Il existe deux types de stationnarité :

-Stationnarité stricte d'un processus stochastique

Définition 2.2.2. *Un processus est dit strictement stationnaire si quelque soit le n uplet $t_1 < t_2 < \dots < t_n$ tel que $t_{i+h} \in T$ et pour tout $h \in T, i = 1..n$ la suite $(X_{t_1+h}, \dots, X_{t_n+h})$ a la même probabilité que la suite $(x_{t_1} \dots x_{t_n})$ de*

$$\forall (x_1 \dots x_n), \forall (t_1 \dots t_n) \text{ et } \forall h : p(X_{t_1} < x_1, \dots, X_{t_n} < x_n) = p(X_{t_1+h} < x_1, \dots, X_{t_n+h} < x_n)$$

Ainsi un processus aléatoire est fortement stationnaire si toutes ses caractéristiques, c'est-à-dire tous ses moments, sont invariants pour tout changement de l'origine du temps.

2.2.3 Stationnarité faible ou large d'un processus Stochastique (au second ordre)

Définition 2.2.3. *Un processus $X_{t \in T}$ est dit faiblement stationnaire si seulement les moments d'ordre 1 et d'ordre 2 sont stationnaires en dépit de la non stationnarité éventuelle des autres moments.*

- $\forall t \in Z, E(X_t^2) < \infty$
- $\forall t \in Z, E(X_t^2) = m$ indépendante de t
- $\forall t \in Z, \text{cov}(X_t^2, X_{t+h}^2) = \delta(h)$ indépendante de t

Donc un tel processus admet une loi dont les variables ont la même variance égale $\delta(0)$.

Remarque 2.2.2. - *La stationnarité au sens strict n'implique pas la stationnarité au sens large excepté dans le cas d'existence des deux premiers moments*

Dans la pratique, la vérification de la stationnarité au sens stricte peut être difficile à démontrer. On se restreint à démontrer la stationnarité au sens large.

Par conséquent, une série stationnaire ne comporte ni tendance ni saisonnalité : sa variance si elle existe est constante.

En résumé, un processus est stationnaire au second ordre si l'ensemble de ses moments sont indépendants du temps.

2.2.4 Les processus aléatoires non stationnaires

Selon la terminologie de Nelson et Plosser (1982), il y a deux classes de processus non stationnaires : les processus TS (Trend Stationary) et les processus DS (Differency Stationary).
processus TS : Cette forme de non stationnarité provient de la présence d'une composante déterministe tendancielle $X_t, t \in Z$ est un processus TS s'il peut s'écrire sous la forme $X_t = f(t) + Z_t$ où $f(t)$ est une fonction du temps et Z_t est processus stochastique stationnaire. Le processus X_t s'écrit comme la somme d'une fonction déterministe du temps et d'une composante stochastique stationnaire. Ce processus ne satisfait plus la définition de la stationnarité du second ordre. On a en effet :

$$E(X_t) = f(t) + E(Z_t)$$

Une des propriétés importantes de ce type de processus réside dans l'influence des innovations stochastiques ϵ_t Lorsqu'un processus TS est affecté par un choc stochastique, l'effet de

ce choc tend à disparaître au fur et à mesure que le temps passe : C'est la propriété de non-persistance des chocs. En effet, pour stationnariser un processus TS, il convient de retirer la composante

2.2.5 Processus DS

Cette forme de non stationnarité est de nature stochastique. Un processus non stationnaire $X_t, t \in Z$ est un processus DS (Differency Stationary) d'ordre d , où d désigne l'ordre d'intégration, si le processus filtré défini par : est stationnaire. On dit aussi que est un processus intégré d'ordre d , noté $I(d)$.

Ainsi, on peut définir une classe de processus stochastiques qui ne satisferont pas les conditions de la stationnarité, mais dont la différence à l'ordre d satisfait elle les propriétés de la stationnarité

La définition des processus DS, repose sur la présence de racines unitaires dans le polynôme associé à la dynamique autorégressive du processus

Dans la classe générale des processus DS, un type de processus apparaît de façon régulière, si bien que l'on lui a attribué un nom particulier : La marche aléatoire. (Random Walk).

L'influence d'une innovation* à une date T sur un processus $I(d)$ * est permanente. On a ainsi une propriété de persistance des cocs. En effet, pour stationnariser un processus DS d'ordre d , il convient d'appliquer le filtre $(1 - L)^d$.

2.3 Tests de racine unitaire

Ces tests permettent, tout d'abord de vérifier que les séries sont non stationnaires, et d'autre part de discriminer entre les processus DS et TS. Il existe un grand nombre de tests de racines unitaires. Les pionniers en la matière sont de Dickey et Fuller. Mais ils souffrent d'un certain nombre de critiques. Celles-ci ont conduit à l'élaboration de plusieurs tests, dont ceux de Philips et Perron (1988).

2.3.1 Test de Dickey Fuller simple (DFS)

Le test de Dickey Fuller simple (1979) est un test de racine unitaire (ou de non stationnarité) dont l'hypothèse nulle est la non stationnarité d'un processus autorégressif d'ordre un. Considérons un processus $(x_t, t) \in \mathbb{Z} \times \mathbb{R}$

Le principe général du test de Dickey Fuller consiste à tester l'hypothèse nulle de la présence d'une racine unitaire :

$$H_0 : \rho = 1$$

$$H_1 : |\rho| < 1$$

En effet, sous l'hypothèse nulle H_0 ; le processus (3.92) se ramène à une pure marche aléatoire (Random Walk Process). L'hypothèse nulle testée correspond ainsi à une hypothèse de non stationnarité stochastique. Ce test, comme tout test non symétrique, peut être réalisé de différentes façons. La plus simple consiste à utiliser une statistique de Student associée à l'hypothèse H_0 :

$$t_{\phi=1} = \frac{\phi - 1}{SE(\phi)}$$

tel que ϕ : l'estimateur de paramètre ρ et $SE(\phi)$: standard deviation d'erreur (ecart-type d'erreur).

En cela, le test de Dickey Fuller ne se distingue pas d'un test quelconque d'une hypothèse non symétrique. Là où l'application du test de Dickey Fuller dans la distribution asymptotique de la statistique de Student associée au test H_0 : En effet, nous allons montrer que la distribution asymptotique de l'estimateur des MCO du paramètre ρ ; sous l'hypothèse de non stationnarité ; est non standard. On n'a plus dans ce cas une distribution asymptotique normale, comme en économétrie de base. De la même façon, la statistique de student associée au tests $\rho = 1$; n'a pas une distribution asymptotique standard (distribution de Student approximée par une distribution normale). C'est pourquoi, l'application du test de Dickey Fuller nécessite que l'on utilise des seuils ditraditionnellement pour des statistiques de Student. De plus, puisqu'un malheur n'arrive jamais seul, nous montrerons que la distribution asymptotique de la statistique de Student associée au test H_0 ; n'est pas la même suivant que dans le modèle (3.92) on inclut ou non, une constante et un trend déterministe. Dès lors, puisque a priori, on se sait pas si l'on doit inclure cette constante et ce trend, il convient d'appliquer non pas un test simple, mais une stratégie de tests de Dickey Fuller.

La stratégie test de Dickey Fuller simple (DFS)

Nous allons à présent proposer une stratégie de tests de Dickey Fuller permettant de tester la non stationnarité conditionnellement à la spécification du modèle utilisé. On considère trois modèles

$$\text{Modèle 1 : } \Delta x_t = \phi x_{t-1} + \varepsilon_t$$

$$\text{Modèle 2 : } \Delta x_t = \phi x_{t-1} + c + \varepsilon_t$$

$$\text{Modèle 3 : } \Delta x_t = \phi x_{t-1} + c + \beta t + \varepsilon_t$$

avec ε_t i.i.d $N(0, \sigma_\varepsilon^2)$.

On cherche à tester l'hypothèse de racine unitaire :

$$H_0 : \phi = 0 \qquad H_1 : \phi < 0$$

Remarque 2.3.1. *Le principe général de la stratégie de tests est le suivant. Il s'agit de partir du modèle le plus général, d'appliquer le test de racine unitaire en utilisant les seuils correspondant à ce modèle, puis de vérifier par un test approprié que le modèle retenu était le "bon". En effet, si le modèle n'était pas le "bon", les seuils utilisés pour le test de racine unitaire ne sont pas valables. On risque alors de commettre une erreur de diagnostic quant à la stationnarité de la série. Il convient dans ce cas, de recommencer le test de racine unitaire dans un autre modèle, plus contraint.*

Et ainsi de suite, jusqu'à trouver le "bon" modèle, les "bons" seuils et bien entendu les

"bons" résultats.

Le déroulement de la stratégie de test est reportée sur la figure suivante :

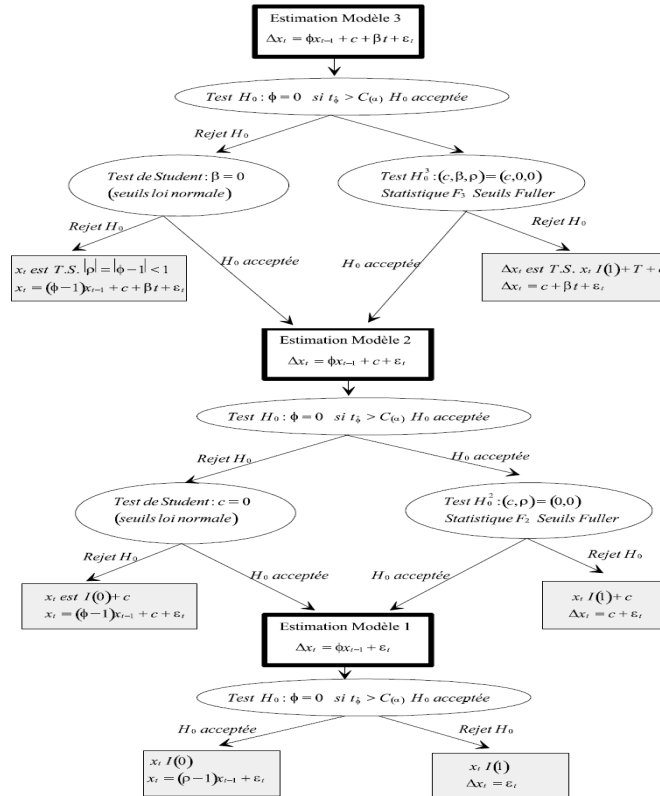


FIGURE 2.1 – la stratégie de test Dickey Fuller

On commence par tester la racine unitaire à partir du modèle le plus général, à savoir le modèle 3.

On compare la réalisation de la statistique de Student $t_{\hat{\phi}=0}$ aux seuils $C_{(\alpha)}^3$ tabulés par Dickey et Fuller, ou McKinnon pour le modèle 3 (par exemple -3.41 à 5% , pour $T \rightarrow \infty$). Si la réalisation de $t_{\hat{\phi}=0}$ est supérieure au seuil $C_{(\alpha)}^3$; on accepte l'hypothèse nulle de non stationnarité. Une fois que le diagnostic est établi, on cherche à vérifier si la spécification du modèle 3, incluant une constante et un trend, était une spécification compatible avec les

données. On teste alors la nullité du coefficient β de la tendance. Deux choses l'une : Soit on a rejeté au préalable l'hypothèse de racine unitaire, dans ce cas on teste la nullité de β par un simple test de Student avec des seuils standards (test symétrique, donc seuil de 1.96 à 5%). Si l'on rejette l'hypothèse $\beta = 0$; cela signifie que le modèle 3 est le "bon" modèle pour tester la racine unitaire, puisque la présence d'une tendance n'est pas rejetée. Dans ce cas, on conclut que la racine unitaire est rejetée, la série est TS ; du fait de la présence de la tendance. En revanche, si l'on accepte l'hypothèse $\beta = 0$, le modèle n'est pas adapté puisque la présence d'une tendance est rejetée. On doit refaire le test de racine unitaire à partir du modèle 2, qui ne comprend qu'une constante.

Soit, au contraire, on avait au préalable, accepté l'hypothèse de racine unitaire, et dans ce cas, on doit construire un test de Fischer de l'hypothèse jointe $\phi = 0$ et $\beta = 0$: On teste ainsi la nullité de la tendance, conditionnellement à la présence d'une racine unitaire :

$$H_0^3 : (c; b; \phi) = (c; 0; 0) \text{ contre } H_1^3$$

La statistique de ce test se construit de façon standard par la relation :

$$F_3 = \frac{(SCR_{3,c} - SCR_3) / 2}{SCR_3 / (T - 3)}$$

où $SCR_{3,c}$ est la somme des carrés des résidus du modèle 3 contraint sous H_0^3 :

$$\Delta x_t = c + \varepsilon_t$$

et SCR_3 est la somme des carrés des résidus du modèle 3 non contraint (équation 3.130). Les

seuils à distance fini de cette statistique sont fournies dans la partie annexe du programme de cours. Si la réalisation de F_3 est supérieure à la valeur ϕ_3 lue dans la table à un seuil $\alpha\%$, on rejette l'hypothèse H_0^3 : Dans ce cas, le modèle 3 est le "bon" modèle et la série x_t est intégrée d'ordre 1, $I(1) + c + T$; le taux de croissance est TS , $\Delta x_t = c + \beta_t + \varepsilon_t$. En revanche, si l'on accepte H_0^3 ; le coefficient de la tendance est nul, le modèle 3 n'est pas le "bon" modèle, on doit donc effectuer à nouveau le test de non stationnarité dans le modèle 2.

Si l'on a accepté la nullité du coefficient β de la tendance, on doit alors effectuer à nouveau les tests de non stationnarité à partir cette fois-ci du modèle 2 (équation 3.129) incluant uniquement une constante. On compare alors la réalisation de la statistique de Student $t_{\hat{\phi}=0}$ aux seuils $C_{(\alpha)}^2$ tabulés par dickey et fuller, ou mcKinnon pour le modèle 2 (par exemple -2.86 à 5%, pour $T \rightarrow \infty$). Si la réalisation de $t_{\hat{\phi}=0}$ est supérieure au seuil $C_{(\alpha)}^2$; on accepte l'hypothèse nulle de non stationnarité. Une fois que le diagnostic est établi, on cherche à vérifier si la spécification du modèle 2, incluant une constante, est une spécification compatible avec les données. On teste alors la nullité du coefficient c de la constante. Deux choses l'une :

Soit on a rejeté au préalable l'hypothèse de racine unitaire, dans ce cas on teste la nullité de c par un simple test de Student avec des seuils standard (test symétrique, donc seuil de 1.96 à 5%). Si l'on rejette l'hypothèse $c = 0$; cela signifie que le modèle 2 est le "bon" modèle pour tester la racine unitaire, puisque la présence d'une constante n'est pas rejetée. Dans ce cas, on conclut que la racine unitaire est rejetée, la série est stationnaire $I(0) + c$. En revanche, si l'on accepte l'hypothèse $c = 0$, le modèle 2 n'est pas adapté puisque la présence d'une constante est rejetée. On doit refaire le test de racine unitaire à partir du modèle 1, qui ne comprend ni constante ni trend.

