A note on testing for purchasing power parity*

Florian Heinen[1]

Institute of Statistics, Faculty of Economics and Management,
Leibniz University of Hannover, D-30167 Hannover, Germany

Abstract
We examine the asymptotic behavior of unit root tests against nonlinear alternatives of the exponential smooth transition type if the data is erroneously nonlinearly transformed.
We show analytically and by a Monte Carlo study that the probability of rejecting the correct null of a random walk depends heavily on the type of data transformation.

JEL-Numbers: C12, C22, F31
Keywords: Unit roots · Misspecification · Nonlinear data transformation · Purchasing Power Parity

1 Introduction

Purchasing power parity (PPP) is arguably one of the most important parities in international economics because "most [economists] instinctively believe in some variant of purchasing power parity as an anchor for long-run real exchange rates" (Rogoff (1996)). At the same time there is an ongoing debate whether PPP holds or not. One popular way to test for PPP is to check whether real exchange rate data follows a mean reverting process and is thus (globally) stationary. The first studies used the classical Dickey-Fuller test (DF test) to check whether a random walk can be found in real exchange rates (see e.g. Adler and Lehmann (1983), Meese and Rogoff (1983), Meese and Rogoff (1988) or Caporale et al. (2003)). As these studies could not reject the null of a random walk and therefore found no evidence for PPP economists have turned to non-

*Financial support by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.
[1]Corresponding author:
Phone: +49-511-762-5383
Fax: +49-511-762-3923
E-Mail: heinen@statistik.uni-hannover.de

linear models to capture a possibly nonlinear adjustment towards PPP (see e.g Taylor et al. (2001)). Especially models that allow for a symmetric adjustment towards PPP have been found useful. The most notable is the exponential smooth transition autoregressive (ESTAR) model (see Teräsvirta (1994)). In this framework Kapetanios et al. (2003) have developed a unit root test against an ESTAR alternative which subsequently has been successfully applied to confirm PPP (see e.g. Liew et al. (2004), Chortareas et al. (2002) and Chortareas and Kapetanios (2004)).

Many papers construct real exchange rates from log data to obtain the log real exchange rate which is consistent with the theory. However, before the analysis is performed the logarithm is applied a second time to the constructed series although, based on theory, it is not a priori clear whether the suspected unit root behavior is in the logs or in the levels (see e.g. Chortareas et al. (2002), Caporale et al. (2003), Bergman and Hansson (2005), Kruse (2009)). It is therefore of interest whether the applied unit root tests maintain their properties if applied to erroneously transformed data. For the linear case this has been done by Krämer and Davies (2002) and Davies and Krämer (2003) (see also de Jong (2010)). Krämer and Davies (2002) are able to analytically confirm parts of the simulation results of Granger and Hallman (1991) that the DF test overrejects a correct null of a random walk. Another striking feature of real exchange rates are the unusual small estimated variances of the innovation term (see e.g. Taylor et al. (2001), Rapach and Wohar (2006) or Smallwood (2005)). An important question is whether or not the unit root tests maintain their properties if the innovation variance becomes very small or very large.

This paper extends the analysis of Krämer and Davies (2002) to the nonlinear unit root test of Kapetanios et al. (2003). We show analytically and by Monte Carlo evidence that the behavior of the test statistic depends heavily on the innovation variance of the underlying process if the data is erroneously nonlinearly transformed.

The rest of the paper is structured as follows: Section 2 presents the considered unit root tests and derives their asymptotic behavior under nonlinear transformation. Section 3 underpins the analytical results with some Monte Carlo evidence before section 4 concludes. All proofs are collected in the appendix A.

2 Unit root tests and their asymptotic behavior under misspecification

The first studies researching the PPP relied on the DF test for a unit root against a linear alternative by Dickey and Fuller (1979). The DF test regression reads

$$y_t = \rho y_{t-1} + \varepsilon_t \,. \tag{1}$$

The pair of hypotheses is given as

$$H_0 : \rho = 1 \quad \text{vs.} \quad H_1 : \rho < 1 \,. \tag{2}$$

The DF test checks if $T(\hat{\rho}-1)$ is sufficiently close to zero. The quantiles of the resulting non-standard distribution of ρ are tabulated e.g. in Hamilton (1994, p. 762).
Kapetanios et al. (2003) extend the linear framework to the nonlinear ESTAR case (KSS test). The prototypical ESTAR process they consider reads

$$y_t = \phi y_{t-1} + \theta y_{t-1} \mathcal{G}(\gamma, c; \cdot) + \varepsilon_t, \text{ with } \varepsilon_t \overset{iid}{\sim} (0, \sigma^2) \,. \tag{3}$$

$\mathcal{G}(\gamma, c; \cdot) : \mathbb{R} \to [0,1]$ is a symmetrically u-shaped transition function governing the regime switching behavior of the process in (3). It reads

$$\mathcal{G}(\gamma, c; y_{t-d}) = 1 - \exp\left(-\gamma (y_{t-d} - c)^2\right) \,. \tag{4}$$

For simplicity Kapetanios et al. (2003) set $c = 0$ and $d = 1$. If $\phi = 1$ the process is still globally mean reverting as long as $-2 < \theta < 0$ which is assumed henceforth.
To derive the unit root test against the model in (3) they approximate the transition function in (4) by a first order Taylor approximation resulting in the following auxiliary regression model

$$\Delta y_t = \delta y_{t-1}^3 + \varepsilon_t \,. \tag{5}$$

The pair of hypotheses in this framework reads

$$H_0 : \delta = 0 \quad \text{vs.} \quad H_1 : \delta < 0 \,, \tag{6}$$

which is also tested via a conventional t-test. Critical values are reported in Kapetanios et al. (2003, p. 364).
We first consider the case when the unit root test is applied to the levels of the time series but the random walk is in the logs.

