LEVEL CROSSING RANDOM WALK TEST ROBUST TO THE PRESENCE OF STRUCTURAL BREAKS.

Preliminary and Incomplete. Please do not cite or distribute.

Vitali Alexeev*And Alex Maynard

Department of Economics
University of Guelph

April 5, 2009

Abstract

We propose a modified version of the nonparametric level crossing random walk test, in which the crossing level is determined locally. This modification results in a test that is robust to unknown multiple structural breaks in the level and slope of the trend function under both the null and alternative hypothesis. No knowledge regarding the number or timing of the breaks is required. A data driven method is suggested to select the extent of the localization based on a trade-off between finite sample power and size distortion in a proximate model.

1 Introduction

Originally examined by Kendall [1953], the random walk hypothesis has been important in numerous disciplines including economics, finance, and international finance, in which random walk models have been used to to model variables such as consumption, stock prices, and exchange rates. Likewise, tests of the random walk hypothesis are frequently employed to test deeper theoretical models, such as the permanent income hypothesis and weak form market efficiency. Tests of weak form market efficiency include serial correlation tests, runs tests, the multiple variance ratio test of Lo and MacKinlay [1988].

It is typical in such tests to allow for a linear trend. This is often necessary on economic grounds as well. For example, economic growth and inflation, give rise

*Corresponding author. Correspondence may be sent to valexeev@uoguelph.ca. We thank Thanasis Stengos and Yiguo Sun for helpful comments and discussion. We take full responsibility for any remaining errors. First Draft January 5, 2009.

to an upward long-run trend in stock prices. On the other hand, few of the tests mentioned above allow for changes or breaks to occur in the trend term. This is arguably a somewhat restrictive assumption, which may result in unreliable inference. For example, changes to the trend growth rates or long-term average inflation rates would imply a break in trend for nominal stock prices.

An alternative perspective on the random walk test is to view it as a special case of a unit root test with uncorrelated errors. In the unit root testing literature there has been a long and ongoing interest in robustifying inference to the presence of structural breaks. Perron [1989], Vogelsang (1990) and Perron and Vogelsang [1992] demonstrate that structural breaks can cause difficulty for unit root tests, by causing an I(0) series with a break to resemble an I(1) process near the break point. Perron [1989] proposed the first unit root test that allows for the possibility of a break. His model allows for a break under both the null and alternative hypothesis, but with exogenously determined break dates. The next generation of tests, for example Zivot and Andrews [1992], Banerjee et al. [1992], allows for endogenously determined break dates under the alternative (but not the null) hypothesis. Most recently, Kim and Perron (2007) extend the earlier test of Perron (1994) to allow for a break in a trend function at an unknown time under both the null and alternative hypotheses.

A second issue that complicates unit root testing in the presence of structural breaks is the risk of misspecification with respect to the number of structural breaks. Vogelsang (1994) shows that the power of a unit root test is non-monotonic when a one-break model is estimated on data that contain two breaks. Lumsdaine and Papell [1997] consider testing the unit-root null against a two endogenous break alternative. Ohara [1999] develops a unit root test that allows for an alternative hypothesis with multiple trend breaks of unknown dates. Kapetanios [2005] provides tests for the unit-root hypothesis against the occurrence of an unspecified number of breaks which may be larger than two but smaller than a user-specified maximum.

We propose a nonparametric level-crossing random walk test based on a modification of a standard Burridge and Guerre [1996] unit root test, in which the global crossing level is replaced by a local average. By localizing the crossing level, the test is rendered robust to multiple trend breaks in either the intercept and/or the slope. These breaks may occur under both the null and alternative hypothesis and the test is nonparametric with respect to the breaks in the sense that we do not require knowledge of either the number or timing of the breaks. The test statistic is shown to have a standard normal null limit distribution, in both the presence and absence of trend breaks. Simulations also indicate that the test retains power in the presence of breaks.

An important practical consideration is the extent to which the crossing level is localized. Too strong a localization can negatively impact the power of the test, whereas too weak a localization may undermine the robustness of both size and power to the presence of trend-breaks. The choice of the localization parameter is also somewhat analogous to the problem of bandwidth selection in density or HAC covariance estimation. Given our focus on testing, we propose a method of selecting this parameter employing a finite sample size-power trade-off, based on a proximate parametric model.

On our approach builds on a rich literature involving level crossings and their application to random walk and unit root testing. Studies on level-crossings by continuous stationary Gaussian processes used in models of physical phenomena date back at least sixty years. Discretization of the continuous schemes and its applications to economics and finance begins with the works of Ho and Sun [1987], Burridge and Guerre [1996], and Ho and Hsing [1997] to name a few. Interestingly, in perhaps the earliest examinations of the random walk (without drift) hypothesis, Cowles and Jones [1937] compared the frequency of sequences and reversals in historical stock returns. This could be interpreted as an informal levels crossing test using a crossing level of zero.

Burridge and Guerre [1996] proposed a non-parametric unit root test based on the standardized number of level crossings of a random walk without deterministic terms (drift or trend). They observed that the number of level crossings in a variable would be larger in the absence of unit root. Because a unit root passes through any chosen level relatively rarely, they were using empirical frequency of this event to distinguish between random walks and stationary processes. They showed that the asymptotic distribution of the test statistic relates to a scaled local time of a Brownian motion and found that the scale factor depends on the long-run variance of the process.

Garcia and Sanso [2006] extended the work of Burridge and Guerre [1996] in two aspects. First, they allow for more general structures of autocorrelated disturbances. Secondly, they allow for a linear time trend and propose a modified crossing statistic, in which the crossings are defined relative to the (estimated) time trend. Put another way, the crossing level follows a trend.

Our approach can be viewed as further extending this approach to allow for breaks in trend. Were the number and timing of these beaks known a priori, one could envision a test based on the standardized crossings relative to a broken trend line. We instead take the view that such a broken trend is difficult to estimate with confidence and calculate crossings relative to a crossing level determined by a local average. This allows the crossing level to respond quickly to breaks, leading to the robustness properties discussed above.

On the other hand, the robustness of the test with respect to structural breaks is not achieved without cost. The more localized is the crossing level, the higher the crossing frequency of the random walk and the less easily it is distinguished a stationary process. For this reason, even in the presence of breaks, we find that the best test power is typically achieved when the crossing parameter is based on a moderately sized local average.

Localizing the crossing parameter also changes the nature of the large sample theory from nonstationary to stationary type asymptotics. Intuitively, deviations from local averages are akin to gradual differences and thus after removing the local average we effectively compare the crossings (about zero) of an appropriately differenced series to those of an over-differenced series. Since the crossing frequency of a stationary series depends on its serial correlation properties, our proposed test of random walk hypothesis, under which this correlation is restricted, has no natural unit root test counterpart.[1]

The plan of the paper is as follows. Section 2 briefly introduces the works previously done in the area of level crossings and considers random walk test based on the modified normalized number of level crossings in a variable, discusses its small-sample behavior under the null hypothesis and develops asymptotic null and alternative distributions of the test. Section 3 studies its power performances and its consistency against a set of alternative hypotheses. Section 4 analyzes the robustness of the test statistic under several departures from standard random walk test assumptions, such as the presence of serially correlated disturbances. In Section 5 we apply our testing methodology to a real-time series and compare the results with those obtained by means of existing random walk or unit root tests. Finally, Section 6 concludes.

2 Level Crossing Random Walk Test

Consider a sample $x_1, x_2, \ldots x_T$ from the process

$$x_t = \mu_t + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}\left(0, \sigma_\varepsilon^2\right), \quad \text{for } t \geq 1 \tag{1}$$

and $u_0 = 0$, where μ_t is a deterministic component, discussed separately below. In the absence of a trend (i.e.$\mu_t = 0$), Burridge and Guerre [1996] define the normalized number of crossing levels as:

$$K_T(z) = T^{-1/2} \sum_{t=1}^{T} \left(\mathbf{1}\left[x_{t-1} \leq z, x_t > z\right] + \mathbf{1}\left[x_{t-1} > z, x_t \leq z\right] \right) \tag{2}$$

where z is a crossing level. They then define a level crossing random walk test statistic as:

$$\varphi = \frac{\sqrt{T^{-1} \sum_{t=1}^{T} \left(\Delta x_t\right)^2}}{\widehat{MAD}} K_T(0) \tag{3}$$

where $\widehat{MAD} = T^{-1} \sum_{t=1}^{T} |\Delta x_t|$ and $\Delta x_t = x_t - x_{t-1}$.

[1] Currently, we also impose an i.i.d on the residuals under the null hypothesis, but expect that this can be weakened in future drafts.

Garcia and Sanso [2006] generalize this procedure to a unit root test by allowing for residual serial correlation. In addition, they allow for the presence of linear trend, specified by $\mu_t = \alpha + \beta t$. They define a modified level crossing, say K_T^{GS}, by replacing x_t in (2) with its detrended version $x_t - x_1 - ct$, where $c = \frac{x_T - x_1}{T}$. Likewise, they generalize the test-statistic in (3) to allow for residual serial correlation using

$$\eta^{(GS)} = \frac{\widehat{\omega}}{\widehat{MAD}} K_T^{(GS)}(0),$$

where $\widehat{\omega}$ is a consistent estimator of long-run variance of ε_t. They show that the null asymptotic distribution does not depend on the level of z.

