Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

On the robustness of cointegration tests when series are fractionally intergrated

Jesús Gonzalo [a] & Tae-Hwy Lee [b]
[a] Department of Statistics and Econometrics , Universidad Carlos III de Madrid , Spain
[b] Department of Economics , University of California , Riverside, USA
Published online: 02 Aug 2010.

To cite this article: Jesús Gonzalo & Tae-Hwy Lee (2000) On the robustness of cointegration tests when series are fractionally intergrated, Journal of Applied Statistics, 27:7, 821-827, DOI: 10.1080/02664760050120515

To link to this article: http://dx.doi.org/10.1080/02664760050120515

Journal of Applied Statistics, Vol. 27, No. 7, 2000, 821–827

On the robustness of cointegration tests when series are fractionally integrated

JESÚS GONZALO[1] **& TAE-HWY LEE**[2], [1]*Department of Statistics and Econometrics, Universidad Carlos III de Madrid, Spain and* [2]*Department of Economics, University of California, Riverside, USA*

ABSTRACT *This paper shows that when series are fractionally integrated, but unit root tests wrongly indicate that they are I(1), Johansen likelihood ratio (LR) tests tend to find too much spurious cointegration, while the Engle–Granger test presents a more robust performance. This result holds asymptotically as well as infinite samples. The different performance of these two methods is due to the fact that they are based on different principles. The Johansen procedure is based on maximizing correlations (canonical correlation) while Engle–Granger minimizes variances (in the spirit of principal components).*

1 Introduction

It is well established that many economic series contain dominant, smooth components, even after the removal of simple deterministic trends. A stochastic process with no deterministic components is defined to be integrated of order d, denoted $\mathrm{I}(d)$, 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 components of X_t are $\mathrm{I}(d)$ and there exists a vector $\alpha(\neq 0)$ such that $\alpha' X_t$ is $\mathrm{I}(d-b)$, $b > 0$. Usually the case with $d = b = 1$ is considered (for more detail see Granger, 1981; Engle & Granger, 1991).

When d is not an integer, the series are said to be fractionally integrated (Granger & Joyeux, 1980; Hosking, 1981). There is considerable evidence that the long memory properties of macroeconomic and financial time series data such as GDP, interest rate spreads, inflation rates, forward premiums, stock returns, exchange

Correspondence: Jesús Gonzalo, Department of Statistics and Econometrics, U. Carlos III de Madrid, 28903 Getafe, Madrid, Spain. E-mail: jgonzalo@elrond.uc3m.es

ISSN 0266-4763 print; 1360-0532 online/00/070821-07

rates, and etc, can be well captured by fractional integrated processes. This paper is concerned with the robustness of cointegration tests when series are fractionally integrated, but based on unit root tests we wrongly consider them as I(1) series.

We investigate two methods to test for cointegration. One method is the one suggested by Engle & Granger (1987, EG hereafter), which looks for a linear combination of level series that minimizes the variance of the linear combination using OLS. Another method is Johansen's (1995) procedure, which maximizes the canonical correlation between the first differenced series and the level series. From the point of view of multivariate analysis, the EG procedure is similar to principal components, while Johansen's method is a canonical correlations technique. The main assumption of both tests is that series are exactly I(1). When series are I(d) with $d \neq 1$, but we wrongly consider them as I(1), this paper finds that Johansen tests tend to find spurious cointegration more often than the EG test does. This result holds asymptotically as well as in finite samples.

Consider a (2×1) I(1) vector $X_t = (y_t x_t)'$. The variance of an I(1) series (given some initial conditions) goes to infinity as $t \to \infty$, while the variance of an I(0) series is finite. Therefore if an I(1) vector X_t is cointegrated, there must exist a vector $\alpha (\neq 0)$ such that the variance of $\alpha' X_t$ is finite. Based on this, EG suggest testing for a unit root on the residuals 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 augmented Dickey–Fuller (DF) statistic (see Dickey & Fuller, 1979) of order k, ADF(k), that is the t-value for $\hat{\rho}$ in the OLS regression

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

Reduced rank regression methods, like the Johansen approach, exploit the fact that I(1) and I(0) variables are asymptotically uncorrelated and look for a vector α that maximizes the correlation between $\alpha' X_t$ and a linear combination of ΔX_t. If that correlation is not zero, $\alpha' X_t$ is I(0) and X_t is cointegrated. According to Granger's representation theorem (Granger, 1983), a cointegrated system admits the following vector error correction model (VECM) representation

