Fractional Integration and the Augmented Dickey–Fuller Test[1]

by

Walter Krämer
Fachbereich Statistik, Universität Dortmund
D - 44221 Dortmund, Germany

Abstract

This note shows that the Augmented Dickey–Fuller test is consistent against fractional alternatives if the order of the autoregression does not tend to infinity too fast.

Keywords: Dickey–Fuller Test, fractional alternatives.

JEL number: C 12

1 The Problem

This note is concerned with testing the hypothesis $H_0 : \beta = 1$ in the model

$$x_t = \beta x_{t-1} + \varepsilon_t \quad (t = 1, \ldots, N), \tag{1}$$

where the disturbances are stationary, but fractionally integrated, $\varepsilon_t \sim I(d)$ with $-\frac{1}{2} < d < \frac{1}{2}$. This situation arises for instance when testing the null hypothesis of no cointegration in a system with $I(d)$–residuals; one would then like

[1] Research supported by Deutsche Forschungsgemeinschaft, SFB 475; I am grateful to Ingolf Dittmann, Uwe Hassler and Jürgen Wolters for constructive criticism and comments.

to know whether residual–based tests are still able to reject the null hypothesis of no cointegration when the residuals are not $I(1)$, but do not have a conventional ARMA–representation either. Sowell (1990) proves the divergence of the standard Dickey–Fuller–test–statistic $N(\hat{\beta}-1)$ and the Dickey–Fuller t–statistic $(\hat{\beta}-1)/s_{\hat{\beta}}$ for the case where $\beta=1$, $x_0=0$, $(1-L)^d\varepsilon_t=u_t$, $u_t\sim iid(0,\sigma^2)$ and $E(|u_t|^r)<\infty$ for $r\geq max[4,-8d/(1+2d)]$ and Diebold and Rudebusch (1991) show via Monte–Carlo that these tests, although consistent, have little power in finite samples.

However, these results are of limited value for real world applications, where one almost always has to allow, under the null hypothesis, for some autocorrelation among the ε's in (1). Therefore the standard Dickey–Fuller test will rarely be appropriate, and there is automatically some implied interest in the power of the Augmented Dickey Fuller test, i.e. the conventional t–test of $H_0:\beta=1$ in

$$x_t=\beta x_{t-1}+\varphi_1\Delta x_{t-1}+\ldots+\varphi_p\Delta x_{t-p}+\varepsilon_{tp}, \tag{2}$$

where $p\to\infty$ as $N\to\infty$ (Said and Dickey 1984).

Hassler and Wolters (1994) show that the power of this test decreases quite drastically as p increases. They also conjecture that this test is not consistent against fractional alternatives, the rationale being that, as $p\to\infty$, the ε_{tp} in (2) are approaching the independent u_t's from $(1-L)^d\varepsilon_t=u_t$: From

$$(1-L)^{d+1}x_t=\sum_{j=0}^{\infty}d_jx_{t-j}\quad\text{and} \tag{3}$$

$$\sum_{j=0}^{\infty}d_j=0,\ d_0=1 \tag{4}$$

one deduces the following relationships for the coefficients and disturbances in (2):

$$\beta=-\sum_{j=1}^{p+1}d_j,\ \varphi_i=\sum_{j=i+1}^{p+1}d_j\quad\text{and} \tag{5}$$

$$\varepsilon_{tp} = u_t + \sum_{j=p+1}^{\infty} x_{t-j}. \tag{6}$$

Therefore, as $p \to \infty$, we have $\beta \to 1$ and $\varepsilon_{tp} \xrightarrow{p} u_t$, and one might expect that the t–test of $H_0 : \beta = 1$ in (2) behaves approximately as a standard t–test does (i.e. it does not reject with increasing probability).

This intuition can be misleading, however, as is shown below. If p does not tend to infinity too fast (e.g., $p = o(N^{\frac{1}{2}+d})$, the conventional t–statistic for $H_0 : \beta = 1$ in (2) still tends to infinity.

2 The limiting rejection probability of the Augmented Dickey–Fuller–Test

The test statistic of the conventional t–test of $H_0 : \beta = 1$ in (2) can be written as

$$ADF = \frac{x_{-1}' Q_p \varepsilon / \sqrt{N\sigma_N^2}}{\sqrt{S^2 x_{-1}' Q_q x_{-1} / \sqrt{N\sigma_N^2}}} \, , \tag{7}$$

where $x_{-1} = (x_0, \ldots, x_{N-1})'$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)'$, $Q_p = (I - X_p(X_p'X_p)^{-1}X_p')$, X_p is the matrix of observations on the p regressors $\Delta x_{t-1}, \ldots, \Delta x_{t-p}$,

$$S^2 \;\; = \;\; \frac{1}{N}\sum_{t=1}^{N}(x_t - \hat{\beta}x_{t-1} - \hat{\varphi}_1 \Delta x_{t-1} - \ldots - \hat{\varphi}_p \Delta x_{t-p})^2, \tag{8}$$

$\sigma_N^2 = var\left(\sum_{t=1}^{N} \varepsilon_t\right)$, and where the $\hat{\beta}$ and $\hat{\varphi}_i$ are the OLS–estimates of the regression coefficients in (2).