Although Garcia and Sanso [2006] allow for a linear trend, neither they nor Burridge and Guerre [1996] allow for the possibility of trend breaks. The statistics constructed in these ways are very susceptible to structural breaks. The larger the magnitude of the break the larger will be the loss in power and the greater will be size distortion. The tests are particularly sensitive to breaks occurring towards the middle of the sample. To address this problem we define the normalized number of crossing levels formula to allow for presence of additive outlier type of structural breaks: sudden level shifts, changes in growth or both.

We suppose that the deterministic process in (1) is given by a broken linear trend with p breaks $(T_{B1}, T_{B2}, \ldots, T_{Bp})$ such that their corresponding break fractions are ordered increasingly, $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_p < 1$. We can generalize deterministic time trend components for the three models: Model 1, multiple jumps; Model 2, multiple kinks; Model 3, multiple jumps-and-kinks:

$$\mu_t = \mathbf{D}U_\mathbf{t} \cdot \mathbf{A} + \mathbf{D}T_\mathbf{t} \cdot \mathbf{B}$$

where $\mathbf{A}$ and $\mathbf{B}$ are $(p+1) \times 1$ vectors of drift and trend parameters respectively; λ is a $(p+1) \times 1$ vector of break fractions, with a first element of zero. $\mathbf{D}U_\mathbf{t}$ and $\mathbf{D}U_\mathbf{t}$ are $1 \times (p+1)$ vectors as follows:

$$\lambda = \begin{bmatrix} 0 & \lambda_1 & \lambda_2 & \ldots & \lambda_p \end{bmatrix}'$$

$$\mathbf{A} = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \ldots & \alpha_p \end{bmatrix}'$$

$$\mathbf{B} = \begin{bmatrix} \beta_0 & \beta_1 & \beta_2 & \ldots & \beta_p \end{bmatrix}'$$

$$\mathbf{D}U_\mathbf{t} = \begin{bmatrix} 1 & d_{1,t} & d_{2,t} & \ldots & d_{p,t} \end{bmatrix}$$

$$\mathbf{D}T_\mathbf{t} = \begin{bmatrix} t & (t - \lambda_1 T)\, d_{1,t} & (t - \lambda_2 T)\, d_{2,t} & \ldots & (t - \lambda_p T)\, d_{p,t} \end{bmatrix}$$

$$d_{i,t} = \begin{cases} 0 & for\ t \leq \lambda_i T \\ 1 & for\ t > \lambda_i T \end{cases}$$

Then the nested equation can be shown to be equal to

$$x_t = \rho x_{t-1} + \mu_t - \rho \mu_{t-1} + \varepsilon_t \tag{4}$$

where $\mu_t = \alpha_0 + \alpha_1 d_{1,t} + \ldots + \alpha_p d_{p,t} + \beta_0 t + \beta_1 (t - \lambda_1 T) d_{1,t} + \ldots + \beta_p (t - \lambda_p T) d_{p,t}$.[2]

Under the null of a random walk with drift:

$$x_t = x_{t-1} + \mu_t - \mu_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right)$$

Let out of sample boundary data points be constructed as follows:

$$x_{1-i} = x_1 + \frac{x_T - x_1}{T}(1-i), \quad \forall i = 1..T-1$$

$$x_{T+i} = x_T + \frac{x_T - x_1}{T}(i), \quad \forall i = 1..T-1$$

This is merely extrapolation of the series on the left and right hand side taking into account the trend of the series. Note that $(x_T - x_1)/T$ is a consistent estimator of β.

Define

$$x_t^{(m)} = x_t - x_{t-m} - \frac{x_{t+m} - x_{t-m}}{2}, \quad \forall t = 1..T, m = 1..(T-1) \tag{5}$$

then normalized number of level crossings can be defined as:

$$\begin{aligned} K_T^{(m)}(0) \quad &= \quad T^{-1/2} \sum_{t=1}^{T-1} \left(\mathbf{1}\left[x_t^{(m)} \leq 0, x_{t+1}^{(m)} > 0\right] + \mathbf{1}\left[x_t^{(m)} > 0, x_{t+1}^{(m)} \leq 0\right] \right) \\ &\quad\ m = 1..(T-1) \end{aligned}$$

Equivalently, it can be expressed in following way:

$$K_T^{(m)}(0) \quad = \quad T^{-1/2} \sum_{t=1}^{T-1} \mathbf{1}\left[x_t^{(m)} x_{t+1}^{(m)} < 0\right], \ m = 1..(T-1). \tag{6}$$

[2]Note that case with no structural breaks is equivalent to $\mu_t = \mathbf{D}U_\mathbf{t} \cdot \mathbf{A} + \mathbf{D}T_\mathbf{t} \cdot \mathbf{B} = \alpha_0 + \beta_0 t$, and case with one structural break $\mu_t = \mathbf{D}U_\mathbf{t} \cdot \mathbf{A} + \mathbf{D}T_\mathbf{t} \cdot \mathbf{B} = \alpha_0 + \alpha_1 d_{1,t} + \beta_0 t + \beta_1 (t - \lambda_1 T) d_{1,t}$.

2.1 Asymptotic properties of the centered partial sum in the absence of the deterministic component

We first establish the asymptotic properties for the random component u_t in this subsection. In the following subsection, we then show that these same properties hold unchanged in the presence of the deterministic component. Similarly, to equations (5) and (6), we define

$$u_t^{(m)} = u_t - u_{t-m} - \frac{u_{t+m} - u_{t-m}}{2}, \quad \forall t = 1..T, \tag{7}$$

where m is a fixed integer less than T and

$$K_{u,T}^{(m)}(0) \quad = \quad T^{-1/2} \sum_{t=1}^{T-1} \mathbf{1}\left[u_t^{(m)} u_{t+1}^{(m)} < 0\right], \; m = 1..(T-1). \tag{8}$$

Likewise, we define the crossing rate for u_t by

$$\mathbf{K}_u^{(m)} = T^{-1} \sum_{t=0}^{T-1} \mathbf{1}\left[u_t^{(m)} u_{t+1}^{(m)} < 0\right]. \tag{9}$$

By application of law of large numbers and Rice [1944]'s formula for a discrete-time, zero-mean, stationary Gaussian process, it can easily be shown $\mathbf{K}_u^{(m)}$ is consistent for the expected crossing rate:

$$\mathbf{K}_u^m \to \mathbb{P}\left(\mathbf{1}\left[u_t^{(m)} u_{t+1}^{(m)} < 0\right]\right) = \frac{1}{\pi}\cos^{-1} r_{u,1} \quad \text{as} \quad T \to \infty, \tag{10}$$

where $r_{u,i} = \mathbb{E}\left[u_t^{(m)} u_{t+i}^{(m)}\right]$.
Next, we consider the centered partial sum. In Section A.2 of the Appendix we show that

$$\sum_{i=0}^{\infty} r_i^2 < \infty \tag{11}$$

With this condition satisfied, the following central limit theorem of Ho and Sun [1987] for non-instantaneous filters of a Gaussian process immediately establishes the asymptotic normality of the centered partial sum. What follows is a special case of Ho and Sun [1987] for the indicator function.[3]

Theorem 1.[Ho and Sun, 1987 (special case)] *Suppose that condition (11) holds. Then*

$$T^{-1/2} \sum_{t=0}^{T-1} \left(\mathbf{1}\left[u_t^{(m)} u_{t+1}^{(m)} < 0\right] - \mathbb{P}\left(\mathbf{1}\left[u_t^{(m)} u_{t+1}^{(m)} < 0\right]\right)\right)$$

[3] Ho and Sun [1987] also provide a more general formula for the asymptotic variance in terms of integrals of functionals of Hermite polynomial approximations with to the spectral density function. We present the asymptotic variance in its more familiar form as the long run variance of the indicator function.

is asymptotically normally distributed with mean zero and variance $\omega_{\mathbf{1}_u,\mathbf{1}_u}$, *where*

$$\omega_{\mathbf{1}_u,\mathbf{1}_u} = \sum_{h=-\infty}^{\infty} \gamma_{\mathbf{1}[\cdot]}(h)$$

where $\gamma_{\mathbf{1}[\cdot]}(h) \equiv Cov\left[\mathbf{1}\left[u_t^{(m)}u_{t+1}^{(m)} < 0\right], \mathbf{1}\left[u_{t+h}^{(m)}u_{t+1+h}^{(m)} < 0\right]\right]$.
Proof. See Ho and Sun [1987].

