Threshold Random Walks in the U.S. Stock Market*

Zisimos Koustas and Jean-François Lamarche
Department of Economics
Brock University
St. Catherines, Ontario
Ontario, L2S 3A1
and
Apostolos Serletis†
Department of Economics
University of Calgary
Calgary, Alberta, T2N 1N4

May 23, 2006

Abstract

This paper extends the work in Serletis and Shintani (2003) and Elder and Serletis (2006) by re-examining the empirical evidence for random walk type behavior in the U.S. stock market. In doing so, it tests the random walk hypothesis by employing unit-root tests that are designed to have more statistical power against nonlinear alternatives. The nonlinear feature of our model is reflected by three regimes, one of which is characterized by a unit root process and the random walk hypothesis while the lower and upper regimes are well captured by a stationary autoregressive process with mean reversion and predictability.

JEL classification: C32, G12, G14.

Keywords: Asymmetric time series; Threshold adjustment; Nonlinear autoregression.

*Lamarche and Serletis gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada (SSHRC).
†Corresponding author. Phone: (403) 220-4092; Fax: (403) 282-5262; E-mail: Serletis@ucalgary.ca; Web: http://econ.ucalgary.ca/serletis.htm.
1 Introduction

Stock market efficiency — see, for example, Fama (1970) — implies stock-price dynamics which are described by a random walk with a drift. Such a process can be decomposed into two nonstationary components: a linear deterministic trend and a stochastic trend. This implies that random shocks have permanent effects on the level of stock prices through the stochastic part of the trend. Simply put, there is no tendency for stock prices to return to their deterministic trend following a shock. This suggests that future returns are unpredictable based on historical observations. Another implication of the random walk model is that the volatility of stock prices can increase without bound over time. By contrast, if stock prices revert to a deterministic trend, past returns can help predict future returns. The latter is typically interpreted as evidence of market inefficiency.

Empirical testing of the random walk hypothesis in stock prices against a linear mean-reverting alternative has generated considerable controversy. Fama and French (1988) and Poterba and Summers (1988) are early examples of studies that report mean reversion in U.S. stock prices. Another body of empirical literature was unable to reject the unit-root hypothesis in U.S. Stock prices. Examples of the latter include Lo and MacKinlay (1988), Kim et al. (1991), and Richardson (1993). However, Perron (1989) has shown that one or more breaks in the deterministic trend during the sample period may bias the unit-root tests in favour of accepting the null hypothesis of a unit root.

Equity markets in emerging economies are likely candidates for structural change as a result of financial market liberalization and/or changes in the underlying economies. Chaudhuri and Wu (2003) have tested the random walk hypothesis in stock-price indexes against the alternative of mean reversion for seventeen emerging markets. Augmented Dickey-Fuller (ADF) tests — see Dickey and Fuller (1981) — and Phillips and Perron (1988) tests yielded very little evidence against the random walk model. However, allowing for structural breaks in the deterministic trend led to the reversal of the results — in favour of mean reversion — in most cases.

The ADF and Phillips and Perron tests used in most of the previous empirical literature are based on linear autoregressions in which the null hypothesis of a unit root is tested against a linear stationary alternative. Such tests, however, have low power against slow mean reversion alternatives, such as the fractionally integrated model, as well as against nonlinear but stationary alternatives. Elder and Serletis (2006) have tested the random walk hypothesis for the Dow Jones Industrial Average against a fractionally integrated alternative hypothesis and found statistical evidence for the random walk hypothesis. In this paper, we follow Pippenger and Goering (1993), Balke and Fomby (1997), and Caner and Hansen (2001) and test the random walk hypothesis for the Dow Jones Industrial Average by employing unit-root tests that are designed to have more statistical power against nonlinear alternatives.

