Journal of Statistical and Econometric Methods, vol.3, no.1, 2014, 97-114
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2014

Granger Causality and Unit Roots

Carlos Vladimir Rodríguez-Caballero[1], **Daniel Ventosa-Santaulària**[2]

Abstract

The asymptotic behavior of the Granger-causality test under stochastic nonstationarity is studied. Our results confirm that the inference drawn from the test is not reliable when the series are integrated to the first order. In the presence of deterministic components, the test statistic diverges, eventually rejecting the null hypothesis, even when the series are independent of each other. Moreover, controlling for these deterministic elements (in the auxiliary regressions of the test) does not preclude the possibility of drawing erroneous inferences. Granger-causality tests should not be used under stochastic nonstationarity, a property typically found in many macroeconomic variables.

Mathematics Subject Classification: 62J05
Keywords: Granger-causality; Unit root; Drift

[1] Facultad de Ciencias, Universidad Nacional Autónoma de México, UNAM,
e-mail: vladimir.rodriguez@ciencias.unam.mx
[2] Centro de Investigación y Docencia Económicas, CIDE,
e-mail: daniel@ventosa-santaularia.com

Article Info: *Received* : October 22, 2013. *Revised* : November 19, 2013.
Published online : February 7, 2014

1 Introduction

Nonstationarity and Granger-causality have been widely studied in econometrics. Both are considered fundamental issues, particularly in applied macroeconomics. On the one hand, cointegration, a technique first developed by [12] and [6] requires (i) the variables to be integrated to the first order, $I(1)$ (nonstationary), and, (ii) the existence of a stationary linear combination of such variables. Cointegration is one of the most widely applied estimation procedure in (macro) econometrics. Its relevance in applied works is due to the well established nonstationary nature of many macroeconomic variables[3] and the also well-known phenomenon of spurious regression. [14] showed that linear regressions (ordinary least squares, OLS) using nonstationary integrated of order one, $I(1)$, are spurious unless the variables are cointegrated.

On the other hand, causality in modern econometrics can be traced back to the works of [37], [10], and [34].[4] The Wiener-Granger-Sims concept is based on the predictability of a variable. Should other variables contain information in past terms (not contained elsewhere) useful to improve the prediction of the former one, then such variables would be said to cause it. The Granger-causality (GC, hereinafter) is ubiquitous in the applied economics literature.[5] [25], [7], [30], *inter alia*, consider GC as a major empirical and theoretical contribution to econometrics.[6]

GC testing has been thoroughly studied. It is well documented, for example, that size distortions and power losses may occur even when the variables are stationary. [1] showed the existence of such effects if one of the variables, y_t or x_t (but not both) is measured with error. [23] also identified size distortions and power losses due to a temporal aggregation bias in the variables. [2] showed that when the assumption of parameter constancy is violated (struc-

---

[3]There is some debate concerning the nature of the trending mechanism, i.e. whether series have a unit root or a deterministic trend with possible level/trend shifts (see [24], [27], and references therein).

[4][21] analyses the difference between Granger and Sims causality concepts.

[5]Granger first developed his results using spectral analysis procedures, such as cross spectrum and the partial cross spectrum, previously developed in [9]. [34] developed a GC-testing procedure based upon a moving average representation; [33] proposed a procedure directly related with the definition of causality given in [10]; [16] and [32] developed a well known statistical procedure to test GC, which uses residuals from univariate models for time series.

[6]Further discussion can be found in [11] and [13].

tural shifts), *GC* tests may provide misleading inference about the underlying relationship of causality.

As for the nonstationary case, there has been some debate concerning the use of differenced variables (see, for example [35] and [17]), the main argument being discussed is whether differencing the series causes a severe power loss. [26] showed, in a Monte Carlo study, the considerable size distortions of the *GC* test occur when the variables behave as unit root processes. Most articles, however, are based on empirical or simulation analysis (see, for example [4] and [5]). However, there are several exceptions: [18] proved that, when the variables used to test *GC* are driftless unit roots, the test statistic has an asymptotic nonstandard distribution under the null hypothesis, free of nuisance parameters that could lead to spurious inference; [36] showed that *GC* tests fail to reject the null hypothesis of no *GC* more often than it should when the data generating process (DGP) of the variables is either Broken-Trend Stationary (BTS) or Broken-Mean Stationary (BMS), even when the former variables are differenced. [38] proved that, when the DGPs of the variables are a mixture of trend stationary processes and unit roots, the inferences drawn from a *GC* test could be misleading as well.

