Abstract

This paper proposes a very simple test of Granger (1969) non-causality for heterogeneous panel data models. Our test statistic is based on the individual Wald statistics of Granger non causality averaged across the groups. 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. A standardized statistic based on an approximation of the moments of Wald statistics is proposed. Monte Carlo experiments show that our panel standardized statistics provide very good small sample properties.

- Keywords : Granger non-causality, Panel data, Wald Test.
- J.E.L Classification : C23

---

*LEO, University of Orleans. email: christophe.hurlin@univ-orleans.fr. A substantial part of the work for this paper was undertaken in the Department of Economics of the University Paris IX Dauphine, EURiScO.

†I am grateful for comments from participants of the econometric seminars at University Paris I, Paris X Nanterre, Marne-la-Vallée and University of Geneva.
1 Introduction

The aim of this paper is to propose a simple Granger (1969) non causality test in heterogeneous panel data models with fixed coefficients. In the framework of a linear autoregressive data generating process, the extension of standard causality tests for panel data implies to test cross sectional linear restrictions on the coefficients of the model. As usually, the use of cross-sectional information may extend the information set on the causality from a given variable to another. Indeed, in many economic problems it is highly probable that if a causal relationship exists for a country or an individual, it exists also for some other countries or individuals. In this case, the causality can be tested with more efficiency in a panel context with $NT$ observations. However, the use of the cross-sectional information implies to take into account the heterogeneity across individuals in the definition of the causal relationship. As discussed in Granger (2003), the usual causality test in panel asks "if some variable, say $X_t$ causes another variable, say $Y_t$, everywhere in the panel [...] This is rather a strong null hypothesis." Then, we propose here a simple Granger non causality test for heterogeneous panel data models. This test allows to take into account both dimensions of the heterogeneity in this context: the heterogeneity of the causal relationships and the heterogeneity of the data generating process ($DGP$).

Let us consider the standard implication of the Granger causality definition\(^1\). For each individual, we say that the variable $x$ is causing $y$ if we are better able to predict $y$ using all available information than if the information apart from $x$ had been used (Granger 1969). If $x$ and $y$ are observed on $N$ individuals, the issue consists in determining the optimal information set used to forecast $y$. Several solutions could be adopted. The most general is to test the causality from the variable $x$ observed on the $i^{th}$ individual to the variable $y$ observed for the $j^{th}$ individual, with $j = i$ or $j \neq i$. The second solution, is more restrictive and is directly derived from the time series analy-

\(^1\) The Granger causality definition is based on the "two precepts that the cause preceded the effect and the causal series had information about the effect that was not contained in any other series according to the conditional distributions" (Granger 2003). The fact that the cause produces a superior forecasts of the effect is just an implication of these statements. However, it does provide suitable post sample tests as discussed in Granger (1980).
sis. It implies to test causal relationship for a given individual. The cross sectional information is then only used to improve the specification of the model and the power of tests as in Holtz-Eakin, Newey and Rosen (1988). The baseline idea is to assume that there exists a minimal statistical representation, which is common to $x$ and $y$ at least for a subgroup of individuals. In this paper we use such a model. Then, causality tests could be implemented and considered as a natural extension of the standard time series tests in the cross sectional dimension.

However, one of the main specific stakes of panel data models is to specify the heterogeneity between cross-section units. In this context, the heterogeneity has two main dimensions. We propose to distinguish between the heterogeneity of the $DGP$ and the heterogeneity of the causal relationships from $x$ to $y$. Indeed, the $DGP$ may be different from an individual to another, whereas there exists a causal relationship from $x$ to $y$ for all individuals. More generally, in a $p$ order linear vectorial autoregressive model, we define four kinds of causal relationships. The first, denoted Homogenous Non Causality ($HNC$) hypothesis, implies that there does not exist any individual causality relationships from $x$ to $y$. The symmetric case is the Homogenous Causality ($HC$) hypothesis, which occurs when there exists $N$ causality relationships, and when the individual predictors of $y$, obtained conditionally to the past values of $y$ and $x$ are identical. The dynamics of $y$ is then totally identical for all the individuals of the sample. The two last cases correspond to heterogeneous process. Under the Heterogenous Causality ($HEC$) hypothesis, we assume that there exists $N$ causality relationships as in the $HC$ case, but the dynamics of $y$ is heterogenous. The heterogeneity does not affect the causality result. Finally, under the Heterogenous Non Causality ($HENC$) hypothesis, we assume that there exists a subgroup of individuals for which there is a causal relationship from $x$ to $y$. Symmetrically, there is at least one and at the most $N-1$ non causal relationships in the model. That is why, in this case, the heterogeneity deals with the causality from $x$ to $y$.

To sum it up, in the $HNC$ hypothesis, there does not exist any individual causality from $x$ to $y$. On the contrary, in the $HC$ and $HEC$ cases, there is a causality relationships for each individual of the sample. In the $HC$ case, the $DGP$ is homogenous,
whereas it is not the case in the \textit{HEC} hypothesis. Finally in the \textit{HENC} hypothesis, there is an heterogeneity of the causality relationships since there is a subgroup of $N_1$ units for which the variable $x$ does not cause $y$.

In this paper, we propose a simple test of the Homogenous Non Causality (\textit{HNC}) hypothesis. Under the null hypothesis, there is no causal relationship for all the units of the panel. However, we do not test this hypothesis against the \textit{HC} hypothesis as Holtz-Eakin, Newey and Rosen (1988). We specify the alternative as the \textit{HENC} hypothesis. There is two subgroups of units: one with causal relationships from $x$ to $y$, but not necessarily with the same DGP, and another subgroup where there is no causal relationships from $x$ to $y$. For that, our test is lead in an heterogenous panel data model with fixed coefficients. Under the null or the alternative, the unconstrained parameters may be different from individual to another. Then, whatever the result on the existence of causal relationships, we assume that the dynamic of the individual variables may be heterogeneous.

As in the literature devoted to panel unit root tests in heterogeneous panels, and particularly in Im, Pesaran and Shin (2003), we propose a statistic of test based on averaging standard individual Wald statistics of Granger non causality tests. Under the assumption of cross-section independence (as used in first generation panel unit root tests), we provide different results. First, this statistic is shown to converge sequentially in distribution to a standard normal variate when the time dimension $T$ tends to infinity, followed by the individual dimension $N$. Second, for a fixed $T$ sample the semi-asymptotic distribution of the average statistic is characterized. In this case, individual Wald statistics do not have a standard chi-squared distribution. However, under very general setting, it is shown that individual Wald statistics are independently distributed with finite second order moments as soon as $T > 5 + 2K$, where $K$ denotes the number of linear restrictions. For a fixed $T$, the Lyapunov central limit theorem is sufficient to get the distribution of the standardized average Wald statistic when $N$ tends to infinity. The two first moments of this normal semi-asymptotic distribution correspond to empirical mean of the corresponding theoretical moments of the
individual Wald statistics. The issue is then to propose an evaluation of the two first moments of standard Wald statistics for small $T$ sample. The first solution consists in using Monte Carlo or Bootstrap simulations. The second solution consists in using an approximation of these moments based on the exact moments of the ratio of quadratic forms in normal variables derived from the Magnus (1986) theorem for a fixed $T$ sample, with $T > 5 + 2K$. Given these approximations, we propose a second standardized average Wald statistic to test the $HNC$ hypothesis in short $T$ sample.

The finite sample properties of our test statistics are examined using Monte Carlo methods. The simulation results clearly show that our panel based tests have very good properties even in samples with very small values of $T$ and $N$. The size of our standardized statistic based the semi-asymptotic moment is reasonably close to the nominal size for all values of $T$ and $N$. Besides, the power of our panel test statistic substantially exceeds that Granger non Causality based on single time series in all experiments and in particular for very small values of $T = 10$, as soon as there are very few cross-section units in the panel ($N = 5$). Finally, approximated critical values are proposed for finite $T$ and $N$ sample.

The paper is organized as follows. Section 2 is devoted to the definition of the Granger causality test in heterogenous panel data models. Section 3 sets out the asymptotic distributions of the average Wald statistic. Section 4 derives the semi-asymptotic distribution for fixed $T$ sample and section 5 proposes some results of Monte Carlo experiments. Section 6 extends the results to a fixed $N$ sample and the last section provides some concluding remarks.

2 A non causality test in heterogenous panel data models

Let us consider two covariance stationary variables, denoted $x$ and $y$, observed on $T$ periods and on $N$ individuals. For each individual $i = 1, \ldots, N$, at time $t = 1, \ldots, T$, we consider the following linear model:

\[ y_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_i^{(k)} y_{i,t-k} + \sum_{k=1}^{K} \beta_i^{(k)} x_{i,t-k} + \varepsilon_{i,t} \]  

(1)
with $K \in \mathbb{N}^*$ and $\beta_i = (\beta_i^{(1)}, \ldots, \beta_i^{(K)})'$. For simplicity, individual effects $\alpha_i$ are supposed to be fixed. Initial conditions $(y_{i,-K}, \ldots, y_{i,0})$ and $(x_{i,-K}, \ldots, x_{i,0})$ of both individual processes $y_{i,t}$ and $x_{i,t}$ are given and observable. We assume that lag orders $K$ are identical for all cross-section units of the panel and the panel is balanced. In a first part, we allow for autoregressive parameters $\gamma_i^{(k)}$ and regression coefficients slopes $\beta_i^{(k)}$ to differ across groups. However, contrary to Weinhold (1996) and Nair-Reichert and Weinhold (2001), parameters $\gamma_i^{(k)}$ and $\beta_i^{(k)}$ are constant. It is important to note that our model is not a random coefficient model as in Swamy (1970): it is a fixed coefficients model with fixed individual effects. We make the following assumptions.