Therefore, in the absence of a deterministic component, one could employ the following to test the null of Gaussian random walk (without deterministic component):

$$\eta_u^{(m)} = \frac{K_{u,T}^{(m)}(0) - \sqrt{T}\mathbb{P}\left(\mathbf{1}\left[u_t^{(m)}u_{t+1}^{(m)} < 0\right]\right)}{\sqrt{\hat{\omega}_{\mathbf{1}_u,\mathbf{1}_u}}} \sim N(0,1), \tag{12}$$

where $\hat{\omega}_{\mathbf{1},\mathbf{1}}$ is a consistent estimator of $\omega_{\mathbf{1},\mathbf{1}}$. One should note however, that this approximation, based on the assumption that m is fixed, deteriorates as m increases relative to T.

2.1.1 Asymptotic invariance to the presence of the deterministic component

The results in the subsection above were defined for the random component u_t. In this section, we first show that when the deterministic component μ_t is a linear trend, without breaks, $x_t^{(m)} = u_t^{(m)}$ and therefore the statistics are numerically unchanged with the addition of a linear trend. Put another way, $x_t^{(m)}$ eliminates the trend, via a gradual differencing. When μ_t is a linear trend with a finite number of breaks we argue that $x_t^{(m)}$ and $u_t^{(m)}$ differ at only a finite number of points. Because the level crossing statistics involve bounded indicator statistics, this implies a tight bound on the discrepancy between the statistics based on the processes with and without the deterministic components with breaks.
Substituting nested equation (4) into (5) yields:

$$\begin{aligned} x_t^{(m)} &= [\rho x_{t-1} + \varepsilon_t] - [\rho x_{t-m-1} + \varepsilon_{t-m}] - \frac{[\rho x_{t+m-1} + \varepsilon_{t+m}] - [\rho x_{t-m-1} + \varepsilon_{t-m}]}{2} \\ &\quad + [\mu_t - \rho\mu_{t-1}] - [\mu_{t-m} - \rho\mu_{t-m-1}] - \frac{[\mu_{t+m} - \rho\mu_{t+m-1}] - [\mu_{t-m} - \rho\mu_{t-m-1}]}{2}. \end{aligned}$$

Equivalently,

$$x_t^{(m)} = u_t - u_{t-m} - \frac{u_{t+m} - u_{t-m}}{2} + DC_t = u_t^{(m)} + DC_t$$

where

$$DC_t = [\mu_t - \rho\mu_{t-1}] - [\mu_{t-m} - \rho\mu_{t-m-1}] - \frac{[\mu_{t+m} - \rho\mu_{t+m-1}] - [\mu_{t-m} - \rho\mu_{t-m-1}]}{2}$$

The test statistic will be invariant to the breaks and trend/drift parameters as long as:

$$DC_t = 0.$$

For a simple case with no structural shits and breaks, $\mu_t = \alpha_0 + \beta_0 t$ and we can show that it is true for any values of α_0, β_0, m and any sample size T:

$$DC_t = (1-\rho)\,\beta_0 m - \frac{2\beta_0 m \ \ - 2\rho\beta_0 m}{2} = 0$$

And thus the test statistics will be numerically equivalent, irrespective of the magnitude of the level and trend parameters.

This exact equivalence no loner holds in the presence of trend-breaks. However, the break only affects the local crossing levels for which the break point is included in the local averaging from which the crossing level is calculated. This bounds the rate of discrepancy between $\mathbf{1}\left[u_t^{(m)} u_{t+1}^{(m)} < 0\right]$ and $\mathbf{1}\left[x_t^{(m)} x_{t+1}^{(m)} < 0\right]$. In particular, it can be shown that deterministic component of the transformed series can be expressed in the following way:

$$\begin{aligned}
DC_t \quad = \quad & (\alpha_1 + \beta_1 (t - \lambda_1 T)) \\
& \left[\left(d_{1,t} - \frac{1}{2} d_{1,t-m} - \frac{1}{2} d_{1,t+m}\right) - \rho\left(d_{1,t-1} - \frac{1}{2} d_{1,t-m-1} - \frac{1}{2} d_{1,t+m-1}\right)\right] \\
& + \frac{1}{2} m\beta_1 \left[(d_{1,t-m} - d_{1,t+m}) - \rho\,(d_{1,t-m-1} - d_{1,t+m-1})\right] \\
& + \rho\beta_1 \left[d_{1,t-1} - \frac{1}{2} d_{1,t-m-1} - \frac{1}{2} d_{1,t+m-1}\right] + ... \\
& \ldots + (\alpha_p + \beta_p (t - \lambda_p T)) \\
& \left[\left(d_{p,t} - \frac{1}{2} d_{p,t-m} - \frac{1}{2} d_{p,t+m}\right) - \rho\left(d_{p,t-1} - \frac{1}{2} d_{p,t-m-1} - \frac{1}{2} d_{p,t+m-1}\right)\right] \\
& + \frac{1}{2} m\beta_p \left[(d_{p,t-m} - d_{p,t+m}) - \rho\,(d_{p,t-m-1} - d_{p,t+m-1})\right] \\
& + \rho\beta_p \left[d_{p,t-1} - \frac{1}{2} d_{p,t-m-1} - \frac{1}{2} d_{p,t+m-1}\right].
\end{aligned}$$

Since the break-points dummy variables, $d_{i,t}$ defined earlier, are indicator functions and can only take values of 0 and 1, DC_t will be equal for values of t satisfying the following conditions:

$$\begin{gathered}
d_{1,t} - \tfrac{1}{2} d_{1,t-m} - \tfrac{1}{2} d_{1,t+m} = 0 \\
d_{1,t-1} - \tfrac{1}{2} d_{1,t-m-1} - \tfrac{1}{2} d_{1,t+m-1} = 0 \\
\vdots \\
d_{p,t} - \tfrac{1}{2} d_{p,t-m} - \tfrac{1}{2} d_{p,t+m} = 0 \\
d_{p,t-1} - \tfrac{1}{2} d_{p,t-m-1} - \tfrac{1}{2} d_{p,t+m-1} = 0
\end{gathered}$$

These conditions are equivalent to

$$\begin{aligned} d_{1,t-m} &= d_{1,t+m} \\ d_{1,t-m-1} &= d_{1,t+m-1} \\ &\vdots \\ d_{p,t-m} &= d_{p,t+m} \\ d_{p,t-m-1} &= d_{p,t+m-1} \end{aligned}$$

It is apparent that for a process with p breaks that these conditions are satisfied at all points outside $[\lambda_i T - m, \lambda_i T + m]$ for $i = 0, 1, 2, \dots p$. Therefore $x_t^{(m)} = u_t^{(m)}$ at all but a maximum of $p(2m+1)$ points. Therefore,

$$\begin{aligned} \left| K_T^{(m)} - K_{u,T}^{(m)} \right| &\leq T^{-1/2} \sum_{t=1}^{T-1} \left| \mathbf{1}\left[x_t^{(m)} x_{t+1}^{(m)} < 0 \right] - \mathbf{1}\left[u_t^{(m)} u_{t+1}^{(m)} < 0 \right] \right| \\ &= T^{-1/2} \sum_{t=1}^{T-1} \left| \mathbf{1}\left[\left(u_t^{(m)} + DC_t \right) \left(u_{t+1}^{(m)} + DC_{t+1} \right) < 0 \right] - \mathbf{1}\left[u_t^{(m)} u_{t+1}^{(m)} < 0 \right] \right| \\ &\leq T^{-1/2} \left(2m+1 \right) p. \end{aligned}$$

Therefore if the local detrending parameter m and the number of breaks p are both finite $K_T^{(m)}$ and $K_{u,T}^{(m)}$ are asymptotically equivalent and the test is robust to the presence of breaks.

2.2 Special polar cases

The analysis underlines the important role of the local detrending parameter m. To gain intuition we now consider two special polar cases: the case in which $m = 1$ and the case for $m = T - 1$.

