Working Paper 96-07
Statistics and Econometrics Series 03
January 1996

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

ON THE ROBUSTNESS OF COINTEGRATION TESTS WHEN SERIES ARE FRACTIONALLY INTEGRATED

Jesús Gonzalo and Tae-Hwy Lee*

Abstract

This paper shows, analytically and numerically, the effects of a misspecification in the degree of integration on testing for cointegration. Johansen LR tests tend to find too much spurious cointegration while the Engle-Granger test shows a more robust performance than the LR tests.

Key Words
Cointegration; fractional unit roots; size of test; Johansen LR test; EG test

*Gonzalo, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid; Lee, Department of Economics, University of California. JEL classification: C32; C22.

1. Introduction

It is well established that many economic series contain dominant, smooth components, even after removal of simple deterministic trends. A stochastic process with no deterministic components is defined to be integrated of order d, denoted I(d), if it has a stationary and invertible ARMA representation after applying the differencing operator $(1 - B)^d$. The components of the vector X_t are said to be cointegrated of order (d, b), if all of components of X_t are I(d) and there exists a vector α ($\neq 0$) such that $\alpha' X_t$ is $I(d - b)$, $b > 0$. Usually the case with $d = b = 1$ is considered.

This paper is concerned with the robustness of testing for cointegration when series are not exactly I(1) but fractionally integrated with $d \neq 1$. When d is not an integer, the series is said to be fractionally integrated and discussed by Granger and Joyeux (1980) and Hosking (1981). The model has been used in applied econometric work where flexible characterization of low frequency dynamics is important, and has yielded new empirical regularities and insights into understanding behavior of many economic time series. There is considerable evidence on long memory properties in macroeconomic and financial time series data such as GDP, interest rate spreads, inflation rates, forward premiums, (power transformations of) stock returns, and exchange rates. See Baillie (1995) for a survey.

We investigate two methods to test for cointegration when the series are fractionally integrated. One method is the one suggested by Engle and Granger (1987, EG hereafter), which looks for a linear combination of level series that minimizes the variance of the linear combination using OLS. Second method is Johansen's (1988, 1991) procedure, which maximizes the canonical correlation between the first differenced series and the level series. The main assumption of the both tests is that series are exactly I(1). When variables are I(d) with $d \neq 1$ but if they are wrongly assumed as I(1), it is found that the Johansen tests tend to find cointegration too often.

Consider a 2 × 1 I(1) vector $X_t = (y_t \; x_t)'$. Since the variance of an I(1) series goes to infinity as t goes to infinity while variance of an I(0) series is finite, if an I(1) vector X_t is cointegrated, there exists a vector α ($\neq 0$) such that the variance of $\alpha' X_t$ is finite. EG suggest to test for a unit root in the residual z_t from the OLS cointegration regression $y_t = \hat{a}_0 + \hat{a}_1 x_t + z_t$. The EG test is based on the Dickey-Fuller (1979, DF) statistic augmented with k lagged differences, the t-value for $\hat{\rho}$, from the OLS regression

$$\Delta z_t = \rho z_{t-1} + \rho_1 \Delta z_{t-1} + \ldots + \rho_k \Delta z_{t-k} + \text{error}. \tag{1}$$

An I(1) series and an I(0) series are not correlated asymptotically. The Johansen method exploits the fact that $\alpha' X_t$ may be correlated with ΔX_t, and look for α that maximizes the canonical correlation between ΔX_t and X_t. If the maximum canonical correlation between ΔX_t and X_t is not zero, $\alpha' X_t$ is I(0) and X_t is cointegrated. A cointegrated 2 × 1 vector X_t can be generated from a vector error correction model (VECM)

$$\Delta X_t = \mu + \Pi X_{t-1} + \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_k \Delta X_{t-k} + \varepsilon_t \tag{2}$$

