Un Test Simple de l'Hypothèse de Non Causalité dans un Modèle de Panel Hétérogène*

Christophe Hurlin†

Ce papier propose un test simple de l'hypothèse de non causalité au sens de Granger (1969) dans un modèle de panel hétérogène avec coefficients fixes. Nous proposons une statistique de test fondée sur la moyenne des statistiques de Wald associées aux tests individuels de non causalité. Cette statistique converge séquentiellement vers une distribution normale. Pour une taille T fixée, une seconde statistique standardisée utilisant une approximation des moments des statistiques de Wald individuelles est proposée. Des simulations montrent que ces deux statistiques possèdent de bonnes propriétés sur des échantillons de faible dimension temporelle.

Testing Granger Causality in Heterogeneous Panel Data Models

This paper proposes a simple test of Granger (1969) non causality hypothesis in heterogeneous panel data models with fixed coefficients. It proposes a statistic of test based on averaging standard individual Wald statistics of Granger non causality tests. First, this statistic is shown to converge sequentially to a standard normal distribution. Second, for a fixed T sample the semi-asymptotic distribution of the average statistic is characterized and a standardized statistic based on general approximations of the moments of individual Wald statistics is proposed. Monte Carlo experiments show that both statistics provide good small sample properties.

JEL Classification : C23, C11

L'intérêt d'étendre les procédures de tests de causalité à des modèles de panel est largement reconnu (Granger, 2003). Sur le plan économique, les grandes problématiques associées à la notion de causalité, comme par exemple la relation entre l'activité réelle et la monnaie (Sims, 1972), ne présentent généralement pas de dimension nationale spécifique. Bien au contraire, l'idée de tester ces relations de causalité dans un contexte international, régional ou sectoriel, est d'autant plus attrayante qu'elle ne fait

*Je tiens à remercier Anindya Banerjee, Baptiste Venet, Gilbert Colletaz et Pierre-Yves Hénin pour leurs commentaires sur une version antérieure de ce travail réalisée en grande partie au CEPREMAP et à EURIsCO, Université Paris IX Dauphine.

†Laboratoire d'Economie d'Orléans (LEO). Université d'Orléans. Rue de Blois, BP 6739, 45067 Orléans Cedex 2. Email: christophe.hurlin@univ-orleans.fr.

que renforcer leur robustesse et leur caractère général. Sur le plan statistique, l'utilisation de données de panel doit permettre d'augmenter l'ensemble d'information et ainsi d'améliorer les propriétés des statistiques de tests notamment dans le cas d'échantillon de faible dimension temporelle. Toutefois, cette assertion pose le problème fondamental de l'hétérogénéité de la relation de causalité.

Dans ce papier, nous proposons un test très simple à mettre en oeuvre de l'hypothèse de non causalité au sens de Granger (1969) dans un modèle de panel hétérogène. Plus précisément, nous testons l'hypothèse de Non Causalité Homogène (NCH) d'une variable x vers une variable y dans un système bivarié. Sous l'hypothèse nulle, on suppose qu'il n'existe aucune relation de causalité au sens de Granger de x vers y pour les N individus de l'échantillon. Toutefois, notre hypothèse alternative n'implique pas nécessairement une relation de causalité pour tous les individus du panel comme dans le cas du modèle homogène de Holtz-Eakin, Newey et Rosen (1988). En effet, nous retenons une spécification hétérogène de l'alternative dans laquelle peuvent coexister deux sous groupes d'individus : un premier sous groupe d'individus pour lesquels il existe une relation de causalité de x vers y et un second sous groupe d'individus pour lesquels, au contraire, il n'existe pas de relation de causalité. La taille respective de ces deux ensembles est a priori inconnue. Enfin, que ce soit sous l'hypothèse nulle ou l'hypothèse alternative, tous les paramètres non contraints sont supposés a priori hétérogènes. On se place ainsi dans le cadre des modèles de panel dynamiques hétérogènes.