Soit, au contraire, on avait au préalable, accepté l'hypothèse de racine unitaire, et dans

ce cas, on doit construire un test de Fischer de l'hypothèse jointe $\phi = 0$ et $c = 0$. On teste ainsi la nullité de la constante, conditionnellement à la présence d'une racine unitaire :

$$H_0^2 : (c; \phi) = (0; 0) \text{ contre } H_1^2$$

La statistique de ce test se construit de façon standard par la relation :

$$F_2 = \frac{(SCR_{2,c} - SCR_2) / 2}{SCR_2 / (T - 2)}$$

où $SCR_{2,c}$ est la somme des carrés des résidus du modèle 2 contraint sous H_0^2 , c'est à dire $SCR_{2,c} = \sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T (\Delta x_t)^2$ et SCR_2 est la somme des carrés des résidus du modèle 2 non contraint (équation 3.129). Les seuils à distance fini de cette statistique sont fournies dans la partie annexe du programme de cours. Si la réalisation de F_2 est supérieure à la valeur ϕ_1 lue dans la table à un seuil α , on rejette l'hypothèse H_0^2 au seuil $\alpha\%$: Dans ce cas, le modèle 2 est le "bon" modèle et la série x_t est intégrée d'ordre 1, $I(1) + c$. En revanche, si l'on accepte H_0^2 ; le coefficient de la constante est nul, le modèle 2 n'est pas le "bon" modèle, on doit donc effectuer à nouveau le test de non stationnarité dans le modèle 1.

Enfin, si l'on a accepté la nullité du coefficient c de la constante, on doit alors effectuer à nouveau les tests de non stationnarité à partir cette fois-ci du modèle 1 (équation 3.128) sans constante ni trend. On compare alors la réalisation de la statistique de Student $t_{\hat{\phi}=0}$ aux seuils $C_{(\alpha)}^1$ tabulés par Dickey et Fuller, ou McKinnon pour le modèle 1 (par exemple -1.95 à 5%, pour $T \rightarrow \infty$). Si la réalisation de $t_{\hat{\phi}=0}$ est supérieure au seuil $C_{(\alpha)}^1$; on accepte l'hypothèse nulle de non stationnarité. Dans ce cas la série x_t est $I(1)$ et correspond à une pure marche

aléatoire, $x_t = x_{t-1} + \varepsilon_t$: Si l'hypothèse nulle est rejetée, la série est stationnaire, $I(0)$ de moyenne nulle $x_t = \rho x_{t-1} + \varepsilon_t$; avec $|\rho| < 1$.

2.3.2 Test de Dickey Fuller Augmenté (ADF)

Il arrive parfois que les résidus ε_t du modèle de Dickey Fuller soient autocorrélés . Or, les distributions asymptotiques des statistiques de test de racine unitaire ont été construites sous l'hypothèse que ε_t est un bruit blanc. Dès que l'on lève cette hypothèse les statistiques des tests de Dickey Fuller ne suivent plus les mêmes distributions asymptotiques, et donc les seuils de significativité des tests de racine Unitaire sont différents. Il apparaît donc nécessaire de tenir compte de l'éventuelle autocorrélation des résidus dans la construction des tests de racine unitaire. Il existe alors deux approches différentes pour tenir de cette éventuelle autocorrélation. La Première approche, proposée par Philips (1987) et Phillips et Perron (1988) consiste à proposer une correction des estimateurs des MCO et des statistiques de Student associées à ces estimateurs prenant en compte la possible autocorrélation des résidus. La seconde approche, développée par Dickey et Fuller (1979), consiste à contrôler directement l'autocorrélation dans le modèle (et non au niveau des estimateurs) en incluant un ou plusieurs termes autorégressifs différenciés. Nous allons montrer qu'une telle approche permet en effet de "blanchir" les résidus et de plus, de se ramener à une représentation similaire à celle du test de Dickey Fuller Simple. Dès lors, l'application de cette nouvelle stratégie est identique à celle présentée précédemment et l'on retrouve les mêmes Distributions asymptotiques.

La prise en compte de l'autocorrélation des résidus

Pour bien comprendre l'argument de Dickey Fuller augmenté, considérons un processus x_t admettant une représentation de type $AR(1)$ dont les innovations ε_t sont autocorrélées d'ordre $p - 1$.

$$x_t = \phi x_{t-1} + \varepsilon_t$$

où ε_t ne satisfait pas les hypothèses d'un bruit blanc et est autocorrélé d'ordre $p - 1$:

$$\varepsilon_t + \theta \varepsilon_{t-1} + \dots + \theta_{p-1} \varepsilon_{t-p+1} + 1 = \mu_t$$

avec μ_t i.i.d $N(0, \sigma_\mu^2)$ En substituant ε_t par l'expression $x_t - \phi x_{t-1}$; l'équation (4.139) se réécrit sous la forme :

$$\begin{aligned} \mu_t &= (x_t - \phi x_{t-1}) + \theta_1 (x_{t-1} - \phi x_{t-2}) + \theta_2 (x_{t-2} - \phi x_{t-3}) \dots + \theta_{p-1} (x_{t-p+1} - \phi x_{t-p}) \\ &= x_t + (\theta_1 - \phi) x_{t-1} + (\theta_2 - \phi \theta_1) x_{t-2} + \dots + (\theta_{p-1} - \phi \theta_{p-2}) x_{t-p+1} - \phi \theta_{p-1} x_{t-p} \end{aligned}$$

Ainsi, la représentation de type $AR(1)$ avec autocorrélation des innovations d'ordre $p - 1$ peut être transformée en une représentation $AR(p)$ où les innovations sont des bruits blancs.

$$x_t = \psi_1 x_{t-1} + \psi_2 x_{t-2} + \dots + \psi_p x_{t-p} + \mu_t$$

avec $\psi_1 = -(\theta_1 - \phi)$, $\psi_i = -(\theta_{i-1} - \phi \theta_{i-2})$ pour $i = 2, \dots, p - 1$ et $\psi_p = \phi \theta_{p-1}$

Remarque 2.3.2. *L'intuition de la démarche du test de Dickey Fuller Augmenté consiste à postuler un modèle de type $AR(p)$ afin de corriger une éventuelle autocorrélation d'ordre $p - 1$*

des innovations d'une représentation de type $AR(1)$.

Toutefois, le test de Dickey Fuller Augmenté n'est pas construit directement à partir de la forme $AR(p)$ avec innovations i.i.d : de l'équation (4.140). Les auteurs privilégient une spécification incluant $p - 1$ termes différenciés retardés, connue sous le nom de représentation de Sims, Stock et Watson (1990).

Proposition 2.3.1. (Représentation de Sims, Stock et Watson 1990) *Tout processus $(x_t, t \in \mathbb{Z})$ satisfaisant une représentation $AR(p)$*

$$x_t = \psi_1 x_{t-1} + \psi_2 x_{t-2} + \dots + \psi_p x_{t-p} + \mu_t$$

avec μ_t i.i.d $N(0, \sigma_{\mu_t}^2)$; peut être exprimé sous la forme suivante :

$$x_t = \rho x_{t-1} + \xi_1 \Delta x_{t-1} + \dots + \xi_{p-1} \Delta x_{t-(p-1)} + \mu_t$$

où ξ_h désigne l'opposé de la somme partielle des coefficients $\forall j = 0, 1, \dots, p - 1$:

$$\xi_j = -[\psi_{j+1} + \psi_{j+2} + \dots + \psi_p]$$

$$\rho = -\xi_0 = \psi_1 + \psi_2 + \dots + \psi_p$$

Cette forme canonique est connue sous le nom de représentation de Sims, Stock et Watson (1990).

Preuve . On considère un processus $AR(p)$:

$$\Psi(L)x_t = (1 - \psi_1 L - \psi_2 L^2 - \dots - \psi_p L^p)x_t = \mu_t$$

avec μ_t i.i.d $N(0, \sigma_{\mu_t}^2)$ Notons que cette équation est formellement identique à l'équation (4.140). Nous allons à présent exprimer ce modèle en introduisant des termes différenciés retardés. On commence par poser :

$$\begin{aligned} & 1 - \psi_1 L - \psi_2 L^2 - \dots - \psi_{p-1} L^{p-1} - \psi_p L^p \\ = & 1 - \psi_1 L - \psi_2 L^2 - \dots - \psi_{p-2} L^{p-2} - [-(\psi_p) + (\psi_{p-1} + \psi_{p-1})] L^{p-1} - \psi_p L^p \\ = & 1 - \psi_1 L - \psi_2 L^2 - \dots - \psi_{p-3} L^{p-3} - [-(\psi_p + \psi_{p-1}) + (\psi_{p-2} + \psi_{p-1} + \psi_p)] L^{p-2} \\ & - [-(\psi_p) + (\psi_{p-1} + \psi_p)] L^{p-1} - \psi_p L^p \end{aligned}$$

En poursuivant cette démarche, on obtient une expression du polynôme $\Psi(L)$ faisant apparaître les sommes partielles des coefficients, notées ξ_j :

$$\begin{aligned} \Psi(L) = & 1 - (\xi_1 - \xi_0) L - (\xi_2 - \xi_1) L^2 - (\xi_3 - \xi_2) L^3 \\ & \dots - (\xi_{p-1} - \xi_{p-2}) L^{p-1} - (-\xi_{p-1}) L^p \end{aligned}$$

avec

$$\xi_j = -[\psi_{j+1} + \psi_{j+2} + \dots + \psi_{j+p}] \quad j = 0, 1, \dots, p-1$$

En arrangeant les termes du polynôme $\Psi(L)$; on parvient alors à la formulation suivante

$$\Psi(L) = (1 + \xi_0 L) - (\xi_1 L + \xi_2 L^2 + \dots + \xi_{p-1} L^{p-1}) (1 - L)$$

C'est cette expression de $\Psi(L)$ qui nous permet de faire apparaître un polynôme retard défini en la différence première $(1 - L)$: En posant $\rho = -\xi_0$; on retrouve alors la représentation de Sims, Stock et Watson (équation 4.143).

Maintenant, envisageons le cas où le processus $AR(p)$ admet au plus une racine unitaire. Si tel est le cas, le polynôme $\Psi(L)$ admet un pour racine.

$$\Psi(1) = 0 \iff 1 - \psi_1 - \psi_2 - \dots - \psi_p = 0$$

Sous cette hypothèse, le terme $\rho = -\xi_0$ de la représentation de Sims, Stock et Watson est alors égal à l'unité. ■

Proposition 2.3.2. *Si le polynôme $\Psi(L)$ associé au processus $(x_t, t \in \mathbb{Z})$, admet une racine unitaire, alors le terme ρ dans la représentation de Sims, Stock et Watson (équation 4.143) est égal à l'unité. Le processus $(x_t, t \in \mathbb{Z})$ se réécrit alors sous la forme :*

$$\Delta x_t = \xi_1 \Delta x_{t-1} + \xi_2 \Delta x_{t-2} + \dots + \xi_{p-1} \Delta x_{t-(p-1)} + \mu_t$$

Dès lors, le test de l'hypothèse de racine unitaire dans un modèle ADF revient à tester la nullité du coefficient ϕ dans le modèle incluant $p - 1$ termes différenciés retardés :

$$\Delta x_t = \phi x_{t-1} + \xi_1 \Delta x_{t-1} + \xi_2 \Delta x_{t-2} + \dots + \xi_{p-1} \Delta x_{t-(p-1)} + \mu_t$$

avec $\phi = p - 1$: Cette structure est identique à celle des tests de Dickey Fuller Simples aux termes retardés près.

Preuve. Voir []

On comprend ainsi que l'idée de Dickey Fuller consiste à se ramener à une représentation similaire à celle du test de Dickey Fuller Simple, mais qui leur permet de traiter le problème de l'autocorrélation des innovations.

Résumé de la démarche :

Partant d'un modèle $AR(1)$ (équation 4.138), on sait que les résidus et sont autocorrélés d'ordre $p - 1$. Donc on peut pas faire le test directement à partir de cette représentation. On cherche à se ramener à une représentation alternative dans laquelle les innovations sont des bruits blancs. Pour ce faire, on se ramène à un AR d'ordre p (équation 3.92). Mais si l'on désire tester la racine unitaire directement à partir de l'équation 3.92, il est nécessaire de proposer une stratégie de test différente de celle du test de Dickey Fuller Simple. Donc, les auteurs ont cherché une autre représentation de l' $AR(p)$ leur permettant de pratiquer un test de racine unitaire dont la spécification est exactement identique à celle du test de Dickey Fuller Simple : c'est la représentation de Sims, Stock et Watson (équation 4.143). Ne reste plus qu'à démontrer que les distributions asymptotiques des statistiques de tests sont identiques à celles des tests de Dickey Fuller Simples. Afin de mieux comprendre cette démarche considérons l'exemple suivant.

Ainsi, pour un choix de p retards, correspondant à une autocorrélation d'ordre $p + 1$ des innovations dans une représentation $AR(1)$; les trois modèles utilisés pour développer le test

ADF sont les suivants :

$$\text{Modèle 1 : } \Delta x_t = \phi x_{t-1} + \sum_{j=1}^p \xi_j \Delta x_{t-j} + \mu_t$$

$$\text{Modèle 2 : } \Delta x_t = \phi x_{t-1} + \sum_{j=1}^p \xi_j \Delta x_{t-j} + c + \mu_t$$

$$\text{Modèle 3 : } \Delta x_t = \phi x_{t-1} + \sum_{j=1}^p \xi_j \Delta x_{t-j} + c + \beta t + \mu_t$$

Remarque 2.3.3. *La stratégie de test ADF consiste en un première étape à déterminer le nombre de retard p nécessaire pour blanchir les résidus. Dans la seconde étape, il suffit d'appliquer la stratégie séquentielle du test de Dickey Fuller Simple aux modèles (4.156), (4.157) et (4.158). Les distributions asymptotiques des statistiques de test $t_{\hat{\phi}}$ obtenues dans ces trois modèles sont alors identiques à celles obtenues dans les modèles de Dickey Fuller Simple correspondants.*

Le principale avantage de démarche de Dickey Fuller (1979). Ces deux auteurs ont ainsi fourni une démarche de tests intégrée, avec des seuils identiques en cas d'autocorrélation ou non des résidus, grâce à l'adoption de la repré

Choix du nombre de retards optimal

Jusqu'à présent nous avons supposé que l'ordre d'autocorrélation des résidus ε_t était connu. Si cet ordre est égal à p , on sait à présent qu'il faut inclure dans le test *ADF* ; p termes différenciés retardés, ce qui correspond alors à un modèle de type $AR(p+1)$. Le problème, c'est que dans la "vraie" vie, on ignore souvent a priori l'ordre d'autocorrélation des résidus.

On doit donc chercher le nombre optimal p : Pour ce faire, plusieurs approches peuvent être envisagées, parmi celles-ci nous n'en retiendrons que deux : le contrôle ex-post de l'absence d'autocorrélation des innovations ou la minimisation de critères d'information.

Les Critères d'information utilisé

Une des manières de choisir le nombre de retards consiste à comparer différents modèles *ADF* incluant différents choix de retards, sur la base de critères d'information. Un critère d'information est un critère fondé sur le pouvoir prédictif du modèle considéré et qui tient du nombre de paramètres à estimer. De façon concrète ces critères sont construits comme des fonctions de la variance des résidus estimés du modèle $\sigma_{\hat{\varepsilon}_t}^2$ et du nombre de paramètres à estimer. L'objectif étant bien entendu de minimiser cette fonction par rapport à ces deux arguments (application du principe de parcimonie). Ces critères s'applique de façon générale à tout type de modèle et pas uniquement aux modèles des tests ADF : Nous en retiendrons que deux : le critère d'Akaike et le critère de Schwarz (1978).