where $\varepsilon_t = (e_{1t} \; e_{2t})'$ is a 2 × 1 Gaussian vector white noise. If X_t is cointegrated, it can be shown that $\Pi = \gamma\alpha'$ where α and γ are 2 × r matrices with r being the rank of Π. Testing for cointegration is equivalent to testing for the hypothesis on the rank of Π. Γ_i's can be eliminated by regressing ΔX_t and X_{t-1} on ΔX_{t-1}, ... , ΔX_{t-k}, and 1. This can give residuals R_{0t} and R_{1t}, respectively. Let $S_{ij} = T^{-1}\Sigma_{t=1}^{T} R_{it} R_{jt}'$ (i, j = 0, 1). The two squared canonical correlations ($1 \geq \hat{\lambda}_1 \geq \hat{\lambda}_2 \geq 0$) between R_{0t} and R_{1t} can be estimated from the eigenvalues of $\hat{M} \equiv S_{11}^{-1} S_{10} S_{00}^{-1} S_{01}$. The Johansen method is designed to test if the maximum squared canonical correlation λ_1 can be positive, which is zero if X_t is not cointegrated.

The likelihood ratio statistics for testing for the null hypothesis H_0: $r = 0$ are

$$Q_1 = - T \ln(1 - \hat{\lambda}_1)(1 - \hat{\lambda}_2) \tag{3a}$$

and

$$Q_2 = - T \ln(1 - \hat{\lambda}_1). \tag{3b}$$

Q_1 is to test the null hypothesis against the alternative hypothesis H_1: $r > 0$, and Q_2 is to test the null hypothesis against H_1: $r = 1$. Note that $Q_2 \geq T\hat{\lambda}_1$ and $Q_1 \geq T(\hat{\lambda}_1 + \hat{\lambda}_2)$.

2. Fractionally integrated processes

Suppose $X_t = (y_t \; x_t)'$ are generated from

$$(1 - B)^d y_t = e_{1t} \tag{4a}$$

$$(1 - B)^d x_t = e_{2t}. \tag{4b}$$

An I(d) process y_t generated from (4a) has the following properties. (a) y_t is covariance stationary if $d < 0.5$. (b) y_t has an invertible moving average representation if $d > - 0.5$. (c) y_t is mean-reverting when $d < 1$. (d) If $d > 0$, y_t has long-memory, the autocovariances of y_t are not absolutely summable, and the power spectrum of y_t is unbounded for frequencies approaching zero. (e) y_t has an infinite variance when $d > 0.5$. (f) The DF-t statistic diverges to $-\infty$ if $d < 1$ and diverges to $+\infty$ if $d > 1$ as $T \rightarrow \infty$ (Sowell, 1990, Theorem 4). Thus if $d > 1$, the DF tests has no power asymptotically to test for a unit root against a stationary root.

The fractional difference operator $(1 - B)^d$ is defined by its Maclaurin series

$$(1 - B)^{d-1} = \sum_{j=0}^{\infty} \frac{\Gamma(-d+j)}{\Gamma(-d)\Gamma(j+1)} B^j = \sum_{j=0}^{\infty} d_j B^j, \quad d_j = \frac{j-1-d}{j} d_{j-1}, \quad d_0 = 1 \tag{5}$$

where $\Gamma(\cdot)$ is the gamma function, and d_j are squared summable if $d > -0.5$. Since $(1 - B)^{d-1}\Delta y_t = e_{1t}$, it can be written

$$\Delta y_t = \Sigma_{j=1}^{\infty} a_j \Delta y_{t-j} + e_{1t} \tag{6}$$

where $a_j = -\Gamma(j - d + 1)/[\Gamma(j + 1)\Gamma(-d + 1)]$. We may then write a VECM of the form

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \sum_{j=1}^{\infty} \begin{bmatrix} a_j & 0 \\ 0 & a_j \end{bmatrix} \begin{bmatrix} \Delta y_{t-j} \\ \Delta x_{t-j} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}. \tag{7}$$

For simplicity, in the following proposition, we consider only the case with $k = 0$ in the VECM to compute the LR statistics. Any finite k that is not sufficiently large enough to make the error ε_t a vector white noise will lead to the same results.

Proposition: *Suppose* $(y_t\ x_t)'$ *are* I(d) *processes generated from* (4), *and we estimate a VECM with* $k = 0$.

a. If $d \geq 1.5$, *then* $\hat{\lambda}_1 \xrightarrow{p} 0$ *as* $T \to \infty$.

b. If $1 < d < 1.5$, *then* $T^{(3-2d)}\hat{\lambda}_1 = O_p(1)$.

c. If $d = 1$, *then* $T\hat{\lambda}_1 = O_p(1)$.

d. If $0.5 < d < 1$, *then* $T^{(2d-1)}\hat{\lambda}_1 = O_p(1)$.