Assumption $(A_1)$ For each cross section unit $i = 1, \ldots, N$, individual residuals $\varepsilon_{i,t}$, 
\[\forall t = 1, \ldots, T \text{ are independently and normally distributed with } E(\varepsilon_{i,t}) = 0 \text{ and } \text{finite heterogeneous variances } E(\varepsilon_{i,t}^2) = \sigma_{\varepsilon,i}^2.\]

Assumption $(A_2)$ Individual residuals $\varepsilon_i = (\varepsilon_{i,1}, \ldots, \varepsilon_{i,T})'$, are independently distributed across groups. Consequently $E(\varepsilon_{i,t}\varepsilon_{j,s}) = 0$, $\forall i \neq j$ and $\forall (t, s)$.

Assumption $(A_3)$ Both individual variables $x_i = (x_{i,1}, \ldots, x_{i,T})'$ and $y_i = (y_{i,1}, \ldots, y_{i,T})'$, 
are covariance stationary with $E(y_{i,t}^2) < \infty$, $E(x_{i,t}^2) < \infty$, $E(x_{i,t}x_{j,z})$, $E(y_{i,t}y_{j,z})$ and $E(y_{i,t}x_{j,z})$ are only function of the difference $t - z$, whereas $E(x_{i,t})$ and $E(y_{i,t})$ are independent of $t$.

This simple two variables model constitutes the basic framework to study the Granger causality in a panel data context. As for time series, the standard causality tests consist in testing linear restrictions on vectors $\beta_i$. However with a panel data model, one must be very careful to the issue of heterogeneity between individuals. The first source of heterogeneity is standard and comes from the presence of individual effects $\alpha_i$. The second source, which is more crucial, is related to the heterogeneity of the parameters $\beta_i$. This kind of heterogeneity directly affects the paradigm of the representative agent and so, the conclusions about causality relationships. It is well known that the estimates of autoregressive parameters $\beta_i$ get under the wrong hypothesis $\beta_i = \beta_j$, $\forall (i, j)$ are biased (see Pesaran and Smith 1995 for an AR(1) process). Then, if we impose the homogeneity of coefficients $\beta_i$, the statistics of causality tests can lead to
a fallacious inference. Intuitively, the estimate $\hat{\beta}$ obtained in an homogeneous model will converge to a value close to the average of the true coefficients $\beta_i$, and that if this mean is itself close to zero, we risk to accept at wrong the hypothesis of no causality.

Beyond these statistical stakes, it is evident that an homogeneous specification of the relation between the variables $x$ and $y$ does not allow to give some interpretation of the relations of causality as soon as at least one individual of the sample has an economic behavior different from that of the others. For example, let us assume that there exists a relation of causality for a set of $N$ countries, for which vectors $\beta_i$ are strictly identical. If we introduce into the sample, a set of $N_1$ countries for which, on the contrary, there is no relation of causality, what are the conclusions? Whatever the value of the ratio $N/N_1$ is, the test of the causality hypothesis is nonsensical.

Given these observations, we now propose to test the Homogenous Non Causality ($HNC$) hypothesis. Under the alternative we allow that there exists a subgroup of individuals with no causality relations and a subgroup of individuals for which the variable $x$ Granger causes $y$. The null hypothesis of $HNC$ is defined as:

$$H_0 : \beta_i = 0 \quad \forall i = 1, \ldots, N$$

with $\beta_i = (\beta_i^{(1)}, \ldots, \beta_i^{(K)})'$. Under the alternative, we allow for $\beta_i$ to differ across groups. We also allow for some, but not all, of the individual vectors to be equal to 0 (non causality assumption). We assume that under $H_1$, there are $N_1 < N$ individual processes with no causality from $x$ to $y$. Then, this test is not a test of the non causality assumption against the causality from $x$ to $y$ for all the individuals, as in Holtz-Eakin, Newey and Rosen (1988). It is more general, since we can observe non causality for some units under the alternative:

$$H_1 : \begin{cases} \beta_i = 0 & \forall i = 1, \ldots, N_1 \\ \beta_i \neq 0 & \forall i = N_1 + 1, N_1 + 2, \ldots, N \end{cases}$$

where $N_1$ is unknown but satisfies the condition $0 \leq N_1/N < 1$. The fraction $N_1/N$ is necessarily inferior to one, since if $N_1 = N$ there is no causality for all the individual of the panel, and then we get the null hypothesis $HNC$. In the opposite case $N_1 = 0$, there
is causality for all the individual of the sample. The structure of this test is similar to the unit root test in heterogenous panels proposed by Im, Pesaran and Shin (2003). In our context, if the null is accepted the variable $x$ does not Granger cause the variable $y$ for all the units of the panel. On the contrary, let us assume that the HNC is rejected and if $N_1 = 0$, we have seen that $x$ Granger causes $y$ for all the individuals of the panel: in this case we get an homogenous result as far as causality is concerned. The DGP may be not homogenous, but the causality relations are observed for all individuals. On the contrary, if $N_1 > 0$, then the causality relationships is heterogeneous: the DGP and the causality relations are different according the individuals of the sample.

In this context, we propose to use the average of individual Wald statistics associated to the test of the non causality hypothesis for units $i = 1, \ldots, N$.

**Definition** The average statistic $W_{N,T}^{Hnc}$ associated to the null Homogenous Non Causality (HNC) hypothesis is defined as:

$$W_{N,T}^{Hnc} = \frac{1}{N} \sum_{i=1}^{N} W_{i,T}$$

where $W_{i,T}$ denotes the individual Wald statistics for the $i^{th}$ cross section unit associated to the individual test $H_0 : \beta_i = 0$.

In order to express the general form of this statistic, we stack the $T$ periods observations for the $i^{th}$ individual’s characteristics into $T$ elements columns as:

$$y_{i,(T,1)}^{(k)} = \begin{bmatrix} y_{i,1-k} \\ \vdots \\ y_{i,T-k} \end{bmatrix}, \quad x_{i,(T,1)}^{(k)} = \begin{bmatrix} x_{i,1-k} \\ \vdots \\ x_{i,T-k} \end{bmatrix}, \quad \varepsilon_{i,(T,1)}^{(k)} = \begin{bmatrix} \varepsilon_{i,1} \\ \vdots \\ \varepsilon_{i,T} \end{bmatrix}$$

and we define two $(T, K)$ matrices:

$$Y_i = \begin{bmatrix} y_i^{(1)} : y_i^{(2)} : \ldots : y_i^{(K)} \end{bmatrix}, \quad X_i = \begin{bmatrix} x_i^{(1)} : x_i^{(2)} : \ldots : x_i^{(K)} \end{bmatrix}$$

Let us denote $Z_i$ the $(T, 2K + 1)$ matrix $Z_i = [e : Y_i : X_i]$, where $e$ denotes a $(T, 1)$ unit vector, and $\theta_i = (\alpha_i' \gamma_i' \beta_i')'$ the vector of parameters of model. The HNC hypothesis test can be expressed as $R\theta_i = 0$ where $R$ is a $(K, 2K + 1)$ matrix with $R = [0 : I_K]$. 
The Wald statistic $W_{i,T}$ associated to the individual test $H_0: \beta_i = 0$ is defined for each $i = 1, \ldots, N$ as:

$$W_{i,T} = \hat{\beta}_i' R' \left[ \hat{\sigma}_i^2 R (Z_i' Z_i)^{-1} R' \right]^{-1} R \hat{\theta}_i = \frac{\hat{\beta}_i' R' \left[ R (Z_i' Z_i)^{-1} R' \right]^{-1} R \hat{\theta}_i}{\hat{\sigma}_i^2 / (T - 2K - 1)}$$

where $\hat{\theta}_i$ is the estimate of parameter $\theta_i$ get under the alternative hypothesis, $\hat{\sigma}_i^2$ the estimate of the variance of residuals. For a small $T$ sample, the corresponding unbiased estimator\(^2\) may be expressed as $\hat{\sigma}_i^2 = \bar{e}_i' \bar{e}_i / (T - 2K - 1)$. It is well known that this Wald statistic can also be expressed as a ratio of quadratic forms in normal variables corresponding to the true population of residuals, with:

$$W_{i,T} = (T - 2K - 1) \left( \frac{\bar{e}_i' \Phi_i \bar{e}_i}{\bar{e}_i' M_i \bar{e}_i} \right) \quad i = 1, \ldots, N \quad (5)$$

where the $(T, 1)$ vector $\bar{e}_i = \varepsilon_i / \sigma_{\varepsilon,i}$ is distributed according $N(0, I_T)$ under assumption $A_1$. The matrix $\Phi_i$ and $M_i$ are positive semi definite, symmetric and idempotent $(T, T)$ matrix.

$$\Phi_i = Z_i (Z_i' Z_i)^{-1} R' \left[ R (Z_i' Z_i)^{-1} R' \right]^{-1} R (Z_i' Z_i)^{-1} Z_i' \quad (6)$$

$$M_i = I_T - Z_i (Z_i' Z_i)^{-1} Z_i' \quad (7)$$

where $I_T$ is the identity matrix of size $T$. The matrix $M_i$ corresponds to the standard projection matrix of the linear regression analysis.

