Localized Level Crossing Random Walk Test Robust to the Presence of Structural Breaks$^\star,\star\star$.

Vitali Alexeev$^\star,\star$, Alex Maynard$^\star$

$^\star$Department of Economics, University of Guelph, Guelph, Ontario N1G 2W1, Canada

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 bootstrap method is suggested to select the extent of the localization in order to maximize power in a proximate model. To control overall test size we propose a second outer bootstrap, in which we replicate the entire procedure, including the inner bootstrap used to select the localization parameter. The test is applied to Canadian nominal inflation and nominal interest rate series with implications for the Fisher hypothesis.

Key words: Level crossing, random walk, structural breaks, unit root, robustness.

JEL: C12, C14, C22

1. Introduction

Originally examined by Kendall [22], 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 model variables such as

$\star$First Draft January 5, 2009.

$\star\star$We thank Ba Chu, Robert de Jong, Lynda Khalaf, Thanasis Stengos, Shinichi Sakata, Yiguo Sun and participants at the 2009 Midwest Econometrics Group, the 3rd International conference on Computational and Financial Econometrics, the 2009 Canadian Econometric Study Group and the 2009 Canadian Economics Association Meetings for helpful comments and discussion. We take full responsibility for any remaining errors. This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET; www.sharcnet.ca).

$^\star$Corresponding author

Email addresses: valexeev@uoguelph.ca (Vitali Alexeev), maynarda@uoguelph.ca (Alex Maynard)

URL: http://www.uoguelph.ca/~valexeev/ (Vitali Alexeev)

January 21, 2010
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, and the multiple variance ratio test of Lo and MacKinlay [24].

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 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 [32] and Perron and Vogelsang [33] 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 [32] 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 [41], Banerjee et al. [4], allows for endogenously determined break dates under the alternative (but not the null) hypothesis. Most recently, Kim and Perron [23] 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 [37] 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. Lumdaine and Papell [25] consider testing the unit-root null against a two endogenous break alternative. Ohara [28] develops a unit root test that allows for an alternative hypothesis with multiple trend breaks of unknown dates. Kapetanios [21] provides tests for the unit-root hypothesis against the occurrence of an unknown, but finite number of breaks.

We propose a level-crossing random walk test based on a modification of a standard Burridge and Guerre [7] 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. Further, 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. Given our focus on testing, we propose a
resampling method for selecting this parameter which maximizes finite sample
power in a proximate parametric model. Naturally, this pre-selection induces
some over-rejection in finite sample. In order to control overall test size, we
propose critical values based on a second outer bootstrap in which the inner
bootstrap (used to select the localization parameter) is repeated on each boot-
strap replicate.

Our approach builds on a rich literature involving level crossings and their ap-
plication 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 [15],
Burridge and Guerre [7], and Ho and Hsing [14] to name a few.

Interestingly, in perhaps the earliest examinations of the random walk (without
drift) hypothesis, Cowles and Jones [8] 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 [7] 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 used the 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.

aspects. First, they allow for more general structures of autocorrelated distur-
bances. 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. Where the number and timing of these breaks known a priori,
one could envision a test based on the standardized crossings relative to a bro-
ken 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 locally. 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 from 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 region.

Localizing the crossing parameter also changes the nature of the large sample theory from nonstationary to stationary type asymptotics. Intuitively, deviations from localized trends are akin to gradual differences and thus after removing the trend component 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. The crossing frequency is also sensitive to the distribution of the errors. Using the arcsine law, we derive the crossing frequencies for the family of elliptically symmetric distributions, which includes as special cases the normal, student-t, logistic, Laplace, and scaled mixed normal distributions. While this allows for the modeling of fat-tails, it is clear that the presence of asymmetries would alter the crossing rates. Finally, to the best of our knowledge, existing central limit theorems applicable to level crossings of stationary linear processes require i.i.d. innovations. However, in our simulations, we demonstrate the good performance of our test under both GARCH and EGARCH innovations, despite their violation of this assumption.

The plan of the paper is as follows: Section 2 briefly introduces the works previously done in the area of level crossings and proposes a random walk test based on the number of localized level crossings. Section 3 introduces the two bootstrap methods: an inner bootstrap to select the localization parameter and an outer bootstrap to select critical values. Section 4 studies finite sample size and power of the test and compares the proposed procedure to some existing representative unit root tests. Section 5 contains empirical applications to both Nelson-Plosser data set and Canadian interest and inflation rates. Finally, Section 6 concludes.
2. Localized Level Crossing Random Walk Test

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

\[
x_t = \mu_t + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t, \quad \text{for } t \geq 1
\]

and \( u_0 = 0 \), where \( \mu_t \) is a deterministic component, discussed separately below. Equivalently, expressed as a nested equation, (1) can be represented as

\[
x_t = \rho x_{t-1} + \mu_t - \rho \mu_{t-1} + \varepsilon_t.
\]

In the absence of a trend (i.e. \( \mu_t = 0 \)) and with \( \varepsilon_t \sim \text{i.i.d.} \left(0, \sigma^2\right)\), Burridge and Guerre [7] define the normalized number of crossing levels as:

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

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

\[
\varphi = \sqrt{\frac{T-1}{\sum_{t=1}^{T} \Delta x_t^2}} K_T(0)
\]

where \( \sum_{t=1}^{T} \Delta x_t = T-1 \sum_{t=1}^{T} |\Delta x_t| \) and \( \Delta x_t = x_t - x_{t-1} \).

Garcia and Sanso [13] generalize this procedure by allowing 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 (3) with its detrended version \( x_t - x_1 - ct \), where \( c = \frac{x_T - x_1}{T} \). In addition, they generalize the test-statistic in (4) to allow for residual serial correlation using

\[
\eta^{(GS)} = \frac{\hat{\omega}_{\varepsilon,\varepsilon} K_T^{(GS)}(0)}{MAD}
\]

where \( \hat{\omega}_{\varepsilon,\varepsilon} \) is a consistent estimator of long-run variance of \( \varepsilon_t \). They show that the null asymptotic distribution is standard Rayleigh and does not depend on the level \( z \).

Although Garcia and Sanso [13] allow for a linear trend, neither they nor Burridge and Guerre [7] allow for the possibility of trend breaks. While these methods work quite well in the no-break case for which they were designed, the crossing statistics constructed in these ways are susceptible to structural breaks. In Figure 1 the number of crossings, on which the test statistic relies, will be equal to the number of crossings of the straight trend line and the stochastic process. It is apparent that for a stationary process with a single break in the middle, as depicted in the figure, the number of crossings will be minimal. 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
stationary process with trend break
trend line
stationary process without break

Figure 1: The dotted black line shows a simulated stationary series with a linear trend. The solid blue line shows a stationary series with a break in the slope of the trend in the middle of the data. The straight solid red line shows an estimated trend line using the approach of Garcia and Sanso [13].

To address this problem we define the normalized number of crossing levels formula to allow for the presence of additive outlier-type 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 such that their corresponding break fractions, \( \lambda_i \), are ordered increasingly: \( 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_p < 1 \). Thus, for any \( i = 1..p \), \( [ \lambda_i T ] \) denotes the observation at which the \( i \)th structural break occurs. Define the indicator for regime \( i = 1, 2, ..., p \) by

\[
d_{i,t} = \begin{cases} 
0 & \text{for } t \leq \lambda_i T \\
1 & \text{for } t > \lambda_i T 
\end{cases}.
\]

Following Perron [32], we generalize the deterministic process \( \mu_t \) in (1) to allow for breaks in either the level or the slope of the linear trend:

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

It is also useful to reformulate (6) in matrix notation. Let \( A = [ \begin{array}{cccc}
\alpha_0 & \alpha_1 & \ldots & \alpha_p 
\end{array} ] \) and \( B = [ \beta_0 \beta_1 \ldots \beta_p ] \) be \( (p + 1) \times 1 \) vectors of level and trend parameters respectively; and \( \lambda = [ \begin{array}{cccc}
0 & \lambda_1 & \ldots & \lambda_p 
\end{array} ] \) be a \( (p + 1) \times 1 \) vector of break

occuring towards the middle of the sample.
fractions, with a first element of zero. Define $\mathbf{D}_t$ and $\mathbf{D}_t^T$ as $1 \times (p + 1)$ vectors as follows:

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

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

Then, in its matrix notation, $\mu_t$ in (6) can be represented as follows:

$$
\mu_t = \mathbf{D}_t \cdot \mathbf{A} + \mathbf{D}_t^T \cdot \mathbf{B}.
$$

(7)

This representation allows for multiple breaks in both level and trend.\(^1\)

Since we assume that the timing, magnitude and number of breaks are unknown, we do not detrend based on (6). Instead, we define the locally detrended series as

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

(8)

Recall that in Garcia and Sanso [13]’s model, discussed earlier, the process in (1) is detrended globally as follows: $x_t - x_1 - ct$, where $c = \frac{x_T - x_1}{T}$. This corresponds to the case in which $m = T - 1$ in (8). We generalize this approach to allow for localized detrending. By introducing a localization parameter, $m$, we control the extent of the detrending of the process in (1). In Figure 2, the number of crossings of a local trend curve with a stochastic process with a single break, would represent the number of intersections of the two curves.

