Documentos de Trabajo

Lan Gaiak

A SIMPLE TEST FOR THE EQUALITY OF INTEGRATION ORDERS

Javier Hualde

D.T. 1206

Departamento de Economía

Ekonomia Saila

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

A simple test for the equality of integration orders

Javier Hualde*
Universidad Pública de Navarra

August 9, 2012

Abstract

A necessary condition for two time series to be nontrivially cointegrated is the equality of their respective integration orders. Thus, it is standard practice to test for order homogeneity prior to testing for cointegration. Tests for the equality of integration orders are particular cases of more general tests of linear restrictions among memory parameters of different time series, for which asymptotic theory has been developed in parametric and semiparametric settings. However, most tests have been developed in stationary and invertible settings, and, more importantly, many of them are invalid when the observables are cointegrated, because they usually involve inversion of an asymptotically singular matrix. We propose a general testing procedure which does not suffer from this serious drawback, and, in addition, it is very simple to compute, it covers the stationary/nonstationary and invertible/noninvertible ranges, and, as we show in a Monte Carlo experiment, it works well in finite samples.

JEL Classification: C32.

Keywords. Integration orders; fractional differencing; fractional cointegration.

*Corresponding author. Department of Economics. Universidad Pública de Navarra. 31006 Pamplona, Spain. Phone. +34948169674. Fax. +34948169721. email: javier.hualde@unavarra.es

1 Introduction

Recently, the concepts of fractionally integration and cointegration have raised the attention of numerous researchers. This new framework introduces additional challenges, because, in practice, those (possibly noninteger) integration orders are unknown, so the traditional way of testing for cointegration, based on ideas like the ones of Dickey and Fuller (1979) or Phillips and Perron (1988), needs to be revised. For example, given two observable series, y_t, x_t, $t = 1, ..., n$, a necessary condition for these processes to be nontrivially cointegrated (so a linear combination of them has a smaller order) is the equality of their respective integration orders. Thus, it is standard practice to test for order homogeneity prior to testing for cointegration. Tests for the equality of integration orders are particular cases of more general tests of linear restrictions among memory parameters of multivariate time series, which have been developed mainly assuming stationarity and invertibility. In the parametric setting rigorous asymptotic theory has been developed by Heyde and Gay (1993) and Hosoya (1997). In the semiparametric setting, under local assumptions, Wald tests of linear restrictions on memory parameters have been proposed for the stationary case by Robinson (1995a) and Lobato (1999), but results in Robinson (1995b) suggest also the use of Lagrange Multiplier and Likelihood Ratio tests. These semiparametric tests enjoy standard asymptotics (feature also shared by the parametric ones), but suffer from a serious drawback, because they are invalid in case there exists cointegration among the series. The reason is that the test statistics involve inversion of an asymptotically singular matrix . This problem was acknowledged by Marinucci and Robinson (2001), and Robinson and Yajima (2002) offered a sensible solution at cost of introducing an additional user-chosen number.

The present paper proposes a testing procedure for the equality of integration orders of two fractionally integrated processes. The test covers the stationary/nonstationary and invertible/noninvertible ranges and it is valid irrespective of whether the time series are cointegrated or not. In addition, its computation just requires estimation of integration orders and of the spectral density of the short memory input series which originate the fractionally integrated processes at frequency zero, and, assuming very mild conditions, it enjoys standard asymptotics under the null hypothesis of equality of orders.

In the next section we present our testing procedure, which is rigorously justified in the Appendix. Section 3 includes a Monte Carlo study of finite-sample behavior and, finally, we conclude in Section 4.

