Testing for Granger Non-causality in Heterogeneous Panels

Christophe Hurlin, Elena Dumitrescu

▶ To cite this version:

Christophe Hurlin, Elena Dumitrescu. Testing for Granger Non-causality in Heterogeneous Panels. 2012. halshs-00224434v2

Testing for Granger Non-causality
in Heterogeneous Panels

Elena-Ivona Dumitrescu[*]   Christophe Hurlin[†‡]

December 2011

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 cross-section units. First, this statistic is shown to converge sequentially to a standard normal distribution. Second, the semi-asymptotic distribution of the average statistic is characterized for a fixed T sample. A standardized statistic based on an approximation of the moments of Wald statistics is hence proposed. Third, Monte Carlo experiments show that our standardized panel statistics have very good small sample properties, even in the presence of cross-sectional dependence.

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

1 Introduction

The aim of this paper is to propose a simple Granger (1969) non causality test in heterogeneous panel data models with fixed (as opposed to time-varying) coefficients. In the framework of a linear autoregressive data generating process, the extension of standard causality tests to panel data implies testing cross sectional linear restrictions on the coefficients of the model. As usual, the use of cross-sectional information may extend the information set on causality from a given variable to another. Indeed, in many economic matters it is highly probable that if a causal relationship exists for a country or an individual, it also exists for some other countries or individuals. In this case, the causality can be more efficiently tested in a panel context with NT observations. However, the use of cross-sectional information involves taking 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.*" Consequently, we propose here a simple Granger non causality test for heterogeneous panel data models. This test allows us to take into account both dimensions of the heterogeneity present in this context: the heterogeneity of the causal relationships and the heterogeneity of the regression model used so as to test for Granger causality.

Let us consider the standard implication of Granger causality [1]. For each individual, we say that variable x causes y if we are able to better predict y using all available information than in the case where the information set used does not include x (Granger 1969). If x and y are observed on N individuals, gauging the presence of causality comes down to determining the optimal information set used to forecast y. Several solutions can be adopted. The most general one consists in testing the causality from variable x observed for the ith individual to the variable y observed for the jth individual, with $j = i$ or $j \neq i$. The second solution is more restrictive and derives directly from the time series analysis. It implies testing the causal relationship for a given individual. The cross-sectional information is then used only 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. In this case, causality tests can be implemented and considered as a natural extension of the standard time series tests in the cross-sectional dimension.

However, one of the main issues specific to panel data models refers to the specification of the heterogeneity between cross-section units. In this Granger causality context, the het-

---

[1]The definition of Granger causality 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 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).

erogeneity has two main dimensions. We hence distinguish between the heterogeneity of the regression model and that of the causal relationship from x to y. Indeed, the model considered may be different from an individual to another, whereas there is a causal relationship from x to y for all individuals. The simplest form of regression model heterogeneity takes the form of slope parameters' heterogeneity. More precisely, in a p order linear vectorial autoregressive model, we define four kinds of causal relationships. The first one, denoted *Homogenous Non Causality* (HNC) hypothesis, implies that no individual causality relationship from x to y exists. The symmetric case is the *Homogenous Causality* (HC) hypothesis, which occurs when N causality relationships exist, and when the individual predictors of y obtained conditionally on the past values of y and x are identical. The dynamics of y is then absolutely identical for all the individuals in the sample. The last two cases correspond to heterogeneous processes. Under the *HEterogenous Causality* (HEC) hypothesis, we assume that N causality relationships exist, as in the HC case, but the dynamics of y is heterogenous. Note, however, that the heterogeneity does not affect the causality result. Finally, under the *HEterogenous Non Causality* ($HENC$) hypothesis, we assume that there is a causal relationship from x to y for a subgroup of individuals. Symmetrically, there is at least one and at most $N-1$ non causal relationships in the model. It is clear that in this case the heterogeneity deals with causality from x to y.

To sum up, under the HNC hypothesis, no individual causality from x to y occurs. On the contrary, in the HC and HEC cases, there is a causality relationship for each individual of the sample. To be more precise, in the HC case, the same regression model is valid (identical parameters' estimators) for all individuals, whereas this is not the case for the HEC hypothesis. Finally, under the $HENC$ hypothesis, the causality relationship is heterogeneous since the variable x causes y only for a subgroup of $N-N_1$ units.

In this context, we propose a simple test of the Homogenous Non Causality (HNC) hypothesis. Under the null hypothesis, there is no causal relationship for any of the units of the panel. Our contribution is three-fold. First, we specify the alternative as the $HENC$ hypothesis. To put it differently, we do not test the HNC hypothesis against the HC hypothesis as Holtz-Eakin, Newey and Rosen (1988), which, as previously discussed, is a strong assumption. Indeed, we allow for two subgroups of cross-section units: the first one is characterized by causal relationships from x to y, but it does not necessarily rely on the same regression model, whereas there is no causal relationships from x to y in the case of the second subgroup. Second, we consider a heterogenous panel data model with fixed coefficients (in time). It follows that both under the null and the alternative hypothesis the unconstrained parameters may be different from one individual to another. The dynamics of the variables may be thus heterogeneous across the cross-section units, regardless of the existence (or not) of causal relationships. Our framework hence relies on less strong assumptions than the ones in Holtz-Eakin, Newey and Rosen (1988), who assume the homogeneity of cross-section units, *i.e.* that the panel vector-autoregressive regression model is valid for all the individuals in the panel. Third, we adapt the Granger causality test-statistic to the case of unbalanced panels

3

and / or different lag orders in the autoregressive process. Most importantly, we propose a block bootstrap procedure to correct the empirical critical values of panel Granger causality tests so as to account for cross-sectional dependence. To our knowledge, these issues have not been tackled before in this context.

Following the literature devoted to panel unit root tests in heterogeneous panels, and particularly Im, Pesaran and Shin (2003), we propose a test statistic based on averaging standard individual Wald statistics of Granger non causality tests[2]. 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. For a fixed T, the Lyapunov central limit theorem is sufficient to establish the distribution of the standardized average Wald statistic when N tends to infinity. The first two moments of this normal semi-asymptotic distribution correspond to the empirical mean of the corresponding theoretical moments of the individual Wald statistics. The issue is then to propose an evaluation of the first two moments of standard Wald statistics for small T samples. A first solution relies on Monte-Carlo or Bootstrap simulations. A second one consists in using an approximation of these moments based on the exact moments of the ratio of quadratic forms in normal variables derived from 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. Then, contrary to Konya (2006), our testing procedure does not require bootstrap critical values generated by simulations. However, a block bootstrap simulation approach similar to theirs is adapted to our framework (group mean Wald-statistic) so as to take into account cross-sectional dependencies.

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 on the semi-asymptotic moments is reasonably close to the nominal size for all the values of T and N considered. Besides, the power of our panel test statistic substantially exceeds that of Granger non Causality tests based on single time series in all experiments and in particular for very small values of T, e.g. $T = 10$, provided that there are at least a

---

[2]The idea of this test was first exposed at the LIIIe annual congress of the French Economic Association (Hurlin, 2005). However, Hurlin (2005) is only a short note, that does not formally develop the asymptotic and semi-asymptotic theory and which relies on only one Monte-Carlo experiment. Moreover, the cross-sectional dependence issue is not tackled and the case of unbalanced panels is not discussed. The present paper thus goes beyond that note from the point of view of both theoretical proofs and empirical results. Besides, as a working paper, it has been very often cited in this literature.

few cross-section units in the panel (e.g. $N = 5$). Furthermore, approximated critical values are proposed for finite T and N samples, as well as a block-bootstrap procedure to compute empirical critical values when taking into account cross-section dependence.