In order to calculate $x_t^{(m)}$ based on (8) when $t \leq m$ or $t \geq T - m$ we require boundary conditions for the out of sample observations. Let the out of sample boundary data points be constructed as follows:

$$
x_{1-i} = x_1 + c (1 - i),
$$

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

(9)

where $c = \frac{x_T - x_1}{T}$ as before. This is merely an extrapolation of the series on the left and right hand side taking into account the trend of the series. One reason for defining the out of sample points this way is that for $m = T - 1$ the set up of the test is identical to Garcia and Sanso [13]. Note that $(x_T - x_1)/T$ is a consistent estimator of $\beta$, when $\mu_t = \alpha + \beta t$ is a linear trend without break.

A localized crossing is then said to occur if the locally detrended series $x_t^{(m)}$ crosses zero. The normalized number of local level crossings can then be defined

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

7
as:

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

$m = 1..(T-1)$.

Equivalently, it can be expressed in following way:

$$K_T^{(m)}(0) = T^{-1/2} \sum_{t=1}^{T-1} 1 \left[x_t^{(m)} x_{t+1}^{(m)} < 0 \right], \quad m = 1..(T-1). \quad (10)$$

We next formalize the relation between this localized level crossing statistic and the crossing statistic of Garcia and Sanso [13], which holds as a special case of (10), in which $m = T - 1$.

**Proposition 1.** For $m = T - 1$ the process in (8) becomes $x_{t}^{(T-1)} = x_t - x_1 - ct$, with $c = \frac{x_T - x_1}{T}$. Then (10) is identical to the normalized number of level crossings, $K_T^{(GS)}$, defined in Garcia and Sanso [13].

**Proof** See appendix. \qed
2.1. Asymptotic properties of the centered partial sum in the absence of the deterministic component

We first establish the asymptotic properties for the localized level crossing statistic when applied to 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, provided that the detrending is sufficiently localized. Similar to equations (8) and (10), we define

\[
u_t^{(m)} = u_t - u_{t-m} - \frac{u_{t+m} - u_{t-m}}{2}, \quad \forall t = 1..T, \tag{11}
\]

where \( m \) is a fixed integer less than \( T \) and

\[
K_{u,T}^{(m)}(0) = T^{-1/2} \sum_{i=1}^{T-1} \mathbf{1}[u_t^{(m)} u_{t+1}^{(m)} < 0], \quad m = 1..(T-1). \tag{12}
\]

Assumptions  Many of the results below rely on one or more of the following assumptions:

\( \mathcal{A}.1 \) \( \varepsilon_t \) is an i.i.d. process satisfying \( \mathbb{E}[\varepsilon_t] = 0 \) and \( \mathbb{E}[\varepsilon_t^2] = \sigma^2_\varepsilon < \infty \).

\( \mathcal{A}.2 \) \( \varepsilon_t \) is a white noise process satisfying

\[
\mathbb{E}[\varepsilon_t] = 0, \quad \mathbb{E}[\varepsilon_t^2] = \sigma^2_\varepsilon < \infty, \quad \mathbb{E}[\varepsilon_t \varepsilon_s] = 0, \forall t \neq s. \tag{13}
\]

\( \mathcal{A}.3 \) \( \varepsilon_t \) is a strictly stationary ellipsoidal process, i.e. its finite dimensional distributions have an elliptically symmetric density of the form \( f(\varepsilon) = |\Sigma|^{-\frac{1}{2}} \phi(\varepsilon^T \Sigma^{-1} \varepsilon) \), where \( \Sigma \) is the characteristic matrix and is proportional to \( \text{Cov}(\varepsilon) \). The functional form of \( \phi(\cdot) \) is assumed to be unknown.

\( \mathcal{A}.4 \) \( m \) is a fixed positive integer.

\( \mathcal{A}.5 \) \( p \) is a fixed positive integer.

\( \mathcal{A}.6 \) The characteristic function of \( \varepsilon_t, \phi(t) = \mathbb{E}[\exp(t\varepsilon \sqrt{-1})] \) satisfies \( \int_{-\infty}^{\infty} |\phi(t)|^r dt < \infty \) for some \( r \in \mathbb{N} \).

The i.i.d. assumption on the innovations in (\( \mathcal{A}.1 \)) is required in existing central limit theorem results on level crossings for linear processes (see for example Ho and Hsing [14], Hsing [20], Ho [16], Wu [38, 39], and Zhao and Wu [40]). (\( \mathcal{A}.1 \)) can be replaced by the white noise assumption (\( \mathcal{A}.2 \)) for most of our other results. The assumption in (\( \mathcal{A}.3 \)) of elliptically symmetric innovations is required in order to employ the arcsine law to calculate the expected crossing rates. Some of the examples of processes satisfying (\( \mathcal{A}.3 \)) include Gaussian, Student t, logistic, Laplace, and scaled mixed normal distributions, such as exponential power family distributions (see Fang et al. [11], Tanaka and Shimizu [36] for further discussion on symmetric type distributions). Therefore, assumption (\( \mathcal{A}.3 \)) allows for parametric distributions widely used to model fat tails. On the other hand (\( \mathcal{A}.3 \)) does not allow for asymmetries. The assumption of elliptical symmetry has been employed previously in areas of econometric theory (Hodgson [19], Brown and Hodgson [6]), asset pricing (Hodgson et al. [18],
Hodgson and Vorkink [17]) and portfolio choice theory (see for example Owen and Rabinovitch [29]). Assumption (A.4) ensures that the detrending is properly localized by preventing \( m \) from growing with the sample size. The large sample behavior can change substantially when this assumption is violated — for the special case of \( m = T - 1 \) see Garcia and Sanso [13]. Assumption (A.5) requires the number of breaks to be finite. Assumption (A.6) is required for asymptotic normality in Wu [38]’s central limit theorem for functionals of linear processes.

We next derive the autocorrelation function for the locally detrended series \( u_t^{(m)} \) in (11).

**Proposition 2.** Provided that (A.2) holds and given the process \( u_t \) defined in (1), the autocorrelation function, \( r_{u,h}^{(m)} \equiv \gamma_h^{(m)}/\gamma_0^{(m)} = \mathbb{E} \left[ u_t^{(m)} u_{t+h}^{(m)} \right]/\mathbb{E} \left[ u_t^{(m)} u_t^{(m)} \right] \), for the process in (11) when \( m < t < T - m \) is given by

\[
\begin{align*}
\rho &< 1: \\
r_{u,h}^{(m)} &= \begin{cases} 
1 & h = 0 \\
\rho^h + \frac{3\rho^h - 2m - h}{2(1 - \rho^m)(3 - \rho^m)} & h = 1..m \\
\rho^h + \frac{3\rho^h - 4m - h}{2(1 - \rho^m)(3 - \rho^m)} & h = m + 1..2m - 1 \\
\rho^h + \frac{3\rho^h - 4m - h}{2(1 - \rho^m)(3 - \rho^m)} & h \geq 2m
\end{cases}
\end{align*}
\]

for \( \rho = 1: \\
r_{u,h}^{(m)} &= \begin{cases} 
1 & h = 0 \\
\frac{2m - 3h}{2m} & h = 1..m \\
\frac{m - h}{2m} & h = m + 1..2m - 1 \\
0 & h \geq 2m.
\end{cases}
\]

**Proof** See appendix. \( \square \)

We next solve for the population crossing rate in terms of the first order autocorrelation.

**Proposition 3.** Provided that (A.2)-(A.4) hold, the average number of zero-crossings for the process \( u_t^{(m)} \) as defined by (11) is:

\[
\mathbb{E} \left( 1 \left[ u_t^{(m)} u_{t+1}^{(m)} < 0 \right] \right) = \frac{1}{\pi} \cos^{-1} r_{u,1}^{(m)}. \tag{14}
\]

This result is an application of the arcsine formula for zero mean symmetric elliptical processes given by Barnett [5] and Tanaka and Shimizu [36], which is a generalization of Rice’s formula for Gaussian processes. The crossing rate
in (14) does not depend on $\sigma_x$, since we only consider crossings about zero, the mean of $u_t^m$, which is also the median due to the symmetry assumption in (A.3). When constructing the test statistic later in this section, we will require an explicit expression for the population crossing rate under the null hypothesis of a random walk ($\rho = 1$).

**Corollary 4.** Provided that (A.2)-(A.4) hold, under the null hypothesis that $\rho = 1$ in (1), the autocorrelation function is given by $r_{u,1}^{(m)} = (2m - 3)/m$ and the average crossing rate for the process $u_t^{(m)}$ equals $\frac{1}{2} \cos^{-1} \left( \frac{2m - 3}{m} \right)$.

The result of Corollary 4 is an immediate application of Propositions 2 and 3.

![Figure 3](image-url)

**Figure 3:** Level crossing rates, $E \left( 1 \left[ u_t^{(m)} u_{t+1}^{(m)} < 0 \right] \right)$, for a set of alternative hypotheses $\rho \in [0,1]$ and for selected values of the localization parameter $m$.