2 Testing for the equality of integration orders

Consider the bivariate process $z_t = (y_t, x_t)'$, prime denoting transposition, $t \in Z$, $Z = \{0, \pm 1, ...\}$, where

$$y_t = \Delta^{-\delta_y} \{v_{1t} 1(t > 0)\}, \quad y_t = 0,\, t \leq 0, \tag{1}$$
$$x_t = \Delta^{-\delta_x} \{v_{2t} 1(t > 0)\}, \quad x_t = 0,\, t \leq 0, \tag{2}$$

$1(\cdot)$ denoting the indicator function. We introduce

Assumption A. *The process $v_t = (v_{1t}, v_{2t})'$, $t \in Z$, has representation*

$$v_t = \sum_{j=0}^{\infty} A_j \varepsilon_{t-j}, \ \sum_{j=0}^{\infty} j \left\| A_j \right\|^2 < \infty, \tag{3}$$

with $\|\cdot\|$ denoting the Euclidean norm, where

(i) ε_t are independent and identically distributed vectors with mean zero, positive definite covariance matrix Ω, $E \left\| \varepsilon_t \right\|^q < \infty$, $q > 2$;

(ii) $f_{ii}(0) > 0$, $i = 1, 2$, where $f_{ij}(0)$ is the (i, j) element of the spectral density of v_t, denoted by $f(\lambda)$.

Model (1), (2) under Assumption A imply that y_t, x_t, are Type II fractionally integrated processes of orders δ_y, δ_x, respectively (see, e.g., Robinson, 2005).

We introduce additional notation. For any vector or scalar sequence ζ_t, $t = 1, ..., n$, define its finite Fourier transform as $w_\zeta(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{n} \zeta_t e^{i\lambda t}$, for any $\lambda \in [-\pi, \pi]$. Also, for any sequence s_t and any real c, let $s_t(c) = \Delta^c \{s_t 1(t > 0)\}$, and related to y_t, x_t in (1), (2), for real c, d, define $z_t(c, d) = (y_t(c), x_t(d))'$. Finally, let " $\sim$ " mean "exact rate of convergence".

Consider certain estimators $\widehat{\delta}_x$, $\widehat{\delta}_y$, $\widehat{f}(0)$ of δ_x, δ_y, $f(0)$ respectively, such that the following condition holds.

Assumption B. *As $n \to \infty$,*

$$\widehat{f}(0) \to_p f(0),$$

and for $\kappa > 0$ and $K < \infty$,

$$\widehat{\delta}_x - \delta_x \sim n^{-\kappa}, \widehat{\delta}_y - \delta_y \sim n^{-\kappa},$$

where

$$\left| \widehat{\delta}_x \right| + \left| \widehat{\delta}_y \right| \leq K. \tag{4}$$

Assumption B, although not primitive, is very mild. (4) is innocuous if $\widehat{\delta}_x$, $\widehat{\delta}_y$, optimize over compact sets. If we assume a parametric structure for v_t, $\sqrt{n}$-consistent estimators of the orders of integration and $f(0)$ are achievable by a multivariate extension of the results in Robinson (2005), which extended results in Velasco and Robinson (2000) in the univariate case to cover our type of processes. This rate is far better than needed, so we might be content by assuming some weak conditions of smoothness of the spectral density of v_t around frequency zero, and estimate the orders and $f(0)$ semiparametrically. Regarding integration orders, the estimates of Robinson (1995a,b), justified by Robinson (2005) for our type of processes, satisfy Assumption B. Also, given estimates $\widehat{\delta}_x$,

$\widehat{\delta}_y$, the nonparametric estimate of $f(0)$ could be based on weighted averages of the periodogram of the proxy $\widehat{v}_t = (y_t(\widehat{\delta}_y), x_t(\widehat{\delta}_x))'$ of v_t.

We introduce now our test statistic. Let $h_n > 0$ be a sequence (whose role will be clarified in Remark 2 below) such that

$$h_n + n^{\kappa} h_n^{-1} \to \infty \text{ as } n \to \infty. \tag{5}$$

Define

$$\widehat{a} = (1(n^{\kappa}(\widehat{\delta}_y - \widehat{\delta}_x) > h_n), 1(n^{\kappa}(\widehat{\delta}_y - \widehat{\delta}_x) \leq h_n))', \tag{6}$$

and let

$$\widehat{t} = \frac{\widehat{a}' w_{z(\widehat{\delta}_x, \widehat{\delta}_y)}(0)}{\left(\widehat{a}' \widehat{f}(0) \widehat{a}\right)^{1/2}}, \tag{7}$$

be the test statistic for $H_0 : \delta_y = \delta_x$ against the alternative $H_1 : \delta_y \neq \delta_x$.