In this case let $z_t := \ln(y_t)$ and assume

$$z_t = \mu + z_{t-1} + \sigma\varepsilon_t, \text{ with } \varepsilon_t \overset{iid}{\sim} (0,1)\,, \tag{7}$$

and $0 < \sigma < \infty$. The DF test is applied to the levels $y_t := \exp(z_t)$. The OLS estimator for ρ in the model

$$y_t = \rho y_{t-1} + \varepsilon_t \tag{8}$$

reads

$$\hat{\rho} = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2}\,. \tag{9}$$

For the linear case the behavior of the DF test for $\sigma^2 \to 0$ and $\sigma^2 \to \infty$ (T fixed) and for $T \to \infty$ (σ^2 fixed) has been studied by Krämer and Davies (2002) and Davies and Krämer (2003). For the case of $T \to \infty$ and fixed σ^2 Davies and Krämer (2003) show that the rejection probability tends to unity.
Krämer and Davies (2002) show that if $\sigma \to \infty$ the probability for rejecting the null tends to zero if the random walk is in the logs but the DF test is erroneously applied to the levels. In the appendix we show that a similar result also holds for the nonlinear unit root test of Kapetanios et al. (2003).

Theorem 1. *If the data is generated by (7), and $\hat{\delta}$ is the OLS estimator in (5) where $y_t = \exp(z_t)$ we have:*

(i) $\sigma^2 \to 0$ *implies* $\hat{\delta} \xrightarrow{D} 0$.
$\sigma^2 \to 0$ *implies* $s.e.\left(\hat{\delta}\right) \xrightarrow{D} 0$.

(ii) $\sigma^2 \to \infty$ *implies* $\hat{\delta} \xrightarrow{D} -1$.
$\sigma^2 \to \infty$ *implies* $s.e.\left(\hat{\delta}\right) \xrightarrow{D} 1$.

The theorem shows that analog to the linear case the probability of rejecting the unit root hypotheses tends to zero as $\sigma \to 0$ because the test statistic defined as

$$\hat{t}_{KSS} = \frac{\hat{\delta}}{s.e.\left(\hat{\delta}\right)}\,, \tag{10}$$

where $s.e.\left(\hat{\delta}\right)$ is the standard error of $\hat{\delta}$ tends to zero as well because $\hat{\delta}$ tends to zero more quickly than $s.e.\left(\hat{\delta}\right)$.
For the part (ii) of the theorem it is important to note that the test statistic formally tends to -1 but at such a slow rate that in finite samples with reasonable sample sizes

and fairly large values of σ the test statistic will be large because $s.e.\left(\hat{\delta}\right)$ approaches 1 from below and much slower than $\hat{\delta}$ approaches -1.

Now we consider the situation when the unit root test is applied to the logs but the unit root process is in the levels. The true data generating process is now given by (7) but the unit root test is applied to $y_t := \ln(z_t)$. Of course this type of misspecification can only happen if $z_t > 0 \;\; (t = 0, 1, \ldots, T)$. Theoretically this only happens if z_0 is rather large and/or a substantial drift component is present in the data generating process. This is however of practical importance because given a finite sample of empirical data, there is a good chance to encounter only positive values although negative values might be possible from a theoretical argument. For the linear case Krämer and Davies (2002) show that $\hat{\rho}$ in (9) tends to one both as $\sigma \to 0$ and as $\sigma \to \infty$.

The next theorem, which is proven in the appendix A, gives the limiting behavior of the OLS estimate for δ in (5).

Theorem 2. *If the data is generated by (7), and $\hat{\delta}$ is the OLS estimator in (5) where $y_t = \log(z_t)$ we have:*

(i) $\sigma^2 \to 0$ *implies* $\hat{\delta} \xrightarrow{D} 0$.
$\sigma^2 \to 0$ *implies* $s.e.\left(\hat{\delta}\right) \xrightarrow{D} 0$.

(ii) $\sigma^2 \to \infty$ *implies* $\hat{\delta} \xrightarrow{D} 0$.
$\sigma^2 \to \infty$ *implies* $s.e.\left(\hat{\delta}\right) \xrightarrow{D} 0$.

This theorem implies that the test statistic of the KSS test in (10) converges to 0 because $\hat{\delta}$ converges faster than its standard error and hence the null is never rejected given that the test is one-sided with negative critical values.

3 Monte Carlo evidence

This section illustrates the analytic results from the previous section by means of Monte Carlo experiments. We start with the case that the random walk is in the logs but the unit root test is applied to the levels.

To illustrate the asymptotic behavior from theorem 1 figure 1 shows plots of $\hat{\delta}$ and $s.e.(\hat{\delta})$ for various values of σ and $T = 100$. Each point is obtained as a mean over 5000 Monte Carlo repetitions.