Figure 3 depicts average crossing rates as functions of $\rho$ for several values of $m$: from a minimum $m = 1$ to a maximum $m = T - 1$ with several cases in between. It is apparent, that higher values of $m$ result in a greater distinction between the null ($\rho = 1$) and alternative ($0 \leq \rho < 1$) hypotheses. To make the point more transparent, consider Figure 4, where we have plotted the average crossing rate as a function of $m$ for $\rho = 1$ and $\rho = 0.85$.

The following result establishes asymptotic normality for the normalized number of crossings.
Figure 4: Level crossing rates as a function of the localization parameter $m$, 
$\mathbb{E}\left(1 \left[u_i^{(m)} u_{i+1}^{(m)} < 0\right]\right)$, under both the null hypotheses of random walk $\rho = 1$ and a single alternative $\rho = 0.85$.

**Proposition 5.** Provided that $(\mathcal{A}.1)$, $(\mathcal{A}.4)$ and $(\mathcal{A}.6)$ hold,

$$K_{u,T}^{(m)}(0) - \mathbb{E}\left[K_{u,T}^{(m)}(0)\right] \overset{d}{\rightarrow} N\left(0, \omega_{1_u,1_u}^{(m)}\right)$$

for some asymptotic variance $\omega_{1_u,1_u}^{(m)} < \infty$.

**Proof** See appendix. \hfill \Box

This result follows as an application of Wu [38] who provides a central limit theorem for functionals of linear processes with i.i.d. innovations, extending the earlier results of Ho and Sun [15] for non-instantaneous filters of a stationary Gaussian processes.

Under the restriction of the null hypothesis ($\rho = 1$), we next show that $\omega_{1_u,1_u}$ can be expressed in terms of the long-run variance of the crossing indicator $1 \left[u_i^{(m)} u_{i+1}^{(m)} < 0\right]$.

**Proposition 6.** Provided that $(\mathcal{A}.1)$, $(\mathcal{A}.4)$ and $(\mathcal{A}.6)$ hold and under the null hypothesis that $\rho = 1$, the asymptotic variance in Proposition 5 is given by

$$\omega_{1_u,1_u}^{(m)} = \sum_{h=-\infty}^{\infty} \gamma_{1_1(h)}$$

where $\gamma_{1_1(h)} = Cov\left(1 \left[u_i^{(m)} u_{i+1}^{(m)} < 0\right], 1 \left[u_{i+h}^{(m)} u_{i+1+h}^{(m)} < 0\right]\right)$.

**Proof** See appendix. \hfill \Box
Remark Under the null hypothesis, when $m = 1$ and \((\mathcal{A}.2)\) and \((\mathcal{A}.3)\) hold, \(\omega_{1_u,1_u}^{(m)} = \frac{s}{\pi}\), can be calculated using binomial probabilities.\(^2\) However, as discussed in Sinn and Keller [35], explicit calculations are no longer available for $m > 1$, since it involves estimation of $\gamma_1(h)$ for any $h > 1$. In the Gaussian case, it follows from Ho and Hsing [14] that

\[
\omega_{1_u,1_u}^{(m)} = \sum_{j=k}^{\infty} \frac{h_j}{j!} \sum_{s=-\infty}^{\infty} \left( \frac{\gamma_1(u,s)}{s} \right)^j
\]

(15)

where $h_j = \int 1[x < 0] H_j(x) d\Phi(x)$ is a truncated Hermite expansion of the indicator function with power rank $k$ and $H_j(x) = (-1)^j e^{x^2/2} \left( d^j e^{-x^2/2}/dx^j \right)$. However, the use of this formula still entails numerical approximation. An alternative numerical approximation is provided by Sinn and Keller [35]. In light of these complications, we instead employ a standard heteroskedasticity autocorrelation consistent (HAC) variance estimator $\hat{\omega}_{1_u,1_u}^{(m)}$ to estimate $\omega_{1_u,1_u}^{(m)}$.

With these results in hand, we provide a $t$-type test statistic with a standard normal null asymptotic distribution.

**Corollary 7.** Let $\hat{\omega}_{1_u,1_u}^{(m)}$ be a consistent estimator of $\omega_{1_u,1_u}^{(m)} > 0$. Provided that \((\mathcal{A}.1)\), \((\mathcal{A}.3)\), \((\mathcal{A}.A)\) and \((\mathcal{A}.6)\) hold, under the null hypothesis that $\rho = 1$ in (1),

\[
i_u^{(m)} = \frac{K_{a,T}^{(m)}(0) - \sqrt{T} \cdot \frac{1}{\pi} \cos^{-1} \left( \frac{2m-3}{m} \right) d}{\sqrt{\hat{\omega}_{1_u,1_u}^{(m)}}} \sim \mathcal{N}(0,1).
\]

(16)

What follows from the above corollary is that in the absence of a deterministic component, one could employ (16) to test the random walk null hypothesis (without deterministic component). One should note, however, that this approximation, based on the assumption \((\mathcal{A}.A)\) that $m$ is fixed, deteriorates as $m$ increases relative to $T$ in finite samples. For example, when $m = T - 1$ Assumption \((\mathcal{A}.A)\) is not met. In this case, using Proposition 1 and the normalized number of zero crossings defined in (10) one can directly apply the result of Garcia and Sanso [13] to show that the underlying test statistic will follow the standard Rayleigh distribution.

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 $\mu_t$ is a linear trend, without breaks, $x_t^{(m)} = u_t^{(m)}$. Therefore the statistics discussed above are numerically unchanged with the addition of a linear trend. Intuitively,

\(^2\)Details of the calculations are available upon request.
$x_t^{(m)}$ eliminates the trend, via a gradual differencing.

We next show that when $\mu_t$ is a linear trend with a finite number of breaks $x_t^{(m)}$ and $u_t^{(m)}$ differ at only $\max(2m + 1)\mu$ number of points. Because the level crossing statistics involve bounded indicator functions, this implies a tight bound on the discrepancy between the statistics based on the processes with and without the structural breaks.

Substituting (2) into (8) yields:

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

Equivalently,

$$x_t^{(m)} = u_t - u_{t-m} - \frac{u_{t+m} - u_{t-m}}{2} + DC_t = u_t^{(m)} + DC_t, \quad (17)$$

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 based on (10) will be invariant to the breaks and trend/level parameters as long as $DC_t = 0$.

For the simple case with no structural changes in the level or linear trend, $\mu_t = \alpha_0 + \beta_0 t$, we can show that $DC_t = (1 - \rho) \beta_0 m - \frac{2 \beta_0 m - 2 \beta_0 m^2}{2} = 0$ for any values of $\alpha_0, \beta_0, m$ and any sample size $T$. And thus the test statistics based on (10) will be numerically equivalent to the test statistics based on (12), irrespective of the magnitude of the level and trend parameters.

This exact equivalence no longer holds in the presence of trend-breaks. However, the break only affects those local crossing levels for which the break point is included in the local detrending. This bounds the rate of discrepancy between $1 \left[ u_t^{(m)} u_{t+1}^{(m)} < 0 \right]$ and $1 \left[ x_t^{(m)} x_{t+1}^{(m)} < 0 \right]$.

**Proposition 8.**

$$\left| K_T^{(m)} - K_{u,T}^{(m)} \right| \leq T^{-1/2} (2m + 1) \mu.$$
Proof See appendix.

Under Assumptions (A.4) and (A.5), both $m$ and $p$ are fixed and therefore $\lim_{T\to\infty} (2m + 1)p/\sqrt{T} = 0$ and $K_T^{(m)}$ is asymptotically equivalent to $K_{u,T}^{(m)}$. If we allow $m$ or $p$ to grow with the sample size, $K_T^{(m)}$ and $K_{u,T}^{(m)}$ will be asymptotically equivalent provided that $mp = o_p(T^{1/2})$. This implies robustness of the test to the presence of breaks.

An important case for which the equivalence of $K_T^{(m)}$ and $K_{u,T}^{(m)}$ is no longer implied by Proposition 8 is when the process is detrended globally by setting $m = T - 1$ as in Garcia and Sanso [13]. This is because $m$ grows too quickly as a function of $T$ for the bound in Proposition 8 to be meaningful.

3. Bootstrap selection of $m$ and bootstrap critical values

The choice of $m$ involves two countervailing influences on power: robustness to breaks and discrimination between the null and alternative hypotheses in the absence of breaks. Smaller values of $m$ will result in a random walk test robust to the presence of multiple structural breaks and their magnitudes. Figure 5 presents empirical rejection rates of the LLC test for a user specified parameter $m$. To illustrate the point, several values of $m$ were selected and several single break alternative hypotheses were considered (with the magnitude of the break, $\beta$, on the x-axis). As can be seen from Figure 5, the LLC test with $m = 1$ or $m = 2$ is robust to the magnitude of the break, but suffers from low power. At the same time, very large values of $m$ (e.g. $m = T - 1$) result in higher rejection rates in the absence of a break. The power drops quickly as the magnitude of the break increases, giving rise to the inverted U-shapes seen in Figure 5.