The rest of 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 distribution of the average Wald statistic. Section 4 derives the semi-asymptotic distribution for fixed T sample and section 5 presents the main results obtained from Monte Carlo experiments. Section 6 extends the results to a fixed N sample and discusses the case with cross-sectional dependence as well as the unbalanced panel framework. The last section provides some concluding remarks.

2 A non causality test in heterogenous panel data models

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

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

with $K \in \mathbb{N}^*$ and $\beta_i = \left(\beta_i^{(1)}, ..., \beta_i^{(K)} \right)'$. For simplicity, the individual effects α_i are supposed to be fixed in the time dimension. Initial conditions $(y_{i,-K}, ..., y_{i,0})$ and $(x_{i,-K}, ..., x_{i,0})$ 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. Besides, we allow the autoregressive parameters $\gamma_i^{(k)}$ and the 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 in time. 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. In the sequel, we make the following assumptions.

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

Assumption (A_2) *Individual residuals $\varepsilon_i = (\varepsilon_{i,1}, .., \varepsilon_{i,T})'$, 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 $x_i = (x_{i,1}, ..., x_{i,T})'$ and $y_i = (y_{i,1}, ..., y_{i,T})'$, are covariance stationary with $E(y_{i,t}^2) < \infty$ and $E(x_{i,t}^2) < \infty$. Besides, $E(x_{i,t}x_{j,z})$, $E(y_{i,t}y_{j,z})$ 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 model with two variables constitutes the basic framework for studying Granger causality in a panel data context. If in a time series context, the standard causality tests consist in testing linear restrictions on the vectors β_i, in 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 α_i. The second source, which is more crucial, is related to the heterogeneity of the parameters β_i. This kind of heterogeneity directly affects the paradigm of the representative agent and hence the conclusions with respect to causality relationships. It is well known that the estimates of autoregressive parameters β_i obtained under the wrong hypothesis, i.e. $\beta_i = \beta_j \ \forall (i,j)$, are biased (see Pesaran and Smith 1995 for an $AR(1)$ process). Then, if we impose the homogeneity of coefficients β_i, the causality test-statistics can lead to fallacious inference. Intuitively, the estimate $\widehat{\beta}$ obtained in an homogeneous model will converge to a value close to the average of the true coefficients β_i, and 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 interpret causality relations if at least one individual from the sample has an economic behavior different from that of the others. For example, let us assume that there is a causality relation for a set of N countries, for which the vectors β_i are strictly identical. What conclusions can be drawn if we introduce into the sample a set of N_1 countries for which, in contrast, there is no relation of causality? Whatever the value of the ratio N/N_1 is, the test of the causality hypothesis is nonsensical.

Given these observations, we propose to test the Homogenous Non Causality (HNC) hypothesis by taking into account both the heterogeneity of the regression model and that of the causal relation. Under the alternative we hence allow for a subgroup of individuals for which there is no causality relation and a subgroup of individuals for which the variable x Granger causes y. The null hypothesis of HNC is defined as:

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

with $\beta_i = \left(\beta_i^{(1)}, ..., \beta_i^{(K)} \right)'$. Additionally, β_i may differ across groups under the alternative (model heterogeneity). We also allow for some, but not all, of the individual vectors β_i 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. It follows that our test is not a test of non-causality assumption against 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:

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

where N_1 is unknown but satisfies the condition $0 \le N_1/N < 1$. The ratio N_1/N is necessarily inferior to one, since if $N_1 = N$ there is no causality for any of the individuals in the panel, which is equivalent to the HNC null hypothesis. Conversely, when $N_1 = 0$ there is causality for all the individuals in 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. By contrast, if we assume that the HNC is rejected and $N_1 = 0$, 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. Indeed, the regression model considered may be not homogenous, i.e. the estimators of the parameters differ across groups, but the causality relations are observed for all individuals. On the contrary, if $N_1 > 0$, the causality relationship is heterogeneous: the regression model and the causality relations are different from one individual from the sample to another.

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

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

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

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

To obtain the general form of this statistic, we stack the observations for the T periods corresponding to the ith individual's characteristics into a T elements vector as:

$$\underset{(T,1)}{y_i^{(k)}} = \begin{bmatrix} y_{i,1-k} \\ \cdot \\ \cdot \\ y_{i,T-k} \end{bmatrix} \quad \underset{(T,1)}{x_i^{(k)}} = \begin{bmatrix} x_{i,1-k} \\ \cdot \\ \cdot \\ x_{i,T-k} \end{bmatrix} \quad \underset{(T,1)}{\varepsilon_i} = \begin{bmatrix} \varepsilon_{i,1} \\ \cdot \\ \cdot \\ \varepsilon_{i,T} \end{bmatrix}$$

and we define two (T, K) matrices:

$$Y_i = \left[y_i^{(1)} : y_i^{(2)} : .... : y_i^{(K)} \right] \quad \text{and} \quad X_i = \left[x_i^{(1)} : x_i^{(2)} : .... : x_i^{(K)} \right].$$

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

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

where $\widehat{\theta}_i$ is the estimate of parameter θ_i obtained under the alternative hypothesis, and $\widehat{\sigma}_i^2$ is the estimate of the variance of the residuals. For a small T sample, the corresponding unbiased estimator[3] takes the form of $\widehat{\sigma}_i^2 = \widehat{\varepsilon}_i'\widehat{\varepsilon}_i / \left(T - 2K - 1 \right)$. 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:

$$W_{i,T} = (T - 2K - 1) \left(\frac{\widetilde{\varepsilon}_i'\Phi_i\widetilde{\varepsilon}_i}{\widetilde{\varepsilon}_i'M_i\widetilde{\varepsilon}_i} \right), \quad i = 1, .., N \tag{5}$$

where the $(T, 1)$ vector $\widetilde{\varepsilon}_i = \varepsilon_i / \sigma_{\varepsilon,i}$ is normally distributed according to $N\left(0, I_T\right)$ under assumption A_1. The (T, T) matrices Φ_i and M_i are positive semi definite, symmetric and idempotent

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

$$M_i = I_T - Z_i \left(Z_i'Z_i \right)^{-1} Z_i' \tag{7}$$

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

Our objective now is 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 tend to infinity, and second we tackle 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, i.e. when T tends to infinity and then N tends to infinity. This sequential convergence result can be deduced from the standard convergence result of the individual 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 $\widehat{\theta}_i$ is convergent, we know that $plim\, \varepsilon_i'M_i\varepsilon_i / \left(T - 2K - 1 \right) = \sigma_{\varepsilon,i}^2$. It implies that:

$$\underset{T \to \infty}{\text{plim}} \frac{\widetilde{\varepsilon}_i'M_i\widetilde{\varepsilon}_i}{T - 2K - 1} = \underset{T \to \infty}{\text{plim}} \frac{1}{\sigma_{\varepsilon,i}^2} \left(\frac{\varepsilon_i'M_i\varepsilon_i}{T - 2K - 1} \right) = 1.$$

---

[3]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 normalization.