The issue is now to determine the distribution of the average statistic $W_{N,T}^{Hnc}$ under the null hypothesis of Homogenous Non Causality. For that, we first consider the asymptotic case where $T$ and $N$ tends to infinity, and in second part the case where $T$ is fixed.

### 3 Asymptotic distribution

We propose to derive the asymptotic distribution of the average statistic $W_{N,T}^{Hnc}$ under the null hypothesis of non causality. For that, we consider the case of a sequential convergence when $T$ tends to infinity and then $N$ tends to infinity. This sequential

\(^2\)It is also possible to use the standard formula of the Wald statistic by substituting the term $(T - 2K - 1)$ by $T$. However, several software (as Eviews) use this normalisation.
convergence result can be deduced from the standard convergence result of the individ-
ual Wald statistic \( W_{i,T} \) in a large \( T \) sample. In a non dynamic model, the normality
assumption in \( A_1 \) would be sufficient to establish the fact for all \( T \), the Wald statistic
has a chi-squared distribution with \( K \) degrees of freedom. But in our dynamic
model, this result can only be achieved asymptotically. Let us consider the expression
(5). Given that under \( A_1 \) the least squares estimate \( \hat{\theta}_i \) is convergent, we know that
\[
\lim_{T \to \infty} \frac{\varepsilon_i' M_i \varepsilon_i}{T - 2K - 1} = \sigma_{\varepsilon_i}^2.
\]
It implies that:
\[
\lim_{T \to \infty} \frac{\varepsilon_i' M_i \varepsilon_i}{T - 2K - 1} = \frac{1}{T - 2K - 1} \left( \frac{\varepsilon_i' M_i \varepsilon_i}{T - 2K - 1} \right) = 1.
\]

Then, if the statistic \( W_{i,T} \) has a limiting distribution, it is the same distribution of
the statistics that results when the denominator is replaced by its limiting value, that
is to say 1. Thus, \( W_{i,T} \) has the same limiting distribution as \( \varepsilon_i' \Phi_i \varepsilon_i \). Under assumption
\( A_1 \), the vector \( \varepsilon_i \) is distributed across a \( N (0, I_T) \). Since \( \Phi_i \) is idempotent, the quadratic
form \( \varepsilon_i' \Phi_i \varepsilon_i \) is distributed as a chi-squared with a number of degrees of freedom equal
to the rank of \( \Phi_i \). The rank of the symmetric idempotent matrix \( \Phi_i \) is equal to its trace,
that is to say \( K \) (cf. appendix ??). Then, under the null hypothesis of non causality,
each individual Wald statistic converges to a chi-squared distribution with \( K \) degrees
of freedom:
\[
W_{i,T} \overset{d}{\to} \chi^2_K \quad \forall i = 1, \ldots, N \tag{8}
\]

In other words, when \( T \) tends to infinity, individual statistics \( \{W_{i,T}\}_{i=1}^N \) are identically distributed. They are also independent since under assumption \( A_2 \), residual \( \varepsilon_i \) and \( \varepsilon_j \) for \( j \neq i \) are independent. To sum it up: if \( T \) tends to infinity individual Wald
statistics \( W_{i,T} \) are \( i.i.d. \) with \( E(W_{i,T}) = K \) and \( V(W_{i,T}) = 2K \). Then, the distribution
of the average Wald statistic \( W_{HNC}^{\text{nc}} \) when \( T \to \infty \) first and then \( N \to \infty \), can be
deduced from a standard Lindberg-Levy central limit theorem.

**Theorem 1** Under assumption \( A_2 \), the individual \( W_{i,T} \) statistics for \( i = 1, \ldots, N \) are
identically and independently distributed with finite second order moments as \( T \to \infty \),
and therefore by Lindberg-Levy central limit theorem under the \( HNC \) null hypothesis,
the average statistic \( W_{HNC}^{\text{nc}} \) sequentially converges in distribution.
\[
Z_{N,T}^{HNC} = \sqrt{\frac{N}{2K}} \left( W_{N,T}^{HNC} - K \right) \overset{d}{\to} N (0, 1) \tag{9}
\]
with $W_{N,T}^{Hnc} = (1/N) \sum_{i=1}^{N} W_{i,T}$, where $T, N \to \infty$ denotes the fact that $T \to \infty$ first and then $N \to \infty$.

For a large $N$ and $T$ sample, if the realization of the standardized statistic $Z_{N,T}^{Hnc}$ is superior to the normal corresponding critical value for a given level of risk, the homogeneous non causality (HNC) hypothesis is rejected. This asymptotic result may be useful in some macro panels. However, it should be extended to the case where $T$ and $N$ tend to infinity simultaneously.

4 Fixed $T$ samples and semi-asymptotic distributions

Asymptotically, individual Wald statistics $W_{i,T}$ for each $i = 1, \ldots, N$, converge toward an identical chi-squared distribution. However, this convergence result can not be achieved for any time dimension $T$, even if we assume the normality of residuals. The issue is then to show that for a fixed $T$ dimension, individual Wald statistics have finite second order moments even they do not have the same distribution and they do not have a standard distribution.

Let us consider the expression (5) of $W_{i,T}$ under assumption $A_1$: this is a ratio of two quadratic forms in a standard normal vector. Magnus (1986) gives general conditions which insure that the expectations of a quadratic form in normal variables exists. Let us consider the moments $E[(x'Ax/x'Bx)^s]$, when $x$ is normally distributed vector $N(0, \sigma^2 I_T)$, $A$ is a symmetric $(T,T)$ matrix and $B$ a positive semi definite $(T,T)$ matrix of rank $r \geq 1$. Let us denote $Q$ a $(T,T-r)$ matrix of full column rank $T-r$ such that $BQ = 0$. If $r \leq T-1$, Magnus (1986)’s theorem identifies three conditions:

(i) If $AQ = 0$, then $E[(x'Ax/x'Bx)^s]$ exists for all $s \geq 0$.

(ii) If $AQ \neq 0$ and $Q'AQ = 0$, then $E[(x'Ax/x'Bx)^s]$ exists for $0 \leq s < r$ and does not exist for $s \geq r$.

(iii) If $Q'AQ \neq 0$, then $E[(x'Ax/x'Bx)^s]$ exists for $0 \leq s < r/2$ and does not exist for $s \geq r/2$. 

11
These general conditions are done in the case where matrices $A$ and $B$ are deterministic. In our case, the corresponding matrices $M_i$ and $\Phi_i$ are stochastic, even we assume that variables $X_i$ are deterministic. However, given a fixed $T$ sample, we propose here to apply these conditions to the corresponding realisations denoted $m_i$ and $\phi_i$. First, in our case the rank of the symmetric idempotent matrix $m_i$ is equal to $T - 2K - 1$. Second, since the matrix $m_i$ is the projection matrix associated to the realization $z_i$ of $Z_i$, we have by construction $m_i z_i = 0$, where $z_i$ of full column rank $2K + 1$; since $T - \text{rank}(m_i) = 2K + 1$. Then, for a given realisation $\phi_i$ by construction, the product $\phi_i z_i$ is different from zero since

$$\phi_i z_i = z_i (z_i' z_i)^{-1} R' \left[ R (z_i' z_i)^{-1} R' \right]^{-1} R \neq 0$$

Besides, the product $z_i' \phi_i z_i$ is also different from zero, since

$$z_i' \phi_i z_i = R' \left[ R (z_i' z_i)^{-1} R' \right]^{-1} R \neq 0$$

Then, the Magnus' theorem allows us to establish that $E \left[ (z_i' \phi_i z_i) / (z_i' m_i z_i) \right]^s$ exists as soon as $0 \leq s < \text{rank}(m_i)/2$. We assume that this condition is also satisfied for $W_{i,T}$:

$$E \left[ (W_{i,T})^s \right] = (T - 2K - 1)^s E \left[ \left( \frac{z_i' \Phi_i z_i}{z_i' m_i z_i} \right)^s \right] \text{ exists if } 0 \leq s < \frac{T - 2K - 1}{2}$$

In particular, given the realizations of $\Phi_i$ and $M_i$, we can identify the condition on $T$ which assures that second order moments ($s = 2$) of $W_{i,T}$ exists.

**Proposition 2** For a fixed time dimension $T \in \mathbb{N}$, the second order moments of the individual Wald statistic $W_{i,T}$ associated to the test $H_{0,i} : \beta_i = 0$, exist if and only if:

$$T > 5 + 2K \quad (10)$$

Hence for a small $T$, individual Wald statistics $W_{i,T}$ are not necessarily identically distributed since matrices $\Phi_i$ and $M_i$ are different from an individual to another. Besides, they do not have standard distribution as in previous section. However, the condition which insure the existence of second order moments are the same for all units. The second order moments of $W_{i,T}$ exist as soon as $T > 5 + 2K$ or equivalently $T \geq 6 + 2K$. 


