Testing for unit roots in the context of misspecified logarithmic random walks[1]

by

Walter Krämer and Laurie Davies

Fachbereich Statistik, Universität Dortmund, D-44221 Dortmund
Fachbereich Mathematik und Informatik, Universität Essen, D-45117 Essen
Germany

1 Introduction and Summary

Testing for unit roots has been among the most heavily researched topics in Econometrics for the last quarter of a century. Much less researched is the equally important issue of the appropriate transformation (if any) of the variable of interest which should preceed any such testing.

In macroeconometrics and empirical finance (stock prices, exchange rates), there are often compelling arguments in favor of a logarithmic transformation. Elsewhere, for instance in the modelling of interest rates, a levels specification automatically suggests itself. In many applications, however, it is not a priori clear, given that one suspects a unit root, whether this unit root is present in the levels or the logs, so there is certainly some interest in the testing for unit roots in the context of an incompletely specified nonlinear transformation of the data.

This issue can be approached from various angles. One is to check which transformations leave the $I(1)$–property of a time series intact, the presumption being that any such transformation could then do little damage to the null distribution of a test for unit roots (Granger and Hallmann 1991, Ermini and Granger 1993, Corradi 1995). A related one is to use tests whose null dis-

[1]Research supported by Deutsche Forschungsgemeinschaft via SFB 475

tribution is robust to monotonic transformations, whether the transformed data are $I(1)$ or not (Granger and Hallmann 1991, Burridge and Guerre 1996, Gourieroux and Breitung 1999), or to embed the levels and log specifications, respectively, in a general Box-Cox-framework and to estimate the transformation parameter before testing (Franses and McAleer 1998, Franses and Koop 1998, Kobayashi and McAleer 1999).

The present paper continues along the lines of Granger and Hallmann (1991) by focussing on a conventional test procedure — the standard Dickey-Fuller-test — and by investigating its properties under a misspecified nonlinear transformation (in particular: investigating whether an existing unit root is still detected, i.e. the null hypothesis of an existing unit root is not rejected when an inappropriate transformation is applied). Given that this test is widely employed, and given that the choice between a linear and a log linear specification is often rather haphazardly done in applications, it is important to know the degree to which the acceptance of the null hypothesis depends on the correctness of the data transformation.

Granger and Hallmann (1991) find through Monte Carlo that the standard Dickey-Fuller-test overrejects a correct null hypotheses of a random walk in the logs, when the test is instead applied to the levels. Below we prove analytically that the rejection probability can take arbitrary values between zero and one for any sample size. An analogous result obtains when the levels follow a random walk, but the Dickey-Fuller-test is applied to the logs. Again, the rejection probability is shown to be depend on both the sample size and the innovation variance, so the null distribution of the DF-test is extremely non-robust to improper data transformations.

2 The null distribution of the Dickey-Fuller-test as applied to levels when the random walk is in the logs

This section is concerned with the case where the true data generating process (DGP) is a random walk in the logs, but the Dickey-Fuller-test is inadvertantly applied to the levels of the data (i.e. to a multiplicative random walk). The common wisdom, as pronounced by Granger and Hallmann (1991) or ...., is that that the null distribution of the DF-test is "more spread out", inducing overrejections of the null hypothesis. Below we show that this is only partially true, as the rejection probability under H_0 depends crucially on the interplay between the innovation variance and the sample size.

Let y_t be the time series under study, let $z_t := ln(y_t)$, and assume that

$$z_t = \delta + z_{t-1} + \varepsilon_t,\ \varepsilon_t \sim nid(0, \sigma^2), \quad t = 1, \ldots, T, \quad z_0 = \sigma \times \text{ constant.} \tag{1}$$

The standard Dickey-Fuller-test, as applied to the levels $y_t := exp(z_t)$, is given by $T(\hat{\rho} - 1)$, where

$$\hat{\rho} = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2} \tag{2}$$

is the OLS-estimator for ρ in the model

$$y_t = \rho y_{t-y} + u_t. \tag{3}$$

It assumes that there is no drift component in the data. Its asymptotic critical values, which can by now be found in many textbooks, are $c = -5,6 (\alpha = 1\%)$, $c = -7,9 (\alpha = 5\%)$ and $c = -13,3 (\alpha = 10\%)$. They are based on the null hypothesis that in eq. (3), $\rho = 1$ and $u_t \sim iid(0, \sigma^2)$ (which, of course, is not quite correct if the true DGP is given by (1)).