A *GC* test could therefore yield spurious inference when the variables are not stationary. In this paper we prove that an asymptotic nuisance-parameter-free distribution under stochastic nonstationarity is an exception rather than the rule, and make a case against the use of the *GC* test under nonstationarity. Independent simple driftless unit-root variables do indeed provide—under the null hypothesis of no *GC*—a nuisance-parameter-free distribution. That said, the sole inclusion of a drift in the DGP specification makes the result no longer valid. Moreover, the inclusion of a constant term or a deterministic trend in the auxiliary regressions of the *GC* test does not elimnate the nuisance parameters in the limit distribution. In other words, the asymptotic evidence presented in the paper shows that the *GC* test is not reliable when the variables are governed by unit roots with drift because of the nonstandard nature of the limit distribution, as in the driftless case, but also because unknown nuisance parameter are found in the later distribution (in contrast with the drifless case).

2 Granger-causality under I(1) processes

There is strong evidence in favour of the presence of unit root process in macroeconomic time series (since the influential work by [24]). During the last thirty years, Nelson and Plosser's historical dataset, for example, has been used as a vehicle for studying the trending nature of U.S. macro data. Table (1) partially summarizes the results available in the literature and shows that most U.S. macro variables are not stationary.

Whether the trend component is stochastic or deterministic is a more controversial issue. Not withstanding the debate, there is evidence that the unit root may adequately represent the properties of some macro variables. This implies that the GC test is frequently used under nonstationarity in macroeconometrics. We test GC based on the classical $\mathcal{F}$ test framework:

$$\mathcal{F} = \frac{SSR^R - SSR^U}{SSR^U/(T-1)}, \tag{1}$$

where SSR^R and SSR^U account for the sum of squared residuals of the restricted (eq. 2) and unrestricted (eq. 3) equations, respectively.

$$y_t = \gamma_{11} y_{t-1} + u_{2t}, \tag{2}$$

$$y_t = \gamma_{21} y_{t-1} + \gamma_{22} x_{t-1} + u_{1t}. \tag{3}$$

Note that both auxiliary regressions correspond to the the simplest GC test specification possible. We first consider the asymptotic properties $\mathcal{F}$ when the underlying variables are independent (i) driftless unit roots, denoted UR_{ND} (eq. 4) and; (ii) unit roots with drift (UR_{WD}) as defined by equation (5). We then study the behavior of $\mathcal{F}$ under such processes.

$$z_t = z_0 + \xi_{zt}, \tag{4}$$

$$z_t = z_0 + \mu_z\, t + \xi_{zt}, \tag{5}$$

where $\xi_{zt} = \sum_{i=1}^{t} u_{zi}$; $u_{zt} \sim iid\mathcal{N}(0, \sigma_z^2)$ for $z = x, y$. The symbol $\xrightarrow{d}$ denotes weak convergence and $W_z \equiv W_z(r)$ denotes a standard Wiener process. The stochastic integral $\int_0^1$ is written as $\int$ for ease of notation.

Table 1: Trending mechanism of US macro variables

Variable/Article	NP 82	P 89	KPSS 92	ZA 92	P 97	LP 97	K 05 (4 br)	CiSS 06	VSG 11
Real GNP	UR	BTS	UR	BTS	BTS	BTSS	BTSS	BTS	BTS
Nominal GNP	UR	BTS	UR	BTS	BTS	BTSS	UR	BTS	BTS
Real per capita GNP	UR	BTS	—	BTS	BTS	BTSS	BTSS	UR	BTS
Industrial production	UR	BTS	—	BTS	BTS	BTSS	UR	BTS	URwd
Employment	UR	BTS	—	BTS	BTS	BTSS	BTSS	UR	URwd
Unemployment rate	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)	BTSS	BTS	I(0)
GDP deflator	UR	BTS	—	UR	UR	UR	UR	UR	URWd+b
Consumer prices	UR	UR	UR	UR	UR	UR	UR	UR	URnd
Wages	UR	BTS	—	UR	BTS	UR	UR	BTS	URwd
Real wages	UR	BTS	UR	UR	BTS	UR	UR	UR	URwd+b
Money stock	UR	BTS	—	UR	BTS	UR	UR	UR	URWd
Velocity	UR	UR	UR	UR	UR	UR	UR	UR	URnd
Bond yield	UR	UR	UR	UR	UR	UR	UR	BTS	URnd
Common stock prices	UR	BTS	UR	BTS	BTS	UR	UR	UR	URnd

Sources: NP 82, Nelson & Plosser (1982); P 89, Perron (1989); KPSS 92, [20]; ZA 92, Zivot & Andrews (1992); P 97, [28]; LP 07, [22] ; K 05, [19]; CiSS, [3]; VSG 11, [8].
Notation: I(0); stationary process, UR; unit root process (without further detail), URnd; driftless UR, URwd; UR with drift, URwd+b; URwd + breaks, TS; trend stationary, BTS; broken-trend stationary (one break), BTSS; broken-trend stationary (multiple breaks).