Case when $m = 1$ Detrended series are obtained by

$$x_t^{(1)} = x_t - x_{t-1} - \frac{x_{t+1} - x_{t-1}}{2}$$

and the normalized number of crossing levels will thus be equal to

$$K_T^{(1)}(0) = T^{-1/2} \sum_{t=0}^{T-1} \mathbf{1}\left[x_t^{(1)} x_{t+1}^{(1)} \leq 0 \right]$$

It can be shown (please refer to the appendix) that under the null of Gaussian random walk, the normalized number of crossing levels can be expressed in the following form:

$$K_T^{(1)}(0) = T^{-1/2} \sum_{t=1}^{T} \mathbf{1}\left[\Delta\varepsilon_t \Delta\varepsilon_{t+1} \leq 0 \right]$$

where condition $\mathbf{1}\left[x_t^{(1)}x_{t+1}^{(1)}\leq 0\right]$ can is transformed to the following representation:

$$\begin{aligned}\mathbf{1}\left[x_t^{(1)}x_{t+1}^{(1)}\leq 0\right] &= \mathbf{1}\left[\left(x_t - x_{t-1} - \frac{x_{t+1}-x_{t-1}}{2}\right)\left(x_{t+1}-x_t-\frac{x_{t+2}-x_t}{2}\right)\leq 0\right]\\ &= \mathbf{1}\left[\left(\frac{2x_t - x_{t-1}-x_{t+1}}{2}\right)\left(\frac{2x_{t+1}-x_t-x_{t+2}}{2}\right)\leq 0\right]\\ &= \mathbf{1}\left[\left(\frac{\varepsilon_t-\varepsilon_{t+1}}{2}\right)\left(\frac{\varepsilon_{t+1}-\varepsilon_{t+2}}{2}\right)\leq 0\right]\\ &= \mathbf{1}\left[\Delta\varepsilon_{t+1}\Delta\varepsilon_{t+2}\leq 0\right]\end{aligned}$$

$$K_T^{(1)}(0) = T^{-1/2}\sum_{t=0}^{T-1}\mathbf{1}\left[x_t^{(1)}x_{t+1}^{(1)}\leq 0\right] = T^{-1/2}\sum_{t=1}^{T}\mathbf{1}\left[\Delta\varepsilon_t\Delta\varepsilon_{t+1}\leq 0\right]$$

Presented in this way, it is apparent that the normalized number of level crossings is independent of a trend parameter β and shift parameter α and asymptotically indifferent to structural breaks of additive outlier dynamics for a finite number of such breaks. This result holds both under the null and alternative hypothesis.

Process $x_t^{(1)}$can be transformed to MA(∞) representation (see Appendix for derivation):

$$0\leq(\rho-1)^2\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-2-j}+\varepsilon_t+(\rho-2)\,\varepsilon_{t-1}$$

Define a new process $y_t\equiv(\rho-1)^2\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-2-j}+\varepsilon_t+(\rho-2)\,\varepsilon_{t-1}$, where $\varepsilon_t\sim N\left(0,\sigma_\varepsilon^2\right)$. Then autocorrelation function for such process, assuming $\gamma_i = \mathbb{E}\left[y_t y_{t+i}\right]$ is given by:

$$r_1 = \frac{\gamma_1}{\gamma_0} = \frac{\rho^2-3\rho+4}{2\rho-6}$$

and givenRice [1944]'s result for a discrete-time, zero-mean, stationary Gaussian process $\left\{x_t^{(1)}\right\}$ [4]:

$$\mathbb{E}\left[\mathbf{K}^{(1)}\right]\equiv\mathbb{P}\left(\mathbf{1}\left[x_t^{(1)}x_{t+1}^{(1)}\leq 0\right]\right) = \frac{1}{\pi}\cos^{-1}r_1 = \frac{1}{\pi}\cos^{-1}\left(\frac{\rho^2-3\rho+4}{2\rho-6}\right)$$

The formula above allows us to look at the different average crossing rates for a wide range of autocorrelation parameter ρ.

[4] Recall that $x_t^{(1)} = -\Delta\varepsilon_{t+1}/2$

Theorem 1 yields the following test statistics:

$$\eta^{(1)} = K_T^{(1)}(0) \sim N\left(T^{-1/2}\mathbb{E}\left[\mathbf{K}^{(1)}\right], \sigma^2\right)$$

where the long run variance can be estimated by $\sigma^2 = \sum_{h=-\infty}^{\infty} \gamma_{\mathbf{1}[\cdot]}(h)$, given that $\gamma_{\mathbf{1}[\cdot]}(h) \equiv Cov\left[\mathbf{1}\left[x_t^{(m)} x_{t+1}^{(m)} < 0\right], \mathbf{1}\left[x_{t+h}^{(m)} x_{t+1+h}^{(m)} < 0\right]\right]$. It can easily be shown that for the case of $m = 1$ under the null of Gaussian random walk, $\sigma^2 = \frac{8}{45} + \frac{1}{60T}$ for a small sample and $\lim_{T\to\infty}\sigma^2 = \frac{8}{45}$.

Case when $m = T - 1$ Similarly, detrended series are obtained by

$$x_t^{(T-1)} = x_t - x_{t-(T-1)} - \frac{x_{t+(T-1)} - x_{t-(T-1)}}{2}$$

And given previously defined boundary conditions, process $x_t^{(T-1)}$ is equivalent to $x_t^{(GS)}$

$$x_t^{(T-1)} = x_t^{(GS)} = x_t - x_1 - ct, \quad c = \frac{x_T - x_1}{T}$$

and

$$\eta^{(T-1)} = \frac{\widehat{\sigma}}{\widehat{MAD}} K_T^{(T-1)}(0) \sim Rayleigh\,(1)$$

For values of $1 < m < T - 1$ distribution of $\eta^{(m)}$ as m increases skews from standard Normal to standard Rayleigh.

3 Selection of m

The choice of m involves a power-robustness tradeoff. Smaller values of m will result in a random walk test robust to the presence of multiple structural breaks and their magnitudes, as well as to misspecification with respect to the number of structural breaks. On the other hand, in the absence of breaks, a reduction of m leads to a reduction in test power. When breaks are present, our simulations indicate a non-monotonic relationship between test power and the magnitude of m. In this case, optimal power is generally obtained for intermediate values of m.

We suggest a data driven method of selecting the parameter m, based on a weighted finite sample power-size criterion. Since the finite sample sample power and size both depend on the true but unknown model for the data generating process, we employ a proximate model in the spirit of Andrews [1991] to perform this comparison. We take as our approximate model an AR(1) with normal errors and up to p breaks in both the shift and trend parameters. In order to estimate our approximate model, we first estimate the number and location of breaks using the procedure of Bai and Perron [1998]. We then estimate autoregressive component of the proximate mode conditional on the estimated number and location of breaks. We use the estimated proximate model to perform a calibrated simulation from which we calculate both power and size distortion (imposing $\rho = 1$). Using a weighted power-size criterion we then select the value of m that maximizes our criterion in the approximate model. When choosing equal weights, as in the simulation results reported below, our power-size criterion is equivalent to maximizing size adjusted power (defined as power-[actual size-nominal size]).

To be specific, we use the following procedure to select m for different sets of parameters.

1. Using the real data series $\{x_t\}_{t=1}^{T}$ estimate the parameters of the proximate model $\left\{\hat{A},\ \hat{B}, \widehat{\lambda}, \widehat{\rho}\right\}$[5]. Note that $\hat{p}$ is implicitly determined by $\widehat{\lambda}$.

2. Simulate n pseudo-time series $(\{\hat{x}_{t,i}\}_{t=1}^{T})_{i=1}^{n}$ from the estimated proximate model using the parameters $\left\{\hat{A},\ \hat{B}, \widehat{\lambda}, \widehat{\rho}, T\right\}$.

3. Find size and power of modified zero-crossings test for each value of $m = 1..(T-1)$.

4. Select m^* as the value of m resulting in the best weighted power-size average in the step above

[5] We use procedure described in Bai and Perron [1998] in oder to estimate the location of breaks.

4 Finite Sample Size, Power of Level Crossing Random Walk Test

In this section we present empirical size and power of our test for number of stationary alternatives. Performance of our test will be compared to performance of DF-GLS test proposed by Elliott, Rothenberg, & Stock (1996) as a benchmark test and RUR test by Aparicio and Sipols [2006], known to be robust to structural changes. We are hopping simulations experiments will confirm that our procedure offers an improvement over commonly used methods in small samples.

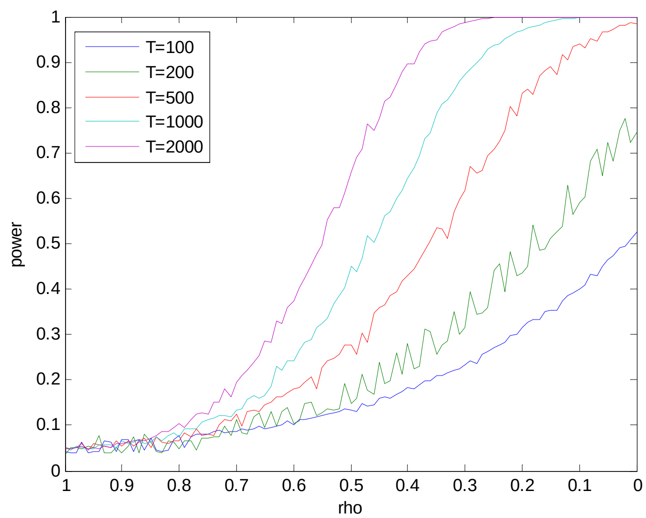

Figure **1**. Estimated power for different sample sizes.

We consider the univariate process $\{x_t\}_{t=1}^{T}$ generated by one of the three additive outlier models. For each model, the series is generated by the sum of a deterministic time trend and an error process: $x_t = \mu_t + u_t$, where $u_t = \rho u_{t-1} + \varepsilon_t$, $\varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right)$ and deterministic time trend component μ_t is defined according to the following three models:

Model 1: (Crush) $\mu_t = \alpha + \alpha_s d_t + \beta t$
Model 2: (Change in Growth) $\mu_t = \alpha + \beta t + \beta_s \left(t - \lambda T\right) d_t$
Model 3: (Both) $\mu_t = \alpha + \alpha_s d_t + \beta t + \beta_s \left(t - \lambda T\right) d_t$

with $0 < \lambda < 1$ denoting the true break fraction and

$$d_t = \begin{cases} 0 & for\ t \leq \lambda T \\ 1 & for\ t > \lambda T \end{cases}$$

Constructed in the above way, the three models allow a single structural break at the single time point.