Notre démarche consiste alors à transposer la méthodologie adoptée dans la littérature sur les tests de racine unitaire dans les panels hétérogènes, et plus particulièrement dans le test bien connu d'Im, Pesaran et Shin (2003), au problème du test de l'hypothèse de non causalité. Il s'agit ici de combiner l'information de N tests individuels supposés indépendants. Dans ce papier, nous proposons ainsi une statistique de test fondée sur la moyenne des statistiques de Wald individuelles associées au test de l'hypothèse de non causalité. Cette statistique converge séquentiellement en distribution vers une loi normale lorsque la dimension T tend vers l'infini, puis la dimension N tend vers l'infini. Lorsque la dimension T est fixée, sa loi semi-asymptotique demeure normale. En effet dans ce cas, les statistiques de Wald individuelles ne possèdent plus une distribution standard du chi-deux. Toutefois, si ces statistiques individuelles sont indépendamment distribuées avec des moments d'ordre deux finis, la statistique moyenne continue à être distribuée selon une loi normale. La difficulté consiste alors à déterminer les moments de cette distribution qui dépendent des moments des statistiques de Wald individuelles, dont la distribution est inconnue. Deux solutions peuvent être retenues : la première consiste à simuler les moments pour différentes tailles d'échantillons et différentes configurations sur les paramètres. La seconde solution consiste à proposer une approximation de ces moments. C'est la voie que nous retenons dans le cadre de ce papier. Ainsi, nous proposons finalement deux statistiques standardisées : la première est fondée sur les moments de la distribution asymptotique des statistiques individuelles, la seconde est fondée quant à elle sur une approximation des moments de la distribution à distance finie. Des simulations de Monte Carlo montrent que ces deux statistiques possèdent de bonnes propriétés dans de petits échantillons. Dans les configurations étudiées, la puissance des tests en panel est notamment supérieure à celle obtenue à partir des tests menés sur séries temporelles.

L'article est organisé de la manière suivante. La première section est consacrée à la présentation du test de l'hypothèse nulle de non causalité homogène et à la définition de la statistique moyenne associée. Dans la deuxième section, nous caractérisons les distributions asymptotique et semi-asymptotique de cette statistique. Dans la troisième section, nous présentons quelques résultats de simulations de Monte Carlo.

1 UN TEST DE NON CAUSALITE DANS UN PANEL HETEROGENE

Considérons deux variables stationnaires, notées x et y, observées sur T périodes et N individus. Pour chaque individu i et chaque période t, on considère le modèle suivant :

$$y_{it} = \alpha_i + \sum_{k=1}^{K} \gamma_{ik} y_{i,t-k} + \sum_{k=1}^{K} \beta_{ik} x_{i,t-k} + \varepsilon_{it} \tag{1}$$

où $K \in \mathbb{N}^*$. On suppose que les effets individuels α_i sont fixes. Ce modèle appelle deux principales remarques. Tout d'abord, nous autorisons ici les paramètres β_{ik} et γ_{ik} a différer selon les individus. Cette hypothèse est primordiale car dans la plupart des applications macroéconomiques il serait totalement fallacieux de vouloir imposer le même ensemble de paramètres à un échantillon de pays, de secteurs, de régions etc. dans le but de tester une notion de causalité à la Granger d'une variable x vers une variable y. Dans de nombreuses applications, les tests d'homogénéité[1] ne permettent pas d'accepter l'hypothèse d'un modèle homogène. Or, on comprend aisément que si le modèle est effectivement hétérogène et que l'on impose à tort l'homogénéité des paramètres, les statistiques de tests ainsi obtenues risquent de conduire à des conclusions erronées. De plus, l'intégration des seuls effets individuels α_i ne suffit généralement pas à capter toute l'hétérogénéité de la relation dynamique entre deux variables économiques. C'est pourquoi nous privilégions dans ce papier une structure de panel dynamique hétérogène. Toutefois, contrairement aux travaux de Weinhold (1996), Nair-Reichert et Weinhold (2001), nous supposons que les paramètres β_{ik} et γ_{ik} sont fixes. Ainsi, il convient de distinguer notre modèle d'un modèle à coefficients aléatoires (Swamy, 1970) ou d'un modèle à coefficients fixes et aléatoires (Hsiao, 1989).