Proof: (a) Following Gourieroux et al (1989), for $d \geq 1.5$, $S_{11} = O_p(T^{2d-1})$, $S_{00} = O_p(T^{2d-3})$, and $S_{10} = O_p(T^{2d-2})$. Therefore $\hat{M} = S_{11}^{-1}S_{10}S_{00}^{-1}S_{01} = O_p(1)$, and the result

follows. (b) For $1 < d < 1.5$, $S_{11} = O_p(T^{2d-1})$, $S_{00} = O_p(1)$, and $S_{10} = O_p(T^{2d-2})$. Therefore $\hat{M} = O_p(T^{2d-3})$. (c) For $d = 1$, $S_{11} = O_p(T)$, $S_{10} = O_p(1)$, $S_{00} = O_p(1)$, and therefore $\hat{M} = O_p(T^{-1})$. (d) For $0.5 < d < 1$, $S_{11} = O_p(T^{2d-1})$, $S_{00} = O_p(1)$, and $S_{10} = O_p(1)$. Therefore $\hat{M} = O_p(T^{1-2d})$, and the result follows.

If $d = 1$, then $T\hat{\lambda}_1 = O_p(1)$. If $d \neq 1$, $T\hat{\lambda}_1 \xrightarrow{p} \infty$ as $T \to \infty$, and the size of the LR tests increases to one as $T \to \infty$ because $Q_2 \geq T\hat{\lambda}_1$ and $Q_1 \geq T(\hat{\lambda}_1 + \hat{\lambda}_2)$. Note that if $1 < d < 1.5$ then $0 < 3 - 2d < 1$, and if $0.5 < d < 1$ then $0 < 2d - 1 < 1$. Thus in these two cases, $\hat{\lambda}_1 \xrightarrow{p} 0$ but at the rate slower than T so that $T\hat{\lambda}_1$ diverges, and therefore the size of the LR tests goes to one.

3. A monte carlo

We generate $X_t = (y_t \; x_t)'$ from (4) assuming that e_{1t} and e_{2t} are i.i.d. N(0, 1), where $t = 1, \ldots, T + q$. The first $q = 2000$ observations are discarded. We approximated $(1 - B)^d = \Sigma_{j=0}^{\infty} d_j B^j$ by assuming $d_j = 0$ for $j > 1000$. In Tables 1 and 2, we report simulated size of the cointegration tests for various values of d.

When $d < 1$, the size is large for both EG and the Johansen tests. Although Sowell (1990) shows that the DF t statistic diverges to $-\infty$ if $d < 1$ as $T \to \infty$, Diebold and Rudebusch (1991a), and Hassler and Wolters (1994) provide evidence that the power of the DF tests are poor against fractional alternatives in finite samples. Table 3 also shows that it is hard to tell I(d) with $d < 1$ from the behavior of I(1) using the augmented DF (ADF) tests in a small sample. Thus if the variables are fractionally integrated with $d < 1$, then it is likely that we proceed assuming the series are I(1) and may get incorrect conclusion that the system has a long-run relationship.

When $d > 1$, the Johansen tests tend to find too much spurious cointegration while the EG test does not. Sowell (1990) shows that if $d > 1$ then the DF test has zero power

asymptotically to test for a unit root against a stationary root. Thus if the variables are fractionally integrated with $d > 1$, then it is likely that we proceed assuming they are I(1) (see Table 3), and the results may suggest cointegration incorrectly. The VECM (7) shows that the true number of lagged differences is infinity. If not enough k is used in computing the Johansen LR tests, all the omitted dynamics will be included in the residual ε_t and the model suffers dynamic misspecification. We thus report the results computed with $k = 3$ and 9, but the problem remains even with $k = 9$.

What are the relevant values of d for economic variables ? For example, the fractional differencing parameter d estimated in the literature are: $d = 1.17$ for annual disposable income (Diebold and Rudebusch (1991b)), $d = 1.29$ for quarterly real GNP (Sowell (1992)), and d is ranged from 1.04 to 1.36 for various nominal spot exchange rates (Cheung (1993)). Also d is estimated about 0.6 for money growth rates (Tieslau (1991)), and d is ranged from 0.40 to 0.57 for inflation rates in several developed countries (Hassler and Wolters(1995)). This indicates money stock and price series may have d greater than one.