Figures 6 and 7 show sensitivity of the test power to changes in break location, $\lambda$, for a range of the localization parameters, $m$, for sample sizes of $T = 100$ (Figure 6) and $T = 500$ (Figure 7). The test is robust to the location of the break for smaller values of $m$, but experiences significant power loss for larger values of $m$ when the break is located close to the middle of the series.

On the other hand, in the absence of breaks, a reduction in $m$ leads to a reduction in test power due to the smaller difference between the average number of crossing rates under the null and alternative hypothesis (see Figure 4). 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 $m$. This is observed in the Figures 5-7 discussed above. The straight lines towards the top of Figure 5 illustrate a range on intermediate values of $m$ for which the test shows good power regardless of the break magnitude ($\beta$s). Likewise, the surfaces in Figures 6 and 7 show a range of intermediate values for $m$ that result in good power regardless of the
break location.

Figure 5: Sensitivity of the test power to the magnitude of the structural break for different choices of the localization parameter $m$. The $y$-axis shows rejection rates under the alternative that $\rho = 0.5$ for $T = 500$, Gaussian errors and $A = [0; 0]$, $B = [0; \beta_1]$ and $\lambda = [0; 0.5]$. The magnitude of the break, $\beta_1$, is varied along the $x$-axis.

We suggest a data driven method of selecting the parameter $m$, based on a finite sample power criterion. Since the finite sample power depends 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.\footnote{We refer to this as a 'proximate' model because we use it to provide an approximation rather than assuming it to be a true model. The better the approximation, the better the choice of $m$ and the more powerful the test. However, the test remains valid even if the approximation is poor and $m$ is chosen sub-optimally. From a pragmatic perspective, one hopes that the proximate model will deliver a reasonable intermediate value of $m$ even if it is misspecified or suffers from estimation error.} For our proximate model we employ an AR(1) with resampled estimated errors and a maximum
Figure 6: Sensitivity of the test power to changes in break location, $\lambda$, for different choices of the localization parameter $m$. Constructed based on the alternative model with a single break in trend using $\rho = 0.8$, $A = [0; 0]$, $B = [0; 5]$, $\lambda = [0; \lambda_1]$ and $T = 100$, where we allow $\lambda_1$ to vary in $[0.05, .95]$. Computed using 5000 replications.

The estimation of the proximate model and the selection of $m$ adds additional estimation error not accounted for in the asymptotic distributions based on the assumption that $m$ is non-random. Selecting $m$ to maximize power may
Figure 7: Sensitivity of the test power to changes in break location, $\lambda$, for different choices of the localization parameter $m$. Constructed based on the alternative model with a single break in trend using $\rho = 0.8$, $A = [0; 0]$, $B = [0; 5]$, $\lambda = [0; \lambda_1]$ and $T = 500$, where we allow $\lambda_1$ to vary in $[0.05, .95]$.) Computed using 5000 replications.

also lead to size distortion. To address these issues, we attempt to control the overall test size using bootstrapped critical values in which we re-estimate the proximate model and re-select $m$ for each bootstrap replication. We perform this bootstrap by resampling the residuals of the following restricted version of our proximate model

$$x_t = x_{t-1} + \hat{\mu}_t - \hat{\mu}_{t-1} + \hat{\epsilon}_t$$
$$\hat{\epsilon}_t = x_t - x_{t-1} - (\hat{\mu}_t - \hat{\mu}_{t-1})$$
$$\hat{\mu}_t = \hat{D}_U \cdot \hat{A} + \hat{D}_T \cdot \hat{B},$$

in which we enforce the null hypothesis by setting $\rho = 1$ in equation (2). In this way we attempt to capture the effect of endogenously selecting $m$ in our bootstrap replications.

To be specific, we use the following procedures to select $m$ and to select critical values in order to control overall test size:
3.1. Procedure to select $m^*$

1. Given the data, we estimate the parameters of the unrestricted proximate model described by equations (2) and (7) to obtain the fitted versions of both the unrestricted, (18), and restricted, (19), models.

2. Then, for every $m = 1, ..., T - 1$ we:
   
   (a) Bootstrap $M$ series of length $T$ from the restricted model (19) where we set $\rho = 1$ and the other parameters are estimated in Step (1), and where $\{\tilde{\varepsilon}\}$ are sampled with replacement from $\{\hat{\varepsilon}\}$.
   
   (b) Calculate the normalized number of crossings, $K_T^{(m)}(0)$ for each of the $M$ series in Step (2a).
   
   (c) Find a one tailed $(1 - \alpha)$ critical value for $K_T^{(m)}(0)$ based on the $\alpha$ quantile of the bootstrap replicates $K_T^{(m)}(0)$.
   
   (d) Bootstrap $M$ series from the unrestricted model using (18) estimated in Step (1), where $\{\tilde{\varepsilon}_t\}$ are sampled with replacement from $\{\hat{\varepsilon}_t\}$ of the same length.
   
   (e) Form the test statistic for each bootstrap replicate and using the critical values in Step (2c), find the rejection rate.

3. Choose $m^*$ corresponding to the value of $m$ with largest rejection rate in Step (2e).

3.2. Selection of critical values to control overall test size

1. Given the data, we estimate parameters of the unrestricted proximate model described by equation (18).

2. Bootstrap $M$ series from the restricted model (19) where we impose the null hypothesis by setting $\rho = 1$. The other parameters are estimated from the unrestricted model in Step (1) and $\{\tilde{\varepsilon}_t\}$ are sampled with replacement from $\{\hat{\varepsilon}_t\}$ of the same length.

3. Select $m^*$ using the procedure in Section 3.1 for each of the $M$ series in (2). Then calculate the normalized number of crossings, $K_T^{(m^*)}(0)$ for each of the $M$ series.

4. Find $(1 - \alpha)$ critical value (one tailed) using the $\alpha$ quantile of the $M K_T^{(m^*)}(0)$ replicates in Step (3).

4. Finite Sample Size and Power

In this section we conduct a small sample simulation study in order to assess the empirical size and power of our test. We use (2) and (7) as the data generating process in our simulations. The sample size is set to 250. We consider three cases: (i) a case without a break where $A = [0]$, $B = [0]$, $\lambda = [0]$; (ii) a case where the break in the trend slope is present in the middle of the series,
for which \( A = [0; 0], B = [0; 0.1], \lambda = [0; 0.5] \); (iii) and a case with two breaks in the trend slope, equally spaced at \( 1/3T \) and \( 2/3T \), specifically \( A = [0; 0; 0], B = [0; 0.1; 0.2], \lambda = [0; 0.33; 0.66] \).

The performance of the Localized Level Crossing test (LLC) proposed above is compared to the performance of the Garcia and Sanso [13] testing procedure (GS) which is not adapted for structural breaks. We also compare to the Zivot and Andrews [41] testing procedure (ZA), which allows for a single endogenous break under the alternative hypothesis. Finally, we report the size and power for the Augmented Dickey-Fuller test as an additional benchmark.\(^5\) We compute our results using 5,000 simulations for the ADF, ZA and GS tests. For the LLC test we employ only 1,000 simulations, since it is more computationally demanding. Within the LLC test procedure we use 1,000 replications for the inner bootstrap used to select \( m^\star \) and 1,000 simulations for the outer bootstrap procedure used for selecting critical values to control overall test size.

The results from our baseline simulations using Gaussian innovations are shown in Table 1. Our simulation experiments show that, in the absence of any structural breaks, the ADF test outperforms all of the other tests, as expected (columns 2-4). The LLC test outperformed the GS test in this scenario. We conjecture that this improvement is due to the fact that the LLC test selects \( m^\star \) to maximize power.

Both the GS and the ADF test did not result in good power when considered for cases with one or two structural breaks.\(^6\) The ZA test, constructed to account for a single endogenous structural break, performed well in the case with one break, but experienced a large power loss and size distortion for the case with two structural breaks. The LLC test did experience some power loss every time a break was added to the DGP. However, the power loss was smaller than that of the ZA procedure when comparing one- and two-break scenarios. The LLC also maintained a relatively constant size across all three cases. Thus, the LLC random walk test proposed in this paper offers an improvement over commonly used methods in moderately sized samples for cases with two or more breaks. Likewise, it may be useful in cases for which the number of breaks is unknown.

The preceding simulations employed i.i.d. Gaussian errors. It is also important to assess the finite sample accuracy of the test under distributions with fatter tails, as this is a common feature of many economic and financial time series. Similarly, in deriving the null asymptotics we required both higher moment independence (\( \mathcal{A}.1 \)) and elliptical symmetry (\( \mathcal{A}.3 \)). We would also like to study

---

\(^5\)One could also consider a recently proposed tests by Aparicio et al. [1] or Kim and Perron [23] which are known to be robust to structural changes.