La deuxième remarque concerne l'ordre des retards K, supposé identique pour tous les individus du panel. Cette hypothèse nous permet de dériver très simplement les distributions asymptotiques de notre statistique. En outre, cette hypothèse se justifie par le fait que les résultats des tests de causalité en séries temporelles sont généralement très sensibles au choix des retards. Or dans un modèle de panel hétérogène, si l'on suppose une structure de retards propre à chaque individu, la détermination de K_i se fera alors uniquement sur la base de l'information temporelle, qui dans de nombreuses applications en panel risque d'être limitée. Tester la causalité en panel dans un modèle sans corrélation inter-individuelle comme le nôtre n'a d'intérêt que lorsque l'on ne dispose que d'un échantillon de faible dimension temporelle. Dès lors, il nous paraît préférable d'étudier la sensibilité des résultats au choix de K plutôt que de recourir à

[1] Voir Hsiao et Pesaran (2004) pour une revue des modèles à coefficients aléatoires et des tests d'homogénéité dans des modèles de panel dynamiques.

l'utilisation de critères peu informatifs pour choisir le nombre de retard pour chaque individu. Signalons enfin que notre approche peut être étendue assez facilement à une structure de retards hétérogène dans le cas asymptotique.

Dans le modèle (1), on impose les contraintes suivantes sur les résidus.

Hypothèses (A_1) *Pour chaque individu* $i = 1, .., N$, *les résidus* ε_{it} *sont indépendamment et normalement distribués avec* $E(\varepsilon_{it}) = 0$ *et* $E\left(\varepsilon_{it}^2\right) = \sigma_{\varepsilon_i}^2$, $\forall t = 1, .., T$. *Les résidus sont en outre indépendamment distribués entre individus. Par conséquent* $E(\varepsilon_{it}\varepsilon_{js}) = 0$, $\forall i \neq j$ *et* $\forall (t, s)$.

Ces différentes hypothèses correspondent à celles généralement adoptées dans le cadre des tests de racine unitaire en panel dits de première génération[2] et en particulier dans le test d'Im, Pesaran et Shin (IPS, 2003). Ce qui est fondamental ici c'est que nous supposons que les résidus sont indépendants dans la dimension individuelle. Cette hypothèse nous permet de caractériser de façon très simple la distribution asymptotique de notre statistique de test moyenne et de mettre en évidence une distribution asymptotique normale. Nous convenons du fait que cette hypothèse est extrêmement restrictive sur le plan économique. Il est évident que lorsque l'on considère par exemple la causalité de la monnaie vers le revenu réel dans un panel de pays fortement intégrés, il est sans aucun doute inopportun de supposer l'indépendance des termes d'erreurs nationaux des équations de revenu. Techniquement, on peut s'attendre suivant la forme des dépendances à une altération des propriétés à distance finie de notre statistique de test. Si l'on veut faire le parallèle avec les tests de racine unitaire, il semble aujourd'hui se dégager un consensus sur le fait que l'existence de ces dépendances peut notamment conduire à de fortes distorsions de taille.

C'est pourquoi, il est dans notre projet d'étendre notre méthode de test de la causalité en panel afin de prendre en compte ces éventuelles dépendances inter-individuelles. Toutefois, puisqu'il n'existe pas de forme "naturelle" pour ses dépendances, hormis peut être sous la forme de mesures de distances géographiques ou économiques, différentes spécifications et approches sont envisageables. Le même problème dans le cadre des tests de racine unitaire, se traduit actuellement par une très importante littérature tendant à rechercher les statistiques de test présentant les meilleurs propriétés à distance finie sous les formes les plus générales de dépendances inter-individuelles. De tels développements dépassent largement le cadre de cet article.

Rappelons enfin que nous reprenons ici la démarche du test de racine unitaire d'IPS (2003). Notre approche souffre donc des mêmes avantages, mais aussi des mêmes défauts que l'on peut adresser à ce test largement utilisé actuellement dans la littérature empirique en macro-économie, en économie internationale ou en économie du développement par exemple.

A partir du modèle (1), on teste l'hypothèse nulle de Non Causalité Homogène (NCH) de la variable x vers la variable y. Différentes extensions du concept de non causalité à la Granger (1969) sont possibles en panel dans le cadre de ce système bivarié. Nous retenons la plus simple : la variable x ne cause pas au sens de Granger la variable y, si la prise en compte des valeurs passées de x pour l'individu i n'améliore pas la

[2] Voir Hurlin et Mignon (2004) pour une revue de la littérature sur la question.