4. Conclusions

This paper shows that when variables are fractionally integrated with $d > 1$, the Johansen tests tend to find too much spurious cointegration while a standard Dickey-Fuller test on the residual of an Engle-Granger regression does not. If the both tests (EG and LR) produce the same results in testing for cointegration, it is likely that $d \leq 1$. If they yield different answers, then it may indicate that $d > 1$.

References

Baillie, R.T. (1995), "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, forthcoming.

Cheung, Y.W. (1993), "Long Memory in Foreign Exchange Rates," *Journal of Business and Economic Statistics*, **11**, 93-101.

Cheung, Y.W. and K.S. Lai (1993), "A Fractional Cointegration Analysis of Purchasing Power Parity," *Journal of Business and Economic Statistics*, **11**, 103-112.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of American Statistical Association*, **74**, 427-31.

Diebold, F.X. and G.D. Rudebusch (1991a), "On the Power of Dickey-Fuller Tests against Fractional Alternatives," *Economics Letters*, **35**, 155-160.

Diebold, F.X. and G.D. Rudebusch (1991b), "Is Consumption Too Smooth ? Long Memory and the Deaton Paradox," *Review of Economics and Statistics*, **71**, 1-9.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, **55**, 251-276.

Gourieroux, C., F. Maurel, and A. Monfort (1989), "Least Squares and Fractionally Integrated Regressors," INSEE, Working Paper No. 8913.

Granger, C.W.J. and R. Joyeux (1980), "An Introduction to Long Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, **1**, 15-29.

Hassler, U. and J. Wolters (1994), "On the Power of Unit Root Tests Against Fractional Alternatives," *Economics Letters*, **45**, 1-5.

Hassler, U. and J. Wolters (1995), "Long Memory in Inflation Rates: International Evidence," *Journal of Business and Economic Statistics*, **13**, 37-45.

Hosking, J.R.M. (1981), "Fractional Differencing," *Biometrika*, **68**, 165-176.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, **12**, 231-254.

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, **59**, 1551-1580.

Sowell, F.B. (1990), "The Fractional Unit Root Distribution," *Econometrica*, **58**, 495-505.

Sowell, F.B. (1992), "Modeling Long Run Behavior with the Fractional ARIMA Model," *Journal of Monetary Economics*, **29**, 277-302.

Tieslau, M.A. (1991), "Long Memory Models and Macroeconomic Time Series," Michigan State University, Working Paper, No. 9005.

Table 1. Fractionally integrated processes (T = 100)

	k = 0			k = 3			k = 9		
	EG	Q_1	Q_2	EG	Q_1	Q_2	EG	Q_1	Q_2
d=0.5	.986	1.000	.997	.361	.651	.432	.083	.233	.171
d=0.6	.900	.991	.947	.223	.406	.262	.060	.180	.141
d=0.7	.636	.832	.652	.146	.224	.167	.043	.152	.132
d=0.8	.341	.447	.340	.094	.142	.105	.042	.129	.129
d=0.9	.143	.147	.132	.065	.087	.076	.038	.132	.122
d=1.0	.047	.048	.052	.048	.068	.072	.032	.133	.130
d=1.1	.016	.058	.069	.037	.077	.078	.034	.153	.149
d=1.2	.010	.152	.166	.028	.090	.085	.036	.163	.165
d=1.3	.031	.337	.351	.021	.105	.113	.030	.171	.171
d=1.4	.053	.563	.592	.024	.110	.128	.027	.188	.184
d=1.5	.076	.753	.774	.019	.135	.144	.032	.221	.206
d=1.6	.100	.860	.873	.016	.165	.185	.037	.236	.230
d=1.7	.111	.921	.931	.012	.185	.196	.048	.274	.267
d=1.8	.135	.952	.957	.016	.191	.213	.045	.312	.294
d=1.9	.150	.969	.973	.029	.231	.243	.053	.357	.336
d=2.0	.176	.982	.981	.039	.271	.254	.054	.400	.385

The frequency of rejecting the null hypothesis in 1000 replications is reported at the 5% level. The critical values for T = 100 are simulated from 90000 replications using the DGP with d = 1.

