SUPPLEMENT TO "EFFICIENT WALD TESTS FOR FRACTIONAL UNIT ROOTS"

(*Econometrica*, Vol. 75, No. 2, March, 2007, 575–589)

BY IGNACIO N. LOBATO AND CARLOS VELASCO[1]

APPENDIX 1

WE PROVIDE HERE the proof of Theorem 1. The proof of part (b) is omitted because it is easily obtained using the same methods as Dolado, Gonzalo, and Mayoral's (DGM) Theorem 3. In addition, because (a) is a particular case ($\delta = 0$) of (c), we just report the proof for (c). For simplicity, and without loss of generality, in this appendix we assume that the variance of ε_t is 1. We start by considering the case where d_2, the input of z_{t-1}, is fixed. The case where it is stochastic (and consistent for some fixed value under condition (10) in the text) is discussed at the end of this appendix.

We begin by introducing some notation. Let the t-test statistic for $\phi_2 = 0$ be

$$t_\phi = t_\phi(d_2) = \frac{\sum_{t=2}^{T} \Delta y_t z_{t-1}(d_2)}{\widehat{S}_T(d_2)\sqrt{\sum_{t=2}^{T}(z_{t-1}(d_2))^2}},$$

where $\widehat{S}_T^2(d_2) = T^{-1}\sum_{t=2}^{T}(\Delta y_t - \widehat{\phi}_2 z_{t-1}(d_2))^2$ and $\widehat{\phi}_2$ denotes the ordinary least squares estimator of ϕ_2 in (7) in the text. Under local alternatives we have that

$$\Delta y_t = \Delta^{-\theta_T}\varepsilon_t 1\{t>0\} = \varepsilon_t + \sum_{i=1}^{t-1}\pi_i(-\theta_T)\varepsilon_{t-i},$$

where $\theta_T := -\delta T^{-1/2}$, $\pi_1(-\theta_T) = \theta_T$, and $\pi_2(-\theta_T) = 0.5\theta_T(1+\theta_T)$, and by Taylor expanding $\pi_i(\cdot)$ around $\pi_i(0) = 0$, $i > 0$, we obtain

$$T^{1/2}\pi_i(-\theta_T) = -i^{-1}\delta + O(T^{-1/2}i^{-1}\log^2 i) \qquad (i = 1, 2, \ldots, T);$$

[1]We thank the co-editor and two referees for very useful comments, J. Arteche, M. Avarucci, M. Delgado, J. Dolado, L. Gil-Alaña, J. Gonzalo, J. Hidalgo, L. Mayoral, P. Perron, W. Ploberger, and P. Robinson for useful conversations, and seminar participants at the London School of Economics and at the 2005 Econometric Society World Congress. This research was financed by the Spanish Ministerio de Educación y Ciencia (SEJ2004-04583/ECON). Part of this research was carried out while Lobato was visiting Universidad Carlos III de Madrid thanks to the Spanish Secretaría de Estado de Universidades e Investigación (SAB2004-0034). Lobato acknowledges financial support from Asociación Mexicana de Cultura and from the Mexican Consejo Nacional de Ciencia y Tecnología (CONACYT) under project grant 41893-S.

see Delgado and Velasco (2005, Lemma 1) and Robinson and Hualde (2003, Lemma D.1). When $d_2 \neq 1$, note that

$$z_{t-1}(d_2) = \frac{\Delta^{-\eta_T} - \Delta^{-\theta_T}}{1-d_2}\varepsilon_t 1\{t>0\} = \varepsilon_{t-1} + \sum_{i=2}^{t-1} \psi_i(\theta_T, \eta_T)\varepsilon_{t-i},$$

where $\eta_T = 1 - d_2 - \delta T^{-1/2}$ and $\psi_i(\theta_T, \eta_T) = (\pi_i(-\eta_T) - \pi_i(-\theta_T))/(1-d_2)$.