For a fixed $T$ sample, the statistic of non causality test $W_{N,T}^{Hnc}$ is the average of non identically distributed variables $W_{i,T}$, but with finite second order moments under the condition of proposition 2. Under assumption $A_2$, residual $\varepsilon_i$ and $\varepsilon_j$ for $j \neq i$ are independent. Consequently, individual Wald $W_{i,T}$ for $i = 1, \ldots, N$ are also independent. Then, the distribution of the non causality test statistic $W_{N,T}^{Hnc}$ can be derived according the Lyapunov central limit theorem.

**Theorem 3** Under assumption $A_2$, if $T > 5 + 2K$ the individual $W_{i,T}$ statistics $\forall i = 1, \ldots, N$ are independently but not identically distributed with finite second order moments, and therefore by Lyapunov central limit theorem under the HNC null hypothesis, the average statistic $W_{HNC}^b$ converges. If

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

the standardized statistic $Z_{N}^{Hnc}$ converges in distribution:

$$
Z_{N}^{Hnc} = \frac{\sqrt{N} \left[ W_{Hnc}^{N} - N^{-1} \sum_{i=1}^{N} E(W_{i,T}) \right]}{\sqrt{N^{-1} \sum_{i=1}^{N} \text{Var} (W_{i,T})}} \xrightarrow{d} N(0,1)
$$

(11)

with $W_{N,T}^{Hnc} = (1/N) \sum_{i=1}^{N} W_{i,T}$, where $E(W_{i,T})$ and $\text{Var} (W_{i,T})$ respectively denote the mean and the variance of the statistic $W_{i,T}$ defined by equation (5).

The decision of rule is the same as in the asymptotic case: if the realization of the standardized statistic $Z_{N}^{Hnc}$ is superior to the normal corresponding critical value for a given level of risk, the homogeneous non causality (HNC) hypothesis is rejected. For large $T$, the moments used in theorem (3) are expected to converge to $E(W_{i,T}) = K$ and $\text{Var} (W_{i,T}) = 2K$ since individual statistics $W_{i,T}$ converge in distribution to a chi-squared distribution with $K$ degrees of freedom. Then, the statistic $Z_{N}^{Hnc}$ converges to $Z_{N,T}^{Hnc}$ and we find the conditions of the theorem 1. However, these asymptotic moment values could lead to poor test results, when we have small values of $T$. The issue is then to evaluate the mean and the variance of the Wald statistic $W_{i,T}$, whereas this statistic does not have a standard distribution for a fixed $T$. 

13
The issue is now to compute the standardized average statistic $Z_{Hnc}^N$. There are two main approaches to compute the two first moments of the individual Wald statistics $W_{i,T}$. Firstly, these moments can be computed via stochastic simulation (Monte Carlo or bootstrap) of the Wald under the null. In this case, for each cross section unit, it is necessary to estimate the parameters of the model ($\gamma_i$, $\sigma_i$ and $\alpha_i$) and the parameters of the DGP for variables $x_{it}$. Then, the variable $y_i$ is simulated under the null with $i.i.d.$ normal residual $\varepsilon_i$ with zero means and variance $\sigma_i^2$ (Monte Carlo) or with re-sampled historical residuals (bootstrap). For each simulation of the processes $y_i$ and $x_i$, the individual Wald statistic $W_{i,T}$ is computed. Finally, using the replications of $W_{i,T}$, the corresponding two first moments are estimated for each cross-section unit. We denote $\tilde{Z}_{N}^{MC}$ the corresponding standardized average statistic. It is obvious that this method can be time consuming, especially if we consider very large $N$ panel.

Secondly, we propose here an approximation of $E(W_{i,T})$ and $Var(W_{i,T})$ based on the results of Magnus (1986) theorem. Let us consider the expression of the Wald $W_{i,T}$ as a ratio of two quadratic forms in a standard normal vector under assumption $A_1$:

$$W_{i,T} = (T - 2K - 1) \left( \frac{\varepsilon_i^t \Phi_i \tilde{\varepsilon}_i}{\varepsilon_i^t M_i \tilde{\varepsilon}_i} \right)$$

(12)

where the $(T, 1)$ vector $\tilde{\varepsilon}_i = \varepsilon_i/\sigma_{\varepsilon,i}$ is distributed according $N(0, I_T)$ where matrices $\Phi_i$ and $M_i$ are idempotent and symmetric (and consequently positive semi-definite). For a given $T$ sample, let us denote respectively $\phi_i$ and $m_i$, the realizations of matrices $\Phi_i$ and $M_i$. We propose here to apply the Magnus (1986) theorem to the quadratic forms in a standard normal vector defined as:

$$\tilde{W}_{i,T} = (T - 2K - 1) \left( \frac{\varepsilon_i^T \phi_i \tilde{\varepsilon}_i}{\varepsilon_i^T m_i \tilde{\varepsilon}_i} \right)$$

(13)

where matrices $\phi_i$ and $m_i$ are idempotent and symmetric (and consequently positive semi-definite).

**Theorem 4 (Magnus 1986)** Let $\tilde{\varepsilon}_i$ be a normal distributed vector with $E(\tilde{\varepsilon}_i) = 0$ and $E(\tilde{\varepsilon}_i \tilde{\varepsilon}_i^T) = I_T$. Let $P_i$ be an orthogonal $(T, T)$ matrix and $\Lambda_i$ a diagonal $(T, T)$ matrix such that

$$P_i^t m_i P_i = \Lambda_i \quad P_i^t P_i = I_T$$

(14)
Then, we have, provided the expectation exists for \( s = 1, 2, 3 \ldots \):

\[
E \left[ \left( \frac{\mathbf{e}_i' \phi_i \mathbf{e}_i}{\mathbf{e}_i' m_i \mathbf{e}_i} \right)^s \right] = \frac{1}{(s-1)!} \sum_v \sum_{s=1}^\infty \gamma_s(v) \times \int_0^\infty \left\{ t^{s-1} |\Delta_i| \prod_{j=1}^s [\text{trace}(R_i)]^{n_j} \right\} \, dt
\]  
(15)

where the summation is over all \((s, 1)\) vectors \( v = (n_1, \ldots, n_s) \) whose elements \( n_j \) are nonnegative integers satisfying \( \sum_{j=1}^s j n_j = s \)

\[
\gamma_s(v) = s! 2^s \prod_{j=1}^s [n_j! (2j)^{n_j}]^{-1}
\]  
(16)

and \( \Delta_i \) is a diagonal positive definite \((T, T)\) matrix and \( R_i \) a symmetric \((T, T)\) matrix given by:

\[
\Delta_i = (I_T + 2t \Lambda_i)^{-1/2} \quad R_i = \Delta_i P_i^t \phi_i P_i \Delta_i
\]  
(17)

In our case, we are interested by the two first moments. For the first order moment \((s = 1)\), there is only one scalar \( v = n_1 \) which is equal to one. Then, the quantity \( \gamma_1(v) \)

is equal to one. For the second order moment \((s = 2)\), there are two vectors \( v = (n_1, n_2) \)

which are respectively defined by \( v_1 = (0, 1) \) and \( v_2 = (2, 0) \). Consequently \( \gamma_2(v_1) = 2 \) and \( \gamma_2(v_2) = 1 \). Given these results, we can compute the exact two corresponding moments of the statistic \( \widetilde{W}_{i,T} \) as:

\[
E \left( \widetilde{W}_{i,T} \right) = (T - 2K - 1) \times \int_0^\infty |\Delta_i| \text{ trace } (R_i) \, dt
\]  
(18)

\[
E \left[ \left( \widetilde{W}_{i,T} \right)^2 \right] = (T - 2K - 1)^2 \times \left\{ 2 \int_0^\infty t |\Delta_i| \text{ trace } (R_i) \, dt + \int_0^\infty t |\Delta_i| [\text{trace}(R_i)]^2 \, dt \right\}
\]  
(19)

where matrices \( \Delta_i \) and \( R_i \) are defined in theorem (4). Both quantities \(|\Delta_i|\) and \(\text{trace}(R_i)\)

can be computed analytically in our model given the properties of these matrices. Since \( \Lambda_i \) is issued from the orthogonal decomposition of the idempotent matrix \( m_i \),

with \( \text{rank}(m_i) = T - 2K - 1 \) (cf. appendix ??), this matrix is a zero except the first block which is equal to the \( T - 2K - 1 \) identity matrix (corresponding to the characteristic roots of \( m_i \) which are non null). Then, for a scalar \( t \in \mathbb{R}^+ \), the matrix

\( \Delta_i = (I_T + 2t \Lambda_i)^{-1/2} \) can be partitioned as:

\[
\Delta_i_{(T,T)} = \begin{pmatrix}
D_i(t) & 0 \\
(T-2K-1, T-2K-1) & (T-2K-1, 2K+1) \\
0 & I_{2K+1} \\
(2K+1, T-2K-1) & (2K+1, 2K+1)
\end{pmatrix}
\]
where \( I_p \) denotes the identity matrix of size \( p \). The diagonal block \( D_i(t) \) is defined as \( D_i(t) = (1 + 2t)^{-\frac{1}{2}} I_{T-2K-1} \). Then, the determinant of \( \Delta_i \) can be expressed as:

\[
|\Delta_i| = (1 + 2t)^{-\left(\frac{T-2K-1}{2}\right)}
\]  