Theorem 1. *Let (1), (2) and Assumptions A, B hold. Then*

$$\widehat{t} \to_d N(0,1) \text{ under } H_0; \widehat{t} \sim n^{|\delta_x - \delta_y|} \text{ under } H_1. \tag{8}$$

Remark 1. As shown in the proof of Theorem 1, $\widehat{a} \to_p a \equiv (1,0)' 1(\delta_y > \delta_x) + (0,1)' 1(\delta_y \leq \delta_x)$, so $\widehat{t}$ is asymptotically equivalent to $w_{y(\widehat{\delta}_x)}(0) / \widehat{f}_{11}^{1/2}(0)$, or, alternatively, to $w_{x(\widehat{\delta}_y)}(0) / \widehat{f}_{22}^{1/2}(0)$, depending on whether $\delta_y > \delta_x$ or $\delta_y \leq \delta_x$, respectively. Thus, asymptotically, $\widehat{t}$ is based on underdifferenced processes under H_1 (which is precisely the source of power), whereas under H_0 it is based on $x_t(\widehat{\delta}_y)$ (although it could have been equally based on $y_t(\widehat{\delta}_x)$ with a slight modification of the definition of $\widehat{a}$).

Remark 2. It would have been natural to set $h_n = 0$ in (6). In this case, the test would have been based on $y_t(\widehat{\delta}_x)$ if $\widehat{\delta}_y > \widehat{\delta}_x$, or $x_t(\widehat{\delta}_y)$ if $\widehat{\delta}_y \leq \widehat{\delta}_x$. However setting $h_n = 0$ in (6) implies that under H_0 the limit of $\widehat{a}$ is random. This leads to a very complicated limit dependence between the numerator and denominator of (7), so the simple and neat result in (8) would no longer hold.

Remark 3. If y_t and x_t were cointegrated $f(0)$ would be singular. This is precisely the reason why the different semiparametric tests considered in the literature are not valid with cointegration, as they require inversion of a matrix which tends in probability to a singular matrix (usually the equivalent to $f(0)$ in a more general framework). As can be inferred from Remark 1, singularity of $f(0)$ does not affect our test procedure as long as $f_{ii}(0) > 0$, $i = 1, 2$.

Remark 4. Although we just consider Type II processes, this was just motivated by the uniform treatment of any value of δ_x and δ_y that this definition allows, all results holding equally for Type I processes.

Remark 5. We do not consider deterministic components in (1), (2). However, there is no loss of generality here, because these components can be eliminated by differencing the observables appropriately, and then applying our procedure to the differenced series. Taking differences might lead to differenced processes with negative integration orders. This complicates the likability of Assumption B, although procedures like Hualde and Robinson (2011) or Hurvich and Chen (2000) are appropriate in these circumstances.

3 Monte Carlo evidence

With the aim of assessing for the finite sample behavior of our test procedure, we performed a small Monte Carlo experiment. We generated y_t and x_t as in (1), (2), v_t being a bivariate Gaussian white noise with $E(v_t) = 0$, $Var(v_{it}) = 1$, $i = 1, 2$, $Cov(v_{1t}, v_{2t}) = 0.5$. We computed $\widehat{t}$ parametrically: δ_x, δ_y were estimated as in Hualde and Robinson (2011) and

$$\widehat{f}(0) = \frac{1}{2\pi n} \sum_{t=1}^{n} \widehat{v}_t \widehat{v}_t'.$$

Using 5000 replications and four different sample sizes $n = 64, 128, 256, 512$, we computed the proportion of rejections of $\widehat{t}$ for nominal sizes $\alpha = .01, .05, .10$, and different combinations of δ_y, δ_x. Denoting $\phi = \delta_y - \delta_x$, we fixed $\delta_y = 1.4$ and considered $\phi = 0, .1, .2, ..., .7$ (note that our test procedure is invariant to the particular values of δ_y, δ_x, depending just on ϕ). Finally, we chose $h_n = \log\left(n^{1/2}\right)$, noting that our estimators of the orders are $\sqrt{n}$-consistent. Other choices of h_n led to very similar results to those presented in Table 1. Looking at sizes first (so $\phi = 0$), $\widehat{t}$ is oversized, although, as n increases, empirical sizes move in all cases in the appropriate direction, so for $n = 512$ and $\alpha = .05, .10$, empirical sizes are very close to the nominal ones. We also looked at power by letting $\phi \neq 0$. As expected, power increases as n and ϕ increase. The test has difficulties to detect the $\phi = 0.1$ situation (although, as n increases, results improve), but performs relatively well in all cases if $\phi \geq 0.2$. Overall, $\widehat{t}$ behaves well in terms of size and enjoys an acceptable power for small values of ϕ (even for relatively short sample sizes).