prévision de la variable y pour ce même individu[3]. Cette hypothèse se traduit par la contrainte suivante :

$$H_0 : \beta_i = 0 \quad \forall i = 1, ..N \tag{2}$$

avec $\beta_i = (\beta_{i1}, ..., \beta_{iK})'$. Cette contrainte implique qu'il n'existe aucune relation de causalité de x vers y pour les N individus de l'échantillon. On considère une spécification hétérogène de l'hypothèse alternative, telle que :

$$H_1 : \quad \begin{array}{ll} \beta_i = 0 & \forall i = 1, .., N_1 \\ \beta_i \neq 0 & \forall i = N_1 + 1, N_1 + 2, .., N \end{array} \tag{3}$$

où la taille N_1 est a priori inconnue mais satisfait la condition $0 \leq N_1/N < 1$. Ainsi, dès lors que $N_1 > 0$, coexistent deux ensembles d'individus : un ensemble de N_1 individus pour lesquels il n'y a pas de causalité de x vers y et un ensemble de $N - N_1$ individus pour lesquels il existe au contraire une relation de causalité de x vers y. La composition des deux ensembles est a priori inconnue. Dès lors, il ne s'agit pas d'un test d'une hypothèse de non causalité pour tous les individus contre une alternative de causalité pour tous les individus. Si l'on rejette l'hypothèse nulle, cela signifie simplement qu'il existe au moins un individu du panel pour lequel la variable x cause la variable y.

La structure de notre test est ainsi totalement similaire à la structure des tests de racine unitaire dans les modèles de panel hétérogènes. Dans ce contexte deux principales approches s'offrent à nous pour construire une statistique en combinant l'information des N tests individuels indépendants. Soit l'on définit cette statistique comme une fonction des statistiques de tests de non causalité individuelles (dans une approche *à la* IPS, 2003), soit l'on combine les seuils de significativité de ces statistiques individuelles (dans une approche *à la* Choi, 2001). Nous avons retenu ici la première solution en raison notamment de sa simplicité de mise en oeuvre. Ainsi, notre statistique de test, notée $W_{N,T}^{Nch}$, est définie tout simplement comme la moyenne des N statistiques de Wald individuelles associées au test de non causalité pour chacun des individus du panel.

Définition *La statistique moyenne $W_{N,T}^{Nch}$ associée à l'hypothèse nulle de Non Causalité Homogène (NCH) est définie par :*

$$W_{N,T}^{Nch} = \frac{1}{N} \sum_{i=1}^{N} W_{iT} \tag{4}$$

où W_{iT} correspond à la statistique de Wald individuelle associée au test de l'hypothèse $H_0 : \beta_i = 0$ pour le $i^{ème}$ individu du panel.

Cette statistique de test présente tout d'abord l'avantage d'être extrêmement simple à mettre en oeuvre et ne nécessite en particulier aucune estimation en panel. Pour construire une réalisation de $W_{N,T}^{Nch}$, il suffit de calculer la moyenne des N réalisations individuelles des statistiques de test usuelles utilisées en séries temporelles pour tester la non causalité de x vers y. Ces statistiques sont aujourd'hui programmées sous la plupart des logiciels (Eviews, Rats..).

[3] Une piste d'extension de ce test consiste à envisager un ensemble d'information prenant en compte les valeurs passées de x des autres individus du panel dans la prévision de la variable y de l'individu i. Mais cela ne présente d'intérêt que si l'on lève l'hypothèse d'indépendance inter-individuelle.

2 DISTRIBUTIONS ASYMPTOTIQUES

Etudions à présent les distributions asymptotique et semi-asymptotique de la statistique moyenne sous l'hypothèse nulle de non causalité homogène. Il existe trois principaux modes de convergence en panel (Phillips et Moon, 1999). Nous nous contenterons ici de retenir le mode de convergence séquentiel en étudiant le comportement de notre statistique lorsque la dimension T, puis la dimension N tendent vers l'infini. Sous les hypothèse A_1, lorsque la dimension T tend vers l'infini, la statistique de Wald individuelle W_{iT} associée au test $H_0 : \beta_i = 0$, où $\beta_i \in \mathbb{R}^K$, converge vers une loi du chi-deux à K degrés de liberté.

$$W_{iT} \underset{T\to\infty}{\overset{d}{\longrightarrow}} \chi^2(K) \qquad \forall i = 1, .., N \tag{5}$$

En d'autres termes sous l'hypothèse nulle de non causalité, lorsque la dimension T tend vers l'infini, les statistiques individuelles $\{W_{iT}\}_{i=1}^N$ sont identiquement distribuées selon une distribution admettant des moments d'ordre deux finis. Parallèlement, on sait que quelle que soit la dimension T les statistiques de Wald individuelles sont indépendantes puisqu'il n'existe aucune dépendance inter-individuelle dans le modèle. Ainsi, lorsque T tend vers l'infini les statistiques indépendantes $\{W_{iT}\}_{i=1}^N$ convergent vers la même distribution vérifiant alors $E(W_{iT}) = K$ et $V(W_{iT}) = 2K$, $\forall i = 1, .., N$. Dès lors, l'utilisation du théorème central limite de Lindberg-Levy suffit à caractériser la distribution de la statistique moyenne $W_{N,T}^{Nch}$ lorsque la dimension individuelle N tend à son tour vers l'infini.