Theorem 1. Let x_t and y_t be: (A) two independent UR_{ND} processes (eq. 4) and; (B) two independent UR_{WD} (eq. 5). Let regressions (2) and (3) be estimated by OLS. Then, as $T \to \infty$:

Case A (driftless unit roots):

$$\mathcal{F} \xrightarrow{d} \frac{\left[\int W_y^2 \int W_x \, dW_y - \frac{1}{2}\left\{[W_y(1)]^2 - 1\right\} \int W_y W_x\right]^2}{\int W_y^2 \left[\int W_y^2 \int W_x^2 - \left[\int W_y W_x\right]^2\right]}. \tag{6}$$

Case B (unit roots with drift):

$$T^{-1}\,\mathcal{F} \xrightarrow{d} \frac{-\mu_y^2\left[\mu_x\,\sigma_y\left(2\int W_y - 3\int rW_y\right) + \mu_y\,\sigma_x\left(3\int rW_x - 2\int W_x\right)\right]}{\lambda}, \tag{7}$$

where λ is defined in appendix 3.

Proof: See appendix 3. □

On the one hand, in case A, the asymptotic distribution of the test statistic does not diverge. To be more precise, $\mathcal{F}$ converges to a nonstandard distribution without nuisance parameters. On the other hand, in case B, $\mathcal{F}$ diverges at rate T. Moreover, even when the test statistic is correctly normalized, the nonstandard asymptotic distribution is not nuisance-parameter free, i.e. it is not a pivotal distribution.

Both cases clearly show that standard $\mathcal{F}$-distribution critical values cannot be used to draw inferences. That said, for case (A), a new set of critical values can be obtained (see the appendix). However, for the second case, the null hypothesis of no GC between variables will eventually be rejected as the sample size grows.

It could be argued that, to eliminate the nuisance parameters in the asymptotic distributions, a more elaborate auxiliary regressions should be employed. For example, both the restricted and unrestricted equations could include a constant term, as in equations 8 and 9;

$$y_t = \gamma_{10} + \gamma_{11} y_{t-1} + \gamma_{12} x_{t-1} + u_{1t}, \tag{8}$$

$$y_t = \gamma_{20} + \gamma_{21} y_{t-1} + u_{2t}. \tag{9}$$

We therefore also examine this case. Results appear in the following Theorem:

Theorem 2. Let x_t and y_t be two independent $I(1)$ processes generated by DGP (5). Let regressions (8) and (9) be estimated by OLS. Then, as $T \to \infty$:

$$\mathcal{F} \xrightarrow{d} \frac{\mu_y\, \sigma_x\, \sigma_y\, \phi_2 - \mu_x\, \sigma_y^2\, \phi_1 + 4\sigma_y^2\, W_y(1) \int W_y}{\mu_x^2\, \sigma_y^2\, \phi_3 + \mu_x\, \mu_y\, \sigma_x\, \sigma_y\, \phi_4 + \mu_y^2\, \sigma_x^2\, \phi_5}, \tag{10}$$

where ϕ_i for $i = 1, 2, \ldots, 5$, is defined in appendix 3.

Proof: See supplementary material.[7] □

The nonstandard distribution of the $\mathcal{F}$ statistic is $O_p(1)$, as in the driftless case. Nevertheless, it is not asymptotically pivotal. In other words, the test statistic remains impractical for applied works and the addition of constant terms is not useful to control for a simple deterministic trend mechanism. A more intuitive approach would be to also include a linear trend in the auxiliary regressions:

$$y_t = \gamma_{11} y_{t-1} + \gamma_{12} x_{t-1} + \gamma_{13} t + u_{1t}, \tag{11}$$

$$y_t = \gamma_{21} y_{t-1} + \gamma_{23} t + u_{2t}. \tag{12}$$

The results are presented in Theorem 3:

Theorem 3. Let x_t and y_t be two independent $I(1)$ processes generated by DGP (5). Let regressions (11) and (12) be estimated by OLS. Then, as $T \to \infty$:

$$T^{-1}\mathcal{F} \xrightarrow{d} \frac{\mu_y^2}{\sigma_x \sigma_y^3} \left[\frac{2 \int W_y \int W_x W_y - \gamma_1 \int W_x - \gamma_2 \int r W_x}{\left[\int W_y^2 - 3 \left(\int r W_y \right)^2 \right] \left\{ \mu_y^2 \rho + \sigma_y^2 \gamma_7 \right\}} \right], \tag{13}$$

where γ_i for $i = 1, 2, 7$, and ρ are defined in appendix 3.

Proof: See supplementary material.[8] □

Theorem 3 clearly shows that the GC test statistic, even correctly normalized, does not converge to a pivotal density. Correct inference is therefore not possible. Including deterministic components in the auxiliary regressions of the GC test does not prevent the presence of unknown nuisance parameters in the limit distribution of the test statistic under $\mathcal{H}_0$.