(20)

Besides, the trace of the matrix \( R_i \) can be computed as follows. Since for any non-singular matrices \( B \) and \( C \), the rank of \( BAC \) is equal to rank of \( A \), we have here:

\[
\text{rank} \left( R_i \right) = \text{rank} \left( \Delta_i P_i^t \phi_i P_i \Delta_i \right) = \text{rank} \left( P_i^t \phi_i P_i \right)
\]

since the matrix \( \Delta_i \) is non-singular. With the same transformation, given the non-singularity of \( P_i \), we get:

\[
\text{rank} \left( R_i \right) = \text{rank} \left( P_i^t \phi_i P_i \right) = \text{rank} \left( \phi_i \right)
\]

Finally, the rank of the realisation \( \phi_i \) is equal to \( K \), the rank of \( \Phi_i \).

\[
\text{trace} \left( R_i \right) = K
\]

Given these results, the two first moments (equations 18 and 19) of the statistic \( \widetilde{W}_{i,T} \) based for a given \( T \) sample on realizations \( \phi_i \) and \( m_i \), can be expressed as:

\[
E \left( \widetilde{W}_{i,T} \right) = (T - 2K - 1) \times K \times \int_0^\infty (1 + 2t)^{-\left(\frac{T-2K-1}{2}\right)} dt
\]

\[
E \left[ \left( \widetilde{W}_{i,T} \right)^2 \right] = (T - 2K - 1)^2 \times (2K + K^2) \times \int_0^\infty t \times (1 + 2t)^{-\left(\frac{T-2K-1}{2}\right)} dt
\]

Then, we get the following results.

**Proposition 5** For a fixed \( T \) sample, where \( T \) satisfies the condition of proposition (2), given realizations \( \phi_i \) and \( m_i \) of matrices \( \Phi_i \) and \( M_i \) (equations 6 and 7), the exact two first moments of the individual statistics \( \widetilde{W}_{i,T} \), for \( i = 1, ..., N \), defined by equation (13) are respectively:

\[
E \left( \widetilde{W}_{i,T} \right) = K \times \frac{(T - 2K - 1)}{(T - 2K - 3)}
\]  

(21)

\[
\text{Var} \left( \widetilde{W}_{i,T} \right) = 2K \times \frac{(T - 2K - 1)^2 \times (T - K - 3)}{(T - 2K - 3)^2 \times (T - 2K - 5)}
\]  

(22)

as soon as the time dimension \( T \) satisfies \( T \geq 6 + 2K \).
The proof of this proposition is done in appendix A. It is important to verify that for large \( T \) sample, the moments of the individual statistic \( \widetilde{W}_{i,T} \) tend to the corresponding moments of the asymptotic distribution of \( W_{i,T} \) since \( \forall i = 1, \ldots, N \):

\[
\lim_{T \to \infty} E\left( \widetilde{W}_{i,T} \right) = K \quad \lim_{T \to \infty} Var\left( \widetilde{W}_{i,T} \right) = 2K
\]

Both moments correspond to the moments of a \( F(K, T - 2K - 1) \). Indeed, in this dynamic model the \( F \) distribution can be used as an approximation of the true distribution of the statistic \( W_{i,T}/K \) for a small \( T \) sample. Then, the use of the Magnus theorem given the realizations \( \phi_i \) and \( m_i \) to approximate the true moments of the Wald statistic is equivalent to assert that the true distribution of \( W_{i,T} \) can be approximated by the \( F \) distribution.

We propose in this paper to approximate the two first moments of the individual Wald statistic \( W_{i,T} \) by the two first moments of the statistics \( \widetilde{W}_{i,T} \) based on the realizations \( \phi_i \) and \( m_i \) of the stochastic matrices \( \Phi_i \) and \( M_i \) (equations 21 and 22). So, for \( T \geq 6 + 2K \), we assume that:

\[
N^{-1} \sum_{i=1}^{N} E\left( W_{i,T} \right) \simeq E\left( \widetilde{W}_{i,T} \right) = K \times \frac{(T - 2K - 1)}{(T - 2K - 3)}
\]

(23)

\[
N^{-1} \sum_{i=1}^{N} Var\left( W_{i,T} \right) \simeq Var\left( \widetilde{W}_{i,T} \right) = 2K \times \frac{(T - 2K - 1)^2 \times (T - K - 3)}{(T - 2K - 3)^2 \times (T - 2K - 5)}
\]

(24)

Given these approximations, we compute an approximated standardized statistic \( \tilde{Z}_{N,T}^{Hnc} \) for the average Wald average statistic \( W_{N,T}^{Hnc} \) of the \( HNC \) hypothesis.

\[
\tilde{Z}_{N,T}^{Hnc} = \frac{\sqrt{N} \left[ W_{N,T}^{Hnc} - E\left( \widetilde{W}_{i,T} \right) \right]}{\sqrt{Var\left( \widetilde{W}_{i,T} \right)}}
\]

(25)

For a large \( N \) sample, under the Homogenous Non Causality (\( HNC \)) hypothesis, we assume that the statistic \( \tilde{Z}_{N,T}^{Hnc} \) follows the same distribution as the standardized average Wald statistic \( Z_{N,T}^{Hnc} \).
Proposition 6 Under assumptions $A_1$ and $A_2$, for a fixed $T$ dimension with $T > 5 + 2K$, the standardized average statistic $\tilde{Z}^N_{Hnc}$ converges in distribution:

$$\tilde{Z}^N_{Hnc} = \sqrt{\frac{N}{2 \times K}} \times \frac{(T - 2K - 5)}{(T - K - 3)} \times \left[ \frac{(T - 2K - 3)}{(T - 2K - 1)} W^N_{Hnc,T,K} - K \right] \xrightarrow{N \to \infty} N(0, 1) \ (26)$$

with $W^N_{Hnc,T,K} = \frac{1}{N} \sum_{i=1}^{N} W_{i,T}$. Consequently, the testing procedure of the $HNC$ hypothesis is very simple and works as follows. For each individual of the panel, we compute the standard Wald statistic $W_{i,T}$ associated to the individual hypothesis $H_{0,i} : \beta_i = 0$ with $\beta_i \in \mathbb{R}^K$. Given these $N$ realizations, we get a realization of the average Wald statistic $W^N_{Hnc,T,K}$. We compute the realization of the approximated standardized\(^3\) statistic $\tilde{Z}^N_{Hnc}$ according to the formula (26) or we compute the statistic $\tilde{Z}^N_{MC}$ based on the Monte Carlo procedure previously described. For a large $N$ sample, if the value of $\tilde{Z}^N_{Hnc}$ (or $\tilde{Z}^N_{MC}$) is superior to the normal corresponding critical value for a given level of risk, the homogeneous non causality ($HNC$) hypothesis is rejected.

When the panel is unbalanced or when the lag order $K_i$ is specific to each cross-unit, the standardized statistic $\tilde{Z}^N_{Hnc}$ must be adapted as follows:

$$\tilde{Z}^N_{Hnc} = \frac{\sqrt{N} \left[ W^N_{Hnc,T,K} - N^{-1} \sum_{i=1}^{N} E \left( \tilde{W}_{i,T} \right) \right]}{\sqrt{N^{-1} \sum_{i=1}^{N} Var \left( \tilde{W}_{i,T} \right)}}$$

$$= \sqrt{N} \left[ W^N_{Hnc,T,K} - N^{-1} \sum_{i=1}^{N} K_i \times \frac{(T_i - 2K_i - 1)}{(T_i - 2K_i - 3)} \right]$$

$$\times \left[ N^{-1} \sum_{i=1}^{N} 2K_i \times \frac{(T_i - 2K_i - 1)^2 \times (T_i - K_i - 3)}{(T_i - 2K_i - 3)^2 \times (T_i - 2K_i - 5)} \right]^{-1/2} \ (27)$$

\(^3\)If one uses the standard definition of the Wald statistic with the $T$ normalization, it is necessary to adapt the formula (26) by substituting the quantity $T - 2K - 1$ by $T$. More precisely, if the Wald individual statistic $W_{i,T}$ is defined as:

$$W_{i,T} = \left\{ \hat{\beta}_i R' \left[ R (Z'_i Z_i)^{-1} R' \right]^{-1} R \hat{\theta} \right\} / \left[ \hat{\varepsilon}'_i R' \hat{\theta} \right]$$

then the standardized average Wald statistic $\tilde{Z}^N_{Hnc}$ is defined as:

$$\tilde{Z}^N_{Hnc} = \frac{\sqrt{N}}{2 \times K} \times \frac{(T - 4)}{(T + K - 2)} \times \left[ \left( \frac{T - 2}{T} \right) W^N_{Hnc,T,K} - K \right]$$
where $T_i > 5 + 2K_i$ denotes the time dimension for the $i^{th}$ cross-section unit.