Résultat 1 *Sous les hypothèses A_1 et l'hypothèse nulle de Non Causalité Homogène, la moyenne $W_{N,T}^{Nch}$ des statistiques de Wald individuelles $\{W_{iT}\}_{i=1}^N$ converge séquentiellement en distribution vers une loi normale. Soit $Z_{N,T}^{Nch}$ la statistique standardisée associée à $W_{N,T}^{Nch}$:*

$$Z_{N,T}^{Nch} = \sqrt{\frac{N}{2K}} \left(W_{N,T}^{Nch} - K \right) \underset{T,N\to\infty}{\overset{d}{\longrightarrow}} N(0,1) \tag{6}$$

avec $W_{N,T}^{Nch} = (1/N) \sum_{i=1}^N W_{iT}$, et où $T, N \to \infty$ désigne une convergence séquentielle où $T \to \infty$ puis $N \to \infty$.

Pour un niveau de risque donné, si la réalisation de la statistique standardisée $Z_{N,T}^{Nch}$ est supérieure en valeur absolue à la valeur critique d'une loi normale centrée réduite, l'hypothèse de Non Causalité Homogène est rejetée. Il existe au moins un individu du panel pour lequel la variable x cause au sens de Granger la variable y. Ce test est donc extrêmement simple à mettre en oeuvre.

Toutefois, notre test ne présente d'intérêt que dans la mesure où il permet d'obtenir de meilleures propriétés dans des panels de petite ou moyenne dimension temporelle. Tout comme pour le cas des tests de racine unitaire, si l'on dispose d'un échantillon de très grande dimension temporelle, on préférera raisonnablement utiliser un test de non causalité sur séries temporelles plutôt que d'utiliser un test en panel fondé sur une hypothèse d'indépendance inter-individuelle[4]. C'est pourquoi nous allons à présent envisager le cas où T est fixé et N tend vers l'infini.

Etant donnée la structure dynamique de notre modèle, le résultat d'une distribution du chi-deux pour la statistique individuelle de Wald ne tient plus dès lors que T est fixé et cela même sous l'hypothèse de normalité. Dans ce cas, les statistiques W_{iT} possèdent des distributions non standards et non identiques, dépendant notamment des paramètres γ_i et β_i. Supposons que les moments d'ordre deux de ces distributions existent. Sous les hypothèses A_1, pour une dimension T finie, les statistiques individuelles de Wald sont alors indépendamment mais non identiquement distribuées. On peut par conséquent appliquer le théorème central limite de Lyapunov et montrer que la statistique moyenne $W_{N,T}^{Nch}$ converge une nouvelle fois vers une loi normale lorsque N tend vers l'infini. Considérons un panel non cylindré où la dimension temporelle T_i varie avec les individus.

Résultat 2 *Sous les hypothèses A_1 et l'hypothèse nulle de Non Causalité Homogène, la moyenne $W_{N,T}^{Nch}$ des statistiques individuelles de Wald $\{W_{i,T_i}\}_{i=1}^{N}$ converge vers une loi normale lorsque N tend vers l'infini, dès lors que :*

$$\lim_{N\to\infty}\left(\sum_{i=1}^{N} Var\left(W_{i,T_i}\right)\right)^{-\frac{1}{2}}\left(\sum_{i=1}^{N} E\left[\left|W_{i,T_i}-E\left(W_{i,T_i}\right)\right|^3\right]\right)^{\frac{1}{3}}=0$$

Soit $Z_{N,T}^{Nch}$ la statistique standardisée associée à $W_{N,T}^{Nch}$:

$$Z_{N,T}^{Nch}=\frac{\sqrt{N}\left[W_{N,T}^{Nch}-N^{-1}\sum_{i=1}^{N}E\left(W_{i,T_i}\right)\right]}{\sqrt{N^{-1}\sum_{i=1}^{N}Var\left(W_{i,T_i}\right)}}\xrightarrow[N\to\infty]{d}N\left(0,1\right) \tag{7}$$

avec $W_{N,T}^{Nch}=(1/N)\sum_{i=1}^{N}W_{i,T_i}$, et où $E\left(W_{i,T_i}\right)$ et $Var\left(W_{i,T_i}\right)$ désignent respectivement l'espérance et la variance de la statistique individuelle W_{iT_i}.