[7]Available at [rgb]0,0,1https://dl.dropboxusercontent.com/u/1307356/JoSEM_SuppMat.pdf.
[8]Available at [rgb]0,0,1https://dl.dropboxusercontent.com/u/1307356/JoSEM_SuppMat.pdf.

3 Concluding remarks

Granger-causality testing under nonstationarity has been thoroughly discussed on empirical grounds. Practitioners are usually reluctant to first-difference the series because they fear a substantial loss of information (contained in the low frequencies) leading to erroneously infer that there is no Granger causality between variables. When the data-generating processes include deterministic components the differencing strategy is even less appealing; differencing a (broken-) trend stationary process, for example, artificially generates non invertible moving average processes.

However, unit root variables in levels neither should be used to test Granger-causality. When the variables are generated as driftless unit root (the simplest stochastic trend), the test statistic has an asymptotic pivotal distribution under the null hypothesis of no Granger-causality. This is no longer true under slightly more complicated data generating processes (unit roots with drifts). Moreover, the unknown nuisance parameters in the limit distribution cannot be eliminated by controlling for deterministic components (such as constant terms or linear trends) in the auxiliary regressions of the test .

The asymptotic results presented in this paper are in line with those previously obtained in the literature, and make clear that Granger-causality inferences should not be drawn when the underlying variables exhibit a nonstationary behavior unless the trending mechanism has been correctly modelled.

References

[1] J. Andersson, Testing for granger causality in the presence of measurement errors, *Economics Bulletin*, **3**,(47),(2005),1-13.

[2] M. Bianchi, Granger causality in the presence of structural changes, IRES-Institut de Recherches Economiques et Sociales *Working Paper*, (1995).

[3] J.L. Carrion-i Silvestre and A. Sanso, Joint hypothesis specification for unit root tests with a structural break, *Econometrics Journal*, **9**(2), (2006), 196-224.

[4] L.J. Christiano and L. Ljungqvist, Money does granger-cause output in the bivariate money-output relation, *Journal of Monetary Economics*, **22**(2), (1988), 217-235.

[5] S. Cook, Further analysis of spurious causality. *Mathematics and Computers in Simulation*, **79**(3), (2008), 647-651.

[6] R.F. Engle and C.W.J. Granger, Co-integration and error correction: representation, estimation, and testing, *Econometrica: Journal of the Econometric Society*, (1987), 251-276.

[7] J. Geweke, Inference and causality in economic time series models, *Handbook of Econometrics*, **2**(2), (1984), 1101-1144.

[8] M. Gomez-Zaldivar and D. Ventosa-Santaulària, Testing for a deterministic trend when there is evidence of unit-root, *Journal of Time Series Econometrics*, **2**(2), (2011), 1-26.

[9] C.W.J. Granger, The typical spectral shape of an economic variable, *Econometrica:Journal of the Econometric Society*, (1966), 150-161.

[10] C.W.J. Granger, Investigating causal relations by econometric models and crossspectral methods, *Econometrica: Journal of the Econometric Society*, (1969), 424-438.

[11] C.W.J. Granger, Testing for causality: A personal viewpoint, *Journal of Economic Dynamics and Control*, **2**, (1980), 329-352.

[12] C.W.J. Granger, Some properties of time series data and their use in econometric model specification, *Journal of econometrics*, **16**(1), (1981), 121-130.

[13] C.W.J. Granger, Some recent development in a concept of causality, *Journal of econometrics*, **39**(1-2), (1988), 199-211.

[14] C.W.J. Granger and P. Newbold, Spurious regressions in econometrics, *Journal of econometrics*, **2**(2), (1974), 111-120.

[15] J.D. Hamilton, *Time series analysis*, Cambridge Univ Press, 1994.

[16] L.D. Haugh, Checking the independence of two covariance-stationary time series: a univariate residual cross-correlation approach, *Journal of the American Statistical Association*, (1976), 378-385.

[17] B. Hayo, Money-output granger causality revisited: an empirical analysis of eu countries, *Applied Economics*, **31**, (1999), 1489-1501.

[18] Z. He and K. Maekawa, On spurious granger causality, *Economics Letters*, **73**(3), (2001), 307-313.

[19] G. Kapetanios, Unit-root testing against the alternative hypothesis of up to m structural breaks, *Journal of Time Series Analysis*, **26**(1), (2005), 123-133.

[20] D. Kwiatkowski, PCB Phillips, P. Schmidt, and Y. Shin, Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root, *Journal of Econometrics*, **54**(1-3), (1992), 159-178.

[21] M. Lechner, The relation of different concepts of causality in econometrics, Institute for the Study of Labor (IZA), *Working Paper*, (2006).