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

where $\varepsilon_t = (e_{1t} e_{2t})'$ is a white noise vector with finite variance. For simplicity we have eliminated all the deterministic components and we only consider a finite k in model (2). If X_t is cointegrated, it can be shown that the matrix Π can be decomposed into $\Pi = \gamma \alpha'$, where α and γ are (2×1) matrices. Testing for cointegration is therefore equivalent to testing the rank of Π (denoted as r) equal to one, and this is exactly what the Johansen method does. Formally the Johansen LR statistics for testing the null hypothesis of no cointegration $H_0 : r = 0$ are

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

and

$$Q_2 = -T \ln(1 - \hat{\lambda}_1) \tag{4}$$

where $(1 \geqslant \hat{\lambda}_1 \geqslant \hat{\lambda}_2 \geqslant 0)$ are the eigenvalues of $\hat{M} \equiv S_{11}^{-1} S_{10} S_{00}^{-1} S_{01}$, and $S_{ij} = T^{-1} \Sigma_{t=1}^{T} R_{it} R_{jt}' \ (i, j = 0, 1)$ are the product moment matrices of the residuals R_{0t} and R_{1t}, from the regressions of ΔX_t and X_{t-1} on the lagged differences, respectively. Q_1 tests the null hypothesis against the alternative hypothesis $H_1 : r > 0$, and Q_2 tests H_0 against $H_1 : r = 1$.

2 Fractionally integrated processes

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

$$(1 - B)^d y_t = e_{1t} \tag{5}$$

$$(1 - B)^d x_t = e_{2t} \tag{6}$$

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

$$(1 - B)^d = \sum_{j=0}^{\infty} \frac{\Gamma(-d + j)}{\Gamma(-d)\Gamma(j + 1)} B^j = \sum_{j=0}^{\infty} d_j B^j, \quad d_j = \frac{j - 1 - d}{j} d_{j-1}, \quad d_0 = 1 \tag{7}$$

where $\Gamma(\cdot)$ is the gamma function. An I(d) process y_t generated from (5) 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 a 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 if $d > 0.5$. (f) The DF-t statistic diverges to $-\infty$ if $d < 1$, and diverges to $+\infty$ if $d > 1$ as $T \to \infty$ (Sowell, 1990, Theorem 4). Thus, if $d > 1$, the standard DF tests have no power asymptotically.

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 white noise vector will lead to the same results.

Proposition:
Suppose $(y_t x_t)'$ are I(d) processes generated from (5)–(6), and we estimate a VECM with $k = 0$.

(a) If $d \geqslant 1.5$, then $\hat{\lambda}_1$ does not converge to zero in probability 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 $d \geqslant 1.5$, $S_{11} = O_p(T^{2d-1})$, $S_{00} = O_p(T^{2d-3})$, and $S_{10} = O_p(T^{2d-2})$. Therefore $\hat{M} = O_p(1)$ 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 $d = 1$, $S_{11} = O_p(T)$, $S_{00} = O_p(1)$, and $S_{10} = O_p(1)$. 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. This completes the proof.

If $d = 1$, $T\hat{\lambda}_1 = O_p(1)$. 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 $Q_2 \geqslant T\hat{\lambda}_1$ and $Q_1 \geqslant T(\hat{\lambda}_1 + \hat{\lambda}_2)$. 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 a slower rate than T so that $T\hat{\lambda}_1$ diverges, and therefore the size of the LR tests goes to one asymptotically.

A sufficiently large k such that the residuals are white noise may solve the problem. But there are many situations in macroeconomics where it is not possible, in practice, to try a large k. As previously mentioned, our Proposition will hold not only for $k = 0$ but also for any $k > 0$ not sufficiently large to make the error a white noise vector.

3 Monte Carlo results

We generate $X_t = (y_t x_t)'$ from (5)–(6) where e_{1t} and e_{2t} are i.i.d. N(0, 1), and $E(\varepsilon_{1i}\varepsilon_{2j}) = 0$ for every i and j. In order to avoid the initial conditions ($x_0 = 0$, $y_0 = 0$) effect, we generate samples of sizes $t = 1, \ldots, T+q$ and discard the first $q = 2000$ observations. We approximated $(1-B)^d = \Sigma_{j=0}^{\infty} d_j B^j$ by assuming $d_j = 0$ for $j > 1000$. It is clear that both variables are not cointegrated in any sense. In Tables 1 and 2, we report the size of the cointegration tests for various values of d.