The objects of our study are the true rejection probabilities under H_0, i.e. $P(T(\hat{\rho} - 1) < c)$. As the joint distribution of $(z_1, \ldots, z_T)$ and therefore also

$(y_1, \ldots, y_t)$ is uniquely determined by σ^2 (given the distribution of the ε's), these probabilities are functions of σ^2 and T. We consider the limits of these probabilities for $\sigma^2 \to 0$ and $\sigma^2 \to \infty$ (T fixed) and for $T \to \infty$ (σ^2 fixed).

To determine the limiting probabilities for a given sample size as $\sigma^2 \to 0$ or $\sigma^2 \to \infty$ we first consider the limiting behaviour of $\hat{\rho}$.

Theorem 1 *If the data are generated by (1), and $\hat{\rho}$ is the OLS-estimator for ρ in (2), we have:*

- *a) $\sigma^2 \to 0$ implies $\hat{\rho} \xrightarrow{d} 1$.*
- *b) $\sigma^2 \to \infty$ implies that $\hat{\rho}$ has a limiting distribution with mass $P(\max_{t=0,\ldots,T-1} y_t^2 > y_T y_{T-1})$ at zero and mass $P(\max_{t=0,\ldots,T-1} y_t^2 < y_T y_{T-1})$ at infinity.*

The proof of the theorem is in the appendix. It does not make any use of the iid–property of the innovations and holds for quite arbitrary joint distributions of $(\varepsilon_1, \ldots, \varepsilon_T)$.

Theorem 1 immediately gives the limiting rejection probability for the standard DW-test, for one sided-tests and arbitrary critical values less than zero: Rejection probabilities tend to zero as the innovation variance becomes small (as $T(\hat{\rho} - 1) \to 0$), and they tend to $P(\max_{t=0,\ldots,T-1} y_t^2 > y_T y_{T-1})$ as the innovation variance increases (assuming that the critical value is larger than $-T$), since $T(\hat{\rho} - 1) \to \infty$, whenever $max y_t^2 > y_T y_{T-1}$.

The probability for the latter event depends only on the sample size T. It is easily seen to tend to the probability that the maximum absolute value of the random walk $\{z_t\}$ exceeds its final value z_T, which in turn converges to ....

Given T, the true rejection probabilities approach their limits (as $\sigma \to \infty$) from below; they are an increasing function of σ, as shown in our Monte Carlo experiments in section 5. This is intuitively obvious, since ....

The next theorem gives the limiting rejection probabilities as $T \to \infty$ and σ^2 fixed.

....

3 The null distribution of the Dickey-Fuller-test as applied to logs when the random walk is in the levels

This section considers the case where the true DGP is given by (1), and the time series to which the Dickey–Fuller test is applied is given $y_t := ln(z_t)$. This type of misspecification appears to be less serious in practice, if only because of possible negative values of z_t. It can therefore occur only if z_0 is rather large and/or a sizable drift component (to be considered later) prevents this from happening.

The next theorem therefore assumes that $z_t > 0$ $(t = 0, \ldots, T)$; it considers the limiting behaviour of

$$\hat{\rho}(c) := \frac{\Sigma ln(cz_{t-1}) ln(cz_t)}{\Sigma [ln(cz_{t-1})]^2} \tag{4}$$

as $c \to 0$ and as $c \to \infty$.

Theorem 2 *If the data are generated by (1), we have* $\hat{\rho}(c) \xrightarrow{d} 1$*, both as* $c \to 0$ *and as* $c \to \infty$*.*

The proof is in the appendix.

Theorem 1 implies that the conditional rejection probability of the DW-test (given that ...) tends to zero when the DW-test is inadvertantly applied to the logs of a random walk. It possibly explains why Nelson and Plosser (1982), in their seminal paper on unit roots in US–macroeconomic time series, found unit roots in log–transformed time series although, as argued by Franses and McAleer (1998, p. 160), the "true" unit root is in the levels rather than in the logs: If the levels series has a unit root, this unit root might even appear more significant if the test is incorrectly applied to logs instead and the innovation variance is extreme enough.

The situation is more complicated if σ is fixed and $T \to \infty$. First of all, if the drift component δ in (1) is ≤ 0, z_t will with probability 1 eventually become negative, so there is no point in investigating the limiting conditional rejection probability when the conditioning event has vanishing probability. If $\delta > 0$, ...

4 Some finite sample Monte Carlo evidence

References

Burridge, P.; Guerre E.: "The limit distribution of level crossings of a random walk, and a simple unit root test." *Econometric Theory* 12, 1996, 705 – 723.

Corradi, V.; Swanson, N.: "Choosing between levels and logs in the presence of deterministic and stochastic trends." Mimeo, University of Pennsylvania 1997.

Corradi, V.: "Nonlinear transformations of integrated time series: a reconsideration." *Journal of Time Series Analysis* 16, 1995, 539 – 549.