\(^6\)In fairness, neither of these tests were designed to accommodate structural breaks. As noted above, they both performed quite well for the cases that they were designed to address.
Table 1: Empirical rejection rates under the null and alternative hypotheses

<table>
<thead>
<tr>
<th>DGP(^a)</th>
<th>No Break(^{(i)})</th>
<th>Single break in the middle(^{(ii)})</th>
<th>Two breaks(^{(iii)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\rho)</td>
<td>LLC(^b) GS(^c) ADF-A(^d)</td>
<td>LLC GS ADF-B(^e) ZA-B(^f)</td>
<td>LLC GS ADF-B ZA-B</td>
</tr>
<tr>
<td>1.00</td>
<td>0.06 0.05 0.06</td>
<td>0.07 0.03 0.05 0.05</td>
<td>0.04 0.01 0.02 0.09</td>
</tr>
<tr>
<td>0.98</td>
<td>0.11 0.10 0.10</td>
<td>0.07 0.05 0.07 0.08</td>
<td>0.06 0.00 0.01 0.09</td>
</tr>
<tr>
<td>0.97</td>
<td>0.14 0.14 0.14</td>
<td>0.09 0.06 0.08 0.09</td>
<td>0.07 0.00 0.01 0.10</td>
</tr>
<tr>
<td>0.95</td>
<td>0.24 0.24 0.29</td>
<td>0.16 0.07 0.11 0.14</td>
<td>0.13 0.00 0.00 0.14</td>
</tr>
<tr>
<td>0.90</td>
<td>0.58 0.57 0.79</td>
<td>0.38 0.07 0.20 0.49</td>
<td>0.29 0.00 0.00 0.31</td>
</tr>
<tr>
<td>0.88</td>
<td>0.70 0.68 0.90</td>
<td>0.49 0.06 0.22 0.67</td>
<td>0.40 0.00 0.00 0.38</td>
</tr>
<tr>
<td>0.87</td>
<td>0.75 0.71 0.94</td>
<td>0.53 0.06 0.24 0.75</td>
<td>0.48 0.00 0.00 0.44</td>
</tr>
<tr>
<td>0.85</td>
<td>0.82 0.77 0.96</td>
<td>0.63 0.05 0.27 0.88</td>
<td>0.61 0.00 0.00 0.57</td>
</tr>
<tr>
<td>0.80</td>
<td>0.91 0.86 0.99</td>
<td>0.80 0.05 0.29 0.99</td>
<td>0.78 0.00 0.00 0.69</td>
</tr>
</tbody>
</table>

Table entries show empirical rejection rates under the null (\(\rho = 1\)) and alternative (\(\rho < 1\)) hypotheses for a nominal five percent test. Computed using 5,000 replications for the ADF, ZA and GS tests and 1,000 replications for LLC test. Sample size is 250. The data is generated by equations (2) and (7) with:

\((i)\) \(A = \{0\}, B = \{0\}, \lambda = \{0\};\)

\((ii)\) \(A = \{0; 0\}, B = \{0; 0.1\}, \lambda = \{0; 0.5\};\)

\((iii)\) \(A = \{0; 0; 0\}, B = \{0; 0.1; 0.2\}, \lambda = \{0; 0.33; 0.66\}.\) In all cases the residuals are \(N(0, 1).\)

\((b)\) Localized Level Crossing (LLC) random walk test.


\((d)\) Augmented Dickey-Fuller test with no trend or constant.

\((e)\) Augmented Dickey-Fuller test with trend and no constant.

\((f)\) Zivot and Andrews [41] test with trend parameter.
Table 2: Empirical crossing rates for different error distributions (no break case)

<table>
<thead>
<tr>
<th>Errors</th>
<th>m = 1</th>
<th>m = 2</th>
<th>m = 3</th>
<th>m = 5</th>
<th>m = 10</th>
<th>m = 25</th>
<th>m = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N(0, 1)$</td>
<td>0.667</td>
<td>0.419</td>
<td>0.333</td>
<td>0.252</td>
<td>0.177</td>
<td>0.111</td>
<td>0.077</td>
</tr>
<tr>
<td>Student $-t_5$</td>
<td>0.666</td>
<td>0.417</td>
<td>0.329</td>
<td>0.247</td>
<td>0.171</td>
<td>0.106</td>
<td>0.075</td>
</tr>
<tr>
<td>GARCH $(1, 1)^a$</td>
<td>0.660</td>
<td>0.420</td>
<td>0.338</td>
<td>0.257</td>
<td>0.179</td>
<td>0.112</td>
<td>0.078</td>
</tr>
<tr>
<td>EGARCH $(1, 1)^b$</td>
<td>0.660</td>
<td>0.419</td>
<td>0.338</td>
<td>0.257</td>
<td>0.177</td>
<td>0.111</td>
<td>0.077</td>
</tr>
</tbody>
</table>

Table entries show empirical crossing rates under the null ($\rho = 1$) hypothesis for different error distributions. Computed using a single replication based on a sample size of $T=1,000,000$. The data is generated by equations (2) and (7) with $A = [0]$, $B = [0]$, $\lambda = [0]$.

(a) GARCH(P,Q) model: $\sigma_t^2 = k + \sum_{i=1}^P G_i \sigma_{t-i}^2 + \sum_{j=1}^Q A_j \varepsilon_{t-j}^2$, where $\varepsilon_t \sim N(0, 1)$. For our simulations we use $P = 1, Q = 1, k = 0.2, G_1 = 0.3, A_1 = 0.5$.

(b) EGARCH(P,Q) model:

$$\log \sigma_t^2 = k + \sum_{i=1}^P G_i \log \sigma_{t-i}^2 + \sum_{j=1}^Q A_j \left[ \frac{\varepsilon_{t-j}}{\sigma_{t-j}} - E \left\{ \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right\} \right] + \sum_{j=1}^Q L_j \left( \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right),$$

where $\varepsilon_t \sim N(0, 1)$. For our simulations we use $P = 1, Q = 1, k = 0.2, G_1 = 0.3, A_1 = 0.5, L_1 = 0.1$.

the behavior of the test in practical situations where these assumptions may be violated. Of particular relevance to financial data are models with conditionally heteroskedastic errors.

Before turning to the test itself, in Table 2 we first examine the null crossing rates on which the LLC test is based. We compare the null crossing rates using standard normal, Student-$t$ (with 5 degrees of freedom), GARCH and EGARCH innovations for different values of $m$.$^7$ For all values of $m$ the simulated null rejection rates for the standard normal and Student-$t$ ($df = 5$) innovation distributions are nearly indistinguishable. Since both belong to the family of elliptically symmetric distributions specified in ($\sigma$.3), this is expected from Proposition 14. More interestingly, crossing rates for both the GARCH and EGARCH models, which lie outside this family, are still quite similar. Therefore, one might consider applying the LLC test to data with conditionally heteroskedastic innovations, even though the test is derived under the assumption of conditional homoskedasticity.

In Table 3 we address this issue more directly by studying the null rejection rates themselves. We consider the same four innovation distributions in both a no break and single break cases and provide a comparison of the LLC test to two benchmarks: the GS and ADF tests. As the table shows, the null rejection rates of the LLC test remain close to the nominal size under all four innovation

---

$^7$In this particular table we use a very large value of $T$ because we are interested in examining the null population crossing rates used to center the test statistic.
Table 3: Empirical rejection rates under the null hypothesis for different error distributions

| DGP\(\text{a}\): |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | No Break\((i)\) | Single break in the middle\((ii)\) |
| Errors         | LLC\(\text{b}\) | GS\(\text{c}\) | ADF-A\(\text{d}\) | LLC | GS | ADF-B\(\text{e}\) |
| \(N(0,1)\)    | 0.06 0.05 0.06  | 0.07 0.03 0.05  |                 | 0.06 | 0.05 | 0.06  |
| Student \(t_5\)| 0.06 0.06 0.054 | 0.06 0.03 0.06  |                 | 0.066 | 0.02 | 0.064 |
| GARCH\((1,1)\) | 0.05 0.06 0.048 | 0.06 0.03 0.06  |                 | 0.07  | 0.01 | 0.075 |
| EGARCH\((1,1)\) | 0.05 0.053 0.046 | 0.07 0.01 0.075 |                 |       |      |        |

Table entries show empirical rejection rates under the null \((\rho = 1)\) hypothesis for a nominal five percent test. Computed using 5,000 replications for the ADF and GS tests and 1,000 replications for LLC test. Sample size is 250. The data is generated by equations (2) and (7) with:

(i) \(A = [0]; \ B = [0]; \ \lambda = [0]\);

(ii) \(A = [0; 0]; \ B = [0; 0.1]; \ \lambda = [0; 0.5]\);

(b) Localized Level Crossing (LLC) random walk test.


d Augmented Dickey-Fuller test with no trend or constant.

e Augmented Dickey-Fuller test with trend and no constant.

(f) GARCH\((P,Q)\) model: \(\sigma_t^2 = k + \sum_{i=1}^{P} G_i \sigma_{t-i}^2 + \sum_{j=1}^{Q} A_j \varepsilon_{t-j}^2\), where \(\varepsilon_t \sim N(0,1)\). For our simulations we use \(P = 1, Q = 1, k = 0.2, G_1 = 0.3, A_1 = 0.5\).