TABLE 1. Size of cointegration tests ($T = 100$)

	$k=0$			$k=3$			$k=9$		
	EG	Q_1	Q_2	EG	Q_1	Q_2	EG	Q_1	Q_2
$d=0.5$	0.986	1.000	0.997	0.361	0.651	0.432	0.083	0.233	0.171
$d=0.6$	0.900	0.991	0.947	0.223	0.406	0.262	0.060	0.180	0.141
$d=0.7$	0.636	0.832	0.652	0.146	0.224	0.167	0.043	0.152	0.132
$d=0.8$	0.341	0.447	0.340	0.094	0.142	0.105	0.042	0.129	0.129
$d=0.9$	0.143	0.147	0.132	0.065	0.087	0.076	0.038	0.132	0.122
$d=1.0$	0.047	0.048	0.052	0.048	0.068	0.072	0.032	0.133	0.130
$d=1.1$	0.016	0.058	0.069	0.037	0.077	0.078	0.034	0.153	0.149
$d=1.2$	0.010	0.152	0.166	0.028	0.090	0.085	0.036	0.163	0.165
$d=1.3$	0.031	0.337	0.351	0.021	0.105	0.113	0.030	0.171	0.171
$d=1.4$	0.053	0.563	0.592	0.024	0.110	0.128	0.027	0.188	0.184
$d=1.5$	0.076	0.753	0.774	0.019	0.135	0.144	0.032	0.221	0.206
$d=1.6$	0.100	0.860	0.873	0.016	0.165	0.185	0.037	0.236	0.230
$d=1.7$	0.111	0.921	0.931	0.012	0.185	0.196	0.048	0.274	0.267
$d=1.8$	0.135	0.952	0.957	0.016	0.191	0.213	0.045	0.312	0.294
$d=1.9$	0.150	0.969	0.973	0.029	0.231	0.243	0.053	0.357	0.336
$d=2.0$	0.176	0.982	0.981	0.039	0.271	0.254	0.054	0.400	0.385

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 90 000 replications using the DGP with $d = 1$.

TABLE 2. Size of cointegration tests ($T = 1000$)

	$k=0$			$k=3$			$k=9$		
	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	0.984	1.000	1.000
$d=0.6$	1.000	1.000	1.000	0.994	1.000	1.000	0.851	0.992	0.936
$d=0.7$	0.998	1.000	1.000	0.879	0.984	0.941	0.572	0.815	0.648
$d=0.8$	0.872	0.978	0.946	0.535	0.708	0.565	0.269	0.370	0.284
$d=0.9$	0.384	0.455	0.376	0.193	0.213	0.186	0.134	0.140	0.116
$d=1.0$	0.063	0.056	0.056	0.059	0.055	0.055	0.059	0.053	0.062
$d=1.1$	0.020	0.118	0.124	0.026	0.057	0.066	0.037	0.051	0.058
$d=1.2$	0.047	0.439	0.464	0.026	0.160	0.175	0.023	0.078	0.104
$d=1.3$	0.102	0.768	0.805	0.029	0.304	0.332	0.023	0.141	0.154
$d=1.4$	0.166	0.925	0.925	0.040	0.441	0.483	0.021	0.210	0.248
$d=1.5$	0.211	0.974	0.977	0.030	0.522	0.546	0.018	0.248	0.283
$d=1.6$	0.253	0.991	0.990	0.017	0.543	0.573	0.014	0.257	0.281
$d=1.7$	0.302	0.994	0.995	0.010	0.483	0.517	0.008	0.245	0.258
$d=1.8$	0.331	0.999	1.000	0.004	0.380	0.390	0.009	0.216	0.215
$d=1.9$	0.345	0.999	0.999	0.007	0.245	0.256	0.015	0.196	0.199
$d=2.0$	0.349	0.999	0.999	0.029	0.204	0.192	0.034	0.208	0.196

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 90 000 replications using the DGP with $d = 1$.