Then, if the statistic $W_{i,T}$ has a limiting distribution, it is the same distribution of the statistics as that resulting from replacing the denominator by its limiting value, that is to say 1. Thus, $W_{i,T}$ has the same limiting distribution as $\widetilde{\varepsilon}_i' \Phi_i \widetilde{\varepsilon}_i$. Under assumption A_1, the vector $\widetilde{\varepsilon}_i$ is normally distributed, i.e. $\widetilde{\varepsilon}_i \sim N(0, I_T)$. Since Φ_i is idempotent, the quadratic form $\widetilde{\varepsilon}_i' \Phi_i \widetilde{\varepsilon}_i$ is chi-squared distributed with a number of degrees of freedom equal to the rank of Φ_i. The rank of the symmetric idempotent matrix Φ_i is equal to its trace, i.e. K (cf. Appendix). Therefore, under the null hypothesis of non causality, each individual Wald statistic converges to a chi-squared distribution with K degrees of freedom:

$$W_{i,T} \xrightarrow[T \to \infty]{d} \chi^2(K), \qquad \forall i = 1, .., N. \tag{8}$$

In other words, when T tends to infinity, the individual statistics $\{W_{i,T}\}_{i=1}^{N}$ are identically distributed. They are also independent since under assumption A_2 the residuals ε_i and ε_j are independent for $j \neq i$. To sum 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 $W_{N,T}^{Hnc}$ 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 $i = 1, .., N$ 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 $W_{N,T}^{Hnc}$ sequentially converges in distribution.*

$$Z_{N,T}^{Hnc} = \sqrt{\frac{N}{2K}} \left(W_{N,T}^{Hnc} - K \right) \xrightarrow[T,N \to \infty]{d} N(0,1) \tag{9}$$

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 large N and T samples, if the realization of the standardized statistic $Z_{N,T}^{Hnc}$ is superior to the corresponding normal 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 simultaneously tend to infinity.

4   Fixed T samples and semi-asymptotic distributions

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

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

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

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

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

These general conditions are established in the case where matrices A and B are deterministic. In our case, the corresponding matrices M_i and Φ_i are stochastic, even though we assume that variables X_i are deterministic. However, given a fixed T sample, we propose here to apply these conditions to the corresponding realizations of the two matrices, denoted m_i and ϕ_i. To be more precise, in our case the rank of the symmetric idempotent matrix m_i is equal to $T - 2K - 1$. Besides, since the matrix m_i is the projection matrix associated with the realization z_i of Z_i, we have by construction $m_i z_i = 0$, where z_i is of full column rank $2K + 1$, as $T - \text{rank}(m_i) = 2K + 1$. Then, for a given realization ϕ_i, by construction, the product $\phi_i z_i$ is different from zero

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

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

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

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

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

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

Proposition 1. *For a fixed time dimension $T \in \mathbb{N}$, the second order moments of the individual Wald statistic $W_{i,T}$ associated with the test $H_{0,i} : \beta_i = 0$, exist if and only if:*

$$T > 5 + 2K \tag{10}$$

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

Under the condition of proposition 1, for a fixed T sample, the non-causality test-statistic $W_{N,T}^{Hnc}$ is the average of non identically distributed variables $W_{i,T}$, but with finite second order moments. At the same time, under assumption A_2, the vectors of residuals ε_i and ε_j are independent for $j \neq i$. Therefore, individual Wald statistics, $W_{i,T}$, are also independent for $i = 1, .., N$. The distribution of the non causality test statistic $W_{N,T}^{Hnc}$ can hence be derived by relying on the Lyapunov central limit theorem.

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

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

the standardized statistic Z_N^{Hnc} converges in distribution:

$$Z_N^{Hnc} = \frac{\sqrt{N} \left[W_{N,T}^{Hnc} - N^{-1} \sum_{i=1}^{N} E\left(W_{i,T} \right) \right]}{\sqrt{N^{-1} \sum_{i=1}^{N} Var\left(W_{i,T} \right)}} \xrightarrow[N \to \infty]{d} N\left(0, 1 \right) \tag{11}$$

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

The decision rule is the same as in the asymptotic case: if the realization of the standardized statistic Z_N^{Hnc} is superior to the corresponding normal critical value for a given level of risk, the homogeneous non causality (HNC) hypothesis is rejected. For large T, the moments used in theorem (2) are expected to converge to $E\left(W_{i,T} \right) = K$ and $Var\left(W_{i,T} \right) = 2K$ since the 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 theorem 1. However, these values of the asymptotic moments can lead to poor test results when T is small. We should then evaluate the mean and the variance of the Wald statistic

$W_{i,T}$, knowing that this statistic does not have a standard distribution for a fixed T sample.

The issue is now to compute the standardized average statistic Z_N^{Hnc}. There are two main approaches to compute the first two moments of the individual Wald statistics $W_{i,T}$. On the one hand, 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 (γ_i, σ_i and α_i) and the parameters β_i associated with the exogenous variables x_{it}. Then, the variable y_i is simulated under the null with *i.i.d.* normal residual ε_i with zero means and variance σ_i^2 (Monte Carlo) or with re-sampled historical residuals (bootstrap). At each simulation of the processes y_i and x_i the individual Wald statistic $W_{i,T}$ is computed. Finally, using the replications of $W_{i,T}$, we estimate the first two moments of the individual test-statistics for each cross-section unit. Denote by $\widetilde{Z}_N^{MC}$ the corresponding standardized average statistic. It is obvious that this method can be time consuming, especially if we consider very large N panel sets. On the other hand, we propose here an approximation of $E(W_{i,T})$ and $Var(W_{i,T})$ based on the results of Magnus's theorem (1986). Let us consider the expression of the Wald statistic $W_{i,T}$ as a ratio of two quadratic forms in a standard normal vector under assumption A_1:

$$W_{i,T} = (T - 2K - 1) \left(\frac{\widetilde{\varepsilon}_i' \Phi_i \widetilde{\varepsilon}_i}{\widetilde{\varepsilon}_i' M_i \widetilde{\varepsilon}_i} \right) \tag{12}$$

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

$$\widetilde{W}_{i,T} = (T - 2K - 1) \left(\frac{\widetilde{\varepsilon}_i' \phi_i \widetilde{\varepsilon}_i}{\widetilde{\varepsilon}_i' m_i \widetilde{\varepsilon}_i} \right), \tag{13}$$

where the matrices ϕ_i and m_i are positive semi-definite.

Theorem 3 (Magnus 1986). *Let $\widetilde{\varepsilon}_i$ be a normal distributed vector with $E(\widetilde{\varepsilon}_i) = 0$ and $E(\widetilde{\varepsilon}_i \widetilde{\varepsilon}_i') = I_T$. Let P_i be an orthogonal (T, T) matrix and Λ_i a diagonal (T, T) matrix such that*

$$P_i' m_i P_i = \Lambda_i \quad and \quad P_i' P_i = I_T. \tag{14}$$

Then, provided that the expectation for $s = 1, 2, 3..$ exists, we have:

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

where the summation is over all $(s, 1)$ vectors $v = (n_1, .., n_s)$ whose elements n_j are nonneg-

ative integers satisfying $\sum_{j=1}^{s} j n_j = s$,

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

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

$$\Delta_i = (I_T + 2\, t\, \Lambda_i)^{-1/2} \;\; and \;\; R_i = \Delta_i\, P_i' \phi_i P_i\, \Delta_i . \tag{17}$$

In our case, we are interested in the first two 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 also equal to one. For the second order moment ($s = 2$), there are two vectors $v = (n_1, n_2)$ which are defined by $v_1 = (0, 1)$ and $v_2 = (2, 0)$, respectively. Consequently, $\gamma_2 (v_1) = 2$ and $\gamma_2 (v_2) = 1$. Given these results, we can compute the exact two moments corresponding to the statistic $\widetilde{W}_{i,T}$ as:

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

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

where matrices Δ_i and R_i are defined in theorem (3). Both quantities $|\Delta_i|$ and $\mathrm{trace}\, (R_i)$ can be computed analytically in our model given the properties of these matrices. Since Λ_i is issued from the orthogonal decomposition of the idempotent matrix m_i with $\mathrm{rank}(m_i) = T - 2K - 1$ (cf. Appendix), it is a zero matrix except for the first block which is equal to the $T - 2K - 1$ identity matrix (corresponding to the characteristic roots of m_i which are not null). Then, for a scalar $t \in \mathbb{R}^+$, the matrix $\Delta_i = (I_T + 2\, t\, \Lambda_i)^{-1/2}$ can be partitioned as:

$$\underset{(T,T)}{\Delta_i} = \begin{pmatrix} \underset{(T-2K-1,T-2K-1)}{D_i (t)} & \underset{(T-2K-1,2K+1)}{0} \\ \underset{(2K+1,T-2K-1)}{0} & \underset{(2K+1,2K+1)}{I_{2K+1}} \end{pmatrix}$$

where I_p denotes the identity matrix of size p. The diagonal block $D_i (t)$ is defined as $D_i (t) = (1 + 2t)^{-\frac{1}{2}} I_{T-2K-1}$. Therefore, the determinant of Δ_i can be expressed as:

$$|\Delta_i| = (1 + 2t)^{-\left(\frac{T-2K-1}{2}\right)} . \tag{20}$$

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 the rank of A, we obtain:

$$\mathrm{rank}\, (R_i) = \mathrm{rank}\, (\Delta_i\, P_i' \phi_i P_i\, \Delta_i) = \mathrm{rank}\, (\, P_i' \phi_i P_i)$$

since the matrix Δ_i is non singular. Using the same transformation, and given the non

13

singularity of P_i,
$$\text{rank}\,(R_i) = \text{rank}\,(\,P_i'\phi_i P_i\,) = \text{rank}\,(\phi_i)\,.$$

Finally, the rank of the realization ϕ_i is equal to K, the rank of Φ_i so that

$$\text{trace}\,(R_i) = K.$$

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

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

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

The following proposition summarizes these results:

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

$$E\left(\widetilde{W}_{i,T}\right) = K \times \frac{(T - 2K - 1)}{(T - 2K - 3)} \tag{21}$$

$$Var\left(\widetilde{W}_{i,T}\right) = 2K \times \frac{(T - 2K - 1)^2 \times (T - K - 3)}{(T - 2K - 3)^2 \times (T - 2K - 5)}, \tag{22}$$

as long as the time dimension T satisfies $T \geq 6 + 2K$.

For a proof of this proposition see Appendix. Besides, it is essential to verify that for large T samples, the moments of the individual statistic $\widetilde{W}_{i,T}$ converge to the corresponding moments of the asymptotic distribution of $W_{i,T}$, since $\forall i = 1, ..., N$:

$$\lim_{T \to \infty} E\left(\widetilde{W}_{i,T}\right) = K \qquad \lim_{T \to \infty} Var\left(\widetilde{W}_{i,T}\right) = 2K.$$

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 Magnus theorem to approximate the true moments of the Wald statistic given the realizations ϕ_i and m_i is equivalent to asserting that the true distribution of $W_{i,T}$ can be approximated by the F distribution.

In this paper, we propose to approximate the first two moments of the individual Wald statistic $W_{i,T}$ by the first two moments of the statistics $\widetilde{W}_{i,T}$ based on the realizations ϕ_i and m_i of the stochastic matrices Φ_i and M_i (equations 21 and 22). Indeed, for $T \geq 6 + 2K$, we

assume that:

$$N^{-1} \sum_{i=1}^{N} E\left(W_{i,T}\right) \simeq E\left(\widetilde{W}_{i,T}\right) = K \times \frac{(T - 2K - 1)}{(T - 2K - 3)} \tag{23}$$

$$N^{-1} \sum_{i=1}^{N} Var\left(W_{i,T}\right) \simeq Var\left(\widetilde{W}_{i,T}\right) = 2K \times \frac{(T - 2K - 1)^2 \times (T - K - 3)}{(T - 2K - 3)^2 \times (T - 2K - 5)} \tag{24}$$

Given these approximations, we compute an approximated standardized statistic $\widetilde{Z}_N^{Hnc}$ for the average Wald statistic $W_{N,T}^{Hnc}$ of the HNC hypothesis

$$\widetilde{Z}_N^{Hnc} = \frac{\sqrt{N}\left[W_{N,T}^{Hnc} - E\left(\widetilde{W}_{i,T}\right)\right]}{\sqrt{Var\left(\widetilde{W}_{i,T}\right)}}. \tag{25}$$

For a large N sample, under the Homogenous Non Causality (HNC) hypothesis, we assume that the statistic $\widetilde{Z}_N^{Hnc}$ follows the same distribution as the standardized average Wald statistic Z_N^{Hnc}.

Proposition 3. *Under assumptions A_1 and A_2, for a fixed T dimension with $T > 5 + 2K$, the standardized average statistic $\widetilde{Z}_N^{Hnc}$ converges in distribution:*

$$\widetilde{Z}_N^{Hnc} = \sqrt{\frac{N}{2 \times K} \times \frac{(T - 2K - 5)}{(T - K - 3)}} \times \left[\frac{(T - 2K - 3)}{(T - 2K - 1)} W_{N,T}^{Hnc} - K\right] \xrightarrow[N \to \infty]{d} N(0, 1) \tag{26}$$

with $W_{N,T}^{Hnc} = (1/N) \sum_{i=1}^{N} W_{i,T}$.

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 statistics $W_{i,T}$ associated with the individual hypothesis $H_{0,i} : \beta_i = 0$ with $\beta_i \in \mathbb{R}^K$. Given these N realizations, we obtain a realization of the average Wald statistic $W_{N,T}^{Hnc}$. We then compute the realization of the approximated standardized[4] statistic $\widetilde{Z}_N^{Hnc}$ according to the formula (26) or we compute the statistic $\widetilde{Z}_N^{MC}$ based on the Monte Carlo procedure previously described. For a large N sample, if the value of $\widetilde{Z}_N^{Hnc}$ (or $\widetilde{Z}_N^{MC}$) is superior to the corresponding normal

---

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

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

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

$$\widetilde{Z}_N^{Hnc} = \sqrt{\frac{N}{2 \times K} \times \frac{(T - 4)}{(T + K - 2)}} \times \left[\left(\frac{T - 2}{T}\right) W_{N,T}^{Hnc} - K\right]$$

critical value for a given level of risk, the homogeneous non causality (HNC) hypothesis is rejected.

5   Monte Carlo simulation results

In this section, we propose three sets of Monte Carlo experiments to examine the finite sample properties of the alternative panel-based non causality tests. The first set focuses on the benchmark model:[5]

$$y_{i,t} = \alpha_i + \gamma_i \ y_{i,t-k} + \beta_i \ x_{i,t-k} + \varepsilon_{i,t} \tag{27}$$

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

The second set of experiments allows for heterogeneity of the causality relationships under the alternative $H_1 : \beta_i = 0$ for $i = 1, .., N_1$ and $\beta_i \neq 0$ for $i = N_1 + 1, .., N$. 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:

$$y_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_i^{(k)} \ y_{i,t-k} + \sum_{k=1}^{K} \beta_i^{(k)} \ x_{i,t-k} + \varepsilon_{i,t}, \tag{28}$$

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 other parameters are calibrated as in the first set of experiments. We consider two cases denoted A and B. In the 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 the lag order is misspecified. To be more precise, 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 by relying on a regression model with two lags ($K = 2$).

---

[5]We also carried out several experiments with other data generating processes. The results are similar to the ones reported in this section and are available from the authors on request.

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 the empirical size and power of the tests at the 5% nominal size. All the parameters' values such as α_i, γ_i, $\sigma_{\varepsilon,i}$ or β_i are generated independently at each simulation.