Définition 2.3.1. *Pour un modèle, incluant k paramètres, estimé sur T périodes et dont la réalisation de l'estimateur de la variance des résidus est $\sigma_{\hat{\varepsilon}_t}^2$; le critère d'Akaike, ou AIC, est :*

$$AIC(k) = T \log(\sigma_{\hat{\varepsilon}_t}^2) + 2(k)$$

Le critère de Schwartz (1978) est défini par :

$$SC(k) = T \log(\sigma_{\hat{\varepsilon}_t}^2) + k \log(T)$$

Dans le cas de notre application au choix du nombre de retards dans les modèles *ADF* ; on cherche le nombre de retard p qui minimise ces deux critères. Prenons par exemple, le cas du modèle *ADF* 3 :

$$\text{Modèle 3 : } \Delta x_t = \phi x_{t-1} + \sum_{j=1}^p \xi_j \Delta x_{t-j} + c + \beta t + \mu_t$$

Dans ce cas, pour un choix de retards p on a un nombre de paramètres à estimer égal à $k = 3 + p$. Si l'on dispose d'un échantillon de taille T et que l'on note $\sigma_{\hat{\varepsilon}_t}^2$ la réalisation de l'estimateur de la variance des résidus obtenue dans le modèle avec p retard, les deux critères sont définis en fonction de p par :

$$AIC(p) = T \log(\sigma_{\hat{\varepsilon}_t}^2) + 2(3 + p)$$

$$SC(p) = T \log(\sigma_{\hat{\varepsilon}_t}^2) + (3 + p) \log(T)$$

On cherche donc le nombre de retards p qui minimise ces deux critères. Sous Eviews, ces deux critères sont fournis dès lors que l'on utilise l'instruction de régression. Nous verrons dans l'application comment obtenir la réalisation de ces deux critères.

Critères standards

- L'erreur absolue moyenne (Mean Absolute Error) :

$$MAE = \frac{1}{n} \sum_t |e_t|$$

où e_t est le résidu du modèle étudié et n le nombre d'observation.

- l'erreur quadratique moyenne (Mean Squared Error) :

$$MSE = \frac{1}{n} \sum_1^n e_t^2.$$

- la racine carrée de l'erreur quadratique moyenne(Root Mean Square Error) :

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n e_t^2}.$$

- Ecart absolu moyen en pourcentage (Mean Absolute Percent Error) :

$$MAPE = 100 \frac{1}{n} \sum_1^n \left| \frac{e_t}{X_t} \right| .$$

2.3.3 Test de Phillips et Perron(PP)

Une telle approche a été développée par Phillips (1987), Phillips et Perron (1988) et Perron (1986, 1988) : les hypothèses faites sur les erreurs sont beaucoup moins restrictives. L'idée est que des erreurs récentes peuvent être dépendantes, mais des erreurs très distantes l'une de l'autre dans le temps sont indépendantes.

Phillips(1987) et Phillips et Perron(1988) adoptent d'emblée une méthodologie radicalement différente. Leur approche est fondée sur une correction non paramétrique pour tenir compte de la structure d'autocorrélation des résidus. Les hypothèses faites sur les erreurs U_t sont ainsi beaucoup moins restrictives : elles peuvent être faiblement dépendantes temporellement et distribuées de manière hétérogène.

Leurs résultats asymptotiques sont fondées sur la théorie de la convergence faible fonctionnelle (Billingsley, 1968) et permettent de généraliser dans un cadre unifié les résultats antérieurs concernant la marche au hasard et des processus "ARMA" plus généraux contenant une racine unitaire.

Une caractéristique particulièrement intéressante des nouvelles statistiques de tests est que leur distribution asymptotique est identique à celles dérivées par Dickey et Fuller sous l'hypothèse d'erreurs i.i.d. Ceci implique que la procédure de test de Phillips et Perron peut être utilisée en se référant aux valeurs critiques asymptotiques tabulées par *dickey et fuller* même si elle permet de spécifier de manière beaucoup plus générale les séries chronologiques étudiées.

L'avantage principal de l'approche de Phillips et Perron est que, même si on peut envisager une grande variété de modèles susceptibles d'engendrer les données, le calcul des statistiques

transformées présentées dans le Tableau B (cf. Annexe) requièrent seulement : l'estimation par les MCO d'un modèle autorégressif du premier ordre (correspondant à l'un des modèles de la procédure de test DF) et le calcul des statistiques DF associées. l'estimation d'un facteur de correction fondé sur la structure des résidus de cette régression.

Lorsque les erreurs ne sont pas i.i.d., Phillips et Perron (1988) montrent en effet que les distributions asymptotiques des statistiques t de la procédure de test DF dépendent du rapport σ^2/σ_u^2 où la variance des résidus est :

$$\sigma_u^2 = \lim \frac{1}{n} \sum_{t=1}^n E(u_t^2)$$

et la variance de long terme est :

$$\sigma^2 = \lim \frac{1}{n} \sum_{t=1}^n E(S_t^2) \quad \text{avec} \quad S_t = \sum_{t=1}^n u_t$$

Un estimateur convergent de est simplement la variance estimée des résidus dans le modèle alternatif considéré. Différents estimateurs convergents de la variance de long terme sont envisageables, mais Phillips et Perron optent pour celui proposé par **Newey et West**(1987) fondé sur la méthode du noyau qui fait intervenir un paramètre de troncature manalogue \tilde{b} celui que nous avons déjà rencontré dans la procédure de test de Said et Dickey :

$$\hat{\sigma}_t^2 = \frac{1}{n} \sum_{t=1}^n \hat{v}_t^2 + 2 \frac{1}{n} \sum_{j=1}^{\tilde{b}} \left(1 - \frac{j}{\tilde{b}+1}\right) \sum_{t=j+1}^n \hat{v}_t \hat{v}_{t-j}$$

et les \hat{v}_t sont les résidus estimés du modèle alternatif considéré. Les pondérations affectées aux autocovariances estimées assurent que la variance estimée est positive. Notons que a une

interprétation naturelle pour stationnaires : il s'agit de $2p$ fois l'estimateur de la densité spectrale de s^2 pour la fréquence zéro ; quant au choix des pondérations, il est équivalent au choix de la fenêtre de retard dans la littérature sur l'estimation de la densité spectrale. Phillips et Perron suggèrent que le paramètre de troncature m soit une fonction croissante de la taille de l'échantillon [en fait $m = o(T^{1/4})$] mais ne donnent aucune indication précise sur la valeur qu'il faut lui attribuer en pratique. Le problème de la détermination de ce paramètre m dans les études empiriques se pose donc à nouveau et longtemps la seule solution envisageable a été d'effectuer les tests pour plusieurs valeurs arbitrairement choisies et d'étudier la sensibilité des résultats par rapport à ces différentes valeurs. Cependant il est aujourd'hui possible d'utiliser une procédure pour fixer ce paramètre de manière optimale en s'inspirant de celle mise au point par Andrews (1991) et Andrews et Monahan (1992) pour l'estimation robuste de la matrice de variances-covariances en présence d'autocorrélation et/ou d'hétéroscédasticité. Considérons les résidus estimés de l'un des modèles autorégressifs du premier ordre de la procédure de Dickey et Fuller et un noyau $k(\cdot)$ satisfaisant les conditions de Andrews (1991, p.822) assurant la non-négativité de la variance estimée de long terme. L'estimateur de la variance de long terme peut se réécrire :

$$\hat{\sigma}^2 = (T/(T-r))[\hat{\Gamma}(0) + 2 \sum_{j=1}^{n-1} K(\frac{j}{S_T}) \hat{\Gamma}(j)]$$

où r est le nombre de paramètres estimés dans la régression donnant les \hat{v}_t , $\hat{\Gamma}(j)$ est l'autocovariance estimée d'ordre j et S_T , la largeur de la fenêtre, devient le paramètre d'intérêt. Ainsi les pondérations utilisées par **Newey et West** sont égales à $\frac{1-j}{S_T}$ pour $j \leq S_T$ et 0 sinon où

S_T est un réel. Si S_T est un entier naturel, alors ces pondérations sont équivalentes lorsque $S_T = m + 1$. Le terme $\frac{T}{T-r}$ est un facteur de correction des degrés de liberté en échantillon de taille finie, il est optionnel. La valeur de S_T peut être déterminée en utilisant les données comme l'indique Andrews (1991). Les formules dépendent des noyaux mais aussi des données par l'intermédiaire de la densité spectrale des vraies erreurs non observées, $f_u(w)$ en $w = 0$. Ceci nécessite l'estimation de $f_u(w)$ la densité spectrale des résidus estimés \hat{v}_t qui est réalisée de manière paramétrique en utilisant des modèles ARMA.

On a ainsi par exemple : pour le noyau de Bartlett $\hat{S}_T = 1.11447(\hat{\alpha}(1)T)_{1/3}$.

pour le noyau de Parzen $\hat{S}_T = 2.6614(\hat{\alpha}(2)T)_{1/5}$.

pour le noyau quadratique spectral préféré par Andrews

$$\hat{S}_T = 1.3221(\hat{\alpha}(2)T)_{1/5}$$

où $\hat{\alpha}(q) = [(2\pi f_{\hat{u}}(0))^{-1} \{2 \sum_{j=1}^{\infty} |j|^q \hat{\Gamma}(j)\}]$ pour $q = 1, 2$ (Andrews, 1991, p.834).

Les formules pour $\hat{\alpha}(q)$ sont fournies par Andrews (1991, p.835) notamment dans le cas de modèles AR(1) et ARMA(1,1). Burke (1996, p.321) fournit une formule dans le cas d'un modèle AR(p).

La procédure à suivre peut alors être résumée de la manière suivante : on estime dans un premier temps par les MCO un modèle autorégressif du premier ordre (correspondant à l'un des modèles de la procédure de test de Dickey et Fuller) et on calcule les résidus estimés

\hat{v}_t ainsi que leurs autocovariances dans l'échantillon. On estime ensuite un modèle AR(1) ou ARMA(1,1) ou même AR(p) sur les résidus estimés \hat{v}_t . On calcule alors $\hat{\alpha}(q)$ et puis \hat{S}_T correspondant au noyau choisi. La variance de long terme est ensuite obtenue en utilisant l'équation (14). On intègre finalement la valeur obtenue dans les formules des statistiques transformées de Phillips et Perron et on effectue les tests correspondants. Il semblerait toutefois, d'après les expériences de simulation de Burke (1996), que cette procédure n'améliore pas significativement la performance des tests de Phillips et Perron qui restent dominés par les tests ADF avec sélection descendante du paramètre de troncature. Il faut finalement noter que toutes ces procédures de tests de la racine unitaire souffrent de distorsions de niveau en présence de certaines formes d'autocorrélation des résidus et de faible puissance au voisinage de l'unité. Il faut donc rester très prudent dans l'interprétation des résultats.

La critique principale qu'on peut adresser à ces procédures de test est la suivante : elles postulent que la composante déterministe de la série considérée suit une tendance déterministe linéaire. Or la mauvaise spécification, linéaire en l'occurrence, de la composante déterministe peut nous conduire à ne pas rejeter l'hypothèse nulle de la racine unitaire et ceci à tort.

Résumé du déroulement du test de Philips-Perron (PP) :

La procédure de test consiste à tester l'hypothèse de racine unitaire $H_0 : \phi = 0$ dans les modèles qui on a utilisé de test DF :

$$\text{Modèle [1] : } \Delta X_t = \phi X_{t-1} + \varepsilon_t$$

$$\text{Modèle [2] : } \Delta X_t = \phi X_{t-1} + c + \varepsilon_t$$

$$\text{Modèle [3] : } \Delta X_t = \phi X_{t-1} + c + \beta t + \varepsilon_t$$

La statistique de test de Phillips-Perron (PP) est une statistique de student corrigée de la présence d'autocorrélation par la prise en compte d'une estimation de la variance de long terme de ε_t (calculé par la densité spectrale de ε_t à la fréquence zéro), robuste à la présence d'autocorrélation et d'hétéroscédasticité.

Doc on peut résumer le déroulement de test Phillips-Perron (PP) dans quatre étapes qui sont :

- 1) Estimation par la méthode des moindres carrés ordinaires les trois modèles du test de dickey-fuller simple et calcul des résidus $\widehat{\varepsilon}_t$
- 2) Détermination de la variance dite de court terme :

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n \widehat{\varepsilon}_t^2$$

- 3) Estimation du facteur correctif appelé variance de long terme :

$$S_t^2 = \frac{1}{n} \sum_{t=1}^n \widehat{\varepsilon}_t^2 + 2 \sum_{j=1}^b \left(1 - \frac{j}{b+1}\right) \frac{1}{n} \sum_{t=j+1}^n \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-j}$$

Où

$$b \approx 4 \left(\frac{n}{100}\right)^{2/9}$$

- 4) Calcul de la statistique de philips-perron :

$$t_{\widehat{\phi}_1}^{pp} = \sqrt{K} \times \frac{(\widehat{\phi}_1 - 1)}{\widehat{\sigma}_{\widehat{\phi}_1}} + \frac{n(k-1)\widehat{\sigma}_{\widehat{\phi}_1}}{\sqrt{K}}$$

avec

$$K = \frac{\hat{\sigma}^2}{S_t^2}$$

Phillips et Perron (1988) montrent que cette correction non-paramétrique apportée à $t_{\hat{\phi}_1}$ ne modifie pas la distribution asymptotique de la statistique qui reste identique à celle qui est observée dans le cas du test de Dickey et Fuller demeurent également valables pour le test Phillips et Perron.

2.3.4 Test Kwiatkowski, Phillips, Schmidt et Shin (KPSS)

On a considéré jusqu'ici des tests où l'hypothèse nulle était la racine unitaire et la non-stationnarité. Dans le modèle $y_t = \rho y_{t-1} + \varepsilon_t$ on testait $H_0 : \phi = 1$

L'alternative était la stationnarité. Supposons maintenant que l'on parte du modèle

$$y_t = u + \theta u_{t-1}$$

où u_t est un processus stationnaire. y_t est donc $I(0)$. Si l'on différencie le processus, on obtiendra

$$\Delta y_t = u_t + \theta u_{t-1}$$

avec $\theta = 1$. Dans ce modèle-ci, l'hypothèse nulle devient $\theta = 1$ et elle est l'équivalente à la stationnarité. Par contre si $\theta \neq 1$, alors $\varepsilon_t = u_t + \theta u_{t-1}$ est stationnaire, et par conséquent le niveau de y_t s'obtient par une accumulation de bruits blancs stationnaires, $y_t = \sum \varepsilon_t$

ce qui donne la définition d'une série intégrée d'ordre 1. Le test de l'hypothèse nulle $H_0 : \phi = 1$ contre l'alternative $H_1 : \phi \neq 1$ permettra de tester cette fois-ci la stationnarité de la série.

Il y a donc une symétrie entre d'un côté racine unité dans la partie AR et non stationnarité

et de l'autre côté racine unitaire dans la partie MA et stationnarité.

Il n'est toutefois pas commode de tester directement la présence d'une racine unitaire dans un processus MA. En effet, même si la vraie valeur de ϕ n'est pas 1.

L'estimateur du maximum de vraisemblance aura tendance à être proche de 1 par un effet dit de pileup. Ceci vient du fait que les deux MA(1) suivants ont les mêmes autocorrélation :

$$y_t = u_t + \theta u_{t-1} \quad \text{et} \quad x_t = u_t + 1/\theta u_{t-1}$$

.