First, consider the numerator of $t_\phi(d_2)$ scaled by $T^{-1/2}$,

$$Q_T(d_2) := T^{-1/2}\sum_{t=2}^{T} \Delta y_t z_{t-1}(d_2)$$

$$= T^{-1/2}\sum_{t=2}^{T}\left(\varepsilon_t + \sum_{i=1}^{t-1}\left(\frac{-\delta}{i\sqrt{T}}\right)\varepsilon_{t-i}\right) \times \left(\varepsilon_{t-1} + \sum_{i=2}^{t-1}\psi_i(\theta_T, \eta_T)\varepsilon_{t-i}\right) \tag{S.1}$$

$$+ T^{-1/2}\frac{\delta^2}{2T}\sum_{t=2}^{T}\left(\sum_{i=1}^{t-1}\pi_i^{(2)}(-\theta^*)\varepsilon_{t-i}\right) \times \left(\varepsilon_{t-1} + \sum_{i=2}^{t-1}\psi_i(\theta_T, \eta_T)\varepsilon_{t-i}\right), \tag{S.2}$$

where $\pi_i^{(2)}$ is the second derivative of $\pi_i(\cdot)$ and θ^* is some point between 0 and θ_T. Note that $|\pi_i^{(2)}(-\theta^*)| \leq Ci^{-1}\log^2 i$, $i = 1, \ldots, T$, by Lemma 1(b) of Delgado and Velasco (2005). Because (S.1) is $O_p(1)$, as is shown next, it is straightforward to show that (S.2) is $o_p(1)$.

The leading term (S.1) of $Q_T(d_2)$ can be written as

$$T^{-1/2}\sum_{t=2}^{T}\left(\varepsilon_t - \frac{\delta}{\sqrt{T}}\varepsilon_{t-1} - \sum_{i=2}^{t-1}\frac{\delta}{i\sqrt{T}}\varepsilon_{t-i}\right)\left(\varepsilon_{t-1} + \sum_{i=2}^{t-1}\psi_i(\theta_T, \eta_T)\varepsilon_{t-i}\right)$$

$$= -T^{-1/2}\sum_{t=2}^{T}\left(\frac{\delta}{\sqrt{T}}\varepsilon_{t-1}^2 + \sum_{i=2}^{t-1}\frac{\delta}{i\sqrt{T}}\psi_i(\theta_T, \eta_T)\varepsilon_{t-i}^2\right) \tag{S.3}$$

$$+ T^{-1/2}\sum_{t=2}^{T}\varepsilon_t\left(\varepsilon_{t-1} + \sum_{i=2}^{t-1}\psi_i(\theta_T, \eta_T)\varepsilon_{t-i}\right) \tag{S.4}$$

$$- T^{-1/2}\sum_{t=2}^{T}\left(\frac{\delta}{\sqrt{T}}\varepsilon_{t-1}\right)\left(\sum_{i=2}^{t-1}\psi_i(\theta_T, \eta_T)\varepsilon_{t-i}\right) \tag{S.5}$$

$$-T^{-1/2}\sum_{t=2}^{T}\left[\sum_{i=2}^{t-1}\frac{\delta}{i\sqrt{T}}\varepsilon_{t-i}\left(\varepsilon_{t-1}+\sum_{j=2,j\neq i}^{t-1}\psi_j(\theta_T,\eta_T)\varepsilon_{t-j}\right)\right]. \tag{S.6}$$

The last two terms, (S.5) and (S.6), in the previous expression are $o_p(1)$ using arguments similar to those in the proof of Theorem 4 in DGM. Using the properties of the fractional difference filter and a weak law of large numbers (see, for instance, the proof of Lemma 1 in DGM), the term (S.3) converges in probability to $-\delta K(d_2)$, where

$$\begin{aligned}K(d_2)&=\lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T}\left(1+\sum_{i=2}^{t-1}\frac{1}{i}\psi_i(\theta_T,\eta_T)\right)\\&=\lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T}\left(1+\sum_{i=2}^{t-1}\frac{\pi_i(-\eta_T)}{i(1-d_2)}\right)\\&=\sum_{i=1}^{\infty}\frac{\pi_i(d_2-1)}{i(1-d_2)}.\end{aligned}$$

Using a standard central limit theorem for martingale difference sequences, the term (S.4) converges in distribution to a $\mathrm{N}(0,V)$, where

$$\begin{aligned}V&=\lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T}E\left(\varepsilon_t\varepsilon_{t-1}+\sum_{i=2}^{t-1}\psi_i(\theta_T,\eta_T)\varepsilon_t\varepsilon_{t-i}\right)^2\\&=\lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T}E\left(\sum_{i=1}^{t-1}\frac{\pi_i(d_2-1)}{(1-d_2)}\varepsilon_t\varepsilon_{t-i}\right)^2\\&=\frac{\sum_{i=1}^{\infty}\pi_i(d_2-1)^2}{(1-d_2)^2}<\infty,\end{aligned}$$

because $1-d_2<0.5$ and $d_2\neq 1$. Hence, $Q_T(d_2)\to_d \mathrm{N}(-\delta K(d_2),\sum_{i=1}^{\infty}(\pi_i(d_2-1)/(1-d_2))^2)$.