All the experiments are carried out using the following two statistics: $Z_{N,T}^{Hnc}$, based on the asymptotic moments (equation 9), and $\widetilde{Z}_N^{Hnc}$, based on the approximation of moments for a fixed T sample (equation 26). The results for the first set of experiments are summarized in Table 1. As a benchmark, in the first row of this table we report the results of the Granger non-causality test based on a Wald statistic for 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 a panel with very few cross-section units. For instance, for a typical panel of macroeconomic annual data ($T = 50$), the power of the non causality test rises from 0.71 in the case of a single time series test ($N = 1$) to 0.99 for a panel test even though 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 compensated by the realizations obtained for other cross-section units since the latter only range from 0 to the chi-squared critical value. Consequently, the cross-section average ($W_{N,T}^{Hnc}$ statistic) 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 units increases. So, for small values of T, the $Z_{N,T}^{Hnc}$ test tends to over-reject the null of non causality and this propensity becomes stronger as N increases.

On the contrary, the size of the standardized $\widetilde{Z}_N^{Hnc}$ 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 $\widetilde{Z}_N^{Hnc}$ statistic substantially augments the power of non-causality tests even for very small values of N. For example, when $T = 10$, the power of our panel test is equal to 0.73 even though only five cross-section units ($N = 5$) are considered. In this case, the test based on time series ($N = 1$) has only a power of 0.43. All in all, the $\widetilde{Z}_N^{Hnc}$ 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 $\widetilde{Z}_N^{Hnc}$ 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' realizations (realizations

below the chi-squared critical value) are annihilated by the realizations on the true side (large).

In the power simulations summarized in Table 1, we assume that there is causality for all the cross-section units of the panel. By contrast, in the second set of experiments we gauge the influence of heterogeneity of causality relationships, i.e. the relative importance of N_1 with respect to N, on the power of our panel tests. Our findings are summarized in Table 2. For $n_1 = 0.5$ and $n_1 = 0.9$, we can verify that the power of the standardized statistics $Z_{N,t}^{Hnc}$ and $\widetilde{Z}_N^{Hnc}$ is slightly reduced compared to the case $n_1 = 0$ (Table 1). Nevertheless, 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 power of our panel tests remains 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 $\widetilde{Z}_N^{Hnc}$ 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 to 0.60.

The results for the third set of experiments are summarized in Table 5. In case A, we consider a model with two lags. The results are quite similar to the ones obtained for the benchmark case with one lag (Table 1): the power of the panel average statistics substantially exceeds that of single times series non-causality test, the $Z_{N,t}^{Hnc}$ statistic is over-sized and $\widetilde{Z}_N^{Hnc}$ has a correct size for all T and N considered. 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 overestimated for all cross-section units, the power of our panel test statistics is reduced but remains reasonable. With $T = 10$, the power of the panel $\widetilde{Z}_N^{Hnc}$ statistic rises from 0.36 with five cross-section units to 0.87 with twenty cross-section units.

6 Further Issues

6.1 Fixed T and Fixed N Distributions

If N and T are fixed, the standardized statistic 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 envisageable: the first consists in using the mean Wald statistic $W_{N,T}^{Hnc}$ and to compute the exact empirical critical values, denoted $c_{N,T}(\alpha)$, for the corresponding sizes N and T via stochastic simulations. The upper panel in table 4 reports the results of an example of such a simulation. As in Im, Pesaran and Shin (2003), the second solution consists in using the approximated standardized statistic $\widetilde{Z}_N^{Hnc}$ and to compute an approximation of the corresponding critical value for a fixed N. Indeed, we can show that:

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

where $\widetilde{z}_N(\alpha)$ is the α-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:

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

where $E\left(\widetilde{W}_{i,T}\right)$ and $Var\left(\widetilde{W}_{i,T}\right)$ denote the mean and the variance of the individual Wald statistic defined by equations (21) and (22). Given the result of proposition (3), we know that the critical value $\widetilde{z}_N(\alpha)$ corresponds to the α-percent critical value of the standard normal distribution, denoted z_α if N tends to infinity whatever the size T. For a fixed N, the use of the normal critical value z_α to built the corresponding critical value $c_{N,T}(\alpha)$ is not founded. Despite this, we can propose an approximation $\widetilde{c}_{N,T}(\alpha)$ based on this value

$$\widetilde{c}_{N,T}(\alpha) = z_\alpha \sqrt{N^{-1}var\left(\widetilde{W}_{i,T}\right)} + E\left(\widetilde{W}_{i,T}\right) \tag{29}$$

or equivalently:

$$\widetilde{c}_{N,T}(\alpha) = z_\alpha \times \frac{(T-2K-1)}{(T-2K-3)} \times \sqrt{\frac{2K}{N} \times \frac{(T-K-3)}{(T-2K-5)}} + \frac{K \times (T-2K-1)}{(T-2K-3)}. \tag{30}$$

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

6.2 Cross-Sectional Dependence

Nowadays, an important issue in panel analysis is related to the existence of cross-sectional dependence. Following the literature on second-generation panel unit-root tests (e.g. Bai et Ng, 2001; Moon et Perron, 2004; Pesaran, 2007), new panel non-causality tests should be developed so as to explicitly take into account general forms of dependencies among the individuals of the panel.

Since specifying a particular form of correlation for the individual residuals has a negative impact on the small-sample properties of a test for alternative forms of dependence, we consider here a very general and simple, although computationally intensive, solution to this issue. To be more precise, we take into account cross-sectional dependence by using bootstrapped critical values instead of asymptotic critical values when performing Granger non-causality tests. The block bootstrap procedure we hence propose implies several steps:

1. *Define the panel model for which we test the Granger non-causality hypothesis (e.g.*
$y_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_{i,k} y_{i,t-1} + \sum_{k=1}^{K} \beta_{i,k} x_{i,t-1}$).

2. *Estimate the model for each cross-sectional unit and compute the standardized test statis-*

tics $Z_{N,T}^{Hnc}$ and $\tilde{Z}_N^{Hnc}$.

3. Estimate the model under the null hypothesis of no Granger causality ($\beta_{i,k}$ are null) for each cross-section unit and compute the N vectors of size (T,1) of residuals.

4. Resample the residuals with replacement by considering a block of size 1 in time-series and size N in the panel dimension. The size of the time-series block can take another shape if we suspect that the residuals are also autocorrelated in time.

5. Next, we construct a resampled series $y_{i,t}$ under the null hypothesis $\tilde{y}_{i,t} = \hat{\alpha}_i + \sum_{k=1}^{K} \hat{\gamma}_{i,k} y_{i,t-1} + \tilde{\varepsilon}_{i,t}$, where $\hat{\alpha}$ is the vector of estimated fixed effects, $\hat{\gamma}$ is the matrix of estimated autoregressive parameters for all the individuals and lags, and $\tilde{\varepsilon}$ is the matrix of resampled residuals.

6. Estimate the model defined at step 1 by using the resampled data $\tilde{y}_{i,t}$ and compute the test statistics for this resampled data.

7. Repeat steps 5 and 6 a large number of times. At each repetition keep the test statistics obtained for the resampled data, so as to compute the empirical critical values as the 95% percentile of the distribution of test-statistics (taken in absolute value) under the null hypothesis of no causality.

8. Compare the test statistics corresponding to the initial dataset (step 2) with the empirical critical values calculated in step 7.

To illustrate this procedure, we assume that the cross-sectional dependence can be mirrored by a correlation coefficient equal to 0.5. Therefore, it is possible to construct the true variance-covariance matrix V by drawing the vector of variances of the residuals from a uniform distribution on [0.5,1.5].