Il faut donc trouver une autre formulation du problème. Considérons le modèle suivant :

$$y_t = \delta t + z_t + u_t \quad V(u_t) = \sigma_u^2$$

$$z_t = x_{t-1} + v_t \quad V(v_t) = \sigma_v^2$$

La variable y_t est décrite sous la forme de la somme d'une composante déterministe comportant un trend δt et d'une composante stochastique z_t qui est un trend stochastique avec comme valeur initiale $z_0 = 0$. Si dans ce modèle on pose $\sigma_v^2 = 0$, alors le trend stochastique se traduit à sa valeur initiale qui est zéro. Ce modèle admet comme forme traduite un processus MA(1).

Effet, appliquons l'opérateur aux deux membres de la première équation et remplaçons z_t par sa valeur. Il vient :

Tester que $\theta = 1$ est équivalent à tester $\sigma_v^2 = 0$.

On va donc chercher une statistique de test où sous H_0 on aura .

On va utiliser ce rapport pour éliminer le paramètre de nuisance.

Kwiatkowski, Phillips, Schmidt et Shin (1992) ont proposé une statistique de test particulière pour tester

$H_0 : \sigma_v^2 = 0$ contre $H_1 : \sigma_v^2 > 0$ C'est un test unilatéral. Considérons tout d'abord la régression auxiliaire y_t

dont on va tirer les résidus estimés ε_t . On retire donc les composantes déterministes du processus. Si y_t ne comporte pas de trend stochastique, les ε_t seront stationnaires. Par contre si y_t n'est pas stationnaire, il y aura une racine unité dans les ε_t . Définissons les sommes partielles

2.4 Les modèles VAR

Un modèle VAR est un outil économétrique particulièrement adapté pour définir l'ensemble des liaisons dynamiques à l'intérieur d'un groupe de variables considérées initialement toutes endogènes. En règle générale, la modélisation VAR consiste à modéliser un vecteur de variables stationnaires à partir de leurs propres historiques, et chaque variable est expliquée par le passé de l'ensemble des variables

2.4.1 Représentation d'un modèle VAR

Définition 2.4.1. *Un processus vectoriel $(X_t, t \in Z)$ de dimension $(n - 1)$ admet une représentation VAR d'ordre P , notée $VAR(P)$, si*

$$X_t = c - \varphi_1 X_{t-1} - \varphi_2 X_{t-2} + \epsilon_t$$

ou de façon équivalente :

$$\varphi_1(B)X_t = c + \epsilon_t$$

Où c 'est le vecteur des constantes de dimension $n - 1$.

2.4.2 Stationnarité d'un modèle VAR

La définition de la stationnarité d'ordre deux (ou stationnarité du second ordre), des modèles multi variés est identique à celle du cas des processus uni variés. En effet, il existe diverses définitions de la stationnarité d'un modèle VAR :

Définition 2.4.2. *un processus vectoriel $(X_t, t \in Z)$ de dimension $(n - 1)$ est dit stationnaire au second ordre, ou stationnaire au sens faible, ou stationnaire d'ordre deux si :*

1. $\forall t \in Z, E(X_t^2) < +\infty$
2. $\forall t \in Z, E(X_t) \in Z, E(X_t) = m_{(n-1)}$ indépendante de t
3. $\forall t \in Z^2, E[(X_{t+h} - m)(X_t - m')] = y(h)_{(n,n)}$ indépendante de t

Définition 2.4.3. *un processus vectoriel $(X_t, t \in Z)$ de dimension (n, n) satisfait une repré-*

sentation VAR (P) telle que $\forall t \in Z$

$$\varphi(B)X_t = X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = c + \epsilon_t$$

Est stationnaire si et seulement si les racines du déterminant du polynôme matriciel $\varphi(B)$ notée $\gamma_i, i = 1 \dots n$ sont toutes supérieures à l'unité en module.

$$\det[\varphi(\gamma_i)] = |\varphi(\gamma_i)| = |I_n \gamma_i^p - I_1 \gamma_i^{p-1} - \dots - \varphi_{p-1} \gamma_i - \varphi_p| = 0$$

$$|\gamma_i| > 1, \forall i \in [1, n]$$

Définition 2.4.4. un processus vectoriel $(X_t, t \in Z)$ de dimension (n, n) satisfait une représentation VAR (P) telle que $\forall t \in Z$

$$\varphi(B)X_t = X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = c + \epsilon_t$$

Est stationnaire si et seulement si les valeurs propres de l'application linéaire $\varphi(B)$ notée $\bar{\gamma}_i, i = 1 \dots n$ sont toutes inférieures à l'unité en module. Ces valeurs propres satisfont l'équation caractéristique associée

$$\det[\varphi(\bar{\gamma}_i)] = |\varphi(\bar{\gamma}_i)| = |I_n \bar{\gamma}_i^p - I_1 \bar{\gamma}_i^{p-1} - \dots - \varphi_{p-1} \bar{\gamma}_i - \varphi_p| = 0$$

$$|\bar{\gamma}_i| > 1, \forall i \in [1, n]$$

2.4.3 Détermination du nombre de décalage

La théorie économique ne donne que très peu d'information sur le nombre de décalage P à retenir dans la modélisation d'un vecteur de variables X sous une forme VAR (P), de plus, il n'existe aucune procédure établie pour déterminer le nombre de décalages adéquat.

Il y a trois méthodes sont plus particulièrement retenues pour estimer le nombre de décalages P dans la pratique des modèles VAR :

Une méthode basée sur l'examen des propriétés statiques des innovations du modèle VAR.

Une méthode basée sur l'utilisation de critères d'information

Une méthode basée sur des tests de nullité emboîtés sur les paramètres associée au dernier décalage du modèle.

2.4.4 Décomposition de la variance et fonction de réponse impulsionnelle

Dans les applications empiriques, une des principales utilisations des processus VAR réside dans la décomposition de la variance et l'analyse de réponse impulsionnelle.

La décomposition de la variance de l'erreur de prévision a pour objectif de calculer pour chacune des innovations sa contribution à la variance de l'erreur. par une technique mathématique, on peut écrire la variance de l'erreur de prévision à un horizon h en fonction de variance de l'erreur attribuée à chacun des variables; il suffit ensuite de rapporter chacune de ces variances à la variance totale pour obtenir son poids relatif en pourcentage.

La fonction de réponse impulsionnelle représente l'effet d'un choc d'une innovation sur les

valeurs courantes et futures des variables endogènes. Un choc sur la i ème variable, mais il se transmet également à l'ensemble des autres variables au travers de la structure dynamique du VAR. Ainsi considérant deux réalisations différentes de notre processus VAR **** supposant que la première réalisation soit telle qu'entre ***le système connaisse un seul choc (ce choc intervenant en t). La deuxième réalisation suppose que le système ne subit pas de choc entre ****. La fonction de réponse impulsionnelle est alors définie comme la différence entre ces deux réalisations qui sont identiques jusqu'en $t - 1$.

Chapitre 3

Tests de cointegration

L'étude des différentes procédures de test de la racine unitaire et des différentes stratégies nous a conduit à la réflexion suivante quant à l'application de ces procédures de test dans les travaux empiriques : il faut souligner d'abord l'importance cruciale que revêt la spécification de la composante déterministe dans les procédures de test de la racine unitaire. Les résultats des tests dépendent en effet de cette spécification. Le principe de base à retenir est le suivant : la stratégie de test doit débiter par le modèle le plus général possible concernant la spécification de la composante déterministe.

Par ailleurs, étant donné l'extrême sensibilité de ces procédures à la structure d'autocorrélation des erreurs et l'absence d'une procédure de sélection réellement efficace du paramètre de troncature, la faiblesse de la puissance des tests et les distorsions de niveau, il paraît nécessaire de chercher systématiquement à évaluer la robustesse des résultats obtenus en appliquant plusieurs d'entre elles et en essayant dans la mesure du possible de ne pas imposer des spécifications trop restrictives sur la composante déterministe ni des hypothèses trop restrictives sur les erreurs.

L'application des tests sur le modèle le plus général possible suppose le recours à une approche heuristique qui requiert, avant tout, l'examen du graphe de la série considérée pour détecter toute forme de non-linéarité et d'éventuels changements structurels majeurs. Cette première étape peut éventuellement être suivie par l'application des procédures de test de Ouliaris, Park et Phillips (1989) ou Perron (1989) ou Zivot et Andrews (1992), Banerjee, Lumsdaine et Stock (1992) ou encore Perron (1997) qui intègrent de manière formelle la spécification polynomiale ou linéaire par morceaux de la composante déterministe.

- la procédure d'estimation en deux étapes développée par Engle et Granger
- les tests basés sur les résidus de Phillips–Ouliaris
- la procédure de Johansen

3.1 Définition de la cointégration

Définition 3.1.1. *On considère un processus $X_t = (x_{1,t}, x_{2,t}, \dots, x_{N,t})$ de dimension $(N, 1)$ intégré d'ordre d . Les processus $(x_{i,t}, t \in Z)$ sont dite cointégrées si et seulement s'il existe un vecteur $\alpha = (\alpha_1, \dots, \alpha_N)' \in \mathbb{R}^N$ tel que la combinaison linéaire $\alpha' X_t$ est stationnaire ou intégré d'ordre 0. Le vecteur α correspond à un vecteur de cointégration.*

3.1.1 Conditions de la cointégration

Deux variable $x_{1,t}$ et $x_{2,t}$ sont dites cointégrées si les deux conditions suivantes sont vérifiées :

1. Les deux séries sont intégrées de même ordre $x_{1,t} \mapsto I(d), x_{2,t} \mapsto I(d)$.
2. Il existe une combinaison linéaire de ces séries permettant de se ramener à une série d'ordre

inferieure. c'est -à-dire :

$$\left\{ \begin{array}{l} x_{1,t} \mapsto I(d) \\ x_{2,t} \mapsto I(d) \end{array} \right. \quad \text{Alors} \quad \alpha_1 x_{1,t} + \alpha_2 x_{2,t} \mapsto I[d - b] \quad \text{avec} \quad d \geq 0, b \geq 0$$

On note $x_{1,t}, x_{2,t} \mapsto CI(d, b)$, avec $[\alpha_1, \alpha_2]$ est le vecteur de cointegration.

Dans le cas généralde N variable, on a :

Soient N processus $:(x_{1,t}, x_{2,t}, \dots, x_{N,t})$ intégrés de même ordre d :

$$\left\{ \begin{array}{l} x_{1,t} \mapsto I(d) \\ x_{2,t} \mapsto I(d) \end{array} \right. \quad \text{nots} \quad X_t = [x_{1,t}, x_{2,t}, \dots, x_{N,t}]; X_t \mapsto I(d).$$

S'il existe un vecteur $\alpha = [\alpha_1, \dots, \alpha_N]$ de dimension $(N \times 1)$ tel que $\alpha X_t \mapsto I(d - b)$ alors les N variables cointégrées et le vecteur de cointegratio est α .

3.2 La méthode de Engle–Granger

La modélisation classique à plusieurs équations structurelles a connu beaucoup de critiques et de défaillances face à un environnement économique très perturbé. Les prévisions élaborées à l'aide de ces modèles sont très médiocres de plus, il a été constaté que dans plusieurs relations de régression, on obtenait des coefficients significatifs avec un coefficient de détermination R2 très proche de 1 (existence de corrélation d'ordre 1 des erreurs). Les staticiens ont commencé donc à prendre conscience que les corrélations entre les séries chronologiques devaient être considérées avec beaucoup de prudence. **Engel et Newbold** (1974) ont introduit dans la littérature économétrique la notion de " spurious régressions " (régression fallacieuse).

Engle et Granger ont développé cette technique cruciale en 1987. Cette technique implique des variables cointégrées qui sont discutées en détail, y compris une preuve de théorème de représentation de Granger, qui relie la moyenne mobile, la autorégressif, et Terreur de représentation de correction pour les systèmes cointégrées. Selon les statisticiens, les étapes permettant de déterminer si deux variables cointégrées de même ordre sont les suivantes :

1. Pré-test chaque variable pour déterminer l'ordre d'intégration.
2. Estimer le modèle à correction d'erreur.

Si les variables intégrées sont trouvés être intégrés au même ordre, alors il doit être testé si ces variables sont cointégrées.

3.2.1 Pré-test chaque variable

Pré-tester chacun variable pour déterminer l'ordre d'intégration de chaque variable. Par définition, la cointégration nécessite que deux variables intégrées du même ordre. Ceci est fait en utilisant le test de racine unitaire "Dickey-Fuller augmente" (ADF) pour déduire le nombre de racines unitaires dans chacune des variables à études. La procédure de test pour le test de racine unitaire ADF est appliquée sur le modèle suivant

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{j=1}^p \delta_j \Delta y_{t-j} + \varepsilon_{it} \quad (3.1)$$

Où α est une constante, β le coefficient sur une série de tendance dans le temps, γ le coefficient de y_{t-1} , p est de l'ordre de retard du processus autorégressif. $\Delta y_t = y_t - y_{t-1}$ sont 1ère différences de y_t , y_{t-1} sont des valeurs de l'ordre une des y_t , y_{t-j} des changements dans les valeurs retardées, et Δ_{it} il est le bruit blanc.

Une fois l'hypothèse du test de racine unitaire est rejetée, nous estimons la relation d'équilibre à long terme sous la forme d'une ligne de régression MCO

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad (3.2)$$

où β_0 est l'ordonnée, β_1 est la pente, et ε_t est le terme d'erreur. Les estimations des paramètres dans l'équation (3.2) sont estimés à partir

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{t=n} (x_t - \bar{x}_t)(y_t - \bar{y}_t)}{\sum_{t=1}^{t=n} (x_t - \bar{x}_t)^2} = \frac{\sum_{t=1}^{t=n} x_t y_t - n \bar{x}_t \bar{y}_t}{\sum_{t=1}^{t=n} x_t^2 - n \bar{x}_t^2} \quad (3.3)$$

où \bar{x}_t et \bar{y}_t sont le moyen de x_t et y_t respectivement. La valeur de β_0 est estimée à partir de

$$\hat{\beta}_0 = \bar{y}_t - \hat{\beta}_1 \bar{x}_t \quad (3.4)$$

La droite de régression estimée est alors donnée sous la forme

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}_t \quad (3.5)$$

Si les variables cointégrés, une équation de régression MCO (3.5) donne un estimateur "super-conforme". Cela signifie qu'il existe une forte relation linéaire entre les variables étudiées. La forte relation linéaire peut être testée dans une ou l'autre des façons suivantes.

(a) La valeur de $\hat{\beta}_1$ se situe entre 0, 5 et 1.

(b) Le terrain de y_t contre x_t affiche les coordonnées figurant dans le sens croissant ou dé-

croissant.

Afin de déterminer si les variables sont cointégrées, nous testons des racines unitaires sur la séquence résiduelle dans l'équation (3.2) en utilisant le test de l'ADF. La séquence résiduelle, notée ε_t et est une série de valeurs estimées de l'écart par rapport à la relation de long terme. Ils sont estimés à partir

$$\varepsilon_t = y_t - \hat{y}_t \quad (3.6)$$

Où \hat{y}_t sont des valeurs de l'équation prédit (3.5). Le test de racine unitaire sur les résidus vise à déterminer si ces écarts sont stationnaires ou non. Si elles sont stationnaire, alors la série est cointégrée. Si les résidus ne sont pas stationnaires. il n'y a pas de cointégration. Le test ADF est effectué sur le modèle suivant

$$\Delta \hat{\varepsilon}_t = \alpha_1 \hat{\varepsilon}_{t-1} + \varepsilon_t \quad (3.7)$$

Où $\Delta \hat{\varepsilon}_t$ sont les premiers résidus différenciés estimés. $\hat{\varepsilon}_{t-1}$ sont les résidus décalés estimés. α_1 est le paramètre d'intérêt représentant la pente de la ligne et ε_t sont des erreurs obtenues en résidus de montage à la fois différenciées.