TABLE 3. Power of ADF tests for a unit root

	$T = 100$					$T = 1000$				
	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})]
$d = 0.5$	0.999	0.553	0.696	1.875	2.805	1.000	1.000	0.963	7.937	3.997
$d = 0.6$	0.941	0.355	0.556	1.926	2.874	1.000	0.998	0.867	7.647	3.860
$d = 0.7$	0.691	0.223	0.399	1.811	2.925	0.999	0.887	0.674	6.602	3.606
$d = 0.8$	0.354	0.129	0.258	1.500	2.807	0.825	0.521	0.412	4.981	3.148
$d = 0.9$	0.141	0.080	0.117	1.258	2.805	0.331	0.177	0.187	2.959	2.705
$d = 1.0$	0.047	0.055	0.069	1.229	2.842	0.055	0.055	0.050	1.021	2.400
$d = 1.1$	0.032	0.049	0.058	1.513	2.846	0.032	0.026	0.030	3.368	2.993
$d = 1.2$	0.047	0.038	0.047	2.150	3.013	0.104	0.036	0.042	5.701	3.394
$d = 1.3$	0.087	0.041	0.056	2.605	2.983	0.205	0.071	0.050	7.322	3.598
$d = 1.4$	0.148	0.047	0.067	2.832	2.916	0.285	0.112	0.064	8.124	3.651
$d = 1.5$	0.214	0.053	0.067	2.842	2.884	0.330	0.133	0.068	8.452	3.688
$d = 1.6$	0.257	0.056	0.080	2.706	2.759	0.368	0.138	0.066	8.045	3.624
$d = 1.7$	0.317	0.052	0.079	2.563	2.674	0.383	0.126	0.069	7.081	3.372
$4 = 1.8$	0.342	0.049	0.077	2.339	2.646	0.388	0.095	0.062	5.698	3.068
$d = 1.9$	0.366	0.052	0.062	2.207	2.706	0.405	0.065	0.052	3.813	2.688
$d = 2.0$	0.387	0.050	0.061	2.301	2.817	0.398	0.055	0.061	2.017	2.359

5% level. 1000 replications. ADF(p) denotes the DF tests augmented with p lagged first differences. $p = 0,3$, or $p_{aic} \cdot p_{aic}$ is chosen using the AIC among $p = 0, 1, \ldots, 19$. 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})].

When $d < 1$, the size is large for both the EG and Johansen tests. These finite sample results match the asymptotic results. For the EG test the asymptotic behaviour is derived from Sowell (1990), where it is shown that the DF t statistic diverges to $-\infty$ if $d < 1$ as $T \to \infty$. For Johansen tests the theoretical result is in our Proposition.

When $d > 1$, Johansen LR tests tend to find too much spurious cointegration while the EG test does not. Again these finite sample results coincide with the asymptotic results. The performance of the EG test is derived from Sowell (1990), where it is shown that if $d > 1$, the DF test has zero power asymptotically. The asymptotic performance of Johansen tests is derived in our Proposition.

Table 3 shows how difficult it is to distinguish in finite samples an I($d, d > 0.5$) variable from an I(1) using the augmented DF (ADF) test. Thus, if the variables are fractionally integrated, it is likely that we will proceed assuming the series are I(1), and therefore get the incorrect conclusion that the variables are related in the long-run (i.e. cointegrated).

In order to avoid the spurious cointegration, one could think that a possible solution is to increase k with T, in a similar way to what Berk (1974) does for stationary and ergodic processes. We are not aware of any result in the literature on how to do this for non-stationary and non-ergodic processes. We suspect the problem must be complicated because the sum of absolute correlations for a fractional integrated process is not bounded, therefore any finite k will produce inconsistent estimates. Moreover, a fractionally integrated process with $d \geqslant 0.5$ is not ergodic. We report the results computed with $k = 3$ and 9, but the problem remains even in the latter case. Based on our Monte Carlo experiment we have to agree with Brockwell & Davis (1991, p. 520) when they say 'While a long memory process can always be approximated by an ARMA(p, q) the orders p and q required to achieve a reasonable good approximation may be so large as to make parameter estimation extremely difficult.'

How often do we have $d > 1$ in practice? Examples of values of estimates of d reported in the literature are: $d = 1.17$ for annual disposable income (Diebold & Rudebusch, 1991), $d = 1.29$ for quarterly real GNP (Sowell, 1992a), 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 is ranged from 0.40 to 0.57 for inflation rates in several developed countries (Hassler & Wolters, 1995), indicating money stock and price series may have d greater than one.

Another important and related question would be to see how precisely one could estimate d with the sample sizes used in applied studies. Several different approaches have been suggested for estimating d: Geweke & Porter-Hudak (1983) suggest a two step estimator from a regression of ordinates of the periodogram on a trigonometric function; Fox & Taqqu (1986) suggest an approximate ML procedure; Sowell (1992b) derives the full ML estimator, and Chung & Baillie (1993) consider the minimum conditional sum of squares estimator. Some simulation evidence on the finite sample performance of these methods has been provided by Agiakoglou *et al.* (1992), Cheung & Diebold (1994), and Chung & Baillie (1993). They show severe biases of these estimators. In our opinion this difficulty on estimating d gives even more relevance to the results obtained in this paper.