We first scrutinize the empirical size of the panel Granger non-causality tests in presence of cross-sectional dependence. For this, we choose as benchmark model the one considered for the first set of Monte-Carlo experiments and follow several steps:

Step A1 - Generate the series of interest $y_{i,t}$ under the null hypothesis of Homogeneous Non Causality.

$$y_{i,t} = \alpha_i + \gamma_i y_{i,t-1} + \varepsilon_{i,t}, \tag{31}$$

where the vectors of disturbances are jointly normal distributed $\varepsilon_i \sim N_n(0, \Sigma)$, with the sample covariance matrix drawn from the Wishart distribution with covariance-matrix V and df degrees of freedom, i.e. $\Sigma \sim W_n(V, df)$, with $i \in [1, N]$ and $t \in [1, T]$. Besides, the autoregressive parameters γ_i are drawn from a uniform distribution on $]-1, 1[$ and the fixed individual effects α_i are generated according to a $N(0, 1)$ distribution.

Step A2 - Define the regression model for each individual i in the panel

$$y_{i,t} = \alpha_i + \gamma_i y_{i,t-1} + \beta_i x_{i,t-1}, \tag{32}$$

where x is normally distributed with mean 0 and unit variance. Note that the specified regression model allows for heterogeneity of the estimated parameters among individuals. The individual Wald statistic for the ith individual $W_{i,T}$ corresponding to the individual test $H0 : \beta_i = 0$, can now be computed. The regression model is hence applied for each $i \in [1, N]$,

20

so that in the end we obtain a vector of N statistics. In this context, we can construct the standardized statistics $Z_{N,T}^{Hnc}$ and $\tilde{Z}_N^{Hnc}$ (see equation 9 and equation 26).

Step A3 - Repeat Steps A1 and A2 a large number of times and count the number of rejections for each test-statistic by relying on the bootstrapped critical values. These nested simulations are very time-consuming. Therefore, we use 5,000 simulations in steps A1-A2, whereas we consider 1,000 repetitions for the empirical critical values.

Second, to account for cross-sectional dependence, the empirical power has to be computed from the rejection rates obtained with the bootstrapped critical values. Indeed, we implement a two-step procedure specially designed to size-adjust the power results, as follows. *First, we generate the series $\{y_{i,t}\}_{t=-100}^T$ under the alternative hypothesis, i.e. $y_{i,t} = \alpha_i + \gamma_i y_{i,t-1} + \beta x_{i,t-1} + \varepsilon_{i,t}$, where β_i are drawn from a standard normal distribution. Next, we estimate equation 32 and implement the panel Granger causality tests so as to obtain the test-statistics under the alternative.* 5,000 repetitions are considered, and for each of them we compute bootstrapped critical values based on 1,000 simulations under the null hypothesis, performed by following the same procedure as for the size simulations. Note that for the critical values the regression model corresponds to equation 31, since these simulations are performed under the null hypothesis.

Table 5 depicts the size and power results obtained for different N and T at the 5% nominal level. Our main finding is that both test-statistics, i.e. $Z_{N,T}^{Hnc}$ and $\tilde{Z}_N^{Hnc}$, are roughly well sized when empirical critical values are used so as to tackle cross-sectional dependence. This result holds regardless of the sample sizes N and T. At the same time, the power of the tests is higher than 0.79 and it increases monotonically with the number of cross-section units in the sample, N, and that of the time-dimension T.

6.3   Unbalanced Panel Data and / or Unit-Specific Lag Order

Two other cases are frequently encountered in economic applications, namely an unbalanced panel dataset and a lag order K_i that is specific to each cross-section unit. In such circumstances, the standardized statistic $\tilde{Z}_N^{Hnc}$ must be adapted as follows:

$$
\begin{aligned}
\widetilde{Z}_N^{Hnc} &= \frac{\sqrt{N}\left[W_{N,T}^{Hnc} - N^{-1}\sum_{i=1}^N E\left(\widetilde{W}_{i,T}\right)\right]}{\sqrt{N^{-1}\sum_{i=1}^N Var\left(\widetilde{W}_{i,T}\right)}} \\
&= \sqrt{N}\left[W_{N,T}^{Hnc} - N^{-1}\sum_{i=1}^N K_i \times \frac{(T_i - 2K_i - 1)}{(T_i - 2K_i - 3)}\right] \\
&\quad \times \left[N^{-1}\sum_{i=1}^N 2K_i \times \frac{(T_i - 2K_i - 1)^2 \times (T_i - K_i - 3)}{(T_i - 2K_i - 3)^2 \times (T_i - 2K_i - 5)}\right]^{-1/2},
\end{aligned}
\tag{33}
$$

where $T_i > 5 + 2K_i$ denotes the time dimension for the ith cross-section unit. Indeed, in the case of unbalanced datasets, the distributions of the test-statistics depends upon the time-series dimension of each cross-sectional unit T_i, apart from the variance of the residuals. This distribution is hence different across individuals and it is not a chi-squared one. Similarly, if the lag-order differs from one individual to another, the distribution of the test-statistics, which depends on the number of restrictions imposed under the null, will vary across groups. However, in both cases the distributions of the test-statistics are independent from one unit to another, and hence the mean of the test-statistics will asymptotically converge in distribution to a normal distribution as long as $T_i > 5 + 2K_i$. For this, it is nevertheless necessary that $N \to \infty$.

A simple experiment, based on Monte-Carlo simulations, looks at the size of the Granger-causality test-statistic $\tilde{Z}_N^{Hnc}$ in unbalanced panels. The time dimension for each individual, T_i, is drawn from a uniform distribution on the interval [8,100]. Besides, $N = 2,000$ cross-sectional units are considered. We find that the test is slightly oversized, the rejection rate obtained after 5,000 simulations being equal to 0.056. This result depends, however, on the frequency of individuals with small (large) time-dimension T_i, which is chosen to be uniform in this exercise. All in all, the asymptotic properties of the test do not change significantly when the dataset is unbalanced or when the lag-order is different across individuals.

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 any of the cross-section units of the panel. Under the alternative, there are two subgroups of cross-section units: one characterized by causal relationships from x to y (even though the regression model is not necessarily the same) and another subgroup for which 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 with the standard Granger causality tests based on single time series. Under the cross-section independence assumption, we show that this average statistic converge to a standard normal distribution when T and N tend sequentially to infinity. The semi-asymptotic distribution is also characterized for fixed T samples. 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 first two 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 these moments of the standard Wald statistics for small T samples. In this paper we hence propose a general approximation of these moments and the corresponding standard-

ized average Wald statistic. Moreover, we tackle the case with cross-sectional dependence by proposing a block-bootstrap procedure so as to obtain empirical critical values for the Granger non-causality tests.

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 substantial increase in the power of the Granger non-causality tests even for samples with very small T and N dimensions. Furthermore, our test statistics (based on cross section average of individual Wald statistics) do not require any particular panel estimation. Finally, the test can be easily implemented in unbalanced panels and / or panels with different lag order K for each individual.