(g) EGARCH\((P,Q)\) model: \(\log \sigma_t^2 = k + \sum_{i=1}^{P} G_i \log \sigma_{t-i}^2 + \sum_{j=1}^{Q} A_j \left[ \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right] - E \left[ \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right] + \sum_{j=1}^{Q} L_j \left( \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right)^2\), where \(\varepsilon_t \sim N(0,1)\). For our simulations we use \(P = 1, Q = 1, k = 0.2, G_1 = 0.3, A_1 = 0.5, L_1 = 0.1\).
distributions and compare favorably with the several of the benchmark cases. For example, the GS test, which was designed for use in the no break case in which it works quite well, is undersized in the presence of the single break, while the LLC procedure shows consistently correct nominal size.

5. Empirical Application

5.1. Nelson-Plosser macroeconomic data set

Nelson and Plosser [27]’s finding that the unit root could not be rejected in most macroeconomic time series has inspired scholars to re-examine the unit root hypothesis and devise tests with greater power against plausible stationary alternatives, including tests accounting for structural breaks. We apply our proposed test procedure to the Nelson-Plosser data set and compare the results to some representative existing unit root test (See Table 4.).

In addition to the results of the LLC test, we include the results based on the ADF test, the Garcia and Sanso [13]’s test (GS), the Zivot and Andrews [41]’s test (ZA), the Park and Fuller [31]’s weighted symmetric DF test (PF-WS), the Elliott et al. [9]’s GLS test (ERS-GLS) and Perron [32]’s original test (P). The LLC and GS both rely on the level crossings of a process. The LLC, ZA, PF-WS and ERS-GLS test allow for endogenous breaks while the P test allows for exogenously determined breaks.

The LLC test rejected the null hypothesis of a random walk for two series at the five percent significance level: the unemployment rate and industrial production. This result is consistent with findings by Elliott et al. [9], Zivot and Andrews [41] and Perron [32]. Although, in comparison to Perron [32] and Park and Fuller [31], the LLC test failed to reject an additional five series (namely real GNP, per-capita real GNP, nominal wages, real wages and money stock). The ZA test, allowing for a single endogenous break and being the closest alternative to the LLC test, rejects the unit root hypothesis for two additional series, real GNP and nominal GNP. Perron [32]’s original test rejects the unit root null for most of the series, but unlike the LLC and ZA tests it assumes a known a priori break location. However, it should be noted, that the LLC testing procedure, which is designed as a random walk (not a unit root) test, does not account for serial correlations, often present in macroeconomic data.

5.2. Application: Inflation and Interest Rates in Canada

The relationship between the inflation and the interest rates has always been a topic of interest in financial economics. The well known Fisher [12] hypothesis posits a positive one-to-one relationship between the inflation and nominal interest rate. A large literature tests this hypothesis using cointegration methods, under the assumption that the nominal interest and inflation rates are non-stationary (see Mishkin [26], Owen [30] and Evans and Lewis [10]). However,
Table 4: Alternative random walk and unit root tests in Nelson-Plosser time series

<table>
<thead>
<tr>
<th>Series</th>
<th>T</th>
<th>ADF-A</th>
<th>ADF-B</th>
<th>P^d</th>
<th>ZA^c</th>
<th>PF-WS</th>
<th>ERS-GLS</th>
<th>GS^h</th>
<th>LLC^i</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Real GNP</td>
<td>62</td>
<td>***</td>
<td>**</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2  Nominal GNP</td>
<td>62</td>
<td>***</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3  Per-capita real GNP</td>
<td>62</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4  Industrial production</td>
<td>111</td>
<td>***</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5  Employment</td>
<td>81</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6  Unemployment rate</td>
<td>81</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7  GNP deflator</td>
<td>82</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8  Consumer prices</td>
<td>111</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9  Nominal wages</td>
<td>71</td>
<td>***</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Real wages</td>
<td>71</td>
<td>**</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Money stock</td>
<td>82</td>
<td></td>
<td>**</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Velocity</td>
<td>102</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Interest rate</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Common stock prices</td>
<td>100</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table entries show the results of the representative test when performed on Nelson-Plosser series: *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

(b) ADF test without constant and trend;
(c) ADF test with a trend and no constant;
(d) Perron [32]'s test;
(e) Zivot and Andrews [41]'s test with trend parameter;
(f) Park and Fuller [31]'s WS test;
(g) Elliott et al. [9]'s GLS test;
(h) Garcia and Sanso [13]'s test;
(i) Localized Level Crossing (LLC) random walk test.
evidence of I(1) behavior could be due to the presence of structural breaks, in which case cointegration tests might not be the most appropriate method to test the Fisher hypothesis.

Atkins and Chan [2], using methodology of Zivot and Andrews [41], Lumsdaine and Papell [25] and Banerjee et al. [4] finds evidence that 90-day Treasury Bill rate and the inflation rate in Canada are stationary series with two breaks. Using the Zivot and Andrews [41] procedure, which only allows for a single break, the authors find the breaks in inflation and interest rates at 1980Q3 and 1972Q3 respectively. Using Lumsdaine and Papell [25] procedure, and allowing for up to two breaks in both intercept and trend (model CC in Lumsdaine and Papell [25]), they find two breaks in inflation at 1972Q3 and 1982Q2 and in the interest rates at 1980Q3 and 1988Q1.

Figure 8 traces the path of Canadian inflation for the period of 1954Q1-2007Q2, showing what are often referred to as the twin peaks of Canadian inflation, when inflation rose to over 14 per cent in 1973 and to almost 13 per cent in 1979. The chart also shows the sharp decline in inflation during the early 1980s, from over 12 per cent to 4 per cent, the modest increase from 4 per cent to 6 per cent over the balance of the 1980s, and then the reduction from just over 6 per cent in 1990 to roughly 2 per cent over the following few years. From 1992 to 2007, the rate of CPI inflation in Canada has fluctuated around 2 per cent. Figure 9 shows nominal interest rate with a peak value of just above 20 per cent in the early ’80s.

Using the LLC procedure, we conduct a random walk test on the interest and inflation rates, that is robust to the potential breaks identified by the Lumsdaine and Papell [25]’s procedure. We use the Bank of Canada 90-day Treasury bill rate as a proxy for the nominal interest rate. We use CPI on all items for the year 2000 basket as a proxy for calculating nominal inflation. The data consists of quarterly observations for the period of 1954Q1-2007Q2 corresponding to a sample size of 216. We reject the null hypothesis of a random walk for both the inflation and interest rates. In contrast, the ADF and GS tests shows that both of these variables contain a unit root. The ZA-C test rejects the unit root hypothesis for the interest rate but fails to reject it for the inflation rate (perhaps due to the presence of the third break in that series).

In selecting the optimal localization parameter, \( m \), we estimated a proximate model in which the number of breaks was selected by the Bai and Perron [3] procedure. Interestingly, it identified three breaks in the inflation rate, corresponding to 1972Q3, 1982Q3 and 1990Q4. The first two breaks are also detected by Lumsdaine and Papell [25]’s procedure. The location of these breaks is consistent with the downturn in the Canadian economy in the early 1980s and the Bank of Canada’s change in it’s monetary policy in late 1979. The break in the early 1970’s could, perhaps, be attributed to the break down of the Bretton Woods system. The third break corresponds to a sharp reduction in inflation corresponding to the introduction of inflation targeting by the Bank
Table 5: Results of the LLC test applied to Canadian data

<table>
<thead>
<tr>
<th></th>
<th>Inflation</th>
<th>Interest Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LLC test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t^{(m)}$</td>
<td>7.2673$^*$</td>
<td>2.0327$^*$</td>
</tr>
<tr>
<td>LLC p-value</td>
<td>0.000</td>
<td>0.044</td>
</tr>
<tr>
<td>$m^*$</td>
<td>6</td>
<td>31</td>
</tr>
<tr>
<td>$\hat{\omega}_{1,1}$</td>
<td>0.1422</td>
<td>0.1302</td>
</tr>
<tr>
<td>$t^{(m)}_{95}$</td>
<td>2.2801</td>
<td>1.8756</td>
</tr>
<tr>
<td>$\hat{\rho}$</td>
<td>0.1260</td>
<td>0.8383</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1972Q3, 1982Q3, 1990Q4$^b$</td>
<td>1981Q3, 1990Q4$^b$</td>
</tr>
</tbody>
</table>

Comparison results from other tests

<table>
<thead>
<tr>
<th></th>
<th>Inflation</th>
<th>Interest Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF-B p-value</td>
<td>0.1934</td>
<td>0.3871</td>
</tr>
<tr>
<td>AIC lag</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>ZA-B statistic$^c$</td>
<td>-3.54</td>
<td>-4.13</td>
</tr>
<tr>
<td>ZA-C statistic$^c$</td>
<td>-5.00</td>
<td>-5.15$^*$</td>
</tr>
<tr>
<td>GS$^d$</td>
<td>1.6367</td>
<td>1.1412</td>
</tr>
</tbody>
</table>

Table entries show estimated parameters of a proximate model as defined in (18).

(a) Within the LLC test procedure we use 1,000 replications for the inner bootstrap used to select $m^*$ and 1,000 simulations for the outer bootstrap procedure used for selecting critical values, $t^{(m)}$.