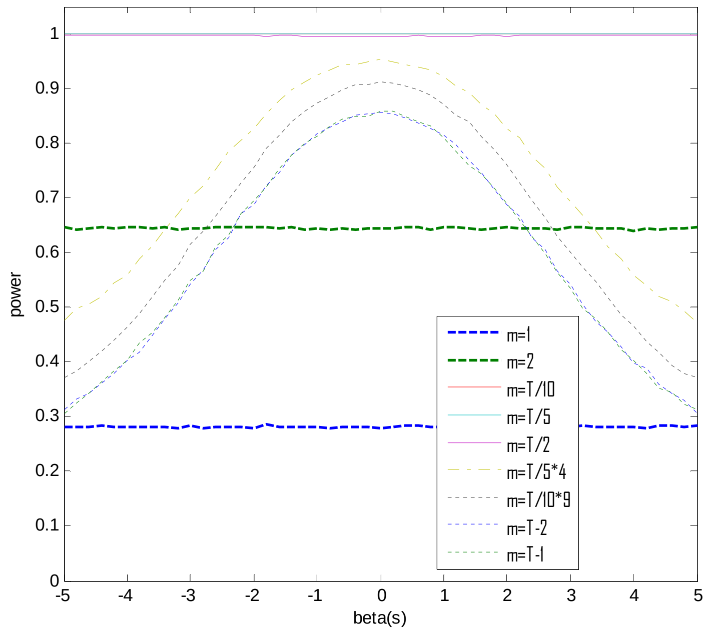

Figure **2**. Power for different levels of trend breaks. *Constructed based on null of Gaussian random walk and alternative of Model 2 with* $\lambda = 0.5,\ \rho = 0.5, T = 500$ *and for* $\beta_s = -5..5$*, based on 5000 simulations.*

For small enough smoothing parameter m the power of the test is unaffected by the location of the break, at least asymptotically. There is a trade off: choosing a small value of m will result in poor power of the test, but the test will be robust to the location of the break, thus eliminating the need for estimation of its position; having a large value of m will increase its power, but the test will be susceptible to the location of the break.

Choosing m will depend of the sample size. We are interested in finding parameter $\gamma \in [0, 1]$ in $m^* = (T-1)^\gamma$ such that m^* will result in optimal power-robustness tradeoff

Figures 3 and 4 below present power of the proposed test for different values of smoothing parameter, m, and location of a single structural break, λ.

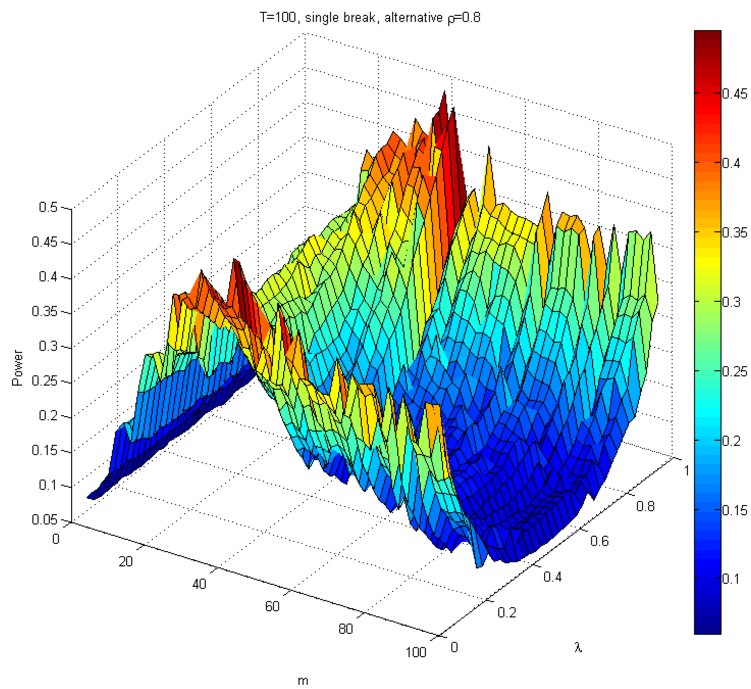

Figure 3. Power of the test. *Constructed based on null of Gaussian random walk and alternative of Model 2 with* ρ=**0.8**, **T**=**1**00 *and for* β=**5**, *based on 1000 simulations. Here* **T** *is a sample size,* γ *is the location of the break and* **m** *is a smoothing parameter.*

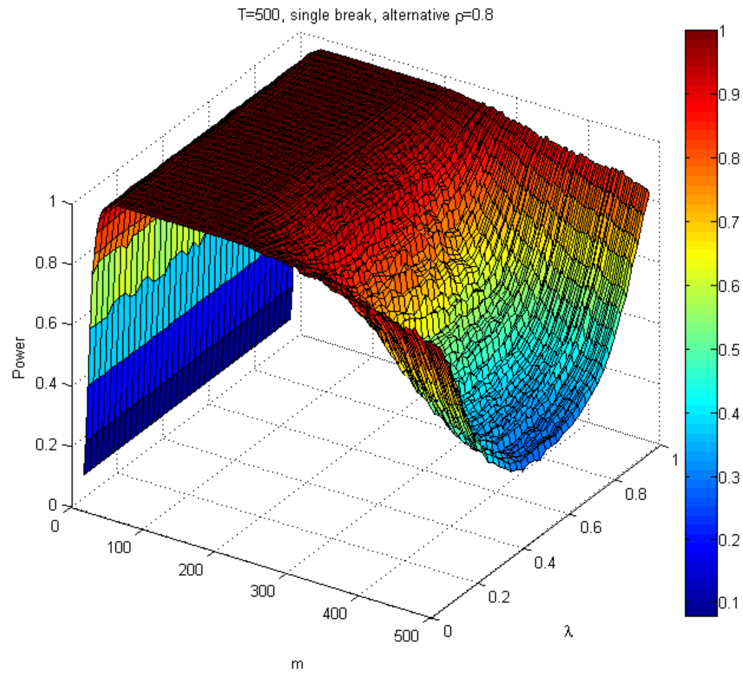

Figure 4. Power of the test. *Constructed based on null of Gaussian random walk and alternative of Model 2 with* ρ=**0.8**, **T**=**5**00 *and for* β=**5**, *based on 5000 simulations. Here* **T** *is a sample size,* γ *is the location of the break and* **m** *is a smoothing parameter.*

5 Departures from the Standard Conditions under H_0

We consider a number of stationary alternatives (serial correlation, heteroskedasticity, structural breaks or shift in a drift parameter...) to compare the performance of our test to the DF-GLS and RUR test of Aparicio and Sipols [2006].

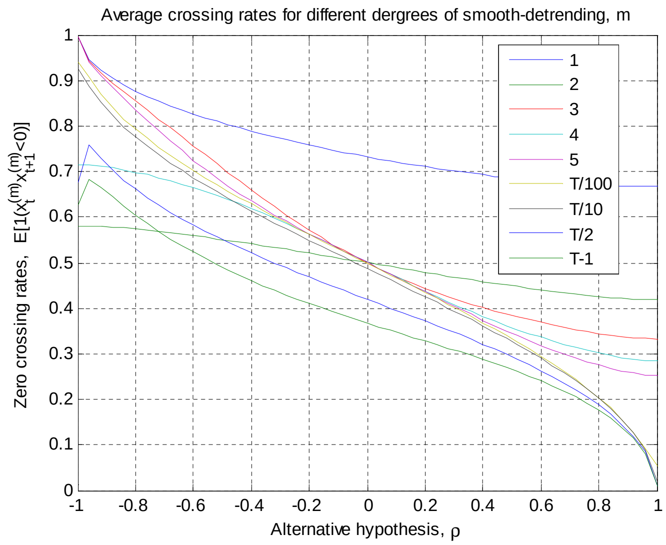

Figure 5. Average zero crossing rates for different values of m.

The power of the test relies on the difference between average crossing rates under the null and alternative. Our conjecture is that adding serial correlation will distort the average level crossings curve (Figure 5), making it difficult to construct a simple unified test statistics. One should note, however, that the presence of serial correlation will affect the test statistics more for smaller values of m and less for larger m.

6 Empirical Application

Formal application of this test to the Nelson-Plosser data set is underway. Application to the post war quarterly real GNP series as in Perron [1989] and Carrion-i Silvestre et al. [2008] are ongoing.

7 Concluding Remarks

In this paper we proposed a new nonparametric random walk test for trending processes that is based on normalized level crossings. The test is a modification of the one proposed in Burridge and Guerre [1996] allowing for trended processes and multiple structural breaks under the null and alternative. Thus, the test proposed in this paper can be considered a good complement to existing level crossing literature.