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. First, the rejection of the null of Homogeneous Non Causality does not provide any guidance with respect to the number or the identity of the particular panel units for which the null of non causality is rejected. Second, the asymptotic distribution of our statistics is established under the assumption of cross-section independence, although a block-bootstrap procedure is proposed to tackle the dependence issue in empirical applications. 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: Moments of individual Wald $\widetilde{W}_{i,T}$

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

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

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

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

$$
\begin{aligned}
E\left(\widetilde{W}_{i,T}\right) &= 2\widetilde{T} \times K \times \int_0^\infty (1 + 2t)^{-\widetilde{T}}\ dt \\
&= 2\widetilde{T} \times K \times \left[\frac{(1 + 2t)^{-\widetilde{T}+1}}{2\left(-\widetilde{T}+1\right)}\right]_0^\infty \\
&= \frac{2\widetilde{T} \times K}{2\left(\widetilde{T}-1\right)}
\end{aligned}
$$

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

$$
E\left(\widetilde{W}_{i,T}\right) = K \times \frac{(T - 2K - 1)}{(T - 2K - 3)}. \tag{34}
$$

At the same time, the definition of the second order moment is:

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

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

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

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

$$
\begin{aligned}
E\left[\left(\widetilde{W}_{i,T}\right)^2\right] &= \frac{4\,\widetilde{T}^2 \times (2K + K^2)}{2\left(\widetilde{T}-1\right)} \times \int_0^\infty (1 + 2t)^{-\widetilde{T}}\ dt \\
&= \frac{4\,\widetilde{T}^2 \times (2K + K^2)}{2\left(\widetilde{T}-1\right)} \times \left[\frac{(1 + 2t)^{-\widetilde{T}+2}}{2\left(-\widetilde{T}+2\right)}\right]_0^\infty \\
&= \frac{4\,\widetilde{T}^2 \times (2K + K^2)}{2\left(\widetilde{T}-1\right)} \times \frac{1}{2\left(\widetilde{T}-2\right)}.
\end{aligned}
$$

After simplifications:

$$
E\left[\left(\widetilde{W}_{i,T}\right)^2\right] = \frac{\widetilde{T}^2 \times (2K + K^2)}{\left(\widetilde{T}-1\right)\left(\widetilde{T}-2\right)} = \frac{(T - 2K - 1)^2 \times (2K + K^2)}{(T - 2K - 3)(T - 2K - 5)}. \tag{35}
$$

Under the condition $T > 5 + 2K$, this second order moment exists as it was previously established in proposition (1).

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

$$
\begin{aligned}
Var\left(\widetilde{W}_{i,T}\right) &= E\left[\left(\widetilde{W}_{i,T}\right)^2\right] - E\left(\widetilde{W}_{i,T}\right)^2 \\
&= \frac{(T-2K-1)^2 \times (2K+K^2)}{(T-2K-3)(T-2K-5)} - \left[\frac{K \times (T-2K-1)}{(T-2K-3)}\right]^2.
\end{aligned}
$$

After simplifications, we have:

$$
Var\left(\widetilde{W}_{i,T}\right) = 2\,K \times \frac{(T-2K-1)^2 \times (T-K-3)}{(T-2K-3)^2 (T-2K-5)}. \tag{36}
$$

References

[1] Granger, C.W.J. 1969. Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37(3), 424-438.

[2] Granger, C.W.J. 1980. Testing for causality. Journal of Economic Dynamics and Control 2, 329-352

[3] Granger C.W.J. 2003. Some aspects of causal relationships, Journal of Econometrics 112, 69-71

[4] Kónya, L. (2006). "Exports and growth: Granger causality analysis on OECD countries with a panel data approach", Economic Modelling, 23, 978-982.

[5] Holtz-Eakin D., Newey W, Rosen H.S. 1988. Estimating vector autoregressions with panel data. Econometrica 56, 1371-1396.

[6] Hurlin C. (2005), Un Test Simple de l'Hypothèse de Non Causalité dans un Modèle de Panel Hétérogène, LIIIe annual congress of the French Economic Association 2004, Revue Economique, 56(3), 799-809

[7] Hsiao, C. 2003. Analysis of panel data, Cambridge University Press

[8] Im, K.S., Pesaran, M.H., Shin, Y. 2003. Testing for Unit Roots in Heterogenous Panels. Journal of Econometrics, 54, 91-115.

[9] Magnus, J.R. 1986. The exact moments of a ratio of quadratic forms in normal variables, Annales d'Economie et de Statistiques, 4, 96-109.

[10] Nair-Reichert, U. Weinhold, D. 2001, Causality tests for cross-country panels: a look at FDI and economic growth in less developed countries, Oxford Bulletin of Economics and Statistics, 63, 153-171.

[11] Pesaran, H.M. Smith, R. 1995. Estimating long-run relationships from dynamic heterogenous panels, Journal of Econometrics, 68, 79-113.

[12] Swamy, P.A. 1970. Efficient inference in a random coefficient regression model, Econometrica, 38, 311-323.

[13] Weinhold, D. 1996. Tests de causalité sur données de panel : une application à l'étude de la causalité entre l'investissement et la croissance, Economie et Prévision, 126, 163-175.

Table 1: Size and Power of Panel Granger Non-causality Tests

N	Test	$T = 10$		$T = 25$		$T = 50$		$T = 100$	
		Size	Power	Size	Power	Size	Power	Size	Power
1	Wald	0.09	0.43	0.06	0.62	0.05	0.71	0.05	0.81
5	Z^{Hnc}	0.16	0.88	0.07	0.98	0.06	0.99	0.05	0.99
	$\widetilde{Z}^{Hnc}$	0.04	0.73	0.04	0.97	0.04	0.99	0.04	0.99
10	Z^{Hnc}	0.21	0.98	0.08	0.99	0.06	1.00	0.05	1.00
	$\widetilde{Z}^{Hnc}$	0.04	0.91	0.04	0.99	0.04	1.00	0.04	1.00
25	Z^{Hnc}	0.31	1.00	0.09	1.00	0.06	1.00	0.05	1.00
	$\widetilde{Z}^{Hnc}$	0.04	0.99	0.04	1.00	0.04	1.00	0.04	1.00
50	Z^{Hnc}	0.44	1.00	0.11	1.00	0.07	1.00	0.06	1.00
	$\widetilde{Z}^{Hnc}$	0.04	1.00	0.04	1.00	0.05	1.00	0.05	1.00

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 $\widetilde{Z}_{N}^{Hnc}$ 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 $i = 1,..,N$ and $t = -100, -99, .., T$. At each replication, the auto-regressive parameters γ_i are drawn from a uniform distribution on $]-1, 1[$ and the fixed individual effects α_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 $[0.5, 1.5]$. The size ($\beta_i = 0$, $i = 1,.,N$) and the power of the tests are computed at the 5% nominal level. Under the alternative (power simulations), β_i is different from 0 for all i, i.e. $N_1 = 0$. The parameters β_i are generated according to a $N(0, 1)$. The number of replications is set to $10,000$.

Table 2: Power of Panel Granger Non-causality Tests: Experiments with Heterogeneity in the Causal Relationship $(n_1 > 0)$

| | | | Power of Panel HNC Tests with $n_1 = 0.5$ |
N	N_1	Test	$T = 10$	$T = 25$	$T = 50$	$T = 100$
6	3	Z^{Hnc}	0.72	0.89	0.96	0.98
		$\widetilde{Z}^{Hnc}$	0.48	0.87	0.95	0.98
10	5	Z^{Hnc}	0.85	0.97	0.99	0.99
		$\widetilde{Z}^{Hnc}$	0.63	0.96	0.99	0.99
20	10	Z^{Hnc}	0.97	0.99	0.99	1.00
		$\widetilde{Z}^{Hnc}$	0.85	0.99	0.99	1.00
50	25	Z^{Hnc}	1.00	1.00	1.00	1.00
		$\widetilde{Z}^{Hnc}$	0.99	1.00	1.00	1.00
		Power of Panel HNC Tests with $n_1 = 0.9$
N	N_1	Test	$T = 10$	$T = 25$	$T = 50$	$T = 100$
6	—	Z^{Hnc}	—	—	—	—
		$\widetilde{Z}^{Hnc}$	—	—	—	—
10	9	Z^{Hnc}	0.40	0.48	0.60	0.71
		$\widetilde{Z}^{Hnc}$	0.16	0.42	0.58	0.71
20	18	Z^{Hnc}	0.58	0.68	0.81	0.90
		$\widetilde{Z}^{Hnc}$	0.22	0.60	0.79	0.89
50	45	Z^{Hnc}	0.85	0.92	0.97	0.99
		$\widetilde{Z}^{Hnc}$	0.38	0.86	0.97	0.99

Notes: This table reports the power of the panel standardized statistic $Z_{N,T}^{Hnc}$ based on asymptotic moments defined by (9) and the panel standardized statistic $\widetilde{Z}_N^{Hnc}$ 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 $i = 1, .., N$ and $t = -100, -99, .., T$. At each replication, the auto-regressive parameters γ_i are drawn according to a uniform distribution on $]-1, 1[$ and the fixed individual effects α_i are generated according to a $N(0,1)$. Individual residuals are $N.i.d.\left(0, \sigma_{\varepsilon,i}^2\right)$. The variances $\sigma_{\varepsilon,i}^2$ are generated according to a uniform distribution on $[0.5, 1.5]$. The power is computed at the 5% nominal level. We consider power simulations with heterogeneous causal relationships. The parameters β_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, β_i are generated according to a $N(0,1)$. The ratio $n_1 = N_1/N$, with $0 \le n_1 < 1$, 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.