[22] R.L. Lumsdaine and D.H. Papell, Multiple trend breaks and the unit-root hypothesis, *Review of Economics and Statistics*, **79**(2), (1997), 212-218.

[23] J.R. McCrorie and M.J. Chambers, Granger causality and the sampling of economic processes, *Journal of econometrics*, **132**(2), (2006),311-336.

[24] C.R. Nelson and C.R. Plosser, Trends and random walks in macroeconomic time series: Some evidence and implications, *Journal of monetary economics*, **10**(2), (1982), 139-162.

[25] P. Newbold, Causality testing in economics, *Time Series Analysis: Theory and Practice*, **1**(1), (1982), 701-716.

[26] L.E. Ohanian, The spurious effects of unit roots on vector autoregressions: A monte carlo study, *Journal of Econometrics*, **39**(3), (1988), 251-266.

[27] P. Perron, The great crash, the oil price shock, and the unit root hypothesis, *Econometrica: Journal of the Econometric Society*, (1989), pages 1361-1401.

[28] P. Perron, Further evidence on breaking trend functions in macroeconomic variables, *Journal of econometrics*, **80**(2), (1997), 355-385.

[29] P.C.B. Phillips, Understanding spurious regressions in econometrics, *Journal of econometrics*, **33**(3), (1986), 311-340.

[30] P.C.B. Phillips, Reflections on Econometric Methodology, *Economic Record*, **64**(4), (1988), 344-359.

[31] P.C.B. Phillips and S.N. Durlauf, Multiple time series regression with integrated processes, *The Review of Economic Studies*, **53**(4), (1986), 473-495.

[32] D.A. Pierce and L.D. Haugh, Causality in temporal systems:: Characterization and a survey, *Journal of Econometrics*, **5**(3), (1977), 265-293.

[33] T.J. Sargent, A classical macroeconometric model for the united states, *The Journal of Political Economy*, (1976), 207-237.

[34] C.A. Sims, Money, income, and causality, *The American Economic Review*, **62**(4), (1972), 540-552.

[35] J.H. Stock and M.W.Watson, Interpreting the evidence on money-income causality, *Journal of Econometrics*, **40**(1), (1989), 161-181.

[36] D. Ventosa-Santaulària and J.E. Vera-Valdés, Granger-causality in the presence of structural breaks, *Economics Bulletin*,**3**(61), (2008), 1-14.

[37] N. Wiener, The theory of prediction, *Modern mathematics for the engineer*, (1956), 165-190.

[38] L. Zhang and X. Zhang, Spurious granger causality between a broken-trend stationary process and a stochastic trend process, *Mathematics and Computers in Simulation*, (2011).

Appendix

Proof of Theorem 1

Notation: the following expressions, λ, θ_i for $i = 1, \ldots, 6$, ϕ_j for $j = 1, \ldots, 5$, ρ and γ_k, were used in the results in Theorems 1, 2, and 3.

$$\begin{aligned}\lambda &= \mu_x^2\,\mu_y^2\,\sigma_y^2\,\theta_1 + \mu_x\,\mu_y^3\,\sigma_x\,\sigma_y\,\theta_2 + \mu_y^4\,\sigma_x^2\,\theta_3 + \mu_x^2\,\sigma_y^4\,\theta_4 + \\ &\quad \mu_x\,\mu_y\,\sigma_x\,\sigma_y^3\,\theta_5 + \mu_y^2\,\sigma_y^2\,\sigma_x^2\,\theta_6,\end{aligned}$$

where

$$\begin{aligned}\theta_1 &= 4\left[\int W_y\right]^2 - \int W_y^2 - 12\int W_y \int r\,W_y + 12\left[\int r\,W_y\right]^2, \\ \theta_2 &= 2\int W_xW_y - 8\int W_xW_y + 12\int r\,W_x \int W_y + 12\int r\,W_y \int W_x \\ &\quad -24\int r\,W_x \int r\,W_y, \\ \theta_3 &= 4\left[\int W_x\right]^2 - \int W_x^2 - 12\int W_x \int r\,W_x + 12\left[\int r\,W_x\right]^2, \\ \theta_4 &= 12\left[\int r\,W_y\right]^2 - 4\int W_y^2, \\ \theta_5 &= \int W_xW_y - 24\int r\,W_x \int r\,W_y,\end{aligned}$$

and

$$\theta_6 = 12\left[\int r\,W_x\right]^2 - 4\int W_x^2.$$

$$\begin{aligned}\phi_1 &= \tfrac{1}{2}\left([W_y(1)]^2 - 1\right) - 6\int W_y \int r\,W_y, \\ \phi_2 &= \int W_x\,dW_y + \int W_x\left[2\,W_y(1) + \int W_y\right] + \int r\,W_x\left[\int W_y - 6W_y(1)\right], \\ \phi_3 &= 4\left(\int W_y\right)^2 - \int W_y^2 - 12\int W_y \int rW_y + 12\left(\int rW_y\right)^2, \\ \phi_4 &= 12\int wW_x\left(\int W_y - 2\int rW_y\right) - 4\int W_x\left(2\int W_y - 3\int rW_y\right),\end{aligned}$$