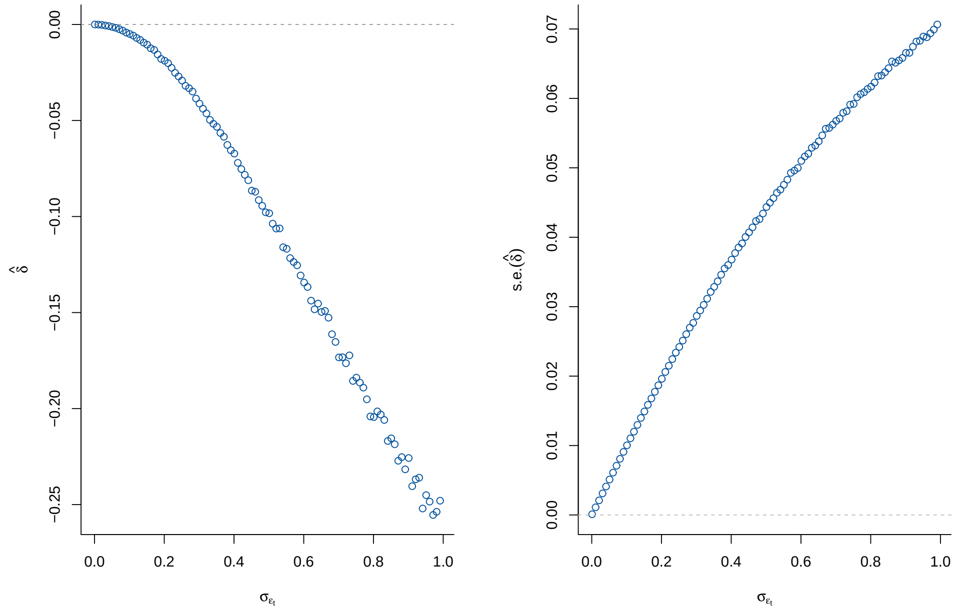

Figure 1: Illustration of theorem 1 part (i) I.

As predicted in the theorem both estimates tend to zero as $\sigma \rightarrow 0$. As a consequence also $\hat{t}_{KSS}$ tends to zero because $\hat{\delta}$ tends to zero more quickly. This is shown in figure 2. The right panel clearly shows a faster convergence of $\hat{\delta}$.

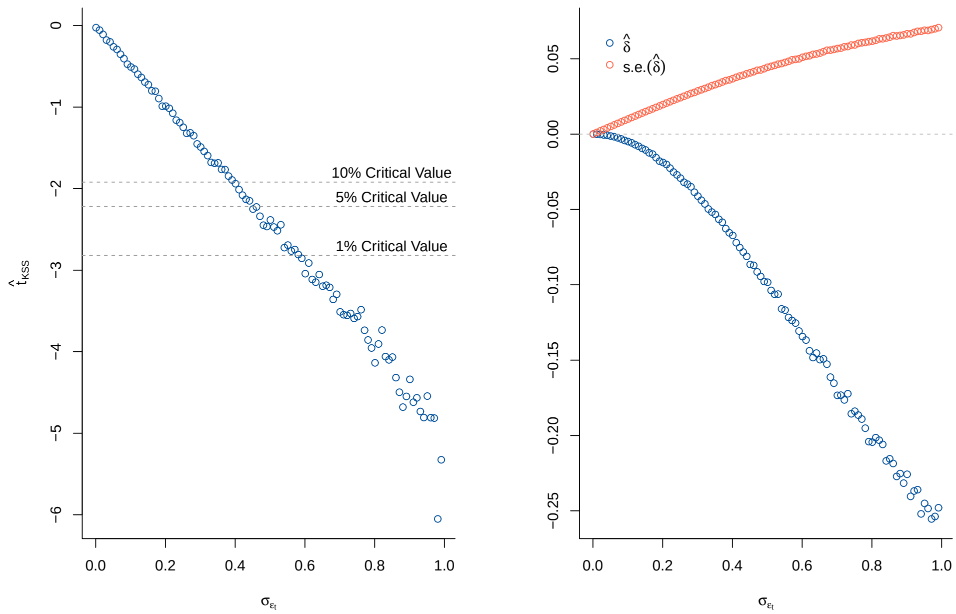

Figure 2: Illustration of theorem 1 part (i) II.

The associated power curves for the KSS test for this situation are depicted in figure 3.

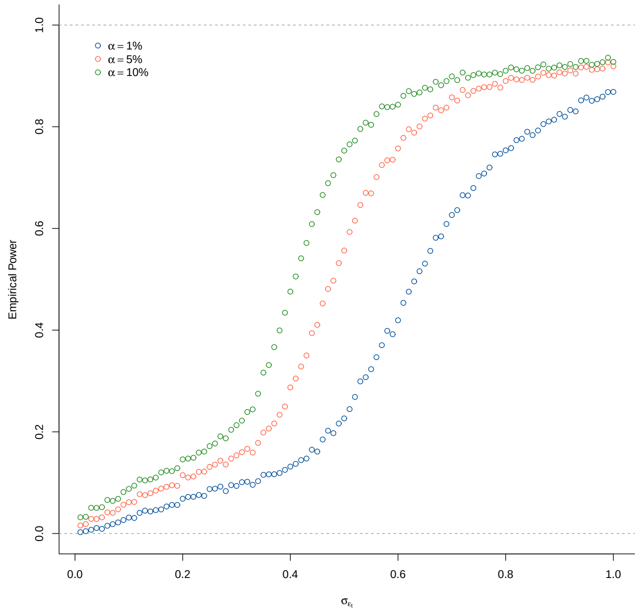

Figure 3: Power curves for exponential random walk I.

As expected the power curves are very low for small values of σ and eventually tend to one as σ increases.