The nonlinear alternative used in this paper is motivated by the presence of transaction costs such that the size of the deviation from the deterministic trend influences the stock
market dynamics. Specifically, following Kapetanios *et al.* (2007) we hypothesize an inner regime in which transaction costs outweigh the profits from trading. Mean reversion in the inner regime is very weak and stock prices follow a random walk. By contrast, the two outer regimes represent larger deviations from trend that are arbitraged away leading to mean reverting dynamics. Kapetanios *et al.* (2007) assume that the speed of adjustment is an increasing function of the size of the deviation from equilibrium but that the sign of the discrepancy does not matter. In this paper, we take into account arguments from behavioral finance models — see Shleifer (2000) — that suggest that the actions of traders differ between rising and falling markets. Therefore, we use a more general model which allows asymmetric adjustment in the outer regimes depending on the sign of disequilibrium.

The rest of the paper is organized as follows. Section 2 outlines the proposed nonlinear model that serves as the alternative for the random walk null hypothesis. Section 3 presents and discusses the empirical results and Section 4 closes with a brief summary and conclusion.

## 2 The Econometric Methodology

Linear unit root tests on the log price index ($z_t$) are based on the following auxiliary regression

$$
\Delta z_t = \mu + \beta t + \rho z_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta z_{t-i} + \epsilon_t
$$

where $\mu$ is a drift term, $t$ is a time trend, and $\epsilon_t$ is a white noise disturbance. The null hypothesis of a unit root is tested by imposing the restriction $\rho = 0$ against the alternative $\rho < 0$ and testing it using the test statistics proposed by Dickey and Fuller (1981), namely $\tau_\mu$ if $\beta = 0$ or $\tau_\tau$ otherwise. If the null hypothesis cannot be rejected using appropriate critical values, it is concluded that deviations of $z_t$ from its mean (or trend) are infinitely persistent.

In contrast, Serletis and Shintani (2003) compare the hypothesis of infinite persistence against the one of chaotic dynamics and find that low-dimensional chaos does not characterize well the logged price, hence supporting the random walk hypothesis. More recently, Elder and Serletis (2006) also found statistical evidence for the random walk hypothesis, this time allowing for fractional integrating dynamics.

In the present paper, we follow a different path and consider testing for the random walk hypothesis while allowing for a stationary but nonlinear autoregressive model under the alternative hypothesis. One of the key motivations behind this testing strategy is that standard unit root tests, which are based on linear autoregressions and assume a linear stationary alternative, have been shown to have low power in distinguishing between the unit root model and a nonlinear but stationary alternative — see Pippenger and Goering (1993), Balke and Fomby (1997), and Caner and Hansen (2001). Here we consider a testing framework that allows for the logged price to be characterized by different regimes.
particular, when the log price is within a certain band (\( \rho \) is close to zero), deviations from equilibrium are extremely persistent. When the logged price moves outside this band we have mean reversion (\( \rho \) becomes more negative in the outer regimes) toward equilibrium. In these regions, we could observe some degree of predictability in the price.

We use the following 3-regime, self exciting threshold autoregressive (SETAR) model, proposed by Kapetanios and Shin (2007),

\[
\Delta z_t = \epsilon_t + \sum_{i=1}^{p} \alpha_i \Delta z_{t-i} + \begin{cases} 
\rho_1 z_{t-d} & \text{if } z_{t-d} \leq \theta_1 \\
\rho_2 z_{t-d} & \text{if } \theta_1 < z_{t-d} \leq \theta_2 \\
\rho_3 z_{t-d} & \text{if } z_{t-d} > \theta_2 
\end{cases}
\]

In (2), \( \epsilon_t \) is a white noise disturbance common across regimes and the common dynamics of the process are captured by \( \Delta z_{t-i} \). The transition variable is \( z_{t-d} \) with a delay, \( d \), that is unknown and could be estimated, in theory. However, as reported by Bec, Salem, and Carrasco (2004), the estimate of \( d \) is not precise. For this reason, and following others, we set \( d = 1 \). The lower and upper asymmetric thresholds are respectively \( \theta_1 \) and \( \theta_2 \) and are unknown. Within the band \([\theta_1, \theta_2]\), deviations are infinitely persistent (\( \rho_2 = 0 \)) while in the outer regimes, \( \rho_i \) will be negative.