and

$$\phi_5 = 4\left(\int W_x\right)^2 - \int W_x^2 - 12\int W - x\int rW_x + 12\left(\int rW_x\right)^2.$$

$$
\begin{aligned}
\rho &= 4\left(\int W_y\right)^2\gamma_3 + 4\gamma_4\int W_y^2 + 12\gamma_5\left(\int rW_y\right)^2 + \gamma_6\int W_xW_y,
\end{aligned}
$$

where

$$
\begin{aligned}
\gamma_1 &= 2\int W_y^2 - 6\left(\int rW_y\right)^2, \\
\gamma_2 &= 6\int rW_y\int W_y - 3\int W_y^2, \\
\gamma_3 &= \int W_x^2 - 3\left(\int rW_x\right)^2, \\
\gamma_4 &= 4\left(\int W_y\right)^2\left[\int W_x^2 - 3\left(\int rW_x\right)^2\right] + \int W_y^2\left[4\left(\int W_x\right)^2 - \int W_x^2\,(1\right. \\
&\quad \left. +12\int rW_y\right) -12\int W_x\int rW_x\left(1-2\int rW_y\right) + 12\left(\int rW_x\right)^2\Big], \\
\gamma_5 &= \int W_x^2 - \left(\int W_x\right)^2, \\
\gamma_6 &= -8\int W_x\int W_y + 12\int rW_x\int W_y + 12\int W_x\int rW_y + \int W_xW_y,
\end{aligned}
$$

and

$$
\begin{aligned}
\gamma_7 &= -4\int W_x^2\left[\int W_y^2 - 3\left(\int rW_y\right)^2\right] + 12\left(\int rW_x\right)\Big[\left(\int rW_x\right)^2\int W_y^2 \\
&\quad -2\int rW_y\int W_xW_y\Big] + 4\left(\int W_xW_y\right)^2.
\end{aligned}
$$

Driftless unit root processes: let x_t and y_t be independently generated as driftless unit roots. Equations (2) and (3) are stacked and written in vector form as $y = X\beta + U$. y is a $T\times 1$ vector and X is (A) a $T\times 2$ matrix when we estimate the unrestricted regression (SSR^U) or, (B) a $T\times 1$ vector when we estimate the restricted regression (SSR^R). U is a $T\times 1$ vector of zero mean disturbances. We first present the procedure to obtain the asymptotics of the unrestricted regression. The OLS estimator is given by

$$
\hat{\beta} = \begin{bmatrix} \hat{\gamma}_{21} \\ \hat{\gamma}_{22} \end{bmatrix} = \left(X'X\right)^{-1}X'y. \tag{14}
$$

We have (all sums run over $t = 1$ to T):

$$
\left(X'X\right) = \begin{bmatrix} \sum y_{t-1}^2 & \sum y_{t-1}\,x_{t-1} \\ \sum y_{t-1}\,x_{t-1} & \sum x_{t-1}^2 \end{bmatrix}, \tag{15}
$$

and

$$(X'y) = \begin{bmatrix} \sum y_{t-1}\ y_t \\ \sum x_{t-1}\ y_t \end{bmatrix}. \tag{16}$$

The order in convergence of the elements appearing in the previous equations can be found in [29], [31] and [15].

$$\begin{aligned}
\sum y_{t-1}^2 &= Y_0\ T + 2\ Y0 \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})} + \underbrace{\sum \xi_{y,t-1}^2}_{Op(T^2)}, \\
\sum y_{t-1}\ x_{t-1} &= Y_0\ X_0\ T + Y_0 \underbrace{\sum \xi_{x,t-1}}_{Op(T^{3/2})} + X_0 \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})} + \underbrace{\sum \xi_{y,t-1}\xi_{x,t-1}}_{Op(T^2)}, \\
\sum x_{t-1}^2 &= X_0\ T + 2\ X0 \underbrace{\sum \xi_{x,t-1}}_{Op(T^{3/2})} + \underbrace{\sum \xi_{x,t-1}^2}_{Op(T^2)}, \\
\sum y_{t-1}\ y_t &= Y_0^2\ T + 2\ Y_0 \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})} + Y_0 \underbrace{\sum u_{y,t}}_{Op(T^{1/2})} + \underbrace{\sum \xi_{y,t-1}^2}_{Op(T^2)} + \underbrace{\sum \xi_{y,t-1}\ u_{y,t}}_{Op(T)},
\end{aligned}$$

and

$$\sum x_{t-1}\ y_t = \sum y_{t-1}\ x_{t-1} + X_0 \underbrace{\sum u_{y,t}}_{Op(T^{1/2})} + \underbrace{\sum \xi_{x,t-1} u_{y,t}}_{Op(T)},$$

where $\xi_{z,t} = \sum_{i=1}^{t} u_{z,i}$ with $z = x, y$. We estimate the unrestricted regression squared residuals:

$$\hat{u}_1^2 = \sum \left(y_t - \hat{\gamma}_{11} y_{t-1} - \hat{\gamma}_{12} x_{t-1}\right)^2. \tag{17}$$