La règle de décision du test est la même que dans le cas précédent. Toute la difficulté consiste alors à donner une valeur aux moments $E\left(W_{i,T_i}\right)$ et $Var\left(W_{i,T_i}\right)$ pour chaque individu du panel afin de construire une réalisation de la statistique standardisée $Z_{N,T}^{Nch}$. Pour cela deux solutions sont envisageables. La première consiste à simuler par Bootstrap pour chaque individu i les deux premiers moments de la distribution de la statistique de Wald pour la dimension T_i. La seconde solution consiste à adopter une approximation de ces deux moments. Nous proposons ici d'approximer à un scalaire près ces deux moments par les moments de la distribution de Fisher que l'on aurait obtenue de façon exacte pour toute taille T_i sous l'hypothèse de normalité A_1, dans un modèle non dynamique. Ainsi, on propose d'approximer les deux premiers moments[5] de W_{i,T_i} à une constante près, par les moments correspondants d'une distribution de

[4] En revanche, l'utilisation d'un test en panel prenant en compte les éventuelles dépendances inter-individuelles sera a priori toujours préférée. Mais il n'existe pas à ce jour de tel test de l'hypothèse de non causalité.

[5] Dans notre approche, l'estimateur de la variance des résidus utilisé dans la construction de la statistique de Wald est normalisé par le nombre de degrés de liberté $T-2K-1$ et non par T. On retrouve ainsi les mêmes formules que celles programmés dans les logiciels usuels comme Eviews.

Fisher $F(K, T_i - 2K - 1)$. Supposons que $T_i > 5 + 2K$, $\forall i = 1, .., N$, alors :

$$E(W_{i,T_i}) \simeq K \times \frac{(T_i - 2K - 1)}{(T_i - 2K - 3)} \tag{8}$$

$$Var(W_{i,T_i}) \simeq 2K \times \frac{(T_i - 2K - 1)^2 \times (T_i - K - 3)}{(T_i - 2K - 3)^2 \times (T_i - 2K - 5)} \tag{9}$$

Naturellement lorsque la dimension T_i tend vers l'infini, on retrouve à partir des formules (8) et (9), les moments exacts d'une distribution du chi-deux à K degrés de libertés, à savoir K et $2K$. Ainsi, la statistique standardisée que nous noterons $\widetilde{Z}_{N,T}^{Nch}$, construite à partir des approximations (8) et (9) des moments pour des tailles T_i finies, correspond à la statistique standardisée $Z_{N,T}^{Nch}$ du résultat 2 lorsque $T_i \to \infty$, $\forall i = 1, .., N$ et l'on retrouve alors le cas d'une distribution normale. Dans tous les autres cas, pour des tailles T_i données, on conjecture que la statistique standardisée $\widetilde{Z}_{N,T}^{Nch}$ définie par :

$$\widetilde{Z}_{N,T}^{Nch} = \sqrt{N} \left[W_{N,T}^{Nch} - \frac{K}{N} \sum_{i=1}^{N} \frac{(T_i - 2K - 1)}{(T_i - 2K - 3)} \right] \left[2\frac{K}{N} \sum_{i=1}^{N} \frac{(T_i - 2K - 1)^2 \times (T_i - K - 3)}{(T_i - 2K - 3)^2 \times (T_i - 2K - 5)} \right]^{-\frac{1}{2}}$$

a même loi, lorsque N tend vers l'infini, que la statistique standardisée $Z_{N,T}^{Nch}$ du résultat 2, à savoir une loi normale centrée réduite. Des simulations de Monte Carlo (Hurlin, 2004) confirment la qualité de cette approximation.

Pour résumer, nous proposons de tester l'hypothèse nulle de Non Causalité Homogène soit à partir de la statistique standardisée $Z_{N,T}^{Nch}$ (équation 6) fondée sur les moments de la distribution asymptotique obtenue lorsque T tend vers l'infini, soit à partir d'une approximation des moments à distance finie avec la statistique $\widetilde{Z}_{N,T}^{Nch}$.