We use the demeaned and logged price in the SETAR model. This implies that this threshold model allows for convergence of the demeaned logged price to the equilibrium of zero outside the band, as well as, non-convergence inside the band. Other specifications are also possible — see for example Balke and Fomby (1997), Bec, Salem, and Carrasco (2004), and Bec, Guay, and Guerre (2004) — but we focus the analysis on this specific SETAR model for two reasons. First, the results are easily tractable and interpretable and second, the threshold is allowed to be asymmetric. Kapetanios and Shin (2007) have shown that, under certain conditions on the autoregressive coefficients, although the SETAR model is locally nonstationary it is globally ergodic which implies that the moments of the distribution can be estimated reliably.

The unit root null hypothesis is characterized by \( \rho_1 = \rho_3 = 0 \) and can be tested against the alternative of threshold stationarity using a Wald test the asymptotic distribution of which has been derived by Kapetanios and Shin (2007). The following steps describe the estimation and testing procedure:

- The data \( z_t \) is demeaned and detrended to take into account the trend. That is, from now on we use the residuals, \( \hat{u}_t \) (which we will still call \( z_t \) for notational simplicity), from the regression \( z_t = a + bt + u_t \). The search for the two thresholds \( \theta_1 \) and \( \theta_2 \) is based on a grid composed of eight equally spaced points between the lower quantile and the mean of \( z_t \) and of another eight points between the mean and the upper quantile generating 64 combinations. Given that the sample size is quite large a trimming parameter of 0.05 is chosen. This ensures that at least 5% of the observations are in a segment. Let the threshold parameter space be \( \Gamma \) and let \( \theta = [\theta_1, \theta_2] \).
• The number of lagged differences in the SETAR model is found using the BIC for the model evaluated under the null hypothesis of linearity.

• Conditional on the thresholds, we compute the following Wald test statistic for \( \rho_1 = \rho_3 = 0 \). First we can re-write the SETAR model in matrix notation as

\[
\Delta z_t = x_t(\theta)\beta + \epsilon_t
\]

where \( x_t(\theta) \) is a row vector containing the desired independent variables and \( \beta \) is a vector of corresponding coefficients. Here \( x_t(\theta) \) contains \( z_{t-1} I [z_{t-d} \in I_1(\theta)] \) for the lower segment, \( z_{t-1} I [z_{t-d} \in I_3(\theta)] \) for the upper segment and \( (\Delta z_{t-1}, \Delta z_{t-2}, \cdots, \Delta z_{t-p}) \) with \( I [\cdot] \) an indicator function taking a value of 1 if its argument is true. The Wald test is computed as

\[
W_T(\theta) = \frac{\left[ R\tilde{\beta}(\theta) \right]'}{\hat{\sigma}^2(\theta)} \left[ R \left( \sum_{t=1}^{T} x_t(\theta)'x_t(\theta) \right)^{-1} R' \right]^{-1} \left[ R\tilde{\beta}(\theta) \right]
\]

where \( \hat{\sigma}^2(\theta) \) is the unrestricted estimate of the variance of the residuals and \( R \) is a selection matrix. Since the test is based on \( (\theta_1, \theta_2) \in \Gamma \), we use the supremum test defined as

\[
\sup W \equiv \sup_{(\theta_1, \theta_2) \in \Gamma} W_T(\theta)
\]

We also commute the exponential version of the test as it can have better finite sample properties. The test statistic is

\[
\exp W \equiv \frac{1}{\#\Gamma} \sum_{j=1}^{\#\Gamma} \exp \left( \frac{W_j^2(\theta)}{2} \right)
\]

where \( \#\Gamma \) is the number of elements in \( \Gamma \).

• The estimate of the \( \sup W \) is compared to its critical value. If the test statistic exceeds the simulated empirical critical value, the unit root hypothesis is rejected.