Depuis les Séquences ε_t sont des résidus d'une équation de régression, il n'y a aucune nécessité d'inclure le terme d'interception dans l'équation (3.7). Pour tester l'hypothèse sur α_1 afin de déterminer si les résidus sont stationnaires, nous suivons les étapes suivantes.

(i) définir à la fois l'hypothèse nulle et alternatives comme

$$H_0 : \alpha_1 = 0 \quad (3.8)$$

$$H_1 : \alpha_1 < 0 \quad (3.9)$$

(ii) Déterminer la statistique de test à l'aide

$$F_{\hat{\varepsilon}_t} = \frac{\hat{\alpha}_1}{SE(\hat{\alpha}_1)} \quad (3.10)$$

Où la valeur de $SE(\hat{\alpha}_1)$ est l'erreur de type $\hat{\alpha}_1$, et $\hat{\alpha}_1$ est l'estimation de α_1 .

(iii)) Comparer la statistique de test calculée (3.10) à la valeur critique de la table de Dickey-Fuller pour rejeter ou ne pas rejeter l'hypothèse nulle.

(iv) Si $F_{\hat{\varepsilon}_t}$ est supérieure à la valeur critique, nous ne rejetons pas l'hypothèse nulle H_0 .

Le rejet de H_0 implique que les résidus sont stationnaires. Cela implique en outre que les variables à l'étude sont cointégrées

Remarque 3.2.1. *Le but dans cette étape est d'examiner la non stationnarité et non pas extraire le type de la non stationnante si elle existe. Le fait de trouver un indicateur de non stationnante la procédure s'arrête c'est-à-dire le résidu ne formule pas une relation de cointégration.*

L'étape suivante consiste à estimer le modèle de correction d'erreurs (ECM), ce qui sera fait dans la section suivante.

3.2.2 Estimation de modèle à correction d'erreur

Un modèle de correction d'erreur est défini comme un modèle dynamique, dans lequel le mouvement d'une variable dans une période quelconque est lié à l'écart de la période

précédente à partir de l'équilibre à long terme. Bien qu'il puisse être possible d'estimer la relation à long terme ou de cointégration, les systèmes économiques $y_t = \beta x_t + \varepsilon_t$ sont rarement en équilibre, car ils sont affectés par des changements institutionnels et / ou structurelles qui pourraient être temporaire ou permanente. Par exemple, un revenu supplémentaire sous la forme d'un bonus d'anniversaire peut soulever la structure des dépenses de quelqu'un dans un ou deux mois, puis ses dépenses va progressivement revenir à l'équilibre. Depuis que l'équilibre est rarement observée, l'évolution à court terme des variables (ajustement dynamique à court terme) est importante. Un modèle dynamique simple d'un modèle d'ajustement à court terme est donné par :

$$y_t = \alpha_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \alpha_1 y_{t-1} + \varepsilon_t \quad (3.11)$$

où y_t est la variable dépendante, x_t est la variable indépendante, y_{t-1} et x_{t-1} sont des valeurs retardées de y_t et x_t respectivement, $\alpha_0, \alpha_1, \gamma_0, \gamma_1$ sont des paramètres, et ε_t est le terme d'erreur être supposé $\varepsilon_t \sim iidN(0, \sigma^2)$. Les problèmes associés à l'utilisation du modèle à court terme sont les suivantes :

- (i) **Multicolinéarité** : Il s'agit d'une situation dans laquelle deux ou plusieurs variables indépendantes dans un modèle de régression multiple sont fortement corrélés.
- (ii) **Corrélation fausse** : Il s'agit d'une situation dans laquelle deux variables n'ont pas de lien de causalité, mais il peut être déduit qu'ils font à la suite d'un certain troisième facteur invisible.

Ces problèmes sont résolus par l'estimation des différences premières de l'équation (3.11) pour obtenir

$$\Delta y_t = \alpha_0 + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \alpha_1 \Delta y_{t-1} + \varepsilon_t \quad (3.12)$$

Ceci, cependant, présente des problèmes de

- perte d'informations à propos de l'équilibre de long terme
- la théorie économique est loin différenciée.

La solution consiste à adopter le mécanisme de correction d'erreur (ECM) formulation de la structure dynamique. Nous mettons en place l'ECM comme suit :

$$y_t = \alpha_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \alpha_1 y_{t-1} + \varepsilon_t \quad (3.13)$$

En soustrayant le terme y_t des deux côtés conduit à

$$\Delta y_t = \alpha_0 + \gamma_0 x_t + \gamma_1 x_{t-1} - (1 - \alpha_1) y_{t-1} + \varepsilon_t \quad (3.14)$$

En soustrayant le terme $\gamma_0 x_{t-1}$ des deux côtés, l'équation (3.14) devient

$$\Delta y_t - \gamma_0 x_{t-1} = \alpha_0 + \gamma_0 x_t + \gamma_0 x_{t-1} + \gamma_1 x_{t-1} - (1 - \alpha_1) y_{t-1} + \varepsilon_t \quad (3.15)$$

C'est égale à

$$\Delta y_t = \alpha_0 + \gamma_0 x_t + (\gamma_0 + \gamma_1) x_{t-1} - (1 - \alpha_1) y_{t-1} + \varepsilon_t \quad (3.16)$$

La reparamétrisation réduit l'équation (3.16) à

$$\Delta y_t = \gamma_0 x_t - (1 - \alpha_1) \left[y_{t-1} - \frac{\alpha_0}{(1 - \alpha_1)} - \frac{(\gamma_0 + \gamma_1)}{(1 - \alpha_1)} x_{t-1} \right] + \varepsilon_t \quad (3.17)$$

Prenant $\beta_0 = \frac{\alpha_0}{(1 - \alpha_1)}$ et $\beta_1 = \frac{(\gamma_0 + \gamma_1)}{(1 - \alpha_1)}$ l'équation (3.17) devient

$$\Delta y_t = \gamma_0 x_t - (1 - \alpha_1) [y_{t-1} - \beta_0 - \beta_1 x_{t-1}] + \varepsilon_t \quad (3.18)$$

Qui est l'ECM avec $-(1 - \alpha_1)$ en tant que la vitesse d'ajustement, et $\varepsilon_{t-1} = y_{t-1} - \beta_0 - \beta_1 x_{t-1}$ comme le mécanisme de correction d'erreur qui mesure la distance du système à une distance à partir de l'équilibre. La taille du coefficient ε_{t-1} est une indication de la vitesse d'ajustement vers l'équilibre en ce que

- les petites valeurs de $-(1 - \alpha_1)$, ce qui tend à -1 . indiquent que les agents économiques supprimer un grand pourcentage de déséquilibre pour chaque période.
- des valeurs plus élevées, tendant à 0 . indiquent que l'ajustement est lent
- des valeurs extrêmement petites, moins de -2 . indiquent un dépassement de l'équilibre économique
- valeurs positives impliquerait que le système s'écarte de la voie de l'équilibre à long terme.

L'ECM satisfait les hypothèses du modèle de régression linéaire normale classique (CXLRM).

Ces hypothèses comprennent :

- Un modèle de régression linéaire
- Les résidus sont normalement distribués.
- Il n'existe aucune corrélation entre les résidus de série
- Le nombre d'observations ne doit pas dépasser le nombre de paramètres à estimer.
- Il n'y a pas de multi-colinéarité parfaite.

Cela signifie que les tests de diagnostic doivent être menés sur le mécanisme de correction d'erreur afin de déterminer si l'une de ces hypothèses n'a pas été violée. Ces tests comprennent

1. un test de normalité
2. un test de l'hétéroscédasticité
3. un test d'Autocorrélation

Test de normalité

Nous utilisons le test Jacque-Bera pour déterminer si l'ECM est normalement distribué. Ce test mesure la différence de coefficient d'asymétrie "*Skwness*" et le coefficient d'aplatissement "*kurtosis*" d'une variable par rapport à ceux de la distribution normale (Jarque et Bera. 1980). Dans le test Jacque-Bera, nous avons mis l'hypothèse nulle et l'alternatif comme suit :

H_0 : Le variable est normalement distribué.

H_1 : Le variable n'est pas normalement distribué.

La statistique de test est

$$JB = \frac{N - k}{6} \left[S^2 + \frac{(K - 3)^2}{4} \right] \quad (3.19)$$

Où N est le nombre d'observations, k est le nombre de paramètres estimés. S est le Skwness d'une variable, et K est le kurtosis d'une variable. Nous rejetons l'hypothèse nulle si la p -valeur $<$ niveau de signification, ou si le $JB > \chi^2(2)$.

Test de l' Hétéroscédasticité

L'Hétéroscédasticité résulte d'une suite de variables aléatoires ayant des variances. Elle implique que lors de l'analyse de régression, il est la variance non-conforme. L'Hétéroscédasticité est testée en utilisant le multiplicateur de Lagrange. également connu comme test d'Engle d'Arch LM. La procédure de test est comme suit :

H_0 : Il n'y a pas d'hétéroscédasticité

H_1 : Il est hétéroscédasticité

La statistique de test est

$$LM_E = nR^2 \quad (3.20)$$

Où n est le nombre d'observations, et R^2 est le coefficient de détermination de la régression résiduelle augmentée.

Nous rejetons l'hypothèse nulle si la p -valeur $<$ niveau de signification et de conclure qu'il va hétéroscédasticité.

Test d'Autocorrélation

l'autocorrélation est la corrélation croisée d'un signal (bruit blanc) avec elle-même. Elle peut être causée par

- La non-stationnarité de la variable dépendante et explicative
- Manipulation des données (moyenne, interpolation et extrapolation)
- Forme fonctionnelle incorrecte.

Ljung et Box (1978) ont suggéré l'utilisation de tests de Ljung-Box pour tester l'hypothèse que les résidus ne contiennent aucune autocorrélation à tout ordre k . La procédure de test

est comme suit :

H_0 : Il n'y a pas d'autocorrélation jusqu'à l'ordre k .

H_1 : L'autocorrélation existe jusqu'à l'ordre k

La statistique de test est

$$Q_{LB} = T(T + 2) \sum_{j=1}^k \frac{r_j^2}{T - j} nR^2 \quad (3.21)$$

Où T est le nombre d'observations, k est l'ordre le plus élevé d'autocorrélation pour lesquels tester, r_j^2 est la j^{ime} autocorrélation.

Nous rejetons l'hypothèse nulle si la p -valeur \leq niveau de signification et de conclure que l'autocorrélation existe jusqu'à l'ordre k .

L'inconvénient majeur de ce test est de décider qui l'ordre de décalage (k) à utiliser. Ljung et Box (1978) ont suggéré que le nombre maximum de décalages à utiliser devrait être $T^{1/3}$ où T est le nombre d'observations.

3.2.3 Limites de la méthode Engle-Granger

Bien que la méthode Engle et Granger est facilement mis en IJuvre. il a plusieurs défauts :

- L'estimation de la régression de l'équilibre à long terme exige que le chercheur lieu une variable sur le côté droit comme variable dépendante et utiliser l'autre variable sur le côté gauche comme variable indépendante. Par exemple, dans le cas de deux variables, il est possible d'exécuter la méthode Engle-Granger pour la cointégration en utilisant

les résidus de l'une des deux équations de régression d'équilibre suivante :

$$y_t = \beta_{10} + \beta_{11}x_t + \varepsilon_{1t} \quad (3.22)$$

ou

$$x_t = \beta_{20} + \beta_{21}y_t + \varepsilon_{2t} \quad (3.23)$$

Comme la taille de l'échantillon augmente infiniment grand, la théorie indique que le test de racine unitaire dans la séquence ε_{1t} Devient équivalente à le test de racine unitaire dans la séquence de ε_{2t} Malheureusement, les propriétés de grands échantillons sur lequel ce résultat est dérivé ne peut pas être applicable à la taille des échantillons habituellement disponibles pour les économistes. Dans de nombreux cas, la taille des échantillons disponibles sont plus petites que la taille d'échantillon requise sur laquelle la théorie est basée.

- La procédure d'estimation à deux étapes est basée sur le principe que, quelle que soit le variable qui est choisi pour la normalisation, les mêmes résultats seront atteints si les variables sont échangés.

Dans la pratique, il est possible de constater que l'une régression indique que les variables sont cointégrées, alors que d'inverser l'ordre indique l'absence de cointégration. Par exemple, dans les enquêtes sur la relation entre le revenu et les dépenses, si le revenu est placé sur le côté gauche comme variable dépendante, il est possible de conclure que

les revenus et les dépenses de cointégration. mais l'inverse n'est pas nécessairement vrai. C'est une caractéristique très souhaitable de la procédure, parce que le test de cointégration doit être invariant au choix de la variable sélectionnée pour la normalisation.

- La procédure d'estimation en deux étapes de Engle-Granger repose sur un estimateur en deux étapes. Rappelons que la première étape du procédé d'estimation en deux étapes de pré-réglage de chaque variable pour déterminer l'ordre d'intégration génère série résiduelle $\hat{\varepsilon}_t$ utilisé dans la deuxième étape. Il est utilisé pour estimer l'équation de la forme de régression $\Delta\hat{\varepsilon}_t = \alpha_1\hat{\varepsilon}_{t-1} + \varepsilon_t$ comme dans l'équation (3.17) Ainsi, le coefficient α_1 est obtenu par l'estimation d'une équation de régression en utilisant les résidus d'une autre régression. Ceci implique que toute erreur introduite dans la première étape est réalisée en plus de la deuxième étape, ce qui rend les résultats peu fiables.

L'avantage de la méthode de Engle-Granger est sa facilité de mise en Ouvre. Toutefois, ses résultats dépendent de la façon dont la relation de long terme est spécifiée. Dans certains cas, il peut être plus facile d'identifier quelle variable entre sur la gauche comme la variable dépendante. Par exemple, pour accéder à la relation entre le revenu et les dépenses, il est facile de dire que les dépenses en fonction du revenu. Malheureusement, cela n'est vrai que dans certains cas. Il est donc recommandé d'utiliser le test de cointégration de Phillips-Ouliaris, qui donne les mêmes résultats quelle que soit la variable qui entre comme variable dépendante, ce qui signifie qu'ils sont invariants à la normalisation.

Heureusement, plusieurs méthodes ont été développées qui peuvent éviter ces défauts. Il

s'agit notamment des méthodes Phillip-Ouliaris. et la procédure de Johansen. Ces tests permettent au chercheur de tester des versions restreintes de vecteur (s) de cointégration et la vitesse de paramètres de réglage. Ils comptent beaucoup sur la relation entre le rang d'une matrice et de ses racines caractéristiques.