## 5 Monte Carlo simulation results

In this section, we propose some Monte Carlo experiments to examine finite sample properties of the alternative panel-based non causality tests. We consider three sets of Monte Carlo experiments. The first set focuses on the benchmark model:\footnote{We also carried out number of other experiments with other data generating processes. The results are similar to the ones reported in this section and are available from the author on request.}

\begin{equation}
y_{i,t} = \alpha_i + \gamma_i y_{i,t-k} + \beta_i x_{i,t-k} + \varepsilon_{i,t} \quad (28)
\end{equation}

The parameters of the model are calibrated as follows. The auto-regressive parameters $\gamma_i$ are drawn according to a uniform distribution on $]-1, 1[$ in order to satisfy the stationarity assumption $A_3$. The fixed individual effects $\alpha_i, i = 1, \ldots, N$ are generated according to a $N(0, 1)$. Individual residuals are drawn in normal distribution with zero means and heterogeneous variances $\sigma_{\varepsilon,i}^2$. The variance $\sigma_{\varepsilon,i}^2$ are generated according to a uniform distribution on $[0.5, 1.5]$. Under the null of $HNC$, $\beta_i = 0$ for all $i$. Under the alternative, $\beta_i$ is different from 0 for all $i$, i.e. $N_1 = 0$. In this case, parameters $\beta_i$ are generated according to a $N(0, 1)$ at each simulation.

The second set of experiments allows for heterogeneity of the causality relationships under the alternative $H_1 : \beta_i = 0$ for $i = 1, \ldots, N_1$ and $\beta_i \neq 0$ for $i = N_1 + 1, \ldots, N$. In these experiments, we evaluate the empirical power of our panel tests for various values of the ratio $n_1 = N_1/N$. We consider a case in which there is no causality for one cross-section unit out of two ($n_1 = 0.5$) and a case with no causality for nine cross-section units out of ten ($n_1 = 0.9$).

The third set of experiments focuses on a model with $K$ lags:

\begin{equation}
y_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_i^{(k)} y_{i,t-k} + \sum_{k=1}^{K} \beta_i^{(k)} x_{i,t-k} + \varepsilon_{i,t} \quad (29)
\end{equation}

where the auto-regressive parameters $\gamma_i^{(k)}$ are drawn according to a uniform distribution on $]-K, K[$ under the constraint that the roots of $\Gamma_i(z) = \sum_{k=1}^{K} \gamma_i^{(k)} z^k$ lie outside the unit circle. The others parameters are calibrated as in the first set of experiments. We consider two cases, denoted $A$ and $B$. In Monte Carlo experiments of case $A$, we
compute the size and the power \((n_1 = 0)\) of our panel tests for a lag order \(K\) equal to 2. In case \(B\), we assume that there is a mis-specification of the lag order. The underlying data are generated by a model with one lag \((K = 1)\), but the individual Wald statistics (and the corresponding standardized average panel statistics) are computed from the simulated series with a regression model with two lags \((K = 2)\).

The second set of experiments were carried out for \(N = 6\) (only for the case \(n_1 = 0.5\)), 10, 20, 50 and \(T = 10, 25, 50, 100\). The other experiments were carried out for \(N = 1, 5, 10, 25, 50\) and \(T = 10, 25, 50, 100\). We used 10,000 replications to compute empirical size and power of the tests at the 5\% nominal size. All the parameters values such as \(\alpha_i, \gamma_i, \sigma_{\varepsilon,i}\) or \(\beta_i\) are generated independently at each simulation.

All the experiments are carried out using the following two statistics: the \(Z_{N,T}^{Hnc}\) based on the asymptotic moments (equation 9) and the \(\tilde{Z}_{N}^{Hnc}\) based on the approximation of moments for a fixed \(T\) sample (equation 26). The results of the first set of experiments are summarized in Table 1. As a benchmark, in the first row of this table we report the results for the Granger non-causality test based on a Wald statistic and single time series \((N = 1)\). For large \(T\) samples, the standardized statistic \(Z_{N,T}^{Hnc}\) based on the asymptotic moments \(K\) and \(2K\) (which are valid if \(T\) tends to infinity) has a correct size. Our panel test is more powerful than tests based on single time series even in panel with very few cross-section units. For instance, for a typical panel of macroeconomic annual data \((T = 50)\), the power of non causality test rises from 0.71 with a single time series test \((N = 1)\) to 0.99 with a panel test as soon as only five cross-section units are included \((N = 5)\). However, for small values of \(T\), the standardized statistic \(Z_{N,T}^{Hnc}\) is oversized and the extent of this over-rejection worsens as \(N\) increases. This over-rejection can be intuitively understood as follows. The Wald statistic based on single time series is slightly over-sized for small values of \(T\). So, under the null, we can observe large values (superior to the chi-squared critical value) of the individual Wald statistics for some cross-section units. For a given value of \(N\), these large values (that range from the chi-squared critical value to infinity) are not annihilated by the realisations obtained for other cross-section units since the latter only range from 0 to
the chi-squared critical value. Consequently, the cross-section average \((W_{N,T}^{Hnc})\) tends to be larger than the corresponding normal critical value. The more \(N\) increases, the more the probability to obtain large values for some cross-section unit increases. So, for small values of \(T\), the \(Z_{N,T}^{Hnc}\) test tends to over-reject the null of non causality and this tendency is stronger as \(N\) is allowed to increase.

On the contrary, the size of the standardized \(\tilde{Z}_{N}^{Hnc}\)-statistic based on the semi-asymptotic moments (defined for fixed values of \(T\)) is reasonably close to the nominal size for all values of \(T\) and \(N\). The semi asymptotic standardized \(\tilde{Z}_{N}^{Hnc}\)-statistic substantially augments the power of non-causality tests applied to single time series even for very small values of \(N\). For example, when \(T = 10\), the power of our panel test is equal to 0.73 even with only five cross-section units \((N = 5)\). In this case, the test based on time series \((N = 1)\) has only a power of 0.43. The \(\tilde{Z}_{N}^{Hnc}\)-statistic has a correct size, and its power rises monotonically and quickly with \(N\) and \(T\). For \(T = 10\), when \(N\) is larger than 10, the power of the \(\tilde{Z}_{N}^{Hnc}\)-test is near to one. This improvement in power can be intuitively understood as follows. Individual statistics are bounded from below (by zero) but may take arbitrarily large value. Hence, when averaging among individual Wald statistics, the ‘abnormal’ realisations (realisations below the chi-squared critical value) are annihilated by the realisations on the true side (large).

In power simulations summarized in Table 1, we assume that there is causality for all the cross-section units of the panel. The issue is now to determine the influence of heterogeneity of causality relationships, i.e. the relative importance to \(N_1\) with respect to \(N\), on the power of our panel tests. The results for the second set of experiments are summarized in Table 2. For \(n_1 = 0.5\) and \(n_1 = 0.9\), we can verify that the powers of the standardized statistics \(Z_{N,T}^{Hnc}\) and \(\tilde{Z}_{N}^{Hnc}\) are slightly reduced compared to the case \(n_1 = 0\) (Table 1). But, even in the worse case studied (in which there is causality for only one cross-section unit out of ten, i.e. \(n_1 = 0.9\)), the powers of our panel tests remain reasonable even for very small values of \(T\) and \(N\). For instance, with \(T = 25\) and \(N = 10\) \((N_1 = 9)\), the power of the \(\tilde{Z}_{N}^{Hnc}\)-statistic is equal to 0.42. With twenty cross-section units (causality for two cross-sections units if \(n_1 = 0.9\)), its power increases.
The results for the third set of experiments are summarized in Table 3. In case A, we consider a model with two lags. The results are quite similar to the results obtained for the benchmark case with one lag (Table 1): the powers of the panel average statistics substantially exceed that of single times series non-causality test, the $Z_{N,t}^{Hnc}$ statistic is over-sized and the $\tilde{Z}_N^{Hnc}$ has a correct size for all $T$ and $N$ values. Similar results (not reported) are obtained when we consider heterogeneous lag orders $K_i$. In case B, we study the influence of a mi-specification of the lag-order. When the lag order is over estimated for all the cross-section units, the power of our panel test statistics is reduced but remains reasonable. With $T = 10$, the power of the panel $\tilde{Z}_N^{Hnc}$ statistic rises from 0.36 with five cross-section units to 0.87 with twenty cross-section units.