Part (ii) of theorem 1 is illustrated in the following figures. Note that for these figures each point is obtained as the median over 5000 Monte Carlo repetitions rather than the mean. This is due to numerical instabilities when dealing with very large standard deviations. In such a case the noise dominates the signal and one obtains larger deviations from the mean estimate. However in order to get a notion about the central value of the distribution of the estimates we choose the more robust median as estimate.

Figure 4 shows the behavior of $\hat{\delta}$ and its standard deviation as $\sigma \to \infty$. As shown in the theorem $\hat{\delta}$ goes to -1 with increasing σ while the standard deviation goes very slowly to 1.

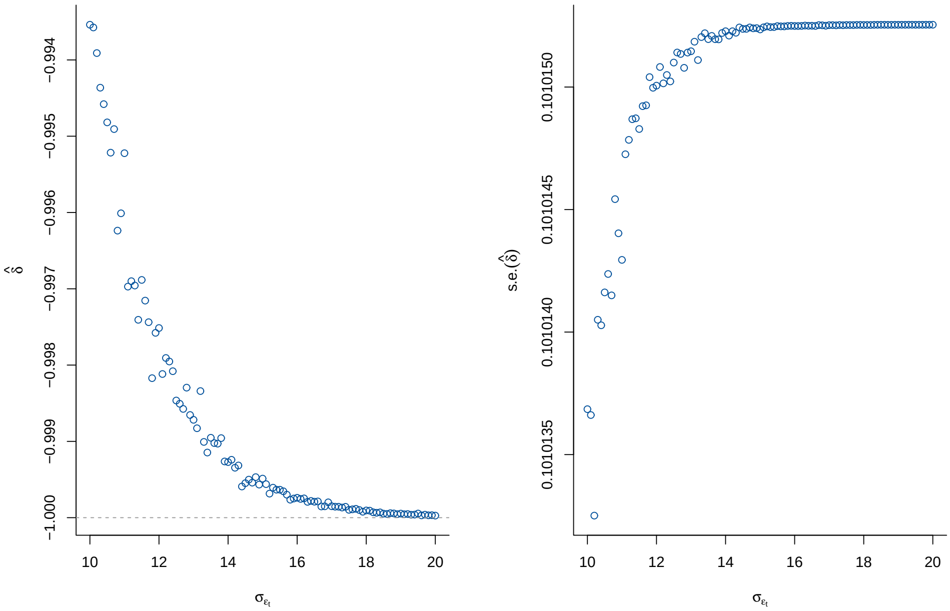

Figure 4: Illustration of theorem 1 part (ii) I.

Because of the slow convergence of $s.e.\left(\hat{\delta}\right)$ the associated test statistic $\hat{t}_{KSS}$ is rather large (see figure 5) leading to a power of 1 displayed in figure 6.

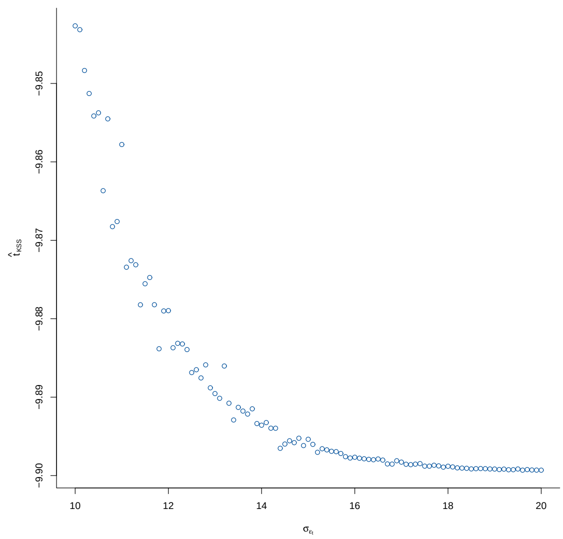

Figure 5: Behavior of $\hat{t}_{KSS}$ under theorem 1 part (ii).

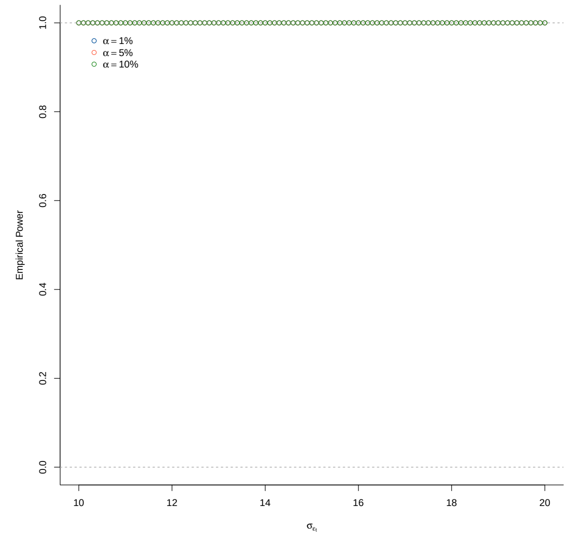

Figure 6: Power curves for exponential random walk II.

For the case that the random walk is in the levels but the unit root test is applied to the logs analog Monte Carlo experiments have been conducted. For part (i) of theorem 2 figure 7 shows the convergence of $\hat{\delta}$ and its associated standard error to zero as $\sigma \rightarrow 0$.