3 Estimation Results

The data set used in this paper is the same as the one used in Elder and Serletis (2006). It consists of daily observations on the Dow Jones Industrial Average from June 30, 1941 to March 13, 2006 for a total of 16,363 observations. Prior to the analysis we take the natural logarithm of the price index which we denote as \( z_t \). In Figure 1, the time series
graph of $z_t$ clearly shows that the price index contains a strong positive trend. If this trend is stochastic, we have a random walk, with drift, otherwise we have some degree of predictability in stock prices as measured by the index. Table 1 presents some standard unit root tests. The auxiliary autoregressions, from which we compute the ADF test, include a constant and a trend and have a lag length that was selected using the BIC criterion. The reported $t$-statistics were less (in absolute terms) than the 1 and 5 percent critical values thus rejecting stationarity. Kwiatkowski et al. (1992) argue that the way in which classical hypothesis testing is carried out ensures that the null hypothesis is accepted unless there is strong evidence against it. They have proposed tests (known as the KPSS tests) of the hypothesis of stationarity against the alternative of the unit root. The null hypothesis of stationarity is strongly rejected. The random walk hypothesis is, therefore, not rejected.

Table 1 also reports the results from unit root tests that are designed to have increased power against nonlinear alternatives. The nonlinear alternative hypothesis is the SETAR model described by equation (2). This model imposes a unit root in the middle regime, $\rho_2 = 0$, and zero drift in all regimes. The number of lagged differences in (2) was set according to the BIC and is reported in column 2. The results of the $\sup W$ and $\exp W$ test statistics indicate that the null hypothesis of nonstationarity is rejected for the Dow Jones Industrial Average. This represents a complete reversal of the results obtained under the assumption of linearity. In Table 2 we present the estimation results for the SETAR model. The values of the lower and upper thresholds are $-0.59$ and $0.40$, respectively, which implies that the unit root middle segment is relatively large. However, as soon as the price level is too low or too high, mean reversion will occur and some degree of predictability in prices will prevail. It is interesting to note that the speed of mean reversion (as measured by the autoregressive parameters $\rho_1$ and $\rho_3$) is extremely slow. This can be explained, in part, by the fact that we are using daily observations. In Figure 2 we show the demeaned and detrended index along with both thresholds. The nonlinear feature of our model is reflected by the three regimes appearing in the Figure. The middle segment is characterized by a unit root process and the random walk hypothesis while the lower and upper segments are well captured by a stationary autoregressive process with mean reversion and predictability.

4 Concluding Remarks

This paper reexamines U.S. stock prices for evidence of random walk type behavior, using unit-root tests that have more statistical power against nonlinear alternatives. It uses a 3-regime, self exciting threshold autoregressive (SETAR) model, and shows that the middle segment is characterized by a unit root process and the random walk hypothesis, while the lower and upper segments are well captured by a stationary autoregressive process with mean reversion and predictability.
References


Figure 1. Logged Dow Jones Industrial Average, 1941-2006
<table>
<thead>
<tr>
<th>Linear alternatives</th>
<th>Nonlinear alternatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF</td>
<td>KPSS</td>
</tr>
<tr>
<td>-1.69</td>
<td>13.31</td>
</tr>
<tr>
<td>supW</td>
<td>expW</td>
</tr>
<tr>
<td>11.35</td>
<td>11.08</td>
</tr>
</tbody>
</table>

Notes: The 1% and 5% critical values for the ADF and KPSS tests are -3.96 and -3.41 and 0.216 and 0.146, respectively. The 1%, 5%, and 10% critical values for the supW and expW tests are 16.28, 12.16, and 10.35 and 3428.92 and 437.03, respectively.
TABLE 2

ESTIMATES OF THRESHOLD MODEL

| $\alpha_1$ | $\alpha_2$ | $\rho_1$ | $\rho_2$ | Thresholds | % of observations in |
|-----|-----|-----|-----|-----|-----|-----|
| 0.08 | -0.05 | -0.0012 | -0.0009 | -0.59/0.40 | 5 | 84 | 11 |

Notes: Robust standard errors are reported in parentheses.
The first entry for the threshold is the lower regime.
Figure 2. Demeaned and Detrended Dow Jones Industrial Average with Thresholds