6 Fixed $T$ and fixed $N$ distributions

If $N$ and $T$ are fixed, the standardized statistic $\tilde{Z}_N^{Hnc}$ and the average statistic $W_{N,T}^{Hnc}$ do not converge to standard distributions under the $HNC$ hypothesis. Two solutions are then possible: the first consists in using the mean Wald statistic $W_{N,T}^{Hnc}$ and to compute the exact sample critical values, denoted $c_{N,T}(\alpha)$, for the corresponding sizes $N$ and $T$ via stochastic simulations. We propose the results of an example of such a simulation in table ?? As in Im, Pesaran and Shin (2003), the second solution consists in using the approximated standardized statistic $\tilde{Z}_N^{Hnc}$ and to compute an approximation of the corresponding critical value for a fixed $N$. Indeed, we can show that:

$$\Pr[\tilde{Z}_N^{Hnc} < \tilde{z}_N(\alpha)] = \Pr[W_{N,T}^{Hnc} < c_{N,T}(\alpha)]$$

where $\tilde{z}_N(\alpha)$ is the $\alpha$ percent critical value of the distribution of the standardized statistic under the $HNC$ hypothesis. The critical value $c_{N,T}(\alpha)$ of $W_{N,T}^{Hnc}$ is defined as:

$$c_{N,T}(\alpha) = \tilde{z}_N(\alpha) \sqrt{N^{-1} \text{var} \left( \tilde{W}_{i,T} \right)} + E \left( \tilde{W}_{i,T} \right)$$

where $E \left( \tilde{W}_{i,T} \right)$ and $\text{Var} \left( \tilde{W}_{i,T} \right)$ respectively denote the mean and the variance of the individual Wald statistic defined by equations (21) and (22). Given the result of proposition (6), we know that the critical value $\tilde{z}_N(\alpha)$ corresponds to the $\alpha$ percent.
critical value of the standard normal distribution, denoted $z_\alpha$ if $N$ tends to infinity whatever the size $T$. For a fixed $N$, the use of the normal critical value $z_\alpha$ to build the corresponding critical value $c_{N,T}(\alpha)$ is not founded, but however we can propose an approximation $\tilde{c}_{N,T}(\alpha)$ based on this value.

$$
\tilde{c}_{N,T}(\alpha) = z_\alpha \sqrt{N^{-1} \text{var}\left(\widetilde{W}_{i,T}\right)} + E\left(\widetilde{W}_{i,T}\right)
$$

(30)

or equivalently:

$$
\tilde{c}_{N,T}(\alpha) = z_\alpha \times \frac{(T - 2K - 1)}{(T - 2K - 3)} \times \sqrt{\frac{2K}{N} \times \frac{(T - K - 3)}{(T - 2K - 5)} + \frac{K \times (T - 2K - 1)}{(T - 2K - 3)}}
$$

(31)

In Table 4 the simulated 5% critical values $c_{N,T}(0.05)$ get from 50 000 replications of the benchmark model under $H_0$ are reproduced. The approximated 5% critical values $\tilde{c}_{N,T}(0.05)$ are also reported. As we can observe, both critical values are very similar and the same result can be obtained for larger lag-order $K$. 

23


7 Conclusion

In this paper, we propose a simple Granger (1969) non-causality test for heterogenous panel data models. Under the null hypothesis of Homogeneous Non Causality (HNC), there is no causal relationship for all the cross-section units of the panel. Under the alternative, there are two subgroups of cross-section units: one with causal relationships from \( x \) to \( y \) (but not necessarily with the same DGP) and another subgroup where there is no causal relationship from \( x \) to \( y \). As in panel unit root test literature, our test statistic is simply defined as the cross-section average of individual Wald statistics associated to the standard Granger causality tests based on single time series. Under a cross-section independence assumption, we show that this average statistic converge to a standard normal distribution when \( T \) and \( N \) tend sequentially to infinity. For fixed \( T \) sample, the semi-asymptotic distribution is characterized. In this case, individual Wald statistics do not have a standard chi-squared distribution. However, under very general setting, Wald statistics are independently distributed with finite second order moments. For a fixed \( T \), the Lyapunov central limit theorem is then sufficient to get the distribution of the standardized average Wald statistic when \( N \) tends to infinity. The two first moments of this normal semi-asymptotic distribution correspond to the cross-section averages of the corresponding theoretical moments of the individual Wald statistics. The issue is then to evaluate the two first moments of standard Wald statistics for small \( T \) samples. In this paper we propose a general approximation of these moments and a corresponding standardized average Wald statistic.

One of the main advantages of our testing procedure is that it is very simple to implement: the standardized average Wald statistics are simple to compute and have a standard normal asymptotic distribution. Besides, Monte Carlo simulations show that our panel statistics lead to substantially augment the power of the Granger non-causality tests applied to single time series even for samples with very small \( T \) and \( N \) dimensions. Finally, our test statistics (based on cross section average of individual Wald statistics) do not require any particular panel estimation.
Our testing procedure has the same advantages, but also the same drawbacks as the approach used by Im, Pesaran and Shin (2003) in the context of panel unit root tests. Firstly, the rejection of the null of Homogeneous Non Causality does not provide any guidance as to the number or the identity of the particular panel members for which the null of non causality is rejected. Secondly, the asymptotic distribution of our statistics is established under the assumption of cross-section independence. As for panel unit root tests, it is now necessary to develop second generation panel non causality tests that allow for general or specific cross-section dependences. This is precisely our objective for further researches.
Appendix

A Exact moments of individual Wald $\tilde{W}_{i,T}$

The two noncentered moments of $\tilde{W}_{i,T}$ are respectively defined as:

\[
E\left(\tilde{W}_{i,T}\right) = (T - 2K - 1) \times K \times \int_0^\infty (1 + 2t)^{-\frac{T-2K-1}{2}} \, dt
\]

\[
E\left[\left(\tilde{W}_{i,T}\right)^2\right] = (T - 2K - 1)^2 \times (2K + K^2) \times \int_0^\infty t \times (1 + 2t)^{-\frac{T-2K-1}{2}} \, dt
\]

Let us denote for simplicity $\bar{T} = (T - 2K - 1)/2$. For the first order moment, we get:

\[
E\left(\tilde{W}_{i,T}\right) = 2\bar{T} \times K \times \int_0^\infty (1 + 2t)^{-\frac{\bar{T}}{2}} \, dt
\]

\[
= 2\bar{T} \times K \times \left[\frac{(1 + 2t)^{-\frac{\bar{T}}{2}+1}}{2(-\bar{T} + 1)}\right]_0^\infty
\]

\[
= \frac{2\bar{T} \times K}{2(\bar{T} - 1)}
\]

Since the quantity $2(\bar{T} - 1) = T - 2K - 3$ is strictly different from zero under the condition of proposition (2), we get

\[
E\left(\tilde{W}_{i,T}\right) = K \times \frac{(T - 2K - 1)}{(T - 2K - 3)}
\]  \hspace{1cm} (32)

For the second order moment, the definition is:

\[
E\left[\left(\tilde{W}_{i,T}\right)^2\right] = 4 \bar{T}^2 \times (2K + K^2) \times \int_0^\infty t \times (1 + 2t)^{-\frac{\bar{T}}{2}} \, dt
\]

By integrating by parts, this expression can be transformed as:

\[
E\left[\left(\tilde{W}_{i,T}\right)^2\right] = 4 \bar{T}^2 \times (2K + K^2) \times \left\{\left[t \times (1 + 2t)^{-\frac{\bar{T}}{2}+1}\right]^{\infty}_0 - \frac{1}{2(-\bar{T} + 1)} \times \int_0^\infty (1 + 2t)^{-\frac{\bar{T}}{2}} \, dt\right\}
\]

Under the condition of proposition (2) we have $\bar{T} > 1$, then:

\[
E\left[\left(\tilde{W}_{i,T}\right)^2\right] = \frac{4 \bar{T}^2 \times (2K + K^2)}{2(\bar{T} - 1)} \times \int_0^\infty (1 + 2t)^{-\frac{\bar{T}}{2}} \, dt
\]

\[
= \frac{4 \bar{T}^2 \times (2K + K^2)}{2(\bar{T} - 1)} \times \left[\frac{(1 + 2t)^{-\frac{\bar{T}}{2}+2}}{2(-\bar{T} + 2)}\right]^{\infty}_0
\]

\[
= \frac{4 \bar{T}^2 \times (2K + K^2)}{2(\bar{T} - 1)} \times \frac{1}{2(\bar{T} - 2)}
\]
After simplifications, we have:

\[
E \left[ \left( \tilde{W}_{i,T} \right)^2 \right] = \frac{\bar{T}^2 \times (2K + K^2)}{(\bar{T} - 1)(\bar{T} - 2)} = \frac{(T - 2K - 1)^2 \times (2K + K^2)}{(T - 2K - 3)(T - 2K - 5)}
\] (33)

Under the condition \( T > 5 + 2K \), this second order moment exists as it was previously established in proposition (2).