Table 2. Fractionally integrated processes (T = 1000)

	k = 0			k = 3			k = 9		
	EG	Q_1	Q_2	EG	Q_1	Q_2	EG	Q_1	Q_2
d=0.5	1.000	1.000	1.000	1.000	1.000	1.000	.984	1.000	1.000
d=0.6	1.000	1.000	1.000	.994	1.000	1.000	.851	.992	.936
d=0.7	.998	1.000	1.000	.879	.984	.941	.572	.815	.648
d=0.8	.872	.978	.946	.535	.708	.565	.269	.370	.284
d=0.9	.384	.455	.376	.193	.213	.186	.134	.140	.116
d=1.0	.063	.056	.056	.059	.055	.055	.059	.053	.062
d=1.1	.020	.118	.124	.026	.057	.066	.037	.051	.058
d=1.2	.047	.439	.464	.026	.160	.175	.023	.078	.104
d=1.3	.102	.768	.805	.029	.304	.332	.023	.141	.154
d=1.4	.166	.925	.925	.040	.441	.483	.021	.210	.248
d=1.5	.211	.974	.977	.030	.522	.546	.018	.248	.283
d=1.6	.253	.991	.990	.017	.543	.573	.014	.257	.281
d=1.7	.302	.994	.995	.010	.483	.517	.008	.245	.258
d=1.8	.331	.999	1.000	.004	.380	.390	.009	.216	.215
d=1.9	.345	.999	.999	.007	.245	.256	.015	.196	.199
d=2.0	.349	.999	.999	.029	.204	.192	.034	.208	.196

The frequency of rejecting the null hypothesis in 1000 replications is reported at the 5% level. The critical values for T = 1000 are simulated from 90000 replications using the DGP with d = 1.

Table 3. Power of ADF tests when a series is I(d)

	T = 100			T = 1000		
	ADF(0)	ADF(3)	ADF(p_{aic}) [mean(p_{aic}), sd(p_{aic})]	ADF(0)	ADF(3)	ADF(p_{aic}) [mean(p_{aic}), sd(p_{aic})]
d=0.5	.999	.553	.696 [1.875, 2.805]	1.000	1.000	.963 [7.937, 3.997]
d=0.6	.941	.355	.556 [1.926, 2.874]	1.000	.998	.867 [7.647, 3.860]
d=0.7	.691	.223	.399 [1.811, 2.925]	.999	.887	.674 [6.602, 3.606]
d=0.8	.354	.129	.258 [1.500, 2.807]	.825	.521	.412 [4.981, 3.148]
d=0.9	.141	.080	.117 [1.258, 2.805]	.331	.177	.187 [2.959, 2.705]
d=1.0	.047	.055	.069 [1.229, 2.842]	.055	.055	.050 [1.021, 2.400]
d=1.1	.032	.049	.058 [1.513, 2.846]	.032	.026	.030 [3.368, 2.993]
d=1.2	.047	.038	.047 [2.150, 3.013]	.104	.036	.042 [5.701, 3.394]
d=1.3	.087	.041	.056 [2.605, 2.983]	.205	.071	.050 [7.322, 3.598]
d=1.4	.148	.047	.067 [2.832, 2.916]	.285	.112	.064 [8.124, 3.651]
d=1.5	.214	.053	.067 [2.842, 2.884]	.330	.133	.068 [8.452, 3.688]
d=1.6	.257	.056	.080 [2.706, 2.759]	.368	.138	.066 [8.045, 3.624]
d=1.7	.317	.052	.079 [2.563, 2.674]	.383	.126	.069 [7.081, 3.372]
d=1.8	.342	.049	.077 [2.339, 2.646]	.388	.095	.062 [5.698, 3.068]
d=1.9	.366	.052	.062 [2.207, 2.706]	.405	.065	.052 [3.813, 2.688]
d=2.0	.387	.050	.061 [2.301, 2.817]	.398	.055	.061 [2.017, 2.359]

5% level. 1000 replications. ADF(p) denotes the DF tests augmented with p lagged first differences. p = 0, 3, or p_{aic}. p_{aic} is chosen using the AIC among p = 0, 1, ... , 19. When $p = p_{aic}$ is used, the mean and the standard deviation of p_{aic} in 1000 replications are reported in brackets, [mean(p_{aic}), sd(p_{aic})].