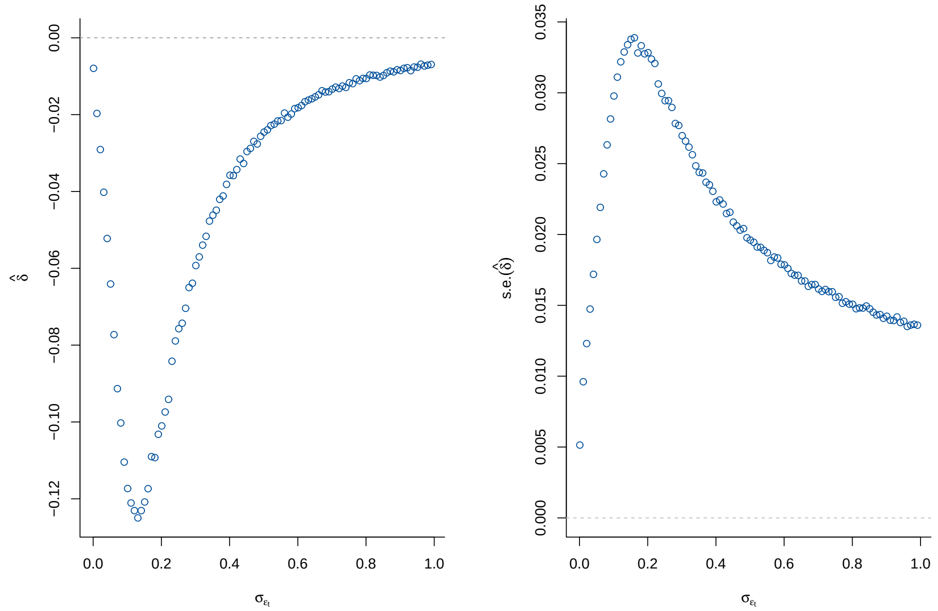

Figure 7: Illustration of theorem 2 part (i) I.

As predicted by theorem 2, $\hat{\delta}$ converges from negative values to zero and $s.e.\left(\hat{\delta}\right)$ converges from positive values to zero. Figure 8 shows the associated test statistic and illustrates the rates of convergence of $\hat{\delta}$ and $s.e.\left(\hat{\delta}\right)$.

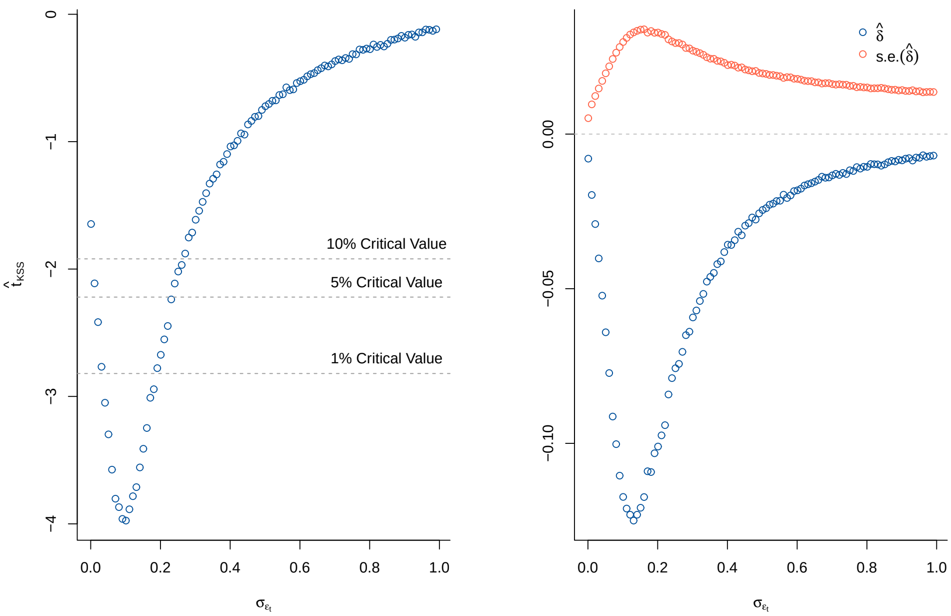

Figure 8: Illustration of theorem 2 part (i) II.

According to the theory the test statistic $\hat{t}_{KSS}$ converges to zero for small values of σ leading to non-rejection of the random walk hypotheses. The figures also clearly show an abnormal non-monotonic behavior of $\hat{\delta}$, $s.e.\left(\hat{\delta}\right)$ and $\hat{t}_{KSS}$. This induces even more uncertainty regarding the test decisions in this situation.

Figure 9 shows the associated power curves. As expected the power decreases with σ tending to zero.

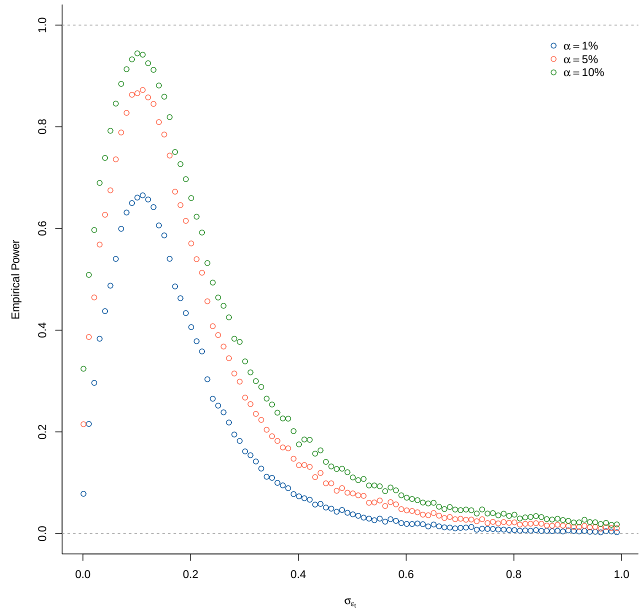

Figure 9: Power curves for logarithmic random walk.