		Case A: DGP with $K = 2$, model with $K = 2$							
		$T = 10$		$T = 25$		$T = 50$		$T = 100$	
N	Test	Size	Power	Size	Power	Size	Power	Size	Power
1	Wald	0.13	0.57	0.07	0.81	0.06	0.91	0.05	0.95
5	Z^{Hnc}	0.52	0.98	0.37	0.99	0.34	1.00	0.31	1.00
	$\widetilde{Z}^{Hnc}$	0.02	0.67	0.04	0.99	0.04	1.00	0.04	1.00
10	Z^{Hnc}	0.61	0.99	0.37	1.00	0.36	1.00	0.33	1.00
	$\widetilde{Z}^{Hnc}$	0.02	0.89	0.03	1.00	0.05	1.00	0.04	1.00
50	Z^{Hnc}	0.92	1.00	0.48	1.00	0.37	1.00	0.34	1.00
	$\widetilde{Z}^{Hnc}$	0.03	1.00	0.05	1.00	0.05	1.00	0.05	1.00
		Case B: DGP with $K = 1$, model with $K = 2$							
		$T = 10$		$T = 25$		$T = 50$		$T = 100$	
N	Test	Size	Power	Size	Power	Size	Power	Size	Power
1	Wald	0.15	0.40	0.07	0.56	0.06	0.69	0.05	0.77
5	Z^{Hnc}	0.52	0.91	0.35	0.98	0.33	0.99	0.32	0.99
	$\widetilde{Z}^{Hnc}$	0.02	0.36	0.04	0.95	0.04	0.99	0.04	0.99
10	Z^{Hnc}	0.62	0.98	0.38	0.99	0.33	1.00	0.32	1.00
	$\widetilde{Z}^{Hnc}$	0.02	0.57	0.04	0.99	0.04	1.00	0.05	1.00
50	Z^{Hnc}	0.91	1.00	0.48	1.00	0.37	1.00	0.34	1.00
	$\widetilde{Z}^{Hnc}$	0.03	0.99	0.04	1.00	0.05	1.00	0.04	1.00

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 $\widetilde{Z}_N^{Hnc}$ 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 $i = 1, .., N$ and $t = -100, -99, .., T$. 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 α_i are generated according to a $N(0,1)$. Individual residuals are $N.i.d. \left(0, \sigma_{\varepsilon,i}^2\right)$. The variances $\sigma_{\varepsilon,i}^2$ are generated according to a uniform distribution on $[0.5, 1.5]$. In case B, the data are generated by a model with one lag ($K = 1$) whereas the individual Wald statistics are computed from a model that includes (at wrong) two lags ($K = 2$). The size ($\beta_i = 0$, $i = 1, .., N$) and the power of the tests are computed at the 5% nominal level. Under the alternative (power simulations), β_i is different from 0 for all i, i.e. $N_1 = 0$. The parameters β_i are generated according to a $N(0,1)$. The number of replications is set to $5,000$.

Table 4: Comparison of Simulated and Approximated Critical Values for Fixed N and T samples.

	Simulated 5% Critical Values $c_{N,T}(0.05)$							
$N \backslash T$	10	15	20	25	30	40	50	100
5	3.54	2.87	2.66	2.53	2.47	2.39	2.36	2.28
10	2.97	2.38	2.19	2.10	2.04	1.98	1.95	1.88
15	2.68	2.15	1.99	1.91	1.85	1.80	1.77	1.71
20	2.49	2.01	1.86	1.79	1.75	1.69	1.67	1.62
25	2.40	1.92	1.78	1.71	1.66	1.62	1.60	1.55
	Approximated 5% Critical Values $\widetilde{c}_{N,T}(0.05)$							
$N \backslash T$	10	15	20	25	30	40	50	100
5	3.46	2.66	2.44	2.34	2.27	2.21	2.17	2.10
10	2.86	2.24	2.06	1.97	1.92	1.87	1.84	1.78
15	2.59	2.05	1.89	1.81	1.77	1.72	1.69	1.64
20	2.43	1.93	1.79	1.72	1.68	1.63	1.61	1.56
25	2.32	1.85	1.72	1.65	1.61	1.57	1.55	1.50

Notes: The approximated critical values for the average statistic $W_{N,T}^{Hnc}$ are computed from equation (30) 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 α_i, $i = 1, .., N$ are drawn from a $N(0, 1)$.

Table 5: Size and Power of Panel Granger Non-causality Tests in the Presence of Cross-sectional Dependence

N	Test	T = 10 Size	T = 10 Power	T = 25 Size	T = 25 Power	T = 50 Size	T = 50 Power	T = 100 Size	T = 100 Power
5	Z^{Hnc}	0.05	0.79	0.05	0.98	0.05	1.00	0.05	1.00
	$\tilde{Z}^{Hnc}$	0.05	0.79	0.05	0.98	0.05	1.00	0.05	1.00
10	Z^{Hnc}	0.05	0.96	0.06	1.00	0.05	1.00	0.05	1.00
	$\tilde{Z}^{Hnc}$	0.05	0.96	0.06	1.00	0.05	1.00	0.05	1.00
25	Z^{Hnc}	0.04	0.98	0.05	1.00	0.06	1.00	0.05	1.00
	$\tilde{Z}^{Hnc}$	0.05	0.98	0.05	1.00	0.06	1.00	0.05	1.00
50	Z^{Hnc}	0.05	0.95	0.06	1.00	0.05	1.00	0.05	1.00
	$\tilde{Z}^{Hnc}$	0.05	0.95	0.06	1.00	0.05	1.00	0.05	1.00

Notes: This table reports the size and power of the standardized panel Wald statistic $Z_{N,T}^{Hnc}$ based on asymptotic moments defined by (9) and the standardized panel statistic $\widetilde{Z}_N^{Hnc}$ 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 $i = 1, .., N$ and $t = -100, -99, .., T$. At each replication, the auto-regressive parameters γ_i are drawn according to a uniform distribution on $]-1, 1[$ and the fixed individual effects α_i are generated according to a $N(0, 1)$. Individual residuals are $N.i.d.(0, \Sigma)$, where Σ is drawn from the Wishart distribution with covariance-matrix V and $1,000$ degrees of freedom. V is obtained by assuming that the dependence takes the form of a correlation coefficient equal to 0.5 and that the variances $\sigma_{\varepsilon,i}^2$ are generated according to a uniform distribution on $[0.5, 1.5]$. The size $(\beta_i = 0,\ i = 1, ., N)$ and the power of the tests are computed at the five percent nominal level. Under the alternative (power simulations), β_i is different from 0 for all i, i.e. $N_1 = 0$. The parameters β_i are generated according to a $N(0, 1)$. The number of replications is set to $5,000$ and $1,000$ simulations are considered at each replication so as to compute the empirical critical values.