3 SIMULATIONS DE MONTE CARLO

Afin d'illustrer les propriétés de notre test de non causalité, nous proposons quelques simulations de Monte Carlo pour le cas $K = 1$. Soit le modèle :

$$y_{it} = \alpha_i + \gamma_i \, y_{i,t-k} + \beta_i \, x_{i,t-k} + \varepsilon_{it} \tag{10}$$

où les paramètres autorégressifs γ_i, pour $i = 1, .., N$, sont tirés dans une distribution uniforme sur $]-1, 1[$ ce qui garantit la stationnarité des processus y_{it}. Les effets fixes α_i sont simulés selon une loi $N(0, 1)$. Les résidus normaux sous les hypothèses A_1 ont une variance $\sigma_{\varepsilon,i}^2$ tirée dans une loi uniforme sur $[0.5, 1.5]$. Le processus $x_{i,t}$, supposé exogène, est tiré dans une loi normale $N(0, 1)$. Pour différents couples (T, N), nous caractérisons la taille empirique ($\beta_i = 0$) et la puissance ($\exists i \,/\, \beta_i \neq 0$) des statistiques de test $Z_{N,T}^{Nch}$ et $\widetilde{Z}_{N,T}^{Nch}$ à partir de 10 000 simulations. Dans le cas du calcul de la puissance, nous pouvons envisager différentes configurations suivant la taille N_1 du sous groupe d'individus pour lesquels il n'y a pas de causalité. Tout comme dans les simulations d'IPS (2003), nous considérerons le cas a priori le plus favorable où il existe une relation de causalité pour tous les individus ($N_1 = 0$, $\beta_i \neq 0$ $\forall i = 1, .., N$). Les paramètres β_i sont alors tirés dans une loi $N(0, 1)$.

Table 1: Puissances et Tailles Empiriques. Cas $K = 1$

		$T = 10$		$T = 25$		$T = 50$		$T = 100$	
N	Test	Taille	Puis.	Taille	Puis.	Taille	Puis.	Taille	Puis.
1	Wald	0.093	0.509	0.067	0.689	0.055	0.780	0.051	0.843
5	Z^{Nch}	0.163	0.947	0.078	0.992	0.056	0.999	0.052	1.00
	$\widetilde{Z}^{Nch}$	0.041	0.861	0.048	0.990	0.044	0.999	0.046	1.00
10	Z^{Nch}	0.210	0.995	0.083	1.00	0.061	1.00	0.055	1.00
	$\widetilde{Z}^{Nch}$	0.039	0.975	0.043	1.00	0.043	1.00	0.046	1.00
25	Z^{Nch}	0.316	1.00	0.100	1.00	0.062	1.00	0.056	1.00
	$\widetilde{Z}^{Nch}$	0.043	1.00	0.043	1.00	0.042	1.00	0.047	1.00
50	Z^{Nch}	0.447	1.00	0.119	1.00	0.072	1.00	0.06	1.00
	$\widetilde{Z}^{Nch}$	0.041	1.00	0.046	1.00	0.047	1.00	0.049	1.00
100	Z^{Nch}	0.644	1.00	0.143	1.00	0.076	1.00	0.058	1.00
	$\widetilde{Z}^{Nch}$	0.041	1.00	0.045	1.00	0.046	1.00	0.044	1.00

Les résultats de ces simulations sont reportés dans le tableau 1. On observe que les deux tests $Z_{N,T}^{Nch}$ et $\widetilde{Z}_{N,T}^{Nch}$ ont des puissances très élevées. Dès lors que $N > 10$, leur puissance excède 90% et ce même pour des panels de taille T relativement faible ($T = 10$). En outre, la puissance croît très rapidement avec la dimension N. A titre de comparaison, le test de l'hypothèse de non causalité mené sur séries temporelles ($N = 1$) à partir de la statistique de Wald présente une puissance inférieure à 70% lorsque $T \leq 25$. Toutefois, il convient de noter une importante distorsion de taille pour la statistique $Z_{N,T}^{Nch}$ lorsque la taille T est faible : cette distorsion croît en outre avec la dimension N. Rappelons que la statistique $Z_{N,T}^{Nch}$ est basée sur les moments asymptotiques. Pour des tailles T faibles, l'utilisation à tort des lois asymptotiques pour chaque individu peut conduire à rejeter trop fréquemment l'hypothèse nulle. Cette distorsion sera accentuée dans notre cas par la dimension de panel : plus l'on dispose d'individus, plus la distorsion sera forte. C'est pourquoi, on préférera généralement la statistique $\widetilde{Z}_{N,T}^{Nch}$ pour laquelle la taille empirique reste dans tous les cas très proche de la valeur nominale de 5%.