3.3 Méthode de Phillips-Ouliaris

Dans la section 1.1 il a été montré que la deuxième étape de la méthode d'Engle- Granger est un test ADF appliqué aux résidus de l'équation de long terme. Phillips et Ouliaris introduit deux tests résiduels à savoir, le test du rapport de la variance et la trace statistique multivariée. Le dernier de ces tests a l'avantage qu'il est invariant pour une normalisation, c'est-à-dire, pour n'importe quelle variable est considérée comme étant la variable dépendante. Ces tests sont utilisés de la même manière que les tests de racine unitaire mais les données sont des résidus des régressions de cointégration. Ils sont mis en oeuvre sur des matrices ou séries multivariée et sont tous les deux basés sur les résidus de l'équation vectorielle d'autorégression de première ordre.

$$z_t = \hat{\Pi}z_t + \hat{\varepsilon}_t \quad (3.24)$$

Où z_t est partitionné comme $z_t = (y_t, x_t')$ avec une dimension de x_t égal à $(m = n + 1)$, et la variable scalaire y_t , $\hat{\Pi}$ est un coefficient de régression, et n est égal au nombre de variables à étudiée; et les partitions suivantes de deux conformables Ω et $B(r)$:

$$\Omega = \begin{bmatrix} w_{11} & w'_{12} \\ w_{21} & w_{22} \end{bmatrix}, \quad B(r) = \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix}$$

Nous supposons $\Omega_{22} > 0$ et utiliser le bloc décomposition triangulaire de Ω :

$$\Omega = L'L, \quad L = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix}$$

Avec

$$l_{11} = (w_{11} - w'_{12}\Omega_{22}^{-1}w_{21})^{1/2}, \quad l_{21} = \Omega_{22}^{-1/2}w_{21}, \quad l_{22} = \Omega_{22}^{1/2}$$

Remarque 3.3.1. Notez que $\det\Omega = w_{11,2}\det\Omega_{22}$ est nulle si $w_{11,2} = 0$ (donnée $\Omega_{22} > 0$).

Notez également que nous pouvons écrire $w_{11,2} = w_{11}(1 - \rho^2)$, $\rho^2 = w'_{12}\Omega_{22}^{-1}w_{21}/w_{11}$, où ρ^2 est un coefficient de corrélation au carré.

Lorsque $w_{11,2} = 0$ ($\rho^2 = 1$) alors Ω est singulier et x_t et y_t sont cointégrées. A Vautre extrême lorsqu'il n'y a pas de corrélation entre les innovations de x_t et y_t . on a $\rho^2 = 0$, ρ non singulière, et une régression de y_t sur x_t est Jaux dans le sens de Granger et Newbold (1914).

3.3.1 Le test du rapport de la variance

Le rapport statistique de la variance \hat{P}_u est défini comme

$$\hat{P}_u = \frac{T\hat{w}_{11,2}}{T^{-1} \sum_{t=1}^T \hat{u}_t^2} \quad (3.25)$$

où T nombre d'observation ; \hat{u}_t sont les résidus de l'équation de régression à long terme

$$y_t = \hat{\beta}x_t + \hat{u}_t \quad (3.26)$$

La variance conditionnelle $\hat{w}_{11.2}$ est dérivée de la matrice de covariance $\hat{\Omega}$ de $\hat{\zeta}$, qui est le résidu de l'équation (3.24), et est défini comme étant

$$\hat{w}_{11.2} = \hat{w}_{11} - \hat{w}'_{21}\hat{\Omega}_{22}^{-1}\hat{w}_{21} \quad (3.27)$$

La matrice de covariance $\hat{\Omega}$ a été partitionnée comme

$$\hat{\Omega} = \begin{bmatrix} \hat{w}_{11} & \hat{w}_{21} \\ \hat{w}_{21} & \hat{\Omega}_{22} \end{bmatrix} \quad (3.28)$$

Et est estimée comme

$$\hat{\Omega} = T^{-1} \sum_{t=1}^T \hat{\zeta}'_t \hat{\zeta}_t + T^{-1} \sum_{s=1}^L w_{sl} \sum_{t=1}^T (\hat{\zeta}_t \hat{\zeta}_{t-s} + \hat{\zeta}_{t-s} \hat{\zeta}'_t) \quad (3.29)$$

Avec la fonction de pondération $w_{sl} = (1 - s)/(l + 1)$

Un test du rapport de la variance est un test fondé sur les résidus qui vise à tester l'hypothèse nulle d'absence de cointégration en utilisant les scalaires des racines unitaires appliquées à l'équation (3.26). L'hypothèse nulle est formulée en termes de paramètre de

variance conditionnelle $w_{11.2}$ comme suit :

$$H_0 : w_{11.2} \neq 0 \quad (3.30)$$

contre

$$H_1 : w_{11.2} = 0 \quad (3.31)$$

Conduisant à $\rho^2 = 1$, c-à-d existence d'une cointégration. comme l'a souligné ci-dessus (Remarque (1.2.1)).

(i) Test de Phillips \hat{Z}_α : Régresse $\hat{u}_t = \hat{\alpha}\hat{u}_{t-1} + \hat{k}_t$, et de calculer :

$$\hat{Z}_\alpha = T(\hat{\alpha} - 1) - \left(\frac{1}{2}\right)(S_{Tl}^2 - S_k^2) \left(T^{-2} \sum_{t=2}^T \hat{u}_{t-1}^2 \right)^{-1}$$

Où

$$S_k^2 = T^{-1} \sum_{t=1}^T \hat{k}_t^2,$$

$$S_{Tl}^2 = T^{-1} \sum_{t=1}^T \hat{k}_t^2 + 2T^{-1} \sum_{s=1}^l w_{sl} \sum_{t=s+1}^T \hat{k}_t \hat{k}_{t-s}$$

pour certains choix de la fenêtre de retard tel que $w_{sl} = (1 - s)/(l + 1)$

(ii) Test de Phillips \hat{Z}_t : Régresse $\hat{u}_t = \hat{\alpha}\hat{u}_{t-1} + \hat{k}_t$, et de calculer :

$$\hat{Z}_t = \left(\sum_{t=2}^T \hat{u}_{t-1}^2 \right)^{\frac{1}{2}} \frac{\hat{\alpha} - 1}{S_{Tl}} - \frac{1}{2} (S_{Tl}^2 - S_k^2) \left[S_{Tl} \left(T^{-2} \sum_{t=2}^T \hat{u}_{t-1} \right)^{\frac{1}{2}} \right]^{-1}$$

Avec S_k^2 et S_{Tl}^2 ; comme dans (i).

Par conséquent, le test du rapport de la variance mesure la taille de la variance résiduelle de la régression de cointégration de y_t sur x_t par rapport à $T^{-1} \sum_{t=1}^T \hat{u}_t^2$, contre celle d'une estimation directe de variance conditionnelle de la population de y_t , étant donné x_t par rapport à $w_{11.2}$. S'il y a une cointégration entre les variables, le rapport de la variance devrait se stabiliser asymptotiquement, alors que si une relation parasite (absurde) est présent, il sera reflété dans la régression de cointégration et le rapport de la variance devrait diverger.

3.3.2 La trace statistique multivariée

La statistique de trace multivariée, notée \hat{P}_z est défini comme

$$\hat{P}_z = T \text{tr}(\hat{\Omega} M_{zz}^{-1}) \quad (3.32)$$

avec T est le nombre d'observations, $M_{zz}^{-1} = t^{-1} \sum_{t=1}^T z_t z_t'$ et $\hat{\Omega}$ est estimé dans l'équation (3.29).

Remarque 3.3.2. Notez que \hat{Z}_α et \hat{Z}_t sont construits en utilisant une estimation S_{T1}^2 qui est basé sur les résidus \hat{k}_t à partir de Vautorégression de \hat{u}_t et \hat{u}_{t-1} . Lorsque l'estimation S_{T1}^2 est basée sur les premières différences $\Delta \hat{u}_t$ à la place de \hat{k}_t (comme suggéré par l'hypothèse nulle d'absence de cointégration) nous noterons les tests résultants par Z_α et Z_t . La distinction est importante, car ces tests ont des propriétés très différentes sous l'hypothèse alternative de cointégration.

Remarque 3.3.3. De la même manière, \hat{P}_u et \hat{P}_z , sont construits en utilisant l'estimation de matrice de covariance $\hat{\Omega}$ ce est sur la base des résidus $\hat{\zeta}_t$ du vecteur d'autorégression d'ordre un au l'équation (3.24). Lorsque l'estimation $\hat{\Omega}$ est basé sur les différences premières ΔZ_t on

a pas des résultats au tests de P_u et P_z . Là encore, la distinction est importante puisque \hat{P}_u et \hat{P}_z ont des propriétés différentes dans le cadre du remplacement de celles de P_u et P_z .

Remarque 3.3.4. *Le test du rapport de la variance \hat{P}_u est nouvelle. Sa construction est intuitivement intéressante. \hat{P}_u mesure la taille de la variance résiduelle de cointégration de la régression de y_t sur x_t . à savoir $T^{-1} \sum_{t=1}^T \hat{u}_t^2$, contre celle d'une estimation directe de la variance conditionnelle de population de y_t donné x_t , à savoir $T\hat{w}_{11.2}$.*

Remarque 3.3.5. *La trace statistique multivariée \hat{P}_z est également nouveau. Son appel est similaire à celle de \hat{P}_u . Ainsi, $t\hat{\Omega}$ est une estimation directe de la matrice de covariance de Z_t . tandis que M_{zz} est simplement la matrice d'instant d'échantillon observé. Une dégénérescence du modèle (3.24), tels que la cointégration finalement se manifestent dans le comportement des M_{zz} , et Par conséquent, celle de la statistique \hat{P}_z .*

Une carence de deux méthodes (procédure en deux étapes et Phillips-Ouliaris) est que l'on ne peut estimer une relation de cointégration unique. Cependant, si on a affaire à plus de deux séries de temps, il est possible que plus d'une relation de cointégration existe, ce qui nécessite le recours à des techniques de cointégration vectoriels comme la procédure de Johansen.

3.3.3 Limitation des méthodes Phillips-Ouliaris

Cette méthode ne peut estimer que les relations de cointégration simples. Si on a affaire à plus de deux séries de temps, il est possible que plus d'une relation de cointégration existe. Le remède à cette limitation est d'utiliser la procédure de Johansen.

3.4 La procédure de Johansen

de tests statistiques pour tester les hypothèses sur le nombre de vecteurs cointégrants existants et leur fonctionnement dans le système. L'estimation du nombre de vecteurs cointégrants est important, car la sous ou la surestimation peut avoir de sérieuses conséquences pour l'estimation et l'inférence. La sous-estimation implique l'omission de termes de correction d'erreur empiriquement pertinents, et la surestimation implique que la distribution des statistiques ne sera pas standard.

Johansen a proposé un cadre général pour considérer la possibilité de vecteurs cointégrants multiples, et ce cadre de travail permet également d'aborder les questions de causalité et les tests des hypothèses générales de manière plus satisfaisante. La procédure prend comme point de départ l'autorégression vectorielle (VAR) d'ordre p donné par :

$$X_t = \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \dots + \Pi_p X_{t-p} + \underline{\mu}_t \quad (3.33)$$

où X_t est un $n \times 1$ vecteur de variables qui sont intégrées de même ordre, c'est-à-dire $I(1)$. $\underline{\mu}_t$ est un $n \times 1$ vecteur d'innovations tout en Π_1 Travers Π_p sont $m \times m$ coefficients des matrices. Reparamétrisation de l'équation (3.33) c'est-à soustraire X_{t-1} sur les deux faces, conduit à

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \Gamma_2 \Delta X_{t-2} + \dots + \Gamma_{p-1} \Delta X_{t-p+1} - \Pi X_{t-p} + \underline{\mu}_t \quad (3.34)$$

Où $\Gamma_1 = \Pi_1 - I$, $\Gamma_2 = \Pi_2 - \Gamma_1$, $\Gamma_3 = \Pi_3 - \Gamma_2$ et $\Pi = I - \Gamma_1 - \Gamma_2 - \dots - \Gamma_p$.

Et : I est la matrice identitaire, et la matrice Π détermine la mesure dans laquelle le

Le système est cointégré et est appelé la matrice d'impact.

La procédure de test de Johansen est un test multivarié du rapport de vraisemblance pour un processus autorégressif avec des erreurs Gaussiennes indépendantes.

Si aucun des éléments du X_t sont cointégrés, ils doivent être de zéro. D'autre part, si elles sont cointégrées, toutes les lignes de Π doivent être cointégrées mais pas nécessairement distinctes. C'est parce que le nombre de vecteurs de cointégration distincts dépend du rang de la ligne de Π .

La matrice Π est d'ordre $m \times m$. Si elle est de rang m , c'est-à-dire m nombre linéairement indépendants des lignes ou des colonnes, il constitue une base pour l'espace vectoriel de dimension m . Cela implique que tous les $m \times 1$ vecteurs peuvent être générés comme des combinaisons linéaires de son rang. N'importe lequel de ces combinaisons linéaires des lignes conduirait à stationnante, ce qui signifie que X_{t-p} a des composants fixes si le rang de Π est $r < m$.

Le cœur de la procédure de Johansen est simplement de décomposer Π en deux matrices. Nous pouvons écrire $\Pi = \beta \alpha'$ convenable pour $m \times r$ matrices, β et α . Ici

$$\underline{\alpha}' = \begin{bmatrix} \alpha'_1 \\ \alpha'_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha'_r \end{bmatrix} \quad (3.35)$$

$$\underline{\beta} = \begin{bmatrix} \beta_1, & \beta_2, & \dots, & \beta_r \end{bmatrix} \quad (3.36)$$

Où les rangs de a peuvent être définis comme les r vecteurs cointégrants distincts (les relations de cointégrantes entre les variables non stationnaires). et les rangées de β montrent comment ces vecteurs cointégrants sont mis dans chaque équation du système. Johansen donne une technique du maximum de vraisemblance pour estimer les matrices et délimiter les tests convenables qui nous permettent de tester le nombre de vecteurs cointégrants distincts qui existent, ainsi que de tester les hypothèses sur les matrices. En testant, α nous pouvons tester les restrictions des paramètres sur les propriétés à long terme des données. En testant β , nous pouvons tester la direction de la causalité au sein du modèle.

Un test de l'hypothèse nulle qu'il y a r vecteurs cointégrants est proposé par Johansen. Il

teste si Π dans l'équation (3.34) a un rang plein¹.

Alors $\Pi X_{t-p} = \underline{\beta} \underline{\alpha}' X_{t-p}$ et toutes les combinaisons linéaires de $\underline{\alpha}' X_{t-p}$ sont stationnaires. Il est à noter que nous avons à effectuer le test de l'ADF pour accéder à l'ordre d'intégration de chaque variable avant d'appliquer la procédure de Johansen. La procédure de Johansen estime le VAR soumis $\Pi = \underline{\beta} \underline{\alpha}'$ pour différentes valeurs du nombre r des vecteurs de cointégration. en utilisant l'estimateur du maximum de vraisemblance en supposant $\mu_t \sim iidN(0, \Sigma)$. Son estimation peut donc être réécrite comme,

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \Gamma_2 \Delta X_{t-2} + \dots + \Gamma_{p-1} \Delta X_{t-p+1} - \underline{\beta} \underline{\alpha}' X_{t-p} + \underline{\mu}_t \quad (3.37)$$

Le rang de la matrice Π , noté r , détermine le nombre de relation de cointégration. Par conséquent si le rang de la matrice est $r = 0$, c'est - à - dire que les éléments de la matrice Π sont nuls, l'hypothèse de cointégration est rejetée.

La question est, comment pouvons-nous détecter le nombre de vecteurs de cointégration ?

Johansen a proposé deux tests du rapport de vraisemblance à savoir :

- Le test de la trace
- La valeur propre maximum.

3.4.1 Le test de la trace

Le rang de la matrice Π , noté r , détermine le nombre de relation de cointégration. Par conséquent si le rang de la matrice est $r = 0$, c'est - à - dire que les éléments de la matrice Π

sont nuls, l'hypothèse de cointégration est rejetée.