(b) Estimated locations of the break in the series.

(c) The 5% critical values for the ZA-B test is -4.42; for the ZA-C test it is -5.08, (Zivot and Andrews [41], Table 4, pp.256-257).

(d) The 5% critical value for the GS test is 2.45.

of Canada in 1991. Two breaks were found in the interest rate at 1981Q3 and 1990Q4. These breaks are marked in figures 8 and 9. The optimal values of $m$ selected by the LLC procedure were 6 and 31 for the inflation and interest rate respectively. This is consistent with our conjecture in Section 3 that the LLC procedure would tend to select a lower value of $m$ for series with a higher number of estimated breaks.

6. Concluding Remarks

In this paper we proposed a new localized level crossing (LLC) random walk test that is robust to an unspecified number, location, and magnitude of breaks in both the level and slope of the deterministic trend. The localization parameter was selected in an inner bootstrap to maximize test power by resampling the residuals from a proximate model. A second outer-bootstrap, in which the localization parameter search was repeated on each bootstrap replicate, was used
Figure 8: Inflation rate (Source: CANSIM, Consumer price index (CPI), 2001 basket content, 1992=100, Canada, quarterly, 1954Q1-2007Q2).

Figure 9: Interest rate (Source: CANSIM, Bank of Canada Rate, quarterly, 1954Q1-2007Q2).
to control the overall test size. Our simulations confirmed that this procedure works well both in controlling for size and in producing good and relatively stable power across models with zero, one or two breaks. Applying our test to Canadian interest rate and inflation data, we are able to reject a random walk hypothesis in both series. This may have implications for tests of the Fisher hypothesis, which are often conducted using cointegration methods.
A. Appendix

A.1. Proofs of Propositions and Lemmas

Proof of Proposition 1 For \( m = T - 1 \) we have:

\[
x_{t-(T-1)} = x_1 + c(t-(T-1))
\]

\[
x_{t+T-1} = x_T + c(t-1)
\]

and therefore

\[
x^{(T-1)}_t = x_t - x_{t-(T-1)} - \frac{x_{t+T-1} - x_{t-(T-1)}}{2}
\]

\[
= x_t - x_1 - c(t-(T-1)) - \frac{x_T + c(t-1) - x_1 - c(t-(T-1))}{2}
\]

\[
= x_t - x_1 - ct - \frac{x_T - x_1 - cT}{2} = x_t - x_1 - ct.
\]

□

Proof of Proposition 2 First, consider the case in which \( |\rho| < 1 \). Using the moving average representation for \( u_t \) we have

\[
u_t = \sum_{j=0}^{\infty} \rho^j \epsilon_{t-m-j}
\]

Then, by recursive backward substitution:

\[
u_t = \rho^m u_{t-m} + \sum_{j=0}^{m-1} \rho^j \epsilon_{t-j} \quad \text{and} \quad u_{t+m} = \rho^{2m} u_{t-m} + \sum_{j=0}^{2m-1} \rho^j \epsilon_{t+m-j}.
\]

This allows us to express \( u^{(m)}_t \) as

\[
u^{(m)}_t = u_t - \frac{1}{2} u_{t-m} - \frac{1}{2} u_{t+m}
\]

\[
= -\frac{1}{2} (\rho^m - 1)^2 \sum_{j=0}^{\infty} \rho^j \epsilon_{t-m-j} + \left(1 - \frac{1}{2} \rho^m \right) \sum_{j=0}^{m-1} \rho^j \epsilon_{t-j}
\]

\[
- \frac{1}{2} \sum_{j=0}^{m-1} \rho^j \epsilon_{t+m-j}.
\]

To simplify notation define \( a = -\frac{1}{2} (\rho^m - 1)^2; \ b = 1 - \frac{1}{2} \rho^m; \ c = -\frac{1}{2} \), then we have the following:

\[
u^{(m)}_t = a \sum_{j=0}^{\infty} \rho^j \epsilon_{t-m-j} + b \sum_{j=0}^{m-1} \rho^j \epsilon_{t-j} + c \sum_{j=0}^{m-1} \rho^j \epsilon_{t+m-j}.
\]
The summands in the expression above can be solved for as follows:

\[ E \left[ \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+h-j} \right] = E \left[ \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m+h-j} \right] = E \left[ \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m+h-j} \right] = 0, \quad \forall h \]

we have the following:

\[
\gamma_h^{(m)} = E \left( u_t^{(m)} u_{t+h}^{(m)} \right) = a^2 E \left[ \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m+h-j} \right] + b^2 E \left[ \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+h-j} \right] + ac E \left[ \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} \sum_{j=0}^{\infty} \rho^j \varepsilon_{t+m+h-j} \right] + bc E \left[ \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+h-j} \right] + c^2 E \left[ \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m+h-j} \right].
\]

The summands in the expression above can be solved for as follows:

\[
E \left[ \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m+h-j} \right] = \sigma_t^2 \sum_{j=0}^{\infty} \rho^j \rho^{j+h}
\]

\[
E \left[ \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} \sum_{j=0}^{\infty} \rho^j \varepsilon_{t+m+h-j} \right] = \begin{cases} 
0 & h = 0 \\
\sigma_t^2 \sum_{j=0}^{h-1} \rho^j \rho^{m-h+j} & \forall h = 1..m \\
\sigma_t^2 \sum_{j=0}^{m-1} \rho^j \rho^{h-j} & \forall h > m
\end{cases}
\]

\[
E \left[ \sum_{j=0}^{m-1} \rho^j \varepsilon_{t-j} \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+h-j} \right] = \begin{cases} 
\sigma_t^2 \sum_{j=0}^{m-1-h} \rho^j \rho^{j+h} & \forall h = 0..m-1 \\
0 & \forall h \geq m
\end{cases}
\]

\[
E \left[ \sum_{j=0}^{m-1} \rho^j \varepsilon_{t+m-j} \sum_{j=0}^{\infty} \rho^j \varepsilon_{t+m+h-j} \right] = \begin{cases} 
0 & \forall h = 0..m-1 \\
\sigma_t^2 \sum_{j=0}^{h-m-1} \rho^j \rho^{2m-h+j} & \forall h = m..2m-1 \\
\sigma_t^2 \sum_{j=0}^{m-1} \rho^j \rho^{h-2m+j} & \forall h \geq 2m
\end{cases}
\]
Using (13) we find
\[
E \left[ \sum_{j=0}^{m-1} \rho^j \epsilon_{t+m-j} \sum_{j=0}^{m-1} \rho^j \epsilon_{t+h-j} \right] = \begin{cases} 
0 & h = 0 \\
\sigma^2 \sum_{j=0}^{h-1} \rho^j \rho^{m-h+j} & h = 1..m \\
\sigma^2 \sum_{j=0}^{h-m-1} \rho^j \rho^{h-m+j} & \forall h = m + 1..2m - 1 \\
0 & \forall h \geq 2m
\end{cases}
\]

The autocorrelation function can then be shown to equal:
\[
E \left[ \sum_{j=0}^{m-1} \rho^j \epsilon_{t+m-j} \sum_{j=0}^{m-1} \rho^j \epsilon_{t+h-m-j} \right] = \begin{cases} 
\sigma^2 \sum_{j=0}^{m-1-h} \rho^j \rho^{h+j} & \forall h = 0..m - 1 \\
0 & \forall h \geq m.
\end{cases}
\]

Next, consider the case when \( \rho = 1 \). Again using the moving average representation and recursive substitution, we have
\[
\begin{align*}
 u_{t-m} &= \sum_{j=0}^{\infty} \epsilon_{t-m-j}, \\
 u_t &= u_{t-m} + \sum_{j=0}^{m-1} \epsilon_{t-j}, \\
 u_{t+m} &= u_{t-m} + \sum_{j=0}^{2m-1} \epsilon_{t+m-j} \\
 u_t^{(m)} &= u_t - \frac{1}{2} u_{t-m} - \frac{1}{2} u_{t+m} = \frac{1}{2} \left( \sum_{j=0}^{m-1} \epsilon_{t-j} - \sum_{j=0}^{m-1} \epsilon_{t+m-j} \right).
\end{align*}
\]

Using (13) we find
\[
\gamma_h^{(m)} \equiv E \left( u_t^{(m)} u_{t+1}^{(m)} \right) = \frac{1}{4} E \left( \sum_{j=0}^{m-1} \epsilon_{t-j} - \sum_{j=0}^{m-1} \epsilon_{t+m-j} \right) \left( \sum_{j=0}^{m-1} \epsilon_{t-h-j} - \sum_{j=0}^{m-1} \epsilon_{t+h+m-j} \right)
\]
\[
= \begin{cases} 
\frac{(2(m-h)-h)}{4} \sigma^2 & \forall h = 0..m \\
\frac{(2m-h)}{4} \sigma^2 & \forall h = m + 1..2m - 1 \\
0 & \forall h \geq 2m.
\end{cases}
\]