4 Conclusion

Nous proposons dans ce papier un test très simple de l'hypothèse de non causalité au sens de Granger dans un modèle de panel hétérogène à coefficients fixes. Nous reprenons la démarche proposée par Im, Pesaran et Shin (2003) pour tester l'hypothèse de racine unitaire en utilisant notamment une statistique moyenne. Ce test de causalité en panel présente de très bonnes propriétés en comparaison des tests menés sur séries temporelles. Une première application est proposée dans Hurlin et Venet (2004) concernant le lien entre développement financier et croissance dans les pays en voie de développement. Cette étude met ainsi en évidence la robustesse du lien de causalité de la sphère réelle vers la sphère financière qui peut être appréhendé dans un cadre bivarié et la difficulté qu'il y a à mettre en évidence la causalité allant du financier vers la croissance dans le même cadre d'analyse. Toutefois, ce test présente deux principales limites. Tout d'abord, il est construit sous l'hypothèse de stationnarité et exclut ainsi le concept de causalité de long terme (Toda et Phillips, 1993). Deuxièmement et plus fondamentalement, ce test suppose l'absence de toute corrélation inter-individuelle et peut être critiqué sur ce point au même titre que les tests de racine unitaire de première génération.

Références

Choi I. [2001], “Unit Root Tests for Panel Data”, *Journal of International Money and Finance,* 20, p. 249-272.

Granger C.W.J. [1969], ”Investigating Causal Relations by Econometric Models and Cross-Spectral Methods”, *Econometrica*, 37(3), p. 424-438.

Granger C.W.J. [2003], ”Some Aspects of Causal Relationships”, *Journal of Econometrics*, 112, p. 69-71.

Holtz-Eakin D., Newey W. et Rosen H.S. [1988], ”Estimating Vector Autoregressions with Panel Data”, *Econometrica*, 56, p. 1371-1396.

Hsiao C. [1989], ”Modeling Ontario Regional Electricity System Demand Using a MFR Coefficient Approaoch”, *Regional Science and Urban Economics,* 19, p. 565-587

Hsiao C. et Pesaran I. [2004], ”Random Coefficient Panel Data Models”, *Mimeo.*

Hurlin C. [2004], ”Testing Granger Causality in Heterogeneous Panel Data Models with Fixed Coefficients”, *Document de Recherche LEO*, 2004-4.

Hurlin C. et Mignon V. [2004], ”Guide Pratique des Séries Stationnaires en Panel. Partie I : Tests de Racine Unitaire”, *Document de Recherche LEO*, 2004-10.

Hurlin C. et Venet B. [2004], ”Financial Development and Growth: A Re-Examination using a Panel Granger Causality Test”, *Document de Recherche LEO*, 2004-18.

Im K.S., Pesaran M.H. et Shin Y. [2003], “Testing for Unit Roots in Heterogeneous Panels”, *Journal of Econometrics*, 115, 1, p. 53-74.

Nair-Reichert U. et Weinhold D. [2001], ”Causality Tests for Cross-Country Panels”, *Oxford Bulletin of Economics and Statistics*, 63, p. 153-171.

Phillips P.C.B. et Moon H. [1999], “Linear Regression Limit Theory for Nonstationary Panel Data”, *Econometrica*, 67, p. 1057-1111.

Sims C.A. [1972], ”Money, Income, and Causality”, American Economic Review, 62, p. 540-552.

Swamy P.A. [1970], ”Efficient Inference in a Random Coefficient Regression Model”, *Econometrica*, 38, p. 311-323.

Toda, H.Y. et Phillips, P.C.B. [1993], ”Vector AutoRegressions and Causality”, *Econometrica,* 61, p. 1367–1393.

Weinhold D. [1996], ”Tests de Causalité sur Données de Panel : Une Application à l'Etude de la Causalité entre l'Investissement et la Croissance”, *Economie et Prévision*, 126, p. 163-175.