Johansen propose un test qui permet de déterminer r , en calculant à partir des valeurs propres de Π , une statistique :

$$\lambda_{trace} = -T \sum_{i=r+1}^n \ln(1 - \lambda_i) \quad (3.38)$$

Qui teste l'hypothèse que le nombre de vecteurs cointégrants est au plus r , contre l'alternative qui'il est grand que r .

Avec :

T : nombre d'observations

λ_i : i^{eme} valeur propres de la matrice Π ;

n : nombre de variables ;

r : rang de la matrice ;

3.4.2 La valeur propre maximale

Le test de valeur propre maximale, d'autre part, teste l'hypothèse nulle de r vecteurs de cointégration contre l'hypothèse alternative de $(r + 1)$ vecteurs de cointégration. Son statistique de test est :

$$\lambda_{max} = -T(1 - \lambda_{r+1}) \quad (3.39)$$

où T est la taille de l'échantillon, et λ_i est la i^{ieme} plus grande corrélation canonique.

La méthode de Johansen consiste à estimer la matrice Π et de voir si on peut rejeter des

restrictions impliquées par le rang réduit de Π : si on a N variables endogènes avec une racine unitaire, on peut avoir $N - 1$ relations de cointégration.

Pour déterminer le nombre de relations de cointégration, Johansen en 1988 a proposé un test fondé sur les valeurs propres de la matrice Π issu d'un calcul en deux étapes :

Étape 1 :

(i) estimer le modèle par la méthode du maximum de vraisemblance (MV) : la fonction de vraisemblance dans ce cas est donnée par :

$$\log(\beta, \alpha, \Gamma_1, \dots, \Gamma_{p-1}, \Omega) = -\frac{NT}{2} \log(2\Pi) - \frac{T}{2} \log[\det(\Omega)] - \frac{1}{2} \sum_{t=1}^T \varepsilon_t' \Omega^{-1} \varepsilon_t \quad (3.40)$$

(ii) Après l'obtention des estimateurs de Γ_i pour $i = 1, \dots, P - 1$, nous calculons deux résidus e_{0t} et e_{pt} , par les deux régressions suivantes :

Première régression

$$\Delta X_t = \hat{\varphi}_1 \Delta X_{t-1} + \dots + \hat{\varphi}_p \Delta X_{t-p} + e_{0t} \quad (3.41)$$

Deuxième régression

$$X_{t-p} = \hat{\varphi}_1 \Delta X_{t-1} + \dots + \hat{\varphi}_p \Delta X_{t-p} + e_{pt} \quad (3.42)$$

Avec $X_t = (x_{1,t}, x_{2,t}, \dots, x_{N,t})'$

Nous avons les mêmes variables explicatives, seule la spécification du bloc de la variable à expliquer est modifiée.

e_{0t} et e_{pt} , sont donc les matrices des résidus de dimension (N, T) avec N est le nombre de variables et T le nombre d'observations. La vraisemblance (3.40z) devient dans ce cas :

$$\log(\beta, \alpha, \Omega) = -\frac{NT}{2} \log(2\Pi) - \frac{T}{2} \log[\det(\Omega)] - \frac{1}{2} \sum_{t=1}^T (e_{0t} + \beta\alpha'e_{pt})' \Omega^{-1} (e_{0t} + \beta\alpha'e_{pt})$$

(iii) On maximise cette expression par rapport à β et Ω pour α fixé, on en déduit donc :

$$\hat{\beta}(\alpha) = -S_{0p}\alpha(\alpha'S_{pp}\alpha)^{-1} \quad (3.43)$$

$$\hat{\Omega}(\alpha) = S_{00} - S_{0p}\alpha(\alpha'S_{pp}\alpha)^{-1}\alpha'S_{p0} \quad (3.44)$$

Où $S_{ij} = \frac{1}{T} \sum_{t=1}^T e_{it}e'_{jt}$ Pour $i, j = 0, P$. l'expression de S_{ij} nous donne les matrices de variance-covariance de e_{0t} et e_{pt} .

Etape 2 :

(i) Calculer la matrice Π permettant de calculer les valeurs propres :

$$\Pi = \hat{S}_{pp}^{-1} \hat{S}_{p0} \hat{S}_{00}^{-1} \hat{S}_{0p} \quad (3.45)$$

(ii) Nous extrayons les N valeurs propres de la matrice de dimension (N, N) .

A partir de ces valeurs propres. Nous calculons la statistique (3.38) :

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^n \ln(1 - \lambda_i)$$

Cette statistique suit une loi de probabilité tabulée par Johansen et Juselius (1990). Ce test fonctionne par exclusion d'hypothèse alternative :

1. Test $H_0 : r = 0$ contre $H_1 : r > 0$.

Test de l'hypothèse aucune relation de cointégration contre au moins une relation.

Si $\lambda_{trace}(0)$ est supérieur à la valeur lue dans la table au seuil α pour cent, on rejette H_0 . il existe au moins une relation, on passe alors à l'étape suivante, sinon on s'arrête et $r = 0$.

2. Test $H_0 : r = 1$ contre $H_1 : r > 1$.

Test de l'hypothèse aucune relation de cointégration contre au moins une relation.

Si $\lambda_{trace}(1)$ est supérieur à la valeur lue dans la table au seuil α pour cent, on rejette H_0 . il existe au moins une relation, on passe alors à l'étape suivante, sinon on s'arrête et $r = 1$.

suite jusqu'à la dernière étape (si elle est nécessaire) :

3. Test $H_0 : r = N - 1$ contre $H_1 : r > N - 1$.

Test de l'hypothèse $N - 1$ relations de cointégration contre au moins $N - 1$ relations.

Si $\lambda_{trace}(N - 1)$ est supérieur à la valeur lue dans la table au seuil α pour cent, on rejette H_0 , il existe N relations (en fait dans ce cas les N variables sont $I(0)$ si non $r = N - 1$).

3.4.3 Limitation de la procédure de Johansen

Cette méthode suppose que le vecteur de cointégration reste constant au cours de la période d'étude. En réalité, il est possible que les relations de long terme entre les variables sous-jacentes changent. La raison en est peut-être le progrès technologique, la crise économique, des changements dans les préférences et les comportements des gens en conséquence, la politique ou le régime altération et le développement institutionnel. Ceci est particulièrement le cas si la période d'échantillonnage est longue.

Pour en tenir compte, Gregory et Hasen (1996) ont mis en place des tests de cointégration avec un et deux ruptures structurelles inconnu (s). Cependant, ces tests ne font pas partie de la présente étude.

Par ailleurs, étant donné l'extrême sensibilité de ces procédures à la structure d'autocorrélation des erreurs et l'absence d'une procédure de sélection réellement efficace du paramètre de troncature, la faiblesse de la puissance des tests et les distorsions de niveau, il paraît nécessaire de chercher systématiquement à évaluer la robustesse des résultats obtenus en appliquant plusieurs d'entre elles et en essayant dans la mesure du possible de ne pas imposer des spécifications trop restrictives sur la composante déterministe ni des hypothèses trop restrictives sur les erreurs.

L'application des tests sur le modèle le plus général possible suppose le recours à une approche heuristique qui requiert, avant tout, l'examen du graphe de la série considérée pour détecter toute forme de non-linéarité et d'éventuels changements structurels majeurs. Cette première étape peut éventuellement être suivie par l'application des procédures de test de Ouliaris, Park et Phillips (1989) ou Perron (1989) ou Zivot et Andrews (1992), Banerjee,

Lumsdaine et Stock (1992) ou encore Perron (1997) qui intègrent de manière formelle la spécification polynomiale ou linéaire par morceaux de la composante déterministe.

Chapitre 4

Causalité au sens de Granger(Cas indices économique Algeria)

Ce travail traite une simulation de la causalité au sens de Granger de deux séries(indices économiques Algeria). Ce travail fait l'objet d'un article soumis

The Granger Causality And Co-Integration Analysis(case of economics indexes in algeria)

A. Hamlili, Z.Chikr El mezouar and M.K. Attouche

ABSTRACT

This paper seeks to investigate the causal relationship between energy consumption and electricity production in Bechar south Algeria over a period from 2003 to 2009. The empirical analysis starts by analysing the time series properties of the data which is followed by examining the nature of causality among the variables. This study analyzes that, how

change in real consumption effects the real production of Bechar south algeria and many other factors differently. The results show that both series are integrated of order one (I(1)).

4.1 Introduction

Electric energy in Algeria plays an essential role in economy in both demand and supply. On the demand side electric energy is one of the products that the consumer decides to buy to maximize his or her utility. On the supply side, electric energy is a key factor of production in addition to capital labour and materials and is seen to play a vital role in the economic and social development of countries. This implies that there should be a causal relationship running from energy consumption or electricity energy production as well as vice versa. In particular, the remaining paper is organized in the following way, section II present the econometric methodology, section III contains empirical results and discussions, and finally conclusions are drawn in section IV.

4.2 Methodology

Granger causality tests. Several studies have been devoted to the study of causality between variables (Granger 1969, Sims 1972). Furthermore, we carried out the Granger causality test where Granger (1969) proposed a time series data based approach in order to determine causality. For example if we want to explore the causal relationship between electricity consumption (x_t) and electricity production (y_t) in Bechar (ALGERIA)

$$x_t = \sum_{i=1}^n \alpha_i y_{t-i} + \sum_{i=1}^n \beta_i x_{t-i} + \varepsilon_{1t} \quad (4.1)$$

$$y_t = \sum_{i=1}^n \lambda_i x_{t-i} + \sum_{i=1}^n \delta_i y_{t-i} + \varepsilon_{2t} \quad (4.2)$$

when n the number of lags.

4.3 Causality test and cointegration variables.

The relationship causality between different time series is based as following :

4.3.1 Unit root tests.

A Stochastic process is stationnary if its first and second moments are constant.

Analytically, y_t is stationnary if

$$E(y_t) = \mu, \quad \forall t \quad (4.3)$$

$$cov(y_t, y_{t+k}) = \delta(h), \quad \forall t \quad (4.4)$$

Dickey-Fuller (DF) proposed a basic models of a unit root test

$$(1) \Delta y_t = (\phi - 1)y_{t-1} + \varepsilon_t$$

$$(2) \Delta y_t = (\phi - 1)y_{t-1} + c + \varepsilon_t$$

$$(3) \Delta y_t = (\phi - 1)y_{t-1} + c + \beta t + \varepsilon_t$$

The hypotheses tests are

$$\begin{cases} H_0 : \phi = 1 \\ H_1 : \phi < 1 \end{cases}$$

To get a broader view, Dickey-Fuller took an autoregressive process of higher order known as the augmented Dickey-Fuller (ADF). This test is represented as a following

$$\begin{aligned}\Delta y_t &= (\phi - 1)y_{t-1} + \sum_i \theta_i \Delta y_{t-i} + \varepsilon_t \\ \Delta y_t &= (\phi - 1)y_{t-1} + \sum_i \theta_i \Delta y_{t-i} + \beta + \varepsilon_t \\ \Delta y_t &= (\phi - 1)y_{t-1} + \sum_i \theta_i \Delta y_{t-i} + \beta + \partial t + \varepsilon_t\end{aligned}$$

4.4 Cointegration.

The main objective of this research is to assess not only the pairwise nature of the causality among the variables, but, also from a dynamic impact as well in using we test for cointegration using two well known approaches : the one developed by Engle and Granger (1987) and the other one by Johansen (1988). Engle - Granger Method : The Engle-Granger test is a procedure that involves an OLS estimation of a pre-specified cointegrating regression between the variables. This was followed by a unit root test performed on the regression residuals previously identified. We applied the Engle- Granger, two-step procedure

Step 1 : Static regression between integrated variables.

Step 2 : Test to verify the residual stationary.

4.5 An error correction model.

The model error corrections reads :

$$\Delta x_t = \alpha_1 z_{t-1} + \text{lagged}(\Delta x_t, \Delta y_t) + \varepsilon_{1t} \quad (4.5)$$

$$\Delta x_t = \alpha_2 z_{t-1} + \text{lagged}(\Delta x_t, \Delta y_t) + \varepsilon_{2t} \quad (4.6)$$

with z_t the error correction term to resulting from estimating the cointegration relationship,

" is the error term stationnary $|\alpha_1| + |\alpha_2| \neq 0$

4.6 Causality test.

The causality test based on the model vector corection has the advantage of providing a causal relationship even if no estimated coefficient of lagged variables used is sign.cant. Thus, an error correction model after precessing can be rewritten as following equations

$$\Delta x_t = \alpha + \sum_{i=1}^k \lambda_i \Delta x_{t-i} + \sum_{i=1}^k \sigma_i \Delta y_{t-i} + \theta z_{t-1} + \varepsilon_t \quad (4.7)$$

$$\Delta y_t = \beta + \sum_{i=1}^k \phi_i \Delta y_{t-i} + \sum_{i=1}^k \varphi_i \Delta x_{t-i} + \psi z_{t-1} + \mu_t \quad (4.8)$$

From these both equations, x_t does not cause y_t the sense of Granger if $\varphi_i = \psi = 0$, y_t does not cause x_t if $\sigma_i = \theta = 0$.

4.7 Empirical results and interpretation

4.7.1 Statistical data proprieties.

The variables that we used in our application are the consumption (x_t) and the production (y_t):

Null Hypothesis: PRODSA has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic based on SIC, MAXLAG=12)		
	t-Statistic	Prob*
<hr/>		
Augmented Dickey-Fuller test statistic	-0.811759	0.8119
Test critical values: 1% level	-3.486064	
5% level	-2.885863	
10% level	-2.579818	
<hr/>		
*MacKinnon (1996) one-sided p-values.		

Table 1 : Test of stationnary for production

Null Hypothesis: CONSSA has a unit root		
Exogenous: Constant		
Lag Length: 1 (Automatic based on SIC, MAXLAG=12)		
	t-Statistic	Prob*
<hr/>		
Augmented Dickey-Fuller test statistic	-1.371915	0.5939
Test critical values: 1% level	-3.486551	
5% level	-2.886074	
10% level	-2.579931	
<hr/>		
*MacKinnon (1996) one-sided p-values.		

Table 2 : Test of stationnary for consumption

Table 1 and 2 presents the test results for stationnary of PROD and CONS. The results showed that the two variables were non stationnary.

4.8 Unit root tests

Table 1 represents the results of unit root tests. The results show that all the series are non stationary at level. Taking the variables in their .rst difference results show that all are I(1) at 1 present level of signi.cance. For consistenary, therefore all the series were considired as I(1) and taken at their .rst difference in the analyses

Null Hypothesis: D(PROD) has a unit root Exogenous: Constant Lag Length: 1 (Automatic based on SIC, MAXLAG=12)		t-Statistic	Prob*
Augmented Dickey-Fuller test statistic		-9.386001	0.0000
Test critical values:	1% level	-3.487046	
	5% level	-2.886290	
	10% level	-2.580046	
*MacKinnon (1996) one-sided p-values.			

Table 3 : Test of stationnary for the difference of PROD

Null Hypothesis: D(CONS) has a unit root Exogenous: Constant Lag Length: 0 (Automatic based on SIC, MAXLAG=12)		t-Statistic	Prob*
Augmented Dickey-Fuller test statistic		-13.69481	0.0000
Test critical values:	1% level	-3.486551	
	5% level	-2.886074	
	10% level	-2.579931	
*MacKinnon (1996) one-sided p-values.			