Further research will be devoted to finding the balance between the power loss and test stability in presence of multiple structural breaks and with added serial correlation, making it the first test in crossing level literature to address the problem of shifts and changing growth present in time series data.

A Appendix

A.1 The normalized number of crossing levels under H_0

For the case $k=1$, we have $z_t\left(1\right) = \frac{x_{t-1}^{(i)}+x_t^{(i)}+x_{t+1}^{(i)}}{3}$ and $z_{t-1}\left(1\right) = \frac{x_{t-2}^{(i)}+x_{t-1}^{(i)}+x_t^{(i)}}{3}$ and therefore

$$
\begin{aligned}
K_T^{(i)}\left(z_t\left(1\right)\right) &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\mathbf{1}\left[x_{t-1}^{(i)} \le z_{t-1}\left(1\right), x_t^{(i)} > z_t\left(1\right)\right] + \mathbf{1}\left[x_{t-1}^{(i)} > z_{t-1}\left(1\right), x_t^{(i)} \le z_t\left(1\right)\right]\right), \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\begin{array}{c} \mathbf{1}\left[x_{t-1}^{(i)} \le \frac{x_{t-2}^{(i)}+x_{t-1}^{(i)}+x_t^{(i)}}{3}, x_t^{(i)} > \frac{x_{t-1}^{(i)}+x_t^{(i)}+x_{t+1}^{(i)}}{3}\right] \\ +\mathbf{1}\left[x_{t-1}^{(i)} > \frac{x_{t-2}^{(i)}+x_{t-1}^{(i)}+x_t^{(i)}}{3}, x_t^{(i)} \le \frac{x_{t-1}^{(i)}+x_t^{(i)}+x_{t+1}^{(i)}}{3}\right] \end{array}\right) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\begin{array}{c} \mathbf{1}\left[2x_{t-1}^{(i)} \le x_{t-2}^{(i)} + x_t^{(i)}, 2x_t^{(i)} > x_{t-1}^{(i)} + x_{t+1}^{(i)}\right] \\ +\mathbf{1}\left[2x_{t-1}^{(i)} > x_{t-2}^{(i)} + x_t^{(i)}, 2x_t^{(i)} \le x_{t-1}^{(i)} + x_{t+1}^{(i)}\right] \end{array}\right), \ i = 0, 1.
\end{aligned}
$$

Under H_0: $x_t = \beta + x_{t-1} + \varepsilon_t$, so:

$$
\begin{aligned}
K_T^{(i)}\left(z_t\left(1\right)\right) &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\begin{array}{l} \mathbf{1}\left[2\left(\beta + x_{t-2}^{(i)} + \varepsilon_{t-1}\right) \le x_{t-2}^{(i)} + \left(\beta + x_{t-1}^{(i)} + \varepsilon_t\right)\right] \\ \times\mathbf{1}\left[2\left(\beta + x_{t-1}^{(i)} + \varepsilon_t\right) > x_{t-1}^{(i)} + \left(\beta + x_t^{(i)} + \varepsilon_{t+1}\right)\right] \\ +\mathbf{1}\left[2x_{t-1}^{(i)} > x_{t-2}^{(i)} + x_t^{(i)}, 2x_t^{(i)} \le x_{t-1}^{(i)} + x_{t+1}^{(i)}\right] \end{array}\right), \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\mathbf{1}\left[2\varepsilon_{t-1}\Delta\varepsilon_{t-1} + \varepsilon_t, 2\varepsilon_t > \varepsilon_t + \varepsilon_{t+1}\right] + \mathbf{1}\left[2\varepsilon_{t-1} > \varepsilon_{t-1} + \varepsilon_t, 2\varepsilon_t\Delta\varepsilon_t + \varepsilon_{t+1}\right]\right) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\mathbf{1}\left[\varepsilon_{t-1}\Delta\varepsilon_t, \varepsilon_t > \varepsilon_{t+1}\right] + \mathbf{1}\left[\varepsilon_{t-1} > \varepsilon_t, \varepsilon_t\Delta\varepsilon_{t+1}\right]\right), \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\mathbf{1}\left[0 \le \varepsilon_t - \varepsilon_{t-1}, 0 > \varepsilon_{t+1} - \varepsilon_t\right] + \mathbf{1}\left[0 > \varepsilon_t - \varepsilon_{t-1}, 0 \le \varepsilon_{t+1} - \varepsilon_t\right]\right) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\mathbf{1}\left[0 \le \Delta\varepsilon_t, 0 > \Delta\varepsilon_{t+1}\right] + \mathbf{1}\left[0 > \Delta\varepsilon_t, 0 \le \Delta\varepsilon_{t+1}\right]\right) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\mathbf{1}\left[\Delta\varepsilon_t < 0, \Delta\varepsilon_{t+1} \le 0\right] + \mathbf{1}\left[\Delta\varepsilon_t \le 0, \Delta\varepsilon_{t+1} < 0\right]\right), \ i = 0, 1.
\end{aligned}
$$

A.2 Average crossing rates for different m

Suppose $T \to \infty$ and m is fixed. We are interested in the probability of crossing of level zero of the process $x_t^{(m)}$, namely $\mathbb{P}\left(\mathbf{1}\left[x_t^{(m)} x_{t+1}^{(m)} < 0\right]\right)$. We have previously shown that the process $x_t^{(m)}$ is independent of the magnitude of parameters α and β. From (5), we have

$$x_t^{(m)} = x_t - x_{t-m} - \frac{x_{t+m} - x_{t-m}}{2} = u_t - x_{t-m}^{(0)} - \frac{x_{t+m}^{(0)} - x_{t-m}^{(0)}}{2} = u_t - \frac{1}{2}u_{t-m} - \frac{1}{2}u_{t+m}$$

where $u_t = \rho x_{t-1}^{(0)} + \varepsilon_t$ and $\varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right)$. For simplicity, assume $u_t = x_t$ for now.

Given the Rice's formula, in order to find the average rate of crossings for a process, we need to find autocorrelation function for that process.

$$r_1 = \frac{\gamma_1}{\gamma_0} = \frac{\mathbb{E}\left(x_t^{(m)} x_{t-1}^{(m)}\right)}{\mathbb{E}\left(x_t^{(m)} x_t^{(m)}\right)}$$

Transformation of crossing condition to MA(∞) representation Recall that x_t can be transformed to represent infinite MA process under the condition of $|\rho| < 1$.

We also derive through subsequent substitution that:

$$x_{t-m} = \rho x_{t-m-1} + \varepsilon_{t-m} = \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j}$$

$$x_t = \rho x_{t-1} + \varepsilon_t = \rho^m x_{t-m} + \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j}$$

$$x_{t+m} = \rho x_{t+m-1} + \varepsilon_{t+m} = \rho^{2m} x_{t-m} + \sum_{j=0}^{2m-1} \rho^j \varepsilon_{t+m-j}$$

$$\begin{aligned} x_t^{(m)} &= x_t - \frac{1}{2}x_{t-m} - \frac{1}{2}x_{t+m} = \\ &\rho^m x_{t-m} + \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} - \frac{1}{2}\left(x_{t-m}\right) - \frac{1}{2}\left(\rho^{2m} x_{t-m} + \sum_{j=0}^{2m-1} \rho^j \varepsilon_{t+m-j}\right) = \\ &\left(\rho^m - \frac{1}{2} - \frac{1}{2}\rho^{2m}\right) x_{t-m} + \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} - \frac{1}{2}\sum_{j=0}^{2m-1} \rho^j \varepsilon_{t+m-j}. \end{aligned}$$

At the same time

$$\sum_{j=0}^{2m-1} \rho^j \varepsilon_{t+m-j} = \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} + \sum_{j=m}^{2m-1} \rho^j \varepsilon_{t+m-j} = \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} + \sum_{j=0}^{m-1} \rho^{j+m} \varepsilon_{t-j}$$

$$= \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} + \rho^m \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j}$$

$$\begin{aligned} x_t^{(m)} &= -\frac{1}{2}\left(\rho^{2m} - 2\rho^m + 1\right) x_{t-m} + \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} - \frac{1}{2}\left(\sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} + \rho^m \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j}\right) \\ &= -\frac{1}{2}(\rho^m - 1)^2 \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} + \left(1 - \frac{1}{2}\rho^m\right) \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} - \frac{1}{2} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} \end{aligned}$$

Since we are assuming no serial correlation, $\varepsilon_t \sim iid\left(0, \sigma_\varepsilon^2\right)$, $\mathbb{E}\left[\varepsilon_t \varepsilon_s\right] = 0$ for all $t \neq s$, and $\mathbb{E}\left[\varepsilon_t \varepsilon_t\right] = \sigma_\varepsilon^2$.