Therefore
\[
\gamma_{u-h}^{(m)} = \frac{\gamma_h^{(m)}}{\gamma_0^{(m)}} = \begin{cases} 
1 & h = 0 \\
\frac{2m-3h}{2m} & h = 1..m \\
\frac{m-h}{2m} & h = m + 1..2m - 1 \\
0 & h \geq 2m.
\end{cases}
\]
Proof of Proposition 3 From (21) it can be seen that \( u_t^{(m)} \) is a linear function of \( \varepsilon_t \) and is, therefore, also elliptically symmetric (see Hodgson [19], pp.262-264). When \(|\rho|<1\), \( u_t \) is also stationary and therefore, by (11), so is \( u_t^{(m)} \). When \( \rho=1 \), since \( m \) is fixed, \( u_t^{(m)} \) is stationary by (22) since it is a finite lag function of \( \varepsilon_t \). Note also that \( \mathbb{E} \left[ u_t^{(m)} \right] = 0 \). The stated result in (14) then follows directly by the application of (Barnett [5], pp.20-23), who generalizes Rice [34]'s formula to elliptically symmetric processes. □

Proof of Proposition 5 We will apply the (Wu [38], p.642, Theorem 4) central limit theorem for zero-crossings directly to \( K_m^{(m)}(0) \). To do so, we must first show that \( u_t^{(m)} \) satisfies the theorem assumptions. For a given \( m \), let \( a_{j,m} = 0, 1, \ldots \) denote the coefficients in the moving average representation of \( u_t^{(m)} \), i.e. \( u_t^{(m)} = \sum_{j=0}^{\infty} a_{j,m} \varepsilon_{t-j} \). Define \( A_k(\delta) \equiv \sum_{j=k}^{\infty} |a_j| \delta \) as in (Wu [38], p.642).

It is necessary to show that \( \sum_{k=1}^{\infty} \left( \frac{A_k(\delta)}{k} \right)^{1/2} < \infty \) for some \( 0 < \delta \leq 2 \) such that \( E \left( |\varepsilon|^\delta \right) < \infty \). Since, by (\(\alpha.1\)), \( E \left( |\varepsilon|^2 \right) < \infty \), we show this below for \( \delta = 2 \).

When \( \rho = 1 \) or \( \rho = 0 \) it can be seen from (22) or (21), respectively that \( u_t^{(m)} \) has a finite order moving average representation, in which case (23) is clearly satisfied, since \( A_k \) is non-zero for only finitely many \( k \). Together with the condition (\(\alpha.4\)), (23) follows immediately.

When \( |\rho|<1 \) and \( \rho \neq 0 \), \( u_t^{(m)} \) can be expressed as in (21). We may ignore the last two terms on the RHS of (21), which again involve finite order moving averages. The infinite moving average in the first term can be re-expressed as

\[
\sum_{j=0}^{\infty} \rho^j \varepsilon_{t-m-j} = \rho^{-m} \left[ \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i} \right] - \rho^{-m} \sum_{i=0}^{m-1} \rho^i \varepsilon_{t-i}.
\]

It suffices to verify (23) for the term in brackets, for which

\[
A_k(2) = \sum_{i=k}^{\infty} \rho^{2i} = \rho^{2k} \sum_{i=0}^{\infty} \rho^{2i} = \left( \frac{1}{1-\rho^2} \right) \rho^{2k} \quad \text{and}
\]

\[
\sum_{k=1}^{\infty} \left( \frac{A_k(2)}{k} \right)^{1/2} \leq \sum_{k=1}^{\infty} \left( A_k(2) \right)^{1/2} = \frac{1}{(1-\rho^2)^{1/2}} \sum_{k=1}^{\infty} |\rho|^k = \frac{1}{(1-\rho^2)^{1/2}(1-|\rho|)} < \infty
\]

Indeed, as pointed out in by Wu [38] in Remark 5, the application of his Theorem 4 is unnecessary in this case because the asymptotic normality follows directly of the \( 2m-1 \)-dependence of \( u_t^{(m)} \). As explained in the same remark, the assumption \#\{i : a_i \neq 0\} = \infty excludes this case only because it already covered by existing results.
for $|\rho| < 1$. The remaining condition of (Wu [38], Theorem 4) is satisfied by assumption given (A4). The stated result then follows by application of Wu [38], Theorem 4.

Proof of Proposition 6 For $t < m$ and $t > T - m$, $u_i^{(m)}$ involves the use of the boundary points defined in (9). However, as these involve only $2m$ points, they can be ignored in the calculation of the limiting variance. Specifically,

$$Var \left( K^{(m)}_{u,T} \right) = T^{-1} Var \left( \sum_{i=1}^{T-1} 1 \left[ u_i^{(m)} u_{i+1}^{(m)} < 0 \right] \right) =$$

$$T^{-1} Var \left( \sum_{i=m}^{T-m} 1 \left[ u_i^{(m)} u_{i+1}^{(m)} < 0 \right] \right) + O_p \left( m/T \right).$$

Therefore, since $m = o_p \left( T \right)$ by assumption,

$$\lim_{T \to \infty} Var \left( K^{(m)}_{u,T} \right) = \lim_{T \to \infty} T^{-1} \sum_{i=m}^{T-m} 1 \left[ u_i^{(m)} u_{i+1}^{(m)} < 0 \right] =$$

$$\lim_{T \to \infty} T^{-1} \sum_{h=-\infty}^{T-2m} (T - 2m - |h|) \gamma_{1\parallel} (h) =$$

$$\lim_{T \to \infty} \sum_{h=-\infty}^{T-2m} \left( 1 - \frac{2m}{T} - \frac{|h|}{T} \right) \gamma_{1\parallel} (h) =$$

$$\sum_{h=-\infty}^{\infty} \gamma_{1\parallel} (h).$$

The last equality follows because $u_i^{(m)}$ is $2m-1$-dependent and hence $1 \left[ u_i^{(m)} u_{i+1}^{(m)} < 0 \right]$ is $2m$-dependent, with the implication that $\gamma_{1\parallel} (h) = 0$ for $h > 2m$ and therefore

$$\lim_{T \to \infty} \sum_{h=-\infty}^{T-2m} \left( \frac{2m + |h|}{T} \right) \gamma_{1\parallel} (h) = \lim_{T \to \infty} \sum_{h=-\infty}^{2m} \left( \frac{2m + |h|}{T} \right) \gamma_{1\parallel} (h) = 0.$$

Proof of Proposition 8 It can be shown that the deterministic component of
the transformed series can be expressed in the following way:

\[ DC_t = (\alpha_1 + \beta_1 (t - \lambda_1 T)) \]

\[
\left[ \left( d_{1,t} - \frac{1}{2} d_{1,t-m} - \frac{1}{2} d_{1,t+m} \right) - \rho \left( d_{1,t-1} - \frac{1}{2} d_{1,t-m-1} - \frac{1}{2} d_{1,t+m-1} \right) \right] \\
+ \frac{1}{2} m \beta_1 \left( d_{1,t-m} - d_{1,t+m} \right) - \rho (d_{1,t-m-1} - d_{1,t+m-1}) \\
+ \rho \beta_1 \left[ d_{1,t-1} - \frac{1}{2} d_{1,t-m-1} - \frac{1}{2} d_{1,t+m-1} \right] + ... \\
\cdots + (\alpha_p + \beta_p (t - \lambda_p T)) \\
\left[ \left( d_{p,t} - \frac{1}{2} d_{p,t-m} - \frac{1}{2} d_{p,t+m} \right) - \rho \left( d_{p,t-1} - \frac{1}{2} d_{p,t-m-1} - \frac{1}{2} d_{p,t+m-1} \right) \right] \\
+ \frac{1}{2} m \beta_p \left( d_{p,t-m} - d_{p,t+m} \right) - \rho (d_{p,t-m-1} - d_{p,t+m-1}) \\
+ \rho \beta_p \left[ d_{p,t-1} - \frac{1}{2} d_{p,t-m-1} - \frac{1}{2} d_{p,t+m-1} \right].
\]

Since the break-points dummy variables, \( d_{i,t} \) defined in (5), are indicator functions and can only take values of 0 and 1, \( DC_t \) will be equal to zero for values of \( t \) satisfying the following conditions:

\[
d_{1,t} - \frac{1}{2} d_{1,t-m} - \frac{1}{2} d_{1,t+m} = 0 \\
d_{1,t-1} - \frac{1}{2} d_{1,t-m-1} - \frac{1}{2} d_{1,t+m-1} = 0 \\
\vdots \\
d_{p,t} - \frac{1}{2} d_{p,t-m} - \frac{1}{2} d_{p,t+m} = 0 \\
d_{p,t-1} - \frac{1}{2} d_{p,t-m-1} - \frac{1}{2} d_{p,t+m-1} = 0.
\]

These conditions are equivalent to

\[
d_{1,t-m} = d_{1,t+m} \\
d_{1,t-m-1} = d_{1,t+m-1} \\
\vdots \\
d_{p,t-m} = d_{p,t+m} \\
d_{p,t-m-1} = d_{p,t+m-1}.
\]

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 \( i = 0, 1, 2, \ldots p \). Therefore \( x_t^{(m)} = u_t^{(m)} \) at all but a maximum of \( p(2m + 1) \) points. Therefore,

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