Second, consider the denominator of $t_\phi(d_2)$ scaled by $T^{-1/2}$. It is straightforward to show that $\widehat{S}_T^2(d_2)\to_p 1$, and, given the preceding expression for $z_{t-1}(d_2)$, by a law of large numbers it is simple to see that the limit in probability of $T^{-1}\sum_{t=2}^{T}(z_{t-1}(d_2))^2$ is given by

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T}E\left(\varepsilon_{t-1}+\sum_{i=2}^{t-1}\psi_i(\theta_T,\eta_T)\varepsilon_{t-i}\right)^2=\frac{\sum_{i=1}^{\infty}\pi_i(d_2-1)^2}{(1-d_2)^2}.$$

So far we have considered the case where $d_2 \neq 1$. The case $d_2 = 1$ follows similarly, the difference being that under the local alternative $z_{t-1}(1)$ is now expressed as

$$z_{t-1}(1) = J(L)\Delta^{-\theta_T}\varepsilon_t 1\{t > 0\}.$$

Note that the filter $\psi_T^*(L) := J(L)\Delta^{-\theta_T}$ can be expressed as $\psi_T^*(L) = \sum_{j=1}^{\infty} \psi_{T,i}^* L^i$, where

$$\psi_{T,i}^* = \sum_{j=1}^{i} \frac{1}{j}\pi_{i-j}(-\theta_T) \qquad (i = 1, 2, 3, \ldots),$$

so that $\psi_{T,i}^* = i^{-1}(1 + O(\log T/\sqrt{T}))$ uniformly in $i = 1, \ldots, T$. Using this definition of $z_{t-1}(1)$, all the previous results can be easily adapted. For instance, we have that $K(1) = \lim_{T\to\infty} T^{-1}\sum_{t=2}^{T}(1 + \sum_{i=2}^{t-1} i^{-1}\psi_{T,i}^*) = \sum_{i=1}^{\infty} i^{-2} = \pi^2/6$.

Next, we analyze briefly the case of a stochastic input $\widehat{d}_2$ that satisfies condition (S.10) in the text. To show that $t_\phi(\widehat{d}_2) \to_p t_\phi(d_2)$, we just analyze here the most critical component of $t_\phi(d_2)$, which is the scaled numerator $Q_T(d_2)$; the analysis for the denominator is similar but simpler. Note that under the null, for $d_2 \neq 1$, $Q_T(d_2)$ simplifies to

$$\begin{aligned} Q_T(d_2) &= T^{-1/2}\sum_{t=1}^{T} \varepsilon_t \left(\frac{\Delta^{d_2-1} - 1}{1 - d_2}\right)\varepsilon_t \\ &= \frac{T^{-1/2}}{1 - d_2}\sum_{t=1}^{T} \varepsilon_t \sum_{j=1}^{t-1} \pi_j(d_2 - 1)\varepsilon_{t-j}. \end{aligned}$$

For $d_2 > 0.5$, $Q_T(d_2)$ converges to a zero mean normal variate in distribution, as we have already seen. Then, proceeding as in Robinson and Hualde (2003, Proposition 9), we just need to prove that, for $d_2 \neq 1$,

$$(1 - d_2)Q_T(d_2) - (1 - \widehat{d}_2)Q_T(\widehat{d}_2) \tag{S.7}$$

$$= T^{-1/2}\sum_{t=1}^{T} \varepsilon_t \sum_{j=1}^{t-1}\{\pi_j(d_2 - 1) - \pi_j(\widehat{d}_2 - 1)\}\varepsilon_{t-j} \tag{S.8}$$

is $o_p(1)$. Note that, for $j = 1, 2, \ldots, T$, the expression $\{\pi_j(d_2 - 1) - \pi_j(\widehat{d}_2 - 1)\}$ equals

$$\sum_{r=1}^{R-1} \frac{1}{r!}(d_2 - \widehat{d}_2)^r \pi_j^{(r)}(d_2 - 1) + \frac{1}{R!}(d_2 - \widehat{d}_2)^R \pi_j^{(R)}(\bar{d}_2 - 1), \tag{S.9}$$

where $\bar{d}_2$ is an intermediate point between d_2 and $\widehat{d}_2$. Using (S.9), (S.8) can be written as

$$T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\sum_{j=1}^{t-1}\left\{\sum_{r=1}^{R-1}\frac{1}{r!}(d_2-\widehat{d}_2)^r\pi_j^{(r)}(d_2-1)\right\}\varepsilon_{t-j} \tag{S.10}$$

$$+T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\sum_{j=1}^{t-1}\left\{\frac{1}{R!}(d_2-\widehat{d}_2)^R\pi_j^{(R)}(\bar{d}_2-1)\right\}\varepsilon_{t-j}. \tag{S.11}$$