Additionally, we need

$$\begin{aligned}
\sum y_t^2 &= Y_0^2\ T + 2\ Y0 \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})} + \underbrace{\sum \xi_{y,t-1}^2}_{Op(T^2)} + \underbrace{\sum u_{y,t}^2}_{Op(T)} \\
&\quad + 2\ Y_0 \underbrace{\sum u_{y,t}}_{Op(T^{1/2})} + 2 \underbrace{\sum \xi_{y,t-1} u_{y,t}}_{Op(T)}.
\end{aligned}$$

The OLS estimator of the SSR^R regression is given by:

$$\hat{\gamma}_{11} = \left(X'X\right)^{-1} X'y.$$

Then,

$$(X'X) = \sum y_{t-1}^2, \tag{18}$$

and

$$(X'y) = \sum y_{t-1}\, y_t. \tag{19}$$

The restricted regression squared residuals sum is:

$$\hat{u}_2^2 = \sum \left(y_t - \hat{\gamma}_{21} y_{t-1}\right)^2. \tag{20}$$

Once we have the respective squared residuals, $\hat{u}_1^2$ and $\hat{u}_2^2$, we compute the $\mathcal{F}_{GC}$ statistic by,

$$\mathcal{F}_{GC} = \frac{(\hat{u}_2^2 - \hat{u}_1^2)}{\hat{u}_1^2/(T-1)}. \tag{21}$$

With the aid of a *Mathematica 7.0* code (available upon request), we derive the expression for the asymptotic nonstandard distribution given in equation (6).

Unit root with drift processes: let y_t and x_t be independent unit roots with drift. The equations (2) and (3) are stacked and written in vector form as $y = X\beta + U$ The OLS estimator is given by (14), where $(X'X^{-1})$ is specified by (15) while $(X'y)$ by (16). The computational-time cost is reduced considerably if we consider first the following linking expressions:

$$\begin{aligned}
\sum y_{t-1} &= Y_0\, T + \mu_y \sum t + \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})}, \\
\sum x_{t-1} &= X_0\, T + \mu_x \sum t + \underbrace{\sum \xi_{x,t-1}}_{Op(T^{3/2})},
\end{aligned}$$

$$
\begin{aligned}
\sum t\, y_{t-1} &= Y_0 \sum t + \mu_y \sum t^2 + \underbrace{\sum t\, \xi_{y,t-1}}_{Op(T^{5/2})}, \\
\sum t\, x_{t-1} &= X_0 \sum t + \mu_x \sum t^2 + \underbrace{\sum t\, \xi_{x,t-1}}_{Op(T^{5/2})}, \\
\sum y_{t-1}\xi_{y,t-1} &= Y_0 \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})} + \mu_y \underbrace{\sum t\, \xi_{y,t-1}}_{Op(T^{5/2})} + \underbrace{\sum \xi^2_{y,t-1}}_{Op(T^2)}, \\
\sum x_{t-1}\xi_{y,t-1} &= X_0 \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})} + \mu_x \underbrace{\sum t\, \xi_{y,t-1}}_{Op(T^{5/2})} + \underbrace{\sum \xi_{x,t-1}\xi_{y,t-1}}_{Op(T^2)}, \\
\sum x_{t-1}u_{y,t} &= X_0 \underbrace{\sum U_{y,t}}_{Op(T^{1/2})} + \mu_x \underbrace{t \sum u_{y,t}}_{Op(T^{3/2})} + \underbrace{\sum \xi_{x,t-1}\, u_{y,t}}_{Op(T)}
\end{aligned}
$$

and

$$
\sum y_{t-1}u_{y,t} = Y_0 \underbrace{\sum U_{y,t}}_{Op(T^{1/2})} + \mu_y \underbrace{t \sum u_{y,t}}_{Op(T^{3/2})} + \underbrace{\sum \xi_{y,t-1}\, u_{y,t}}_{Op(T)} .
$$