Finally, we can compute the second order centered moment, \( Var(\tilde{W}_{i,T}) \) as:

\[
Var(\tilde{W}_{i,T}) = E \left[ \left( \tilde{W}_{i,T} \right)^2 \right] - E \left( \tilde{W}_{i,T} \right)^2
\]

\[
= \frac{(T - 2K - 1)^2 \times (2K + K^2)}{(T - 2K - 3)(T - 2K - 5)} - \left[ \frac{K \times (T - 2K - 1)}{(T - 2K - 3)} \right]^2
\]

After simplifications, we have:

\[
Var(\tilde{W}_{i,T}) = 2K \times \frac{(T - 2K - 1)^2 \times (T - K - 3)}{(T - 2K - 3)^2(T - 2K - 5)}
\] (34)
References


<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test</td>
<td>Size</td>
<td>Power</td>
<td>Size</td>
<td>Power</td>
<td>Size</td>
<td>Power</td>
</tr>
<tr>
<td>----</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>1</td>
<td>Wald</td>
<td>0.09</td>
<td>0.43</td>
<td>0.06</td>
<td>0.62</td>
<td>0.05</td>
<td>0.71</td>
</tr>
<tr>
<td>5</td>
<td>Z_{Nch}</td>
<td>0.16</td>
<td>0.88</td>
<td>0.07</td>
<td>0.98</td>
<td>0.06</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.04</td>
<td>0.73</td>
<td>0.04</td>
<td>0.97</td>
<td>0.04</td>
<td>0.99</td>
</tr>
<tr>
<td>10</td>
<td>Z_{Nch}</td>
<td>0.21</td>
<td>0.98</td>
<td>0.08</td>
<td>0.99</td>
<td>0.06</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.04</td>
<td>0.91</td>
<td>0.04</td>
<td>0.99</td>
<td>0.04</td>
<td>1.00</td>
</tr>
<tr>
<td>25</td>
<td>Z_{Nch}</td>
<td>0.31</td>
<td>1.00</td>
<td>0.09</td>
<td>1.00</td>
<td>0.06</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.04</td>
<td>0.99</td>
<td>0.04</td>
<td>1.00</td>
<td>0.04</td>
<td>1.00</td>
</tr>
<tr>
<td>50</td>
<td>Z_{Nch}</td>
<td>0.44</td>
<td>1.00</td>
<td>0.11</td>
<td>1.00</td>
<td>0.07</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.04</td>
<td>1.00</td>
<td>0.04</td>
<td>1.00</td>
<td>0.05</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Notes: This table reports the size and power of the Wald statistic based on time series \(N = 1\), the panel standardized statistic \(Z_{N,T}^{HNC}\) based on asymptotic moments defined by (9) and the panel standardized statistic \(\tilde{Z}_{N,T}^{HNC}\) based on semi-asymptotic moments defined by (26). The underlying data are generated by \(y_{i,t} = \alpha_i + \gamma_i y_{i,t-k} + \beta_i x_{i,t-k} + \varepsilon_{i,t}\), for \(i = 1, \ldots, N\) and \(t = -100, -99, \ldots, T\). At each replication, the autoregressive parameters \(\gamma_i\) are drawn according to a uniform distribution on \([-1, 1]\) and the fixed individual effects \(\alpha_i\) are generated according to a \(N(0, 1)\). Individual residuals are \(N.i.d. \left(0, \sigma^2_{\varepsilon_i}\right)\). The variance \(\sigma^2_{\varepsilon_i}\) are generated according to a uniform distribution on \([0.5, 1.5]\). The size (\(\beta_i = 0, i = 1, \ldots, N\)) and the power of the tests are computed at the five percent nominal level. Under the alternative (power simulations), \(\beta_i\) is different from 0 for all \(i\), i.e. \(N_1 = 0\). The parameters \(\beta_i\) are generated according to a \(N(0, 1)\). Number of replications is set to 10,000.
Table 2: Power of Panel Granger Non-causality Tests: Experiments with Heterogeneity in Causality Relationships ($n_1>0$)

<table>
<thead>
<tr>
<th>$N$</th>
<th>$N_1$</th>
<th>Test</th>
<th>$T = 10$</th>
<th>$T = 25$</th>
<th>$T = 50$</th>
<th>$T = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>3</td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.72</td>
<td>0.89</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.48</td>
<td>0.87</td>
<td>0.95</td>
<td>0.98</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.85</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.63</td>
<td>0.96</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.85</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>50</td>
<td>25</td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N$</th>
<th>$N_1$</th>
<th>Test</th>
<th>$T = 10$</th>
<th>$T = 25$</th>
<th>$T = 50$</th>
<th>$T = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>—</td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.40</td>
<td>0.48</td>
<td>0.60</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.16</td>
<td>0.42</td>
<td>0.58</td>
<td>0.71</td>
</tr>
<tr>
<td>20</td>
<td>18</td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.58</td>
<td>0.68</td>
<td>0.81</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.22</td>
<td>0.60</td>
<td>0.79</td>
<td>0.89</td>
</tr>
<tr>
<td>50</td>
<td>45</td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.85</td>
<td>0.92</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{N_{ch}}^{H}$</td>
<td>0.38</td>
<td>0.86</td>
<td>0.97</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Notes: This table reports the power of the panel standardized statistic $Z_{N,T}^{H_{N_{ch}}}$ based on asymptotic moments defined by (9) and the panel standardized statistic $\tilde{Z}_{N,T}^{H_{N_{ch}}}$ based on semi-asymptotic moments defined by (26). The underlying data are generated by $y_{i,t} = \alpha_i + \gamma_i y_{i,t-k} + \beta_i x_{i,t-k} + \varepsilon_{i,t}$, for $i = 1,..,N$ and $t = -100,-99,...,T$. At each replication, the auto-regressive parameters $\gamma_i$ are drawn according to a uniform distribution on $]-1,1[$ and the fixed individual effects $\alpha_i$ are generated according to a $N(0,1)$. Individual residuals are $N.i.d. \left(0, \sigma_{i}^{2} \right)$. The variance $\sigma_{i}^{2}$ are generated according to a uniform distribution on $[0.5,1.5]$. The power is computed at the five percent nominal level. We consider power simulations with heterogeneity of causality relationships. The parameters $\beta_i$ are equal to 0 (non-causality) for $i = 1,..,N_1$ and different from 0 (causality) for $i = N_1+1,..,N$. In this case, $\beta_i$ are generated according to a $N(0,1)$. The ratio $n_1 = N_1/N$, with $0 \leq n_1 < 1$, denotes the fraction of cross-section units for which there is no causality under the alternative.
Table 3: Size and Power of Panel Non-causality Tests: Influence of Lag Order K.

<table>
<thead>
<tr>
<th>Case</th>
<th>DGP with K = 2, model with K = 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( T = 10 )</td>
</tr>
<tr>
<td></td>
<td>( N )</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case</th>
<th>DGP with K = 1, model with K = 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( T = 10 )</td>
</tr>
<tr>
<td></td>
<td>( N )</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports the size and power of the Wald statistic based on time series (\( N = 1 \)), the panel standardized statistic \( Z_{\text{Nch}}^{Hnc} \) based on asymptotic moments defined by (9) and the panel standardized statistic \( \tilde{Z}_{\text{Nch}}^{Hnc} \) based on semi-asymptotic moments defined by (26). The underlying data are generated by \( y_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_i^{(k)} y_{i,t-k} + \sum_{k=1}^{K} \beta_i^{(k)} x_{i,t-k} + \varepsilon_{i,t} \), for \( i = 1, \ldots, N \) and \( t = -100, -99, \ldots, T \). At each replication, the auto-regressive parameters \( \gamma_i^{(k)} \) are drawn according to a uniform distribution on \([-K, K]\) under the constraint that the roots of \( \Gamma_i(z) = \sum_{k=1}^{K} \gamma_i^{(k)} z^k \) lie outside the unit circle. The fixed individual effects \( \alpha_i \) are generated according to a \( N(0,1) \). Individual residuals are \( N.i.d. \left( 0, \sigma_{\varepsilon,i}^2 \right) \). The variance \( \sigma_{\varepsilon,i}^2 \) are generated according to a uniform distribution on \([0.5, 1.5]\). In case B, the data are generated by a model with two lags (\( K = 2 \)) whereas the individual Wald statistics are computed from a model that includes (at wrong) only one lag (\( K = 1 \)). The size \( (\beta_i = 0, i = 1, \ldots, N) \) and the power of the tests are computed at the five percent nominal level. Under the alternative (power simulations), \( \beta_i \) is different from 0 for all \( i \), i.e. \( N_1 = 0 \). The parameters \( \beta_i \) are generated according to a \( N(0,1) \). Number of replications is set to 5,000.
Table 4: Comparison of Simulated and Approximated Critical Values for Fixed N and T samples.

<table>
<thead>
<tr>
<th></th>
<th>Simulated 5% Critical Values $c_{N,T} (0.05)$</th>
<th></th>
<th>Approximated 5% Critical Values $\tilde{c}_{N,T} (0.05)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N \backslash T$</td>
<td>10</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>3.54</td>
<td>2.87</td>
<td>2.66</td>
</tr>
<tr>
<td>10</td>
<td>2.97</td>
<td>2.38</td>
<td>2.19</td>
</tr>
<tr>
<td>15</td>
<td>2.68</td>
<td>2.15</td>
<td>1.99</td>
</tr>
<tr>
<td>20</td>
<td>2.49</td>
<td>2.01</td>
<td>1.86</td>
</tr>
<tr>
<td>25</td>
<td>2.40</td>
<td>1.92</td>
<td>1.78</td>
</tr>
</tbody>
</table>

Notes: The approximated critical values for the average statistic $W_{N,T}^{Hnc}$ are computed from equation (31) for the case $K = 1$. The simulated critical values are computed via stochastic simulations with 50,000 replications. The individual Wald statistics $W_{i,T}$ are built under the $HNC$ hypothesis, where the auto-regressive parameters $\gamma_i^{(k)}$ are drawn according to a uniform distribution on $]-1,1[$. The fixed individual effects $\alpha_i$, $i = 1, .., N$ are simulated according a $N (0,1)$. 

32