The second part of theorem 2 analyzes the behavior of the KSS test for the case of $\sigma \to \infty$.
Figure 10 shows the convergence of the OLS estimates $\hat{\delta}$ and $s.e.\left(\hat{\delta}\right)$ to zero as $\sigma \to \infty$.

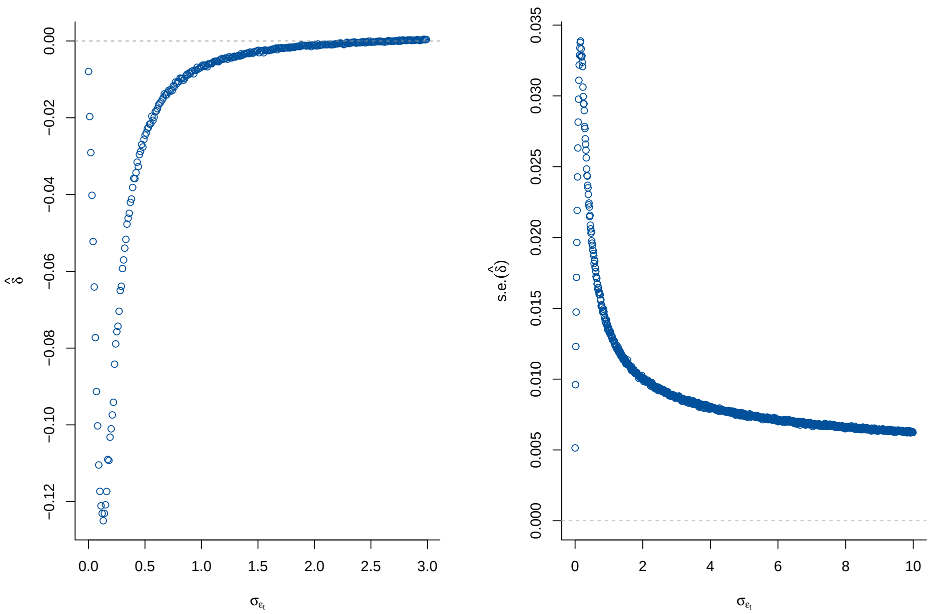

Figure 10: Illustration of theorem 2 part (ii).

The OLS estimates $\hat{\delta}$ is much faster than the convergence of $s.e.\left(\hat{\delta}\right)$ causing the test statistic $\hat{t}_{KSS}$ to rapidly converge to zero as well. This leads the KSS test to lose its power properties as the null cannot be rejected even for moderate values of σ. This is also supported by figure 11 that shows the power of the KSS test for large variances.

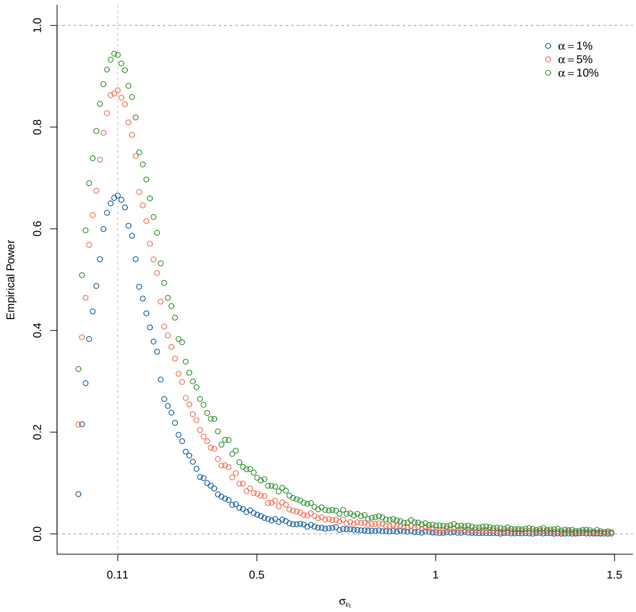

Figure 11: Power curves for logarithmic random walk and large variances.

The power decreases rapidly to zero. The vertical dashed line shows the maximal power value for the $\alpha = 5\%$ level. Except in the region around $\sigma \approx 0.11$ the power is virtually zero and thus the unit root hypothesis is never rejected in this case.

4 Conclusion

This paper studies the asymptotic behavior of the unit root test against a globally stationary ESTAR model proposed by Kapetanios et al. (2003). In particular we study the behavior of the test statistic if the innovation variance becomes very small or very large and the data has been erroneously nonlinearly transformed. I.e. the unit root is in the logs but the unit root test is applied to the levels or vice versa. We find that the behavior of the test depends heavily on the innovation variance if the data is transformed. If the unit root is in the logs but the unit root test is applied to the levels the test never rejects the random walk hypotheses if $\sigma \to 0$.

The situation when the unit root is in the levels but the unit root test is applied to the logs is even more severe. In this situation the test virtually never rejects the null. Only in a very small region of around $\sigma \approx 0.11$ the test rejects the null. This leads to a non-monotonic power of the KSS test.

The consequences for testing for PPP are that one should be very careful if nonlinear data transformation should be applied or not because in both considered cases the KSS test loses power and thus yields unreliable results.