4 Conclusions

We have presented a simple test for the equality of the integration orders of two fractionally integrated processes. Our procedure covers the stationary/nonstationary and invertible/noninvertible ranges, it remains valid under cointegration, it is very simple to compute and enjoys standard asymptotics.

Appendix. Proof of Theorem 1

First, we show that $\widehat{a} \to_p a$, which noting that

$$1(n^{\kappa}(\widehat{\delta}_y - \widehat{\delta}_x) > h_n) + 1(n^{\kappa}(\widehat{\delta}_y - \widehat{\delta}_x) \leq h_n) = 1,$$

follows on showing that, as $n \to \infty$,

$$1(n^{\kappa}(\widehat{\delta}_y - \widehat{\delta}_x) \leq h_n) = o_p(1), \text{ if } \delta_y > \delta_x, \quad (9)$$

$$1(n^{\kappa}(\widehat{\delta}_y - \widehat{\delta}_x) > h_n) = o_p(1), \text{ if } \delta_y \leq \delta_x. \quad (10)$$

Denote $Q_n = n^{\kappa}(\widehat{\delta}_y - \widehat{\delta}_x - (\delta_y - \delta_x))$, noting that under Assumption B, $|Q_n| = O_p(1)$. We show (9) first. The left hand side of (9) equals

$$1(Q_n + n^{\kappa}(\delta_y - \delta_x) \leq h_n) \leq 1(n^{\kappa}(\delta_y - \delta_x) \leq h_n + |Q_n|) \leq \frac{h_n + |Q_n|}{n^{\kappa}(\delta_y - \delta_x)} = o_p(1),$$

by (5). Next, the left hand side of (10) equals

$$1(Q_n > h_n + n^{\kappa}(\delta_x - \delta_y)) \leq \frac{|Q_n|}{h_n + n^{\kappa}(\delta_x - \delta_y)} = o_p(1),$$

by (5). Next we show that

$$\begin{aligned} & \widehat{a}' w_{z(\widehat{\delta}_x, \widehat{\delta}_y)}(0) - a' w_{z(\delta_x, \delta_y)}(0) \\ = \; & (\widehat{a} - a)' w_{z(\delta_x, \delta_y)}(0) + \widehat{a}'(w_{z(\widehat{\delta}_x, \widehat{\delta}_y)}(0) - w_{z(\delta_x, \delta_y)}(0)) && (11) \\ = \; & o_p(1) \text{ under } H_0, && (12) \\ = \; & o_p\left(n^{|\delta_x - \delta_y|}\right) \text{ under } H_1. && (13) \end{aligned}$$

We show (12) first. Under Assumption A, noting that (3) implies that $f(\lambda)$ is $Lip(\varkappa)$, $\varkappa > 0$, by Central Limit Theorem (see Hannan, 1970), $w_{z(\delta_x, \delta_y)}(0) = O_p(1)$, so given that $\widehat{a} - a = o_p(1)$, the first term in (11) is $o_p(1)$. Next, noting that $\|\widehat{a}\| \leq \sqrt{2}$, (12) holds if $w_{z(\widehat{\delta}_x, \widehat{\delta}_y)}(0) - w_{z(\delta_x, \delta_y)}(0) = o_p(1)$ under H_0. By Taylor's expansion, for certain constant R to be defined subsequently, $w_{z(\widehat{\delta}_x, \widehat{\delta}_y)}(0) - w_{z(\delta_x, \delta_y)}(0)$ equals