Again, the order in convergence of the elements that appear in the previous equations are:

$$
\begin{aligned}
\sum y^2_{t-1} = {} & Y_0^2\, T + \mu_y^2 \sum t^2 + \underbrace{\sum \xi^2_{y,t-1}}_{Op(T^2)} + 2\, Y0\, \mu_y \sum t + 2\, Y_0 \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})} \\
& + 2\, \mu_y \underbrace{\sum t\, \xi_{y,t-1}}_{Op(T^{5/2})},
\end{aligned}
$$

$$
\begin{aligned}
\sum x^2_{t-1} &= X_0^2\, T + \mu_x^2 \sum t^2 + \underbrace{\sum \xi^2_{x,t-1}}_{Op(T^2)} + 2\, X0\, \mu_x \sum t + 2\, X_0 \underbrace{\sum \xi_{x,t-1}}_{Op(T^{3/2})} \\
&\quad + 2\, \mu_y \underbrace{\sum t\, \xi_{x,t-1}}_{Op(T^{5/2})}, \\
\sum y_{t-1}\, y_t &= Y_0 \sum y_{t-1} - Y_0\, \mu_y\, T - Y_0 \underbrace{\sum u_{y,t}}_{Op(T^{1/2})} - \mu_y^2 \sum t + \mu_y \sum t\, y_{t-1} \\
&\quad - \mu_y \underbrace{\sum t\, u_{y,t}}_{Op(T^{3/2})} + \sum y_{t-1}\xi_{y,t-1} - \mu_y \underbrace{\sum \xi_{y,t-1}}_{Op(T^{3/2})} - \underbrace{\sum \xi_{y,t-1}\, u_{y,t}}_{Op(T)}, \\
\sum y_{t-1}x_{t-1} &= Y_0 \sum x_{t-1} + \mu_y \sum t\, x_{t-1} + \sum x_{t-1}\xi_{y,t-1}, \\
\sum x_{t-1}y_t &= \sum y_{t-1}x_{t-1} - \mu_y \sum x_{t-1} - \sum x_{t-1}u_{y,t}
\end{aligned}
$$

and

$$\begin{aligned}\sum y_t^2 &= \sum y_{t-1}^2 + \mu_y^2\, T + \underbrace{\sum u_{y,t}^2}_{Op(T)} - 2\,\mu_y\, \sum y_{t-1} - 2\, \sum y_{t-1} u_{y,t} \\ &\quad + 2\,\mu_y \underbrace{\sum u_{y,t}}_{Op(T^{1/2})}.\end{aligned}$$

All of the above expressions are required to compute SSR^R (2) and SSR^U (3). As in the previous section, the asymptotics are computed using a *Mathematica 7.0*. □

The proofs of Theorems 2 and 3 follow the same steps as those of Theorem 1. They are available as supplementary material at:
[rgb]0,0,1https://dl.dropboxusercontent.com/u/1307356/JoSEM_SuppMat.pdf.

Critical values of the GC test under driftless unit roots

Table 2 presents the asymptotic critical values of the $\mathcal{F}$ Granger-causality test, when the processes are independent driftless unit roots (see Theorem 1 part A). Table 2 also shows finite sample Monte Carlo evidence of the test for different sample sizes and variance parameters. The rejection rates (using 5% critical values) remain fairly close to the nominal 5% notwithstanding the value of parameters σ_x and σ_y.

Table 2: Critical values and finite sample behavior of $\mathcal{F}$. DGPs: independent UR_{ND}. Rejection rate of $\mathcal{F}$ using $\alpha = 5\%$. Number of Replications: $10,000$.

	Sample size	25	50	100	250	500	1000	10000
	Critical value	8.09	7.17	6.97	6.94	6.61	6.55	6.52
σ_x	σ_y							
0.5	0.5	4.32	4.96	4.62	4.47	5.02	5.23	5.08
	1.0	4.33	4.85	4.54	4.02	5.32	5.13	4.74
	1.5	4.08	5.03	4.70	4.32	4.54	5.18	4.64
	2.0	4.20	4.87	4.52	4.46	5.44	4.68	5.19
1.0	0.5	4.19	4.45	4.53	4.39	5.03	4.62	4.96
	1.0	4.51	4.82	4.63	4.09	5.29	5.16	4.89
	1.5	4.07	4.77	4.68	4.15	4.58	5.37	5.06
	2.0	4.37	5.05	4.62	4.52	5.04	5.18	4.86
1.5	0.5	4.27	4.91	4.84	4.6	4.71	4.93	4.87
	1.0	4.21	4.50	4.69	4.54	4.65	4.86	5.13
	1.5	4.11	4.64	4.54	4.08	4.89	5.04	5.00
	2.0	4.53	4.79	4.40	4.53	4.76	5.00	4.6
2.0	0.5	4.22	4.75	5.06	4.37	4.94	4.95	4.94
	1.0	4.11	4.81	4.75	4.31	4.97	4.77	4.80
	1.5	4.13	4.73	4.88	4.04	4.86	5.13	5.28
	2.0	4.96	4.84	4.40	4.35	4.99	5.08	4.59