Table 4 : Test of stationnary for the difference of CONS

Table 3 and 4 presents the test results for stationnary of difference production and consumption. The results showed that the two variables were stationnary.

Null Hypothesis: RES has a unit root Exogenous: Constant Lag Length: 0 (Automatic based on SIC, MAXLAG=12)		t-Statistic	Prob*
Augmented Dickey-Fuller test statistic		-11.02460	0.0000
Test critical values:	1% level	-3.487046	
	5% level	-2.886290	
	10% level	-2.580046	
*MacKinnon (1996) one-sided p-values.			

Table 5 : Test of stationnary for the residual series

Table 5 presents the test results for stationnary of the residual series.

Vector Error Correction Estimates		
Date: 07/08/13 Time: 11:31		
Sample (adjusted): 2000M03 2009M12		
Included observations: 118 after adjustment		
Statistic		
Standard errors in () & t-statistics in [] -3.486551		
Cointegrating Eq: CointEq1		
PROD(-1)	1.000000	
CONS(-1)	-448.9530	
	(21.0664)	
	[-21.3113]	
C	7.66E+08	
Error Correction:	D(PROD)	D(CONS)
CointEq1	0.437198	0.001863
	(0.15445)	(0.00046)
	[2.83075]	[4.08236]
D(PROD(-1))	-0.262695	0.000323
	(0.24244)	(0.00072)
	[-1.08357]	[0.45116]
D(CONS(-1))	153.0013	0.128658
	(82.4388)	(0.24360)
	[1.85594]	[0.52816]
C	13601053	24168.81
	(2.1E+07)	(62201.5)
	[0.64611]	[0.38856]

Table 6 : Johansen cointegration test

6 presents the test results for the number of cointegrating vectors.

4.9 Johansen cointegration tests.

Table 6 presents the test results for the number of cointegrating vectors. The results show that the trace statistic suggests the presence of one cointegration equation among the two variables at 5 percent level

4.10 An error correction model estimate

Next we apply the Johansen procedure to obtain the long run coefficients of model. Table 6 presents the normalized coefficient of the variables in the model. The coefficient was correctly signed and statistically significant at 1 percent level. Thus, we can write the cointegration equation from the above results

$$PROD = -448,9530 + 7,66E + 08CONS$$

4.11 Granger causality test

Causality tests between consumption and production by the vector error correction model shows the existence of unidirectional causality from production to consumption.

4.12 Pairwise Granger causality test

Table 7 presents the results of pairwise Granger causality among the production and consumption.

Pairwise Granger Causality Tests			
Date: 10/02/13 Time: 09:02			
Sample: 2000M01 2009M12			
Lags: 2			
Null Hypothesis:	obs	F-Statistic	Probability
CONS does not Granger Cause PROD	118	1.66994	0.19287
PROD does not Granger Cause CONS		4.34406	0.01522

Table 7 : Test of causality

4.13 Conclusion

This paper employs an empirical analysis to examine of production fluctuations on the level of real electric energy in Bechar south Algeria using a sample of observations from 2003 to 2009. The first step in the empirical analysis in values testing the time series characteristics of the data series using ADF test and running the pairwise Granger causality test. This was followed by applying the Johansen cointegration test. The analysis was capped with the estimation of short run vector error correction model (VECM). Finally the Granger pairwise causality test showed that the null hypotheses that consumption do not Granger cause electricity production. In other words electricity production Granger cause energy consumption in Bechar (Algeria). The results from the short run VECM showed the coefficient is correctly signed and statically significant. This the electricity that long run equilibration conditions influences the short run dynamics. Electricity production in Bechar (Algeria) has an automatic adjustment mechanism of the energy consumption.

Finally, as a conclusion, we tried to give the same challenges a statistico-economic analysis to the problem cited above and this course in collaboration with our fellows economists

of our research laboratory in the University of Bechar in southern Algeria.

In addition , we are very confident that this causal study will take us a very interesting econometric about the linearity condition and nonlinearity models.

In the end , in the outlook, we will consider a very serious discussion to solve this kind of problem.

4.14 Conclusion

L'objectif de la présente étude était d'utiliser certains développements de la causalité dans le champ d'économétrie des séries temporelles non stationnaires, notamment la théorie de cointegration pour explorer la liaison causale entre deux séries chronologiques , particulièrement notre cas sur la production et la consommation d'électricité en Algérie .

Empiriquement , l'application de cette théorie nécessité la démarche suivante (Jacquinot, 1989) de tester l'ordre d'intégration des séries (Test de racine unitaire)pour s'assurer qu'elles suivent une marche aléatoire (seul domaine d'application du théorème de représentation de Granger)

- d'utiliser le modèle a correction d'erreur qui vise a rendre compte et révèle que la production d'électricité <cause> au sens de Granger la consommation d'électricité en Algérie

Notre domaine de recherche reste ouvert a des perspectives surtout sur le plan d'économétrie , à savoir la signifiance et l'efficacité du modèle reste liée a la nature des séries en question la linéarité ou non linéarité du processus étudié notamment pour nos collègues praticiens

économistes qui ne donnent pas une grande importance au sujet de linéarité ou non

Bibliographie

- [1] **Akarca, A.T., Long, T.V.**, (1979). On the relationship between Energy and GNP : are-examination, *Journal of Energy and Development*, Vol 5, *n*2, pp 326-331
- [2] **Aqeel, A., Butt M.S.**, (2001). The relationship between energy consumption and economic growth in Pakistan. *Asia-Pacific Development Journal*. Vol. 8, December pp 101-110.
- [3] **Antille, G.**, (1984). Utilisation des tableaux entrées-sorties pour l'analyse des interdépendances entre économie et énergie. *Revue Suisse d'Economie Politique et de Statistique*. Vol 120, *n* 3, pp 261-279.
- [4] **Berrah, N.E.**, (1983). Energie et développement : l'effet tunnel. *Revue de l'Energie*, *n* 356, août-septembre, pp 409-415.
- [5] **Bruneau, C.**, (1996). Analyse économétrique de la causalité : un bilan de la littérature. *Revue d'Economie Politique*. Vol 106, *n*3, mai-juin, pp 325-353
- [6] **Bruneau, C., Nicolai, J-P.**, (1989). Une étude de causalité. Concepts et applications. L'exemple du marché pétrolier. Document de travail *n* 1989-21. Banque de France.
- [7] **Campbell, J.Y., Perron, P.**, (1991). Pitfalls and opportunities : what macroeconomists should know about unit roots. *NBER Macroeconomics Annual*, Vol 6 pp 141-201.

- [8] **Darne, O., Terraza, M.**, (2002). Les tests de non-stationnarité des séries temporelles non saisonnières en économie : une revue de littérature. Document de Travail *n* 2002-05. Laboratoire Montpelliérain d'Économie Théorique et Appliquée. Université de Montpellier
- [9] **Davidson, J.E.H., Hendry, D.F., Srba, F., Yeo, S.**, (1978). Econometric modelling of the aggregate time series relationship between consumer's expenditure in the U.K. *Economic Journal*. Vol 88, pp 661-692.
- [10] **Destais, G.**, (1989). La modélisation des interactions énergie-économie. Thèse de Doctorat. IEPE, Grenoble.
- [11] **Dickey, D., Fuller, W.**, (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, Vol. 74, *n* 366, Part 1, June, pp 427-43.
- [12] **Dickey, D., Fuller, W.**, (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*. Vol 49, pp 1057-1072
- [13] **Enders, W.**, (2004). *Applied Econometric Time Series*. Wiley
- [14] **Engle, R.F., Granger, C.W.J.**, (1987). Co-integration and error correction : representation, estimation and testing. *Econometrica*, Vol 55, *n*2, March, pp 251-276.
- [15] **Engle, R.F., Yoo, B.S.**, (1987). Forecasting and testing in cointegrated systems. *Journal of Econometrics*. Vol. 35, pp 143-159
- [16] **Ertur, K.C.**, (1998). Méthodologie des tests de racine unitaire. Document de Travail *n*9813, LATEC, Université de Bourgogne.

- [17] **Fatai, K., Oxley, L., Scrimgeour F.**, (2002). Energy consumption and employment in New Zealand : searching for causality. Paper presented at NZAE Conference, Wellington, 26-28 June
- [18] **Geweke, J., Meese, R., Dent, W.**, (1983). Comparing alternative tests of causality in temporal systems : analytic results and experimental evidence. *Journal of Econometrics*. Vol 21, pp 161-194.
- [19] **Girod, J., Percebois, J.**, (1985). Energie : Fonctions de production et modèles de demande. Dans *Énergie : Modélisation et Économétrie*. *Economica*.
- [20] Glasure, Y.U., (2002). Energy and national income in Korea : further evidence on the role of omitted variables. *Energy Economics*. Vol. 24, *n4*, pp 355-365.
- [21] **Gourieroux, C., Monfort, A.**, (1990). Séries temporelles et modèles dynamiques. *Economica*
- [22] **Granger, C.W.J.**, (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*. Vol 37, *n3*, July, pp 424-438.
- [23] **Granger, C.W.J.**, (1986). Developments in the study of cointegrated variables. *Oxford Bulletin of Economics and Statistics*. Vol 48. pp 213-228.
- [24] **Granger, C.W.J.**, (1991), Developments in the study of cointegrated economic variables. In *Long-run economic relation Readings in cointegration* (Eds.) Engle RF, Granger C.W.J., Oxford University Press, Oxford, 65-80.
- [25] **Granger, C.W.J., Newbold, P.**, (1974). Spurious regressions in econometrics. *Journal of Econometrics*, Vol 2. pp 111-120

- [26] **Granger, C.W.J., Weiss, A.A.**, (1983). Time series analysis of error correcting models. Studies in Econometrics, Time Series and Multivariate Statistics, New York : Academic Press, pp 255-278.
- [27] **Guttormsen, A.G.**, (2004). Causality between energy consumption and economic growth. Discussion Paper D-24/2004. Department of Economics and Resource Management. Agricultural University of Norway.
- [28] **Hendry, D.F.**, (1986). Econometric modelling with cointegrated variables : an overview. Oxford Bulletin of Economics and Statistics. Vol 48. pp 201-212.
- [29] **Hondroyannis, G., Lolos, S., Papapetrou, E.**, (2002). Energy consumption and economic growth : assessing the evidence from Greece. Energy Economics. Vol. 24. pp319-336.
- [30] **Houcarde, J.Ch., Ben Chaabane, N.**, (1991). Les liens énergie croissance dans les pays en développement (1960-1985) : propositions méthodologiques pour une formalisation des liens énergie-modes de développement. Dans *Politiques Economiques et Marchés internationaux de matières premières*. Economica.
- [31] **Jacquinot, P.**, (1989). Une approche économétrique à la mode : la cointégration. S.E.E.R. 89-41 Direction Générale des Etudes, Banque de France.
- [32] **Jobert, T.**, (1992). Test de racine unitaire : une stratégie et sa mise en IJuvre. Cahiers Eco Maths *n* 92.44, MAD Université de Paris I, Pantheon-Sorbonne.
- [33] **Johansen, S.**, (1988). Statistical analysis of cointegrating vectors. Journal of Economic Dynamics and control. Vol 12. pp 231-254.

- [34] **Johansen, S.**, (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vectors autoregressive models. *Econometrica*. Vol 59, pp 1551-1580.
- [35] **Johansen, S.**, (1995). Likelihood-based inference in cointegrated vector autoregressive models. Oxford University Press.
- [36] **Johansen, S., Juselius, K.**, (1990). Maximum likelihood estimation and inferences on cointegration with application to the demand for money. *Oxford Bulletin of Economics and Statistics*. Vol 52, pp 160-210.
- [37] **Jumbe, C.B.L.**, (2004). Cointegration and causality between electricity consumption and GDP : empirical evidence from Malawi. *Energy Economics*. Vol 26, pp 61-68.
- [38] **Kraft, J., Kraft, A.**, (1978). On the relationship between energy and GNP. *Journal of Energy and Development*. pp 401-403.
- [39] **Lardic, S., Mignon, V.**, (2002). *Econométrie des séries temporelles macroéconomiques et financières*. Economica.
- [40] **Martin, J-M.**, (1992). *Economie et politique de l'énergie*. Armand Colin
- [41] **Masih, A.M.M., Masih, R.**, (1996). Energy consumption, real income and temporal causality : results from a multi-country study based on cointegration and errorcorrection modeling techniques. *Energy Economics*. Vol 18, pp 165-183.
- [42] **Masih, A.M.M., Masih, R.**, (1997). On the temporal causal relationship between energy consumption, real income and prices : some new evidence from Asian-energy dependent NICs based on a multivariate cointegration/vector error correction approach. *Journal of Policy Modeling*. Vol 19, pp 417-440.

- [43] **Masih, A.M.M., Masih, R.**, (1998). A multivariate cointegration modeling approach in testing temporal causality between energy consumption, real income and prices with an application to two Asian LDCs. *Applied Economics*. Vol 30, pp 1287-1298.
- [44] **Mouttet, F., Plateau, C., Brillet, J.L., Morand, J.P.**, (1983). Mini-DMS-énergie. *Modèle des interactions économie-énergie*. Archives et Documents *n* 74, INSEE
- [45] **Nachane, D.M., Nadkarni, R.M., Karnik, A.V.**, (1988). Co-integration and causality testing of the energy-GDP relationship : a cross-country study. *Applied Economics*. Vol 20, pp 1511-1531.
- [46] **Nelson, C.R., Plosser, C.I.**, (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*. Vol 10, pp 139-162.
- [47] **Njiki, R.P.**, (1998). Cointégration et test d'efficacité sur les marchés dérivés. *Journal de la Société Statistique de Paris*, Vol 139, pp 35-59.
- [48] **Oh, W., Lee, K.**, (2004). Causal relationship between energy consumption and GDP revisited : the case of Korea 1970-1999. *Energy Economics*. Vol 26, pp 51-59.
- [49] **Osterwald-Lenum, M.**, (1992). A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics*, Vol 54, pp 461-471.
- [50] **Perron, P.**, (1993). Racines unitaires en macroéconomie : le cas d'une variable. In *Macroéconomie : développements récents*, Malgrange P. et Salvas-Bronsard L. eds., pp 327-358, *Economica*.
- [51] **Perron, P., Campbell, J.Y.**, (1992). Racines unitaires en macroéconomie : le cas multidimensionnel. *Annales d'Économie et de Statistique*. Vol 27, pp 1-50.

- [52] **Phillips, P., Perron, P.**, (1988). Testing for a unit root in time series regression. *Biometrika*, Vol 75, pp 335-346.
- [53] **Pizzaro Rios, J.V.**, (1993). The links between money and prices. Causality tests for the Peruvian and Chilean cases. *Cahiers du Département d'Économétrie n 93.11 Faculté des sciences économiques et sociales. Université de Genève.*
- [54] **Salmon, M.**, (1982). Error correction mechanisms. *The Economic Journal*, Vol 92, pp 615-629.
- [55] **Sekkat, K.**, (1989). L'analyse de causalité comme méthode de détermination des filières industrielles. *Annales d'Économie et de Statistique*, Vol 14, pp 191-224.
- [56] **Sims, C.A.**, (1972). Money, income and causality. *American Economic Review*. Vol 62, pp 540-552.
- [57] **Sims, C.A.**, (1980). Macroeconomics and reality. *Econometrica*, Vol 48, pp 1-48
- [58] **Yu, E., Choi, J.**, (1985). The causal relationship between energy and GDP, an international comparison. *Journal of Energy and Development*. Vol 10, pp 249-272.