$$\begin{aligned} & \frac{1}{\sqrt{2\pi n}} \sum_{r=1}^{R-1} \frac{1}{r!} \begin{pmatrix} (\delta_x - \widehat{\delta}_x)^r & 0 \\ 0 & (\delta_y - \widehat{\delta}_y)^r \end{pmatrix} \sum_{t=2}^{n} g^{(r)}(v_t; \delta_y - \delta_x, \delta_x - \delta_y) \\ & + \frac{1}{\sqrt{2\pi n} R!} \begin{pmatrix} (\delta_x - \widehat{\delta}_x)^R & 0 \\ 0 & (\delta_y - \widehat{\delta}_y)^R \end{pmatrix} \sum_{t=2}^{n} g^{(R)}\left(v_t; \delta_y - \overline{\delta}_x, \delta_x - \overline{\delta}_y\right), \end{aligned} \quad (14)$$

where $\left|\overline{\delta}_x - \delta_y\right| \leq \left|\widehat{\delta}_x - \delta_y\right|$, $\left|\overline{\delta}_y - \delta_x\right| \leq \left|\widehat{\delta}_y - \delta_x\right|$, for any scalar or vector sequence ψ_t and any real b, $g^{(r)}(\psi_t; b) = \sum_{s=1}^{t-1} a_s^{(r)}(b) \psi_{t-s}$, with $a_s^{(r)}(b) = d^r a_s(b) / db^r$, and for any p-dimensional vector $\xi_t = \left(\xi_{1t}, ..., \xi_{pt}\right)'$ and real $b_1, ..., b_p$,

$$g^{(r)}(\xi_t; b_1, ..., b_p) = \left(g^{(r)}(\xi_{1t}; b_1), ..., g^{(r)}(\xi_{pt}; b_p)\right)'.$$

First, by a simple modification of the proof of (C.8) in Lemma C.2 of Robinson and Hualde (2003) (using the bounds in Lemma D.4 of Robinson and Hualde, 2003), it can be easily shown that

$$Var\left(\sum_{t=2}^{n} g^{(r)}\left(v_t;0,0\right)\right) = O\left(n\left(\log n\right)^{2r}\right),$$

implying that the first term in (14) is $O_p\left(n^{-\kappa}\log n\right)$. Next, by Lemma C.4 of Robinson and Hualde (2003)

$$g^{(R)}\left(v_t;\delta_y-\overline{\delta}_x,\delta_x-\overline{\delta}_y\right) = O_p\left(t^{\frac{1}{2}}\right),$$

so the second term in (14) is $O_p\left(n^{1-R\kappa}\right)$, and choosing $R > (1+\kappa)/\kappa$, (14) is $O_p\left(n^{-\kappa}\log n\right)$, so (12) holds because $\kappa > 0$.

Next we show (13). First, by Theorem 1 of Marinucci and Robinson (2000), $w_{z(\delta_x,\delta_y)}(0) = O_p\left(n^{|\delta_x-\delta_y|}\right)$, so the result for the first term in (11) follows by previous arguments. Next we show $w_{z\left(\widehat{\delta}_x,\widehat{\delta}_y\right)}(0) - w_{z(\delta_x,\delta_y)}(0) = o_p\left(n^{|\delta_x-\delta_y|}\right)$ under H_1. Noting (14), again by a simple modification of the proof of (C.8) in Lemma C.2 of Robinson and Hualde (2003), it can be shown that

$$Var\left(\sum_{t=2}^{n} g^{(r)}\left(v_t;\delta_y-\delta_x,\delta_x-\delta_y\right) = O\left(\left(\log n\right)^{2r} n^{2|\delta_x-\delta_y|+1}\right)\right).$$

Next, as in Lemma C.4 of Robinson and Hualde (2003), for any $\epsilon > 0$,

$$\left\|g^{(R)}\left(v_t;\delta_y-\overline{\delta}_x,\delta_x-\overline{\delta}_y\right)\right\| \le K\left(t\sum_{s=1}^{t}\left(\log s\right)^{2R} s^{2(|\delta_x-\delta_y|+\epsilon-1)}\right)^{\frac{1}{2}},$$

where K is a finite positive constant, implying that

$$\sum_{t=2}^{n} g^{(R)}\left(v_t;\delta_y-\overline{\delta}_x,\delta_x-\overline{\delta}_y\right) = O_p\left(\left(\log n\right)^{R} n^{\max\left\{|\delta_x-\delta_y|,\frac{1}{2}\right\}+\epsilon+1}\right),$$

and choosing $R > \left(\max\left\{|\delta_x-\delta_y|,\frac{1}{2}\right\}+1/2+\epsilon+\kappa\right)/\kappa$, (14) is $O_p\left(n^{|\delta_x-\delta_y|-\kappa}\log n\right)$ under H_1, to conclude the proof of (13).