Because $|\pi_j^{(r)}(d_2-1)|\le Cj^{-d_2}\log^r j,\ j=1,2,\dots,T$ (see Robinson and Hualde (2003, Lemma D1)), it is straightforward to check that

$$T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\sum_{j=1}^{t-1}\pi_j^{(r)}(d_2-1)\varepsilon_{t-j}=O_p(1) \qquad (r=1,2,\dots,R-1),$$

because it has zero mean and finite variance because the sequence $\pi_j^{(r)}(d_2-1)$ is square summable when $d_2>0.5$. Then, using condition (10) in the text, we derive that (S.10) is $o_p(1)$. To analyze (S.11), note that $|\pi_j^{(R)}(\bar{d}_2-1)|\le Cj^{-\bar{d}_2}\log^R j\le Cj^{-1/2},\ j=1,2,\dots,T$, because $\bar{d}_2>0.5$. Therefore, the remainder term

$$T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\sum_{j=1}^{t-1}\pi_j^{(R)}(\bar{d}_2-1)\varepsilon_{t-j} \tag{S.12}$$

has first absolute moment bounded by

$$\begin{aligned}
&T^{-1/2}\sum_{t=1}^{T}(E|\varepsilon_t|^2)^{1/2}\left\{E\left[\left(\sum_{j=1}^{t-1}\pi_j^{(R)}(\bar{d}_2-1)\varepsilon_{t-j}\right)^2\right]\right\}^{1/2}\\
&\quad\le CT^{-1/2}\sum_{t=1}^{T}\left\{\sum_{j=1}^{t-1}\sum_{k=1}^{t-1}E\left|\pi_j^{(R)}(\bar{d}_2-1)\pi_k^{(R)}(\bar{d}_2-1)\varepsilon_{t-j}\varepsilon_{t-k}\right|\right\}^{1/2}\\
&\quad\le CT^{-1/2}\sum_{t=1}^{T}\left\{\sum_{j=1}^{t-1}\sum_{k=1}^{t-1}(jk)^{-1/2}E|\varepsilon_{t-j}\varepsilon_{t-k}|\right\}^{1/2}\\
&\quad\le T^{-1/2}\sum_{t=1}^{T}t^{1/2}\le CT.
\end{aligned}$$

Therefore, (S.12) is $O_p(T)$, and if we choose R such that $R\tau>1$, so that $(d_2-\widehat{d}_2)^R=o_p(T^{-1})$, (S.11) is of order $o_p(1)$ and Theorem 1 follows.

APPENDIX 2

In this appendix we give a sketch of the proof of Theorem 2(c). The proof of part (b) is omitted because it can be easily derived using the same methods as DGM's Theorem 7. We assume that the true d is known and that the proof when d is consistently estimated is similar but lengthier. We employ techniques similar to those explained at the end of Appendix 1.

The key idea is to use the basic equation of multivariate regression

$$t_{\phi} = \sqrt{T}\frac{R_T}{\sqrt{1-R_T^2}}, \qquad \text{(S.13)}$$

where R_T denotes the sample partial correlation coefficient between $Y_t := \Delta y_t$ and $X_t := \alpha(L)z_{t-1}(d)$ given the p lags of Δy_t, and $Z_t := (Z_{t,1}, \ldots, Z_{t,p})'$ with $Z_{t,k} = \Delta y_{t-k}$, $k = 1, \ldots, p$, to derive the drift of the asymptotic distribution of t_{ϕ}. Note that the denominator in (S.13) tends to 1 in probability under local alternatives for which the DGP is given by

$$\Delta y_t = \alpha(L)^{-1}\Delta^{\delta/\sqrt{T}}\varepsilon_t$$

and where the operator $\Delta^{\delta/\sqrt{T}}$ can be written as

$$\Delta^{\delta/\sqrt{T}} = 1 - \frac{\delta}{\sqrt{T}}J(L) + \frac{1}{T}H_T(L)$$

with $H_T(L) = \sum_{j=1}^{\infty} h_{T,j}L^j$, so that $|h_{T,j}| \leq Cj^{-1}\log^2 j$, $j \geq 1$, uniformly in T. Then we can write the series involved in t_{ϕ} in terms of the independent and identically distributed variables ε_t, as follows: $Y_t = \alpha(L)^{-1}\Delta^{\delta/\sqrt{T}}\varepsilon_t$, $X_t = [\alpha(L)J(L)]\Delta y_t = J(L)\Delta^{\delta/\sqrt{T}}\varepsilon_t$, and $Z_{t,k} = \alpha(L)^{-1}\Delta^{\delta/\sqrt{T}}L^k\varepsilon_t$, $k = 1, \ldots, p$.