Ermini, L.; Granger, C.W.J.: "Some generalizations on the algebra of $I(1)$-processes." *Journal of Econometrics* 58, 1993, 369 – 384.

Ermini, L.; Hendry, D.F.: "Log income versus linear income: An application of the encompassing principle." Mimeo, Nuffield College, Oxford 1995.

Franses, P.H.; Koop, G.: "On the sensitivity of unit root inference to nonlinear data transformations." *Economics Letters* 59, 1998, 7 – 15.

Franses, P.H.; McAleer, M.: "Testing for unit roots and non-linear transformations." *Journal of Time Series Analysis* 19, 1998, 147 – 164.

Granger, C.W.J.; Hallmann, J.: "Nonlinear transformations of integrated time series." *Journal of Time Series Analysis* 12, 1991, 207 – 224.

Guerre, F.; Jouneau, F.: "Geometric versus arithmetic random walk: The case of trended variables." INSEE discussion paper 9556, 1995.

Nelson, C.R.; Plosser, C.I.: "Trends and random walks in macroeconomic time series." *Journal of Monetary Economics* 10, 1982, 139 – 162.

Park, J.Y.; Phillips, P.C.B.: "Asymptotics for nonlinear transformations of integrated time series." Mimeo, Yale University, 1998.

Stock, J.: "Deciding between $I(1)$ and $I(0)$." *Journal of Econometrics* 63, 1994, 105 – 131.

Tabayashi, M.; McAleer, M.: "Tests of linear and logarithmic transformations for integrated processes." *Journal of the American Statistical Association* 94, 1999, 860 – 868.

A Appendix: Proof of Theorems

Proof of Theorem 1:

Consider the distribution of

$$\hat{\rho} = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2} \tag{A.1}$$

for some given $\sigma > 0$. The distribution of $\hat{\rho}$, when σ is replaced by $\tilde{\sigma} := c\sigma$, is then identical to the distribution of

$$\hat{\rho}(c) := \frac{\sum_{t=1}^{T} (y_{t-1} y_t)^c}{\sum_{t=1}^{T} (y_{t-1}^2)^c}. \tag{A.2}$$

As both the numerator and the denominator in (A1) tend to T as $c \to \infty$, part(a) of the theorem follows.

To determine the limiting distribution of $\hat{\rho}(c)$ as $c \to \infty$, keep $y_1 \ldots, y_T$ fixed and let $y_{k-1}^2 = max_{t=1,\ldots,T} y_{t-1}^2$. Then

$$\hat{\rho}(c) = \frac{\sum\limits_{t=1}^{T} \left(\frac{y_{t-1} y_t}{y_{k-1}^2} \right)^c}{\sum\limits_{t=1}^{T} \left(\frac{y_{t-1}^2}{y_{k-1}^2} \right)^c}, \tag{A.3}$$

and this expression can have only two limits as $c \to \infty$ (assuming without loss of generality that $y_{k-1}^2 \neq y_T y_{T-1}$). If $y_{k-1}^2 > y_T y_{T-1}$, then $y_{k-1}^2 > y_t y_{t-1}$ for all $t = 1, \ldots, T$, so the numerator tends to zero as $c \to \infty$. Since the denominator tends to unity (neglecting the possibility that k is not unique), we have $\hat{\rho}(c) \to 0$.

If $y_{k-1}^2 < y_T y_{T-1}$, the denominator still tends to unity, but the numerator tends infinity, so $\hat{\rho}(c) \to \infty$. This implies that the limiting distribution of $\hat{\rho}(c)$ as $c \to \infty$ is degenerate, with mass $P(y_{k-1}^2 > y_T y_{T-1})$ at zero, and the rest at infinity (in the sense that, for all $m \in \mathbf{N}$, $\delta > 0$, and $\varepsilon > 0$ there is an n^* such that

$$P\left(0 \leq \hat{\rho}(c) < \varepsilon \text{ or } \hat{\rho}(c) > m\right) > 1 - \varepsilon \quad \text{for all } n > n^*. \tag{A.4}$$

Proof of Theorem 3:

We have

$$\begin{aligned}
\hat{\rho}(c) &= \frac{\Sigma ln(cz_{t-1})ln(cz_t)}{\Sigma[ln(cz_{t-1})]^2} \\
&= \frac{\Sigma lnz_{t-1}lnz_t + lnc(\Sigma lnz_{t-1} + lnz_t) + Tln(c)^2}{\Sigma(lnz_{t-1})^2 + 2ln(c)\Sigma ln(z_{t-1}) + Tln(c)^2}.
\end{aligned}$$

and this expression tends to 1 both as $c \to 0$ and $c \to \infty$.