4 Conclusions

In applied research, once a pair of variables are considered to be I(1), the next step is to investigate if there exists a long-run equilibrium relationship between them. Because it is very difficult to distinguish an $I(d, d > 0.5)$ from an I(1), this paper shows that, asymptotically, as well as in finite samples, Johansen LR tests tend to find spurious cointegration more often than EG does. Therefore, a recommendation in order to detect this problem is to run both tests. If they produce different cointegration results, then proceed with a more exhaustive univariate analysis than a simple unit root test.

Acknowledgements

J. Gonzalo thanks financial support from the Spanish Secretary of Education (PB 98-0026). T. Lee thanks the Academic Senate of UCR for the research support.

REFERENCES

Agiakloglou, C., Newbold, P. & Wohar, M. (1992) Bias in an estimator of the fractional difference parameter, *Journal of Time Series Analysis*, 14, pp. 235–246.

Berk, K. (1974) Consistent autoregressive spectral estimates, *Annals of Statistics*, 2, pp. 489–502.

Brockwell, P. & Davis, R. A. (1991) *Time Series: Theory and Methods*, 2nd edn (Springer Verlag).

Cheung, Y. W. (1993) Long memory in foreign exchange rates, *Journal of Business and Economic Statistics*, 11, pp. 93–101.

Cheung, Y. W. & Diebold, F. X. (1994) On maximum likelihood estimation of the differencing parameter of fractionally integrated noise with unknown mean, *Journal of Econometrics*, 62, pp. 301–316.

CHEUNG, Y. W. & LAI, K. S. (1993) A fractional cointegration analysis of purchasing power parity, *Journal of Business and Economic Statistics*, 11, pp. 103–112.

CHUNG, C. F. & BAILLIE, R. T. (1993) Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA models, *Empirical Economics*, 18, pp. 791–806.

DICKEY, D. A. & FULLER, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root, *Journal of American Statistical Association*, 74, pp. 427–431.

DIEBOLD, F. X. & RUDEBUSCH, G. D. (1991) Is consumption too smooth? Long memory and the Deaton paradox, *Review of Economics and Statistics*, 71, pp. 1–9.

ENGLE, R. F. & GRANGER, C. W. J. (1987) Cointegration and error correction: representation, estimation, and testing, *Econometrica*, 55, pp. 251–276.

ENGLE, R. F. & GRANGER, C. W. J. (1991) *Long-Run Economic Relationships, Readings in Cointegration* (Oxford University Press).

FOX, R. & TAQQU, M. S. (1986) Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series, *Annals of Statistics*, 14, pp. 517–532.

GEWEKE, J. & PORTER-HUDAK, S. (1983) The estimation and application of long memory time series models, *Journal of Time Series Analysis*, 4, pp. 221–238.

GOURIEROUX, C., MAUREL, F. & MONFORT, A. (1989) Least squares and fractionally integrated regressors, INSEE, Working Paper No. 8913.

GRANGER, C. W. J. (1981) Some properties of time series data and their use in econometric model specification, *Journal of Econometrics*, 16, pp. 121–130.

GRANGER, C. W. J. (1983) Co-Integrated variables and error-correcting models, unpublished UCSD Discussion Paper 83-13.

GRANGER, C. W. J. & JOYEUX, R. (1980) An introduction to long memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1, pp. 15–29.

HASSLER, U. & WOLTERS, J. (1995) Long memory in inflation rates: international evidence, *Journal of Business and Economic Statistics*, 13, pp. 37–45.

HOSKING, J. R. M. (1981) Fractional differencing, *Biometrika*, 68, pp. 165–176.

JOHANSEN, S. (1995) *Likelihood-based Inference in Cointegrated Vector Autoregressive Models* (Oxford University Press).

SOWELL, F. B. (1990) The fractional unit root distribution, *Econometrica* 58, pp. 495–505.

SOWELL, F. B. (1992a) Modeling long run behavior with the fractional ARIMA model, *Journal of Monetary Economics*, 29, pp. 277–302.

SOWELL, F. B. (1992b) Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of Econometrics*, 53, pp. 165–188.

TIESLAU, M. A. (1991) Long memory models and macroeconomic time series, Michigan State University, Working Paper, No. 9005.