$$\begin{aligned} \gamma_0 &\equiv \mathbb{E}\left(x_t^{(m)} x_t^{(m)}\right) \\ &= \mathbb{E}\left(\begin{array}{c} \left[-\frac{1}{2}(\rho^m - 1)^2 \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} + \left(1 - \frac{1}{2}\rho^m\right) \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} - \frac{1}{2} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j}\right] \\ \times \left[-\frac{1}{2}(\rho^m - 1)^2 \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} + \left(1 - \frac{1}{2}\rho^m\right) \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} - \frac{1}{2} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j}\right] \end{array}\right) \\ &= \frac{1}{4}(\rho^m - 1)^4 \mathbb{E}\left[\sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j}\right] - (\rho^m - 1)^2 \left(1 - \frac{1}{2}\rho^m\right) \mathbb{E}\left[\sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j}\right] \\ &\quad + \frac{1}{2}(\rho^m - 1)^2 \mathbb{E}\left[\sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j}\right] + \left(1 - \frac{1}{2}\rho^m\right)^2 \mathbb{E}\left[\sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j}\right] \\ &\quad - \left(1 - \frac{1}{2}\rho^m\right) \mathbb{E}\left[\sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j}\right] + \frac{1}{4}\mathbb{E}\left[\sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j}\right] \\ &= \frac{1}{4} \frac{(\rho^m - 1)^4}{1 - \rho^2} \sigma_\varepsilon^2 + \left(1 - \frac{1}{2}\rho^m\right)^2 \sigma_\varepsilon^2 \sum_{j=0}^{m-1} \rho^{2j} + \frac{1}{4}\sigma_\varepsilon^2 \sum_{j=0}^{m-1} \rho^{2j} \\ &= \left(\frac{1}{4} \frac{(\rho^m - 1)^4}{1 - \rho^2} + \left[\left(1 - \frac{1}{2}\rho^m\right)^2 + \frac{1}{4}\right] \sum_{j=0}^{m-1} \rho^{2j}\right) \sigma_\varepsilon^2 \end{aligned}$$

Likewise,

$$\begin{aligned} \gamma_1 &\equiv \mathbb{E}\left(x_t^{(m)} x_{t-1}^{(m)}\right) \\ &= \mathbb{E}\left(\begin{array}{c} \left[-\frac{1}{2}(\rho^m - 1)^2 \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} + \left(1 - \frac{1}{2}\rho^m\right) \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} - \frac{1}{2} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j}\right] \\ \times \left[-\frac{1}{2}(\rho^m - 1)^2 \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j-1} + \left(1 - \frac{1}{2}\rho^m\right) \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j-1} - \frac{1}{2} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j-1}\right] \end{array}\right) \end{aligned}$$

$$
\begin{aligned}
= \quad & \frac{1}{4}(\rho^m-1)^4\mathbb{E}\left[\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-m-j}\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-m-j-1}\right] \\
& -\frac{1}{2}\left(\rho^m-1\right)^2\left(1-\frac{1}{2}\rho^m\right)\mathbb{E}\left[\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-m-j}\sum_{j=0}^{m-1}\rho^j\varepsilon_{t-j-1}\right] \\
& +\frac{1}{4}(\rho^m-1)^2\mathbb{E}\left[\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-m-j}\sum_{j=0}^{m-1}\rho^j\varepsilon_{t+m-j-1}\right] \\
& -\frac{1}{2}(\rho^m-1)^2\left(1-\frac{1}{2}\rho^m\right)\mathbb{E}\left[\sum_{j=0}^{m-1}\rho^j\varepsilon_{t-j}\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-m-j-1}\right] \\
& +\left(1-\frac{1}{2}\rho^m\right)^2\mathbb{E}\left[\sum_{j=0}^{m-1}\rho^j\varepsilon_{t-j}\sum_{j=0}^{m-1}\rho^j\varepsilon_{t-j-1}\right] \\
& -\frac{1}{2}\left(1-\frac{1}{2}\rho^m\right)\underbrace{\mathbb{E}\left[\sum_{j=0}^{m-1}\rho^j\varepsilon_{t-j}\sum_{j=0}^{m-1}\rho^j\varepsilon_{t+m-j-1}\right]}_{=\rho^{m-1}\sigma_\varepsilon^2} \\
& +\frac{1}{4}(\rho^m-1)^2\underbrace{\mathbb{E}\left[\sum_{j=0}^{m-1}\rho^j\varepsilon_{t+m-j}\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-m-j-1}\right]}_{=0} \\
& -\frac{1}{2}\left(1-\frac{1}{2}\rho^m\right)\underbrace{\mathbb{E}\left[\sum_{j=0}^{m-1}\rho^j\varepsilon_{t+m-j}\sum_{j=0}^{m-1}\rho^j\varepsilon_{t-j-1}\right]}_{=0} \\
& +\frac{1}{4}\mathbb{E}\left[\sum_{j=0}^{m-1}\rho^j\varepsilon_{t+m-j}\sum_{j=0}^{m-1}\rho^j\varepsilon_{t+m-j-1}\right] \\
= \quad & \frac{1}{4}\frac{\rho\left(\rho^m-1\right)^4}{1-\rho^2}\sigma_\varepsilon^2-\frac{1}{2}\left(\rho^m-1\right)^2\left(1-\frac{1}{2}\rho^m\right)\rho^{m-1}\sigma_\varepsilon^2 \\
& +\left(1-\frac{1}{2}\rho^m\right)^2\rho\sigma_\varepsilon^2\sum_{j=0}^{m-2}\rho^{2j}-\frac{1}{2}\left(1-\frac{1}{2}\rho^m\right)\rho^{m-1}\sigma_\varepsilon^2+\frac{1}{4}\rho\sigma_\varepsilon^2\sum_{j=0}^{m-2}\rho^{2j} \\
= \quad & \sigma_\varepsilon^2\left(\begin{array}{c}\frac{1}{4}\frac{\rho(\rho^m-1)^4}{1-\rho^2}-\frac{1}{2}\left(\rho^m-1\right)^2\left(1-\frac{1}{2}\rho^m\right)\rho^{m-1} \\ +\left(1-\frac{1}{2}\rho^m\right)^2\rho\sum_{j=0}^{m-2}\rho^{2j}-\frac{1}{2}\left(1-\frac{1}{2}\rho^m\right)\rho^{m-1}+\frac{1}{4}\rho\sum_{j=0}^{m-2}\rho^{2j}\end{array}\right).
\end{aligned}
$$

We define the first-order correlation and its limit under the null hypothesis as

$$\begin{aligned} r_1 \equiv \frac{\gamma_1}{\gamma_0} &= \frac{\rho^{2m+1} - 4\rho^{m+1} + 6\rho + \rho^{2m-1} - 4\rho^{m-1}}{2\left(1-\rho^m\right)\left(3-\rho^m\right)} \\ \lim_{\rho \to 1} r_1 &= \frac{2m-3}{2m}. \end{aligned}$$

Similarly, we find:

$$\begin{aligned} \gamma_m &\equiv \mathbb{E}\left(y_t y_{t-m}\right) \\ &= \frac{\left(-\rho^m + 2\rho^{m-1} - \rho^{m-2}\right)(1-\rho)}{1+\rho}\sigma_\varepsilon^2, \quad \forall m \geq 2 \\ r_m &= \frac{-(\rho-1)^4(1-\rho)}{2\rho-6}\rho^{m-2}, \quad \forall m \geq 2. \end{aligned}$$

In order to apply Theorem 1, we show that the following necessary condition is satisfied:

$$\sum_{i=0}^{\infty} r_i^2 = r_0^2 + r_1^2 + \sum_{i=2}^{\infty} r_i^2 = 1 + \frac{\left(\rho^2 - 3\rho + 4\right)^2}{4(\rho-3)^2} + \frac{(\rho-1)^6}{4(\rho-3)^2\left(1-\rho^2\right)} < \infty.$$

The above quantity is undefined under, the special cases of $\rho = 1$ and $\rho = -1$. However employing L'Hopital's rule, it can be seen that it still converges to a finite number.
Given Rice [1944]'s formula, we have the following

$$\begin{aligned} \mathbb{E}\left(\mathbf{1}\left[y_t y_{t+1} < 0\right]\right) &= cos^{-1}\left(r_1\right)/\pi \quad \text{and therefore} \\ \gamma_{\mathbf{1}[\cdot]}(1) &\equiv Cov\left(\mathbf{1}\left[y_t y_{t+1} < 0\right], \mathbf{1}\left[y_{t+1} y_{t+2} < 0\right]\right) \\ &= \mathbb{E}\left(\mathbf{1}\left[y_t y_{t+1} < 0\right] \cdot \mathbf{1}\left[y_{t+1} y_{t+2} < 0\right]\right) - \mathbb{E}\left(\mathbf{1}\left[y_t y_{t+1} < 0\right]\right)\mathbb{E}\left(\mathbf{1}\left[y_{t+1} y_{t+2} < 0\right]\right) \\ &= \mathbb{E}\left(\mathbf{1}\left[y_t y_{t+1} < 0\right] \cdot \mathbf{1}\left[y_{t+1} y_{t+2} < 0\right]\right) - \left(cos^{-1}\left(r_1\right)/\pi\right)^2, \end{aligned}$$

where

$$\mathbb{E}\left(\mathbf{1}\left[y_t y_{t+1} < 0\right] \cdot \mathbf{1}\left[y_{t+1} y_{t+2} < 0\right]\right) = \mathbb{P}\left(y_t < 0, y_{t+1} \geq 0, y_{t+2} < 0\right) + \mathbb{P}\left(y_t \geq 0, y_{t+1} < 0, y_{t+2} \geq 0\right).$$