Consider first the denominator in (7). Along the lines of Sowell (1990, proof of theorem 4), it is easily seen that $S^2 \xrightarrow{p} E(\varepsilon_t^2)$. From

$$x_{-1}' Q_p x_{-1} = x_{-1}' x_{-1} - x_{-1}' P_p x_{-1}, \tag{9}$$

where $P_p = X_p(X_p'X_p)^{-1}X_p'$ and where

$$\begin{aligned} x_{-1}'x_{-1} &= O_p(N\sigma_N^2), \qquad (10) \\ x_{-1}'P_px_{-1} &= o_p(N\sigma_N^2), \qquad (11) \end{aligned}$$

we therefore deduce that $\sqrt{S^2x_{-1}'Q_px_{-1}/N\sigma_N^2}$ has the same nondegenerate limiting distribution as $\sqrt{S^2x_{-1}'x_{-1}/N\sigma_N^2}$ as $N \to \infty$, to be found e.g. in Sowell (1990, p. 505).

The crucial step here is equation (11), i.e. the fact that $x_{-1}'P_px_{-1}$ is stochastically of smaller order than $x_{-1}'x_{-1}$. To see this, note that $x_{-1}'P_px_{-1}$ is the explained sum of squares when regressing x_{-1} on $\Delta x_{-1}, \ldots, \Delta x_{-p}$, i.e. $x_{-1}'P_px_{-1} = \hat{\theta}'X_p'X_p\hat{\theta}$, where the p components of $\hat{\theta}$ are $O_p(1)$ and where the $p \times p$ components of $X_p'X_p$ are $O_p(N)$. Therefore,

$$x_{-1}'P_px_{-1} = O_p(p^2N), \qquad (12)$$

which in view of (10) is stochastically of a smaller order than $x_{-1}'x_{-1}$ whenever $p^2 = o_p(S_N^2)$, which in turn, noting $S_N = O(N^{\frac{1}{2}+d})$, is guaranteed whenever $p = o(N^{\frac{1}{2}+d})$.

The numerator in (7) can be expressed as

$$\frac{1}{\sqrt{N\sigma_N^2}}x_{-1}'Q_p\varepsilon = \frac{1}{\sqrt{N\sigma_N^2}}(x_{-1}'\varepsilon - x_{-1}'P_p\varepsilon), \qquad (13)$$

where again the second term on the right can be asymptotically neglected. In view of

$$x_{-1}'\varepsilon = \frac{1}{2}x_N^2 - \frac{1}{2}\sum_{t=1}^{N}\varepsilon_t^2 \qquad (14)$$

we can therefore reexpress the numerator as

$$\frac{\sigma_N}{\sqrt{N}}\left[\frac{X_n}{\sigma_N}\right]^2 - \frac{\sqrt{N}}{\sigma_N}\left[\frac{1}{N}\sum_{t=1}^{N}\varepsilon_t^2\right] + O_p(1), \qquad (15)$$

where the terms in brackets are $O_p(1)$ and where from $\sigma_N = O(N^{\frac{1}{2}+d})$, for $-\frac{1}{2} < d < \frac{1}{2}$, $d \neq 0$, one of the factors in front of the brackets tends to infinity as $N \to \infty$: for $d < 0$, $\sqrt{N}/\sigma_N \to \infty$, and for $d > 0$, $\sigma_N/\sqrt{N} \to \infty$, so the test statistic diverges irrespective of the particular value of d.

Whether or not the ADF is consistent therefore depends on d and lag length and on the alternative considered: if one follows Schwert (1989) by choosing $p = 0(N^{\frac{1}{4}})$, the test is inconsistant when $d < -\frac{1}{4}$. It is also inconsistant against the alternative $\beta > 1$ (i.e. when a one-sided rejection region $(-\infty, c]$ is used), as the test statistic then tends to $+\infty$. It is consistent for $p > 0(N^{\frac{1}{2}+d})$ when we use a two–sided test and for $p > 0(N^{\frac{1}{2}+d})$ plus $d < 0$ when the conventional one–sided version is employed.

References

Dickey, D.A. and Fuller, W.A. (1979): "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American Statistical Association* 74, 427 – 431.

Diebold, F.X. and Rudebusch, G. (1991): "On the power of Dickey–Fuller tests against fractional alternatives." *Economics Letters* 35, 155 – 160.

Hassler, U. and Wolters, J. (1994): "On the power of unit root tests against fractional alternatives." *Economics Letters* 45, 1 – 5.

Said, S.E. and Dickey, D.A. (1984): "Testing for unit roots in autoregressive moving average models of unknown order." *Biometrika* 71, 599 – 607.

Schwert, G. (1989): "Tests for unit roots: A Monte Carlo investigation." *Journal of Business and Economic Statistics* 7, 147 – 159.

Sowell, F. (1990): "The fractional unit root distribution." *Econometrica* 58, 495 – 505.