Regarding data transformation in nonlinear time series future work should focus on discriminating between linear or logarithmic model formulation for nonlinear time series. A possible road is to extend the work of Kobayashi and McAleer (1999) or Spanos et al. (2008) to nonlinear time series models.

Possibly the asymptotic distributions of unit root tests for the case of erroneously transforming the time series according to $\log(|z_t|)$ where z_t is $I(1)$ could be derived using the results in de Jong (2004) (see also de Jong and Schmidt (2002)).

A Appendix

A.1 Proof of Theorem 1

Assume without loss of generality that $z_0 = 0$ and set the drift to $\mu = 0$. Then write

$$z_t = z_{t-1} + \sigma\varepsilon_t = \sigma \sum_{i=1}^{t} \varepsilon_i \,, \tag{11}$$

and notice that

$$\Delta y_t = y_t - y_{t-1} = \exp(z_t) - \exp(z_{t-1})\,. \tag{12}$$

With this the OLS estimator of $\hat{\delta}$ in (5) reads

$$\begin{aligned}
\hat{\delta} &= \frac{\sum_{t=1}^{T} y_{t-1}^3 \Delta y_t}{\sum_{t=1}^{T} y_{t-1}^6} \\
&= \frac{\sum_{t=1}^{T} [\exp(3z_{t-1} + z_t) - \exp(4z_{t-1})]}{\sum_{t=1}^{T} \exp(6z_{t-1})} \\
&= \frac{\sum_{t=1}^{T} \left[\exp\left(3\sigma \sum_{i=1}^{t-1} \varepsilon_i + \sigma \sum_{j=1}^{t} \varepsilon_i\right) - \exp\left(4\sigma \sum_{i=1}^{t-1} \varepsilon_i\right)\right]}{\sum_{t=1}^{T} \exp\left(6\sigma \sum_{i=1}^{t-1} \varepsilon_i\right)}\,.
\end{aligned} \tag{13}$$

As the numerator tends to zero and the denominator tends to T as $\sigma \to 0$ the first part of (i) of the Theorem follows.
Now consider the OLS estimator of the standard deviation of $\hat{\delta}$

$$\begin{aligned}
s.e.\left(\hat{\delta}\right) &= \frac{\frac{1}{T-1} \sum_{t=1}^{T} \left(\Delta y_t - \hat{\delta} y_{t-1}^3\right)^2}{\sum_{t=1}^{T} y_{t-1}^6} \\
&= \frac{\sum_{t=1}^{T} \left[(\exp(z_t) - \exp(z_{t-1}))^2 - 2\hat{\delta}(\exp(z_t) - \exp(z_{t-1}))\exp(3z_{t-1}) + \hat{\delta}^2 \exp(6z_{t-1})\right]}{\sum_{t=1}^{T} \exp(6z_{t-1})}\,.
\end{aligned} \tag{14}$$

Given the representation from above we immediately see that the numerator tends to zero and the denominator tends to T as $\sigma \to 0$. Consequently $s.e.\left(\hat{\delta}\right)$ tends to zero as claimed in part the second part of (i) in the theorem. Note that $\hat{\delta}$ tends to zero more

quickly so that $\hat{t}_{KSS}$ tends to zero as $\sigma \to 0$.
To prove the part (ii) note that for the case $\sigma \to \infty$ the slope of the exponential function becomes much steeper with increasing positive values as for negative values where the difference in the slopes eventually becomes negligible. We have from the representation in (13) that the numerator tends to $-\infty$ because of the higher power of the second term. As the denominator tends to ∞ with $\sigma \to \infty$ the whole expression for $\hat{\delta}$ tends to -1 as claimed. In the representation (14) for the standard deviation the denominator clearly tends to ∞. In the numerator the crucial term is the last one. This expression also goes to ∞ but much slower than the denominator (because $|\hat{\delta}| < 1$). Thus the whole term will eventually go to one but very slowly. □

A.2 Proof of Theorem 2

Assume without loss of generality that $z_0 = 0$ and set the drift to $\mu = 0$. Then write

$$z_t = z_{t-1} + \sigma\varepsilon_t = \sigma \sum_{i=1}^{t} \varepsilon_i \,, \tag{15}$$

and notice that

$$\begin{aligned}
\Delta y_t = y_t - y_{t-1} &= \log(z_t) - \log(z_{t-1}) \\
&= \log\left(\sigma \sum_{i=1}^{t} \varepsilon_i\right) - \log\left(\sigma \sum_{i=1}^{t-1} \varepsilon_i\right) \\
&= \log(\sigma) + \log\left(\sum_{i=1}^{t} \varepsilon_i\right) - \log(\sigma) - \log\left(\sum_{i=1}^{t-1} \varepsilon_i\right) \\
&= \log\left(\frac{\sum_{i=1}^{t} \varepsilon_i}{\sum_{i=1}^{t-1} \varepsilon_i}\right) \\
&= \log\left(1 + \frac{\varepsilon_t}{\sum_{i=1}^{t-1} \varepsilon_i}\right) \\
&=: \log(\eta) \,.
\end{aligned} \tag{16}$$

With this the OLS estimator reads

$$\hat{\delta} = \frac{\sum_{t=1}^{T} y_{t-1}^3 \Delta y_t}{\sum_{t=1}^{T} y_{t-1}^6} = \frac{\sum_{t=1}^{T}\left[\log(z_{t-1})^3 \log(\eta)\right]}{\sum_{t=1}^{T}\left[\log(z_{t-1})^6\right]} = \frac{\sum_{t=1}^{T}\left[\left\{\log(\sigma) + \log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)\right\}^3 \log(\eta)\right]}{\sum_{t=1}^{T}\left[\log(\sigma) + \log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)\right]^6}.$$