Finally, noting that by Assumptions A, B and previous arguments, $\widehat{a}'\widehat{f}(0)\widehat{a} \to_p a'f(0)a > 0$, in view of (12), (13), (8) follows easily by simple application of Central Limit Theorem and Theorem 1 of Marinucci and Robinson (2000), noting that under H_1 the overdifferenced process has smaller order.

Acknowledgements. I am grateful to P.M. Robinson, J. Hidalgo, C. Velasco, L. Giraitis for several comments and suggestions. The author's research was supported by the Spanish Ministerio de Ciencia e Innovación through ref. ECO2011-24304.

REFERENCES

Dickey, D.A., Fuller, W.A., 1979. Distribution of estimators of autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427-431.

Hannan, E.J., 1970. Multiple time series analysis. Wiley, New York.

Heyde, C.C., Gay, R., 1993. Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. Stochastic Processes and their Applications 45, 169-182.

Hosoya, Y., 1997. A limit theory for long-range dependence and statistical inference on related models. Annals of Statistics 25, 105-137.

Hualde, J., Robinson, P.M., 2011. Gaussian pseudo-maximum likelihood estimation of fractional time series models. Annals of Statistics 39, 3152-3181.

Hurvich, C.M., Chen, W.W., 2000. An efficient taper for potentially overdiferenced long-memory time series. Journal of Time Series Analysis 21, 155–180.

Lobato, I.G., 1999. A semiparametric two-step estimation on a multivariate long memory model. Journal of Econometrics 90, 129-153.

Marinucci, D., Robinson, P.M., 2000. Weak convergence of multivariate fractional processes. Stochastic Processes and their Applications 86, 103-120.

Marinucci, D., Robinson, P.M., 2001. Semiparametric fractional cointegration analysis. Journal of Econometrics 105, 225-247.

Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335-346.

Robinson, P.M., 1995a. Log-periodogram regression of time series with long range dependence. Annals of Statistics 23, 1048-1072.

Robinson, P.M., 1995b. Gaussian semiparametric estimation of long-range dependence. Annals of Statistics 23, 1630-1661.

Robinson, P.M., 2005. The distance between rival nonstationary fractional processes. Journal of Econometrics 128, 283–300

Robinson, P.M., Hualde, J., 2003. Cointegration in fractional systems with unknown integration orders. Econometrica 71, 1727-1766.

Robinson, P.M., Yajima, Y., 2002. Determination of cointegrating rank in fractional systems. Journal of Econometrics 106, 217-241.

Velasco, C., Robinson, P.M., 2000. Whittle pseudo-maximum likelihood estimation for nonstationary time series. Journal of the American Statistical Association 95, 1229-1243.

Table 1
Proportion of rejections of $\widehat{t}$

n		64			128			256			512	
ϕ/α	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10
0	.053	.110	.165	.035	.085	.139	.023	.074	.125	.018	.062	.110
.1	.066	.101	.130	.046	.071	.092	.045	.066	.087	.058	.086	.109
.2	.168	.210	.235	.199	.246	.275	.309	.382	.428	.410	.518	.578
.3	.359	.419	.451	.477	.560	.599	.587	.675	.724	.658	.729	.768
.4	.550	.622	.656	.655	.730	.768	.755	.815	.840	.810	.857	.880
.5	.693	.757	.792	.777	.829	.856	.837	.876	.897	.885	.912	.926
.6	.785	.837	.864	.861	.894	.911	.906	.929	.940	.935	.951	.959
.7	.856	.890	.907	.903	.927	.941	.938	.955	.961	.964	.973	.977