A.3 Indifference to pre-detrending

Below we show that detrending u_t in the manner of Garcia and Sanso [2006] has no effect on the number of crossing levels $K_T^{(i)}(m)$.

$$K_T^{(i)}(m) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\begin{array}{c} \mathbf{1}\left[x_{t-1}^{(i)} \leq z^{(i)}(t-1;m), x_t^{(i)} > z^{(i)}(t;m)\right] \\ +\mathbf{1}\left[x_{t-1}^{(i)} > z^{(i)}(t-1;m), x_t^{(i)} \leq z^{(i)}(t;m)\right] \end{array} \right), \ i = 0, 1$$

where $z^{(i)}(t;m) = (2m+1)^{-1}\sum_{j=t-m}^{t+m} x_j^{(i)}$. Given $x_t^{(1)} = x_t^{(0)} - x_1 - ct$ and $x_t^{(0)} = x_t$, we can express the first part of the crossing condition $\mathbf{1}\left[x_{t-1}^{(1)} \leq z^{(1)}(t-1;m),\quad x_t^{(1)} > z^{(1)}(t;m)\right]$ as

$$\begin{aligned}
&\mathbf{1}\left[x_{t-1}^{(0)} - x_1 - c(t-1) \leq \frac{\sum_{j=t-1-m}^{t-1+m}\left(x_j^{(0)} - x_1 - c(t-1)\right)}{2m+1},\right] \\
&\times\mathbf{1}\left[x_{t-1}^{(0)} - x_1 - ct > \frac{\sum_{j=t-m}^{t+m}\left(x_j^{(0)} - x_1 - ct\right)}{2m+1}\right] \\
&= \mathbf{1}\left[x_{t-1}^{(0)} - x_1 - c(t-1) \leq \frac{\sum_{j=t-1-m}^{t-1+m} x_j^{(0)} - \sum_{j=t-1-m}^{t-1+m} x_1 - \sum_{j=t-1-m}^{t-1+m} c(j-1)}{2m+1}\right] \\
&\times\mathbf{1}\left[x_{t-1}^{(0)} - x_1 - ct > \frac{\sum_{j=t-1-m}^{t-1+m} x_j^{(0)} - \sum_{j=t-1-m}^{t-1+m} x_1 - \sum_{j=t-1-m}^{t-1+m} cj}{2m+1}\right] \\
&= \mathbf{1}\left[x_{t-1}^{(0)} - x_1 - c(t-1) \leq \frac{\sum_{j=t-1-m}^{t-1+m} x_j^{(0)} - (2m+1)x_1 - c(2m+1)(t-1)}{2m+1}\right] \\
&\times\mathbf{1}\left[x_{t-1}^{(0)} - x_1 - ct > \frac{\sum_{j=t-1-m}^{t-1+m} x_j^{(0)} - (2m+1)x_1 - c(2m+1)t}{2m+1}\right] \\
&= \mathbf{1}\left[x_{t-1}^{(0)} \leq \frac{\sum_{j=t-1-m}^{t-1+m} x_j^{(0)}}{2m+1}, x_{t-1}^{(0)} > \frac{\sum_{j=t-1-m}^{t-1+m} x_j^{(0)}}{2m+1}\right].
\end{aligned}$$

Similarly, the second part of the crossing condition $\left[x_{t-1}^{(1)} > z^{(1)}(t-1;m), x_t^{(1)} \leq z^{(1)}(t;m)\right]$ can be shown to be equivalent to $\left[x_{t-1}^{(0)} > z^{(0)}(t-1;m), x_t^{(0)} \leq z^{(0)}(t;m)\right]$. This implies that $K_T^{(1)}(m) = K_T^{(0)}(m)$.

A.4 Indifference to trend level and slope in the absence of breaks

Below we demonstrate that the addition of a linear trend leaves the level crossing condition unchanged. With the addition of the linear trend, the process for x_t becomes $x_t = a + bt \quad + \rho x_{t-1} + \varepsilon_t$ and the standardized level crossing, defined as,

$$T^{-1/2}\sum_{t=0}^{T-1}\left(\mathbf{1}\left[x_t^{(m)}x_{t+1}^{(m)} < 0\right] - \mathbb{P}\left(\mathbf{1}\left[x_t^{(m)}x_{t+1}^{(m)} < 0\right]\right)\right),$$

uses the input

$$x_t^{(m)} = x_t - x_{t-m} - \frac{x_{t+m} - x_{t-m}}{2},$$

which is itself a function of x_t. However, we show below that $x_t^{(m)}$ relies only on u_t, the autoregressive component and not the deterministic trend component.

Specifically,

$$
\begin{aligned}
x_t^{(m)} &= a + bt \quad + \rho x_{t-1} + \varepsilon_t - [a + b\,(t-m) \quad + \rho x_{t-m-1} + \varepsilon_{t-m}] \\
&\quad - \frac{a + b\,(t+m) \quad + \rho x_{t+m-1} + \varepsilon_{t+m} - [a + b\,(t-m) \quad + \rho x_{t-m-1} + \varepsilon_{t-m}]}{2} \\
&= \rho x_{t-1} + \varepsilon_t - [-bm \quad + \rho x_{t-m-1} + \varepsilon_{t-m}] \\
&\quad - \frac{bm \quad + \rho x_{t+m-1} + \varepsilon_{t+m} - [-bm \quad + \rho x_{t-m-1} + \varepsilon_{t-m}]}{2} \\
&= \rho x_{t-1} + \varepsilon_t - [\rho x_{t-m-1} + \varepsilon_{t-m}] \\
&\quad - \frac{\rho x_{t+m-1} + \varepsilon_{t+m} - [\rho x_{t-m-1} + \varepsilon_{t-m}]}{2} \\
&= u_t - u_{t-m} - \frac{u_{t+m} - u_{t-m}}{2} \\
&= u_t^{(m)},
\end{aligned}
$$

where $u_t = \rho u_{t-1} + \varepsilon_t$.

References

Donald W. K. Andrews. Heteroskedasticity and autocorrelation consistent covariance estimation. *Econometrica*, 59:817–858, 1991.

Escribano A. Aparicio, F. and A. Sipols. Range unit root (rur) tests: Robust against nonlinearities, error distributions, structural breaks and outliers. *Journal of Time Series Analysis*, 27(4):545–576, 2006.

Jushan Bai and Pierre Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66:47–78, 1998.

A. Banerjee, R.L. Lumsdaine, and J.H. Stock. Recursive and sequential tests of the unit-root and trend-break hypotheses: Theory and international evidence. *Journal of Business and Economic Statistics*, 10(2):271–287, 1992.

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

J. L. Carrion-i Silvestre, D. Kim, and P. Perron. Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. *Journal of Applied Econometrics*, 2008. forthcoming.

A. Cowles and H. Jones. Some a posteriori probabilities in stock market action. *Econometrica*, 5(3):280–294, 1937.

A. Garcia and A. Sanso. A generalization of the burridge-guerre nonparametric unit root test. *Econometric Theory*, 22:756–761, 2006.

H. Ho and T. Hsing. Limit theorems for functionals of moving averages. *The Annals of Probability*, 25(4):1636–1669, 1997.

H. Ho and T. Sun. A central limit theorem for non-instantaneous filters of a stationary gaussian process. *Journal of Multivariate Analysis*, 22:144–155, 1987.

G. Kapetanios. Unit-root testing against the alternative hypothesis of up to m structural breaks. *Journal of Time Series Analysis*, 26(1):123–133, 2005.

M. G. Kendall. The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A, General.*, 116(1):11–34, 1953.

A. Lo and C. MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, 1(1):41–66, 1988.

R. Lumsdaine and D. Papell. Multiple trend breaks and the unit-root hypothesis. *The Review of Economics and Statistics*, 79(2):212–218, 1997.

H. Ohara. A unit root test with multiple trend breaks: A theory and an application to us and japanese macroeconomic time-series. *The Japanese Economic Review*, 50(3):266–290, 1999.

P. Perron. The great crush, the oil price shock, and the unit root hypothesis. *Econometrica*, 57:1361–1401, 1989.

P. Perron and T. Vogelsang. Nonstationary time series and level shifts with an application to purchasing power parity. *Journal of Business and Economic Statistics*, 10:301–320, 1992.

S. Rice. The mathematical analysis of random noise. *Bell Systems Technical Journal*, 23:282–332, 1944.

E. Zivot and D.W.K. Andrews. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, 10(2):251–270, 1992.