Next we obtain the residuals Y_t^* and X_t^* of projecting Y_t and X_t, respectively, onto the vector Z_t. It is simple to show that $Y_t^* = \Delta^{\delta/\sqrt{T}}\varepsilon_t$ plus a term due to the estimation of the projection on Z_t that contributes to the drift of t_{ϕ} at a smaller order of magnitude because it is orthogonal to the residuals X_t^*. To study X_t^*, notice that

$$\begin{aligned}\operatorname*{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}X_t Z_{t,k} &= E[J(L)\varepsilon_t \cdot \alpha(L)^{-1}\varepsilon_{t-k}] \\ &= \sum_{j=k}^{\infty} j^{-1}c_{j-k} = \kappa_k \qquad (k = 1, \ldots, p),\end{aligned}$$

whereas

$$\operatorname*{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}Z_{t,k}Z_{t,j} = E[\alpha(L)^{-1}\varepsilon_{t-k} \cdot \alpha(L)^{-1}\varepsilon_{t-j}]$$

$$= \sum_{t=0}^{\infty} c_t c_{t+|k-j|} = \Phi_{k,j} \qquad (k, j = 1, \ldots, p).$$

Then the (population) least squares projection coefficients of X_t onto Z_t are given by $\Phi^{-1}\kappa$ and, therefore, $X_t^* = J(L)\varepsilon_t - \kappa'\Phi^{-1}\alpha(L)^{-1}\varepsilon_{t,p}$, $\varepsilon_{t,p} = (\varepsilon_{t-1}, \ldots, \varepsilon_{t-p})'$, plus smaller order terms. Next we have that

$$\operatorname*{plim}_{T\to\infty} \sqrt{T}\frac{1}{T}\sum_{t=1}^{T} Y_t^* X_t^* = E\big[-\delta J(L)\varepsilon_t \cdot \{J(L)\varepsilon_t - \kappa'\Phi^{-1}\alpha(L)^{-1}\varepsilon_{t,p}\}\big]$$

$$= -\delta\left(\sum_{j=1}^{\infty} j^{-2} - \kappa'\Phi^{-1}\kappa\right) = -\delta\omega^2$$

and also $\operatorname{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T}(Y_t^*)^2 = \operatorname{Var}[\varepsilon_t] = 1$. Therefore, $\operatorname{plim}_{T\to\infty} T^{-1} \times \sum_{t=1}^{T}(X_t^*)^2$ is given by

$$\begin{aligned}
&\operatorname{Var}(J(L)\varepsilon_t - \kappa'\Phi^{-1}\alpha(L)^{-1}\varepsilon_{t,p}) \\
&\quad = \operatorname{Var}(J(L)\varepsilon_t) + \operatorname{Var}(\kappa'\Phi^{-1}\alpha(L)^{-1}\varepsilon_{t,p}) \\
&\qquad - 2\operatorname{Cov}(J(L)\varepsilon_t, \kappa'\Phi^{-1}\alpha(L)^{-1}\varepsilon_{t,p}) \\
&\quad = \pi^2/6 + \kappa'\Phi^{-1}\kappa - 2\kappa'\Phi^{-1}\kappa = \omega^2,
\end{aligned}$$

so that the drift of t_ϕ is given by $-\delta\omega$ and the theorem follows.

Centro de Investigación Económica, Instituto Tecnológico Autónomo de México, Av. Camino Sta. Teresa 930, México D.F. 10700, Mexico; ilobato@itam.mx

and

Departamento de Economía, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe (Madrid), Spain; carlos.velasco@uc3m.es.

REFERENCES

DELGADO, M. A., AND C. VELASCO (2005): "Sign Tests for Long Memory Time Series," *Journal of Econometrics*, 128, 215–251. [2]

ROBINSON, P. M., AND J. HUALDE (2003): "Cointegration in Fractional Systems with Unknown Integration Orders," *Econometrica*, 71, 1727–1766. [2,4,5]