Using the binomial theorem we can expand the powers of the binomials in the numerator and the denominator. Leading to the following expression for the numerator

$$\sum_{t=1}^{T}\left[\left\{\log(\sigma)^3 + 3\log(\sigma)^2\log\left(\sum_{i=1}^{t-1}\varepsilon_i\right) + 3\log(\sigma)\log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^2 + \log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^3\right\}\log(\eta)\right].$$

And for the denominator

$$\sum_{t=1}^{T}\left[\log(\sigma)^6 + 6\log(\sigma)^5\log\left(\sum_{i=1}^{t-1}\varepsilon_i\right) + 15\log(\sigma)^4\log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^2 + 20\log(\sigma)^3\log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^3 + \right.$$
$$\left. 15\log(\sigma)^2\log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^4 + 6\log(\sigma)\log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^5 + \log\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^6\right].$$

The numerator tends to $-\infty$ as $\sigma \to 0$ and the denominator eventually tends to ∞. Thus $\hat{\delta}$ tends to 0 as claimed. Note that, because of the complex interplay of the powers of the logarithms involved and the associated changes of signs and the different slopes of the logarithmic functions in the interval $(0,1]$ the convergence is expected to be slow. However because the denominator converges at a higher rate than the numerator (because of the higher powers involved) the OLS estimate will eventually converge to zero.

Analyzing the standard deviation of $\hat{\delta}$ in (14) for $y_t = \log(z_t)$ we see for the same reasons as above that numerator slowly converges to zero (as $\hat{\delta}$ goes to zero) and the denominator diverges to ∞. Thus $s.e.\left(\hat{\delta}\right) \to 0$ as claimed.

For the first part of (ii) notice that the numerator as well as the denominator diverge to ∞. But again the denominator diverges much faster due to the high powers of $\log(\sigma)$ involved. Therefore $\hat{\delta}$ will eventually diverge to zero as claimed. The result for the standard error of $\hat{\delta}$ is obtained with the same reasoning. □

References

Adler, M. and Lehmann, B. (1983). Deviations from the purchasing power parity in the long run. *Journal of Finance*, 38:1471–1487.

Bergman, U. M. and Hansson, J. (2005). Real exchange rates and regime switching regimes. *Journal of International Money and Finance*, 24:121–138.

Caporale, G. M., Pittis, N., and Sakellis, P. (2003). Testing for PPP: The erratic behaviour of unit root tests. *Economics Letters*, 80:277–284.

Chortareas, G. and Kapetanios, G. (2004). The Yen real exchange rate may be stationary after all: Evidence from non-linear unit-root tests. *Oxford Bulletin of Economics and Statistics*, 66:113–131.

Chortareas, G., Kapetanios, G., and Shin, Y. (2002). Nonlinear mean reversion in real exchange rates. *Economics Letters*, 77:411–417.

Davies, L. and Krämer, W. (2003). The Dickey-Fuller test for exponential random walks. *Econometric Theory*, 19:865–877.

de Jong, R. (2010). Exponentials of unit root processes. Working Paper Ohio State University.

de Jong, R. M. (2004). Addendum to "Asymptotics for nonlinear transformations of integrated time series". *Econometric Theory*, 20:627–635.

de Jong, R. M. and Schmidt, P. (2002). Spurious logarithms and the KPSS statistic. *Economics Letters*, 76:383–391.

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.

Granger, C. and Hallman, J. (1991). Nonlinear transformations of integrated time series. *Journal of Time Series Analysis*, 12:207–224.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Kapetanios, G., Shin, Y., and Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics*, 112:359–379.

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

Krämer, W. and Davies, L. (2002). Testing for unit roots in the context of misspecified logarithmic random walks. *Economics Letters*, 75:313–319.

Kruse, R. (2009). A new unit root test against ESTAR based on a class of modified statistics. *Statistical Papers*, 52:71–85.

Liew, V., Baharumshah, A., and Chong, T. (2004). Are Asian real exchange rates stationary? *Economics Letters*, 83:313–316.

Meese, R. and Rogoff, K. (1983). Emprical exchange rate models of the seventies – Do they fit out of sample? *Journal of International Economics*, 14:3–24.

Meese, R. and Rogoff, K. (1988). Was it real? The exchange rate-interest differential relation over the modern floating-rate period. *Journal of Finance*, 43:933 – 948.

Rapach, D. E. and Wohar, M. E. (2006). The out-of-sample forecasting performance on nonlinear models of real exchange rate behavior. *International Journal of Forecasting*, 22:341–261.

Rogoff, K. (1996). The purchasing power parity puzzle. *Journal of Economic Literature*, 34:647–668.

Smallwood, A. D. (2005). Joint tests for non-linearity and long memory: The case of purchasing power parity. *Studies in Nonlinear Dynamics and Econometrics*, 9:1–28.

Spanos, A., Hendry, D., and Reade, J. (2008). Linear vs. log-linear unit-root specification: An application of misspecification encompassing. *Oxford Bulletin of Economics and Statistics*, 70:829–847.

Taylor, M. P., Peel, D. A., and Sarno, L. (2001). Nonlinear mean-reversion in real exchange rates: Toward a solution to the purchasing power parity puzzles. *International Economic Review*, 42:1015–1042.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89:208–218.