Unit-Root Tests and Asymmetric Adjustment With an Example Using the Term Structure of Interest Rates

Walter Enders & C. W. J. Granger

Unit-Root Tests and Asymmetric Adjustment With an Example Using the Term Structure of Interest Rates

Walter ENDERS
Department of Economics, Iowa State University, Ames, IA 50011 (enders@iastate.edu)

C. W. J. GRANGER
Department of Economics, University of California at San Diego, La Jolla, CA 92093

This article develops critical values to test the null hypothesis of a unit root against the alternative of stationarity with asymmetric adjustment. Specific attention is paid to threshold and momentum threshold autoregressive processes. The standard Dickey–Fuller tests emerge as a special case. Within a reasonable range of adjustment parameters, the power of the new tests is shown to be greater than that of the corresponding Dickey–Fuller test. The use of the tests is illustrated using the term structure of interest rates. It is shown that the movements toward the long-run equilibrium relationship are best estimated as an asymmetric process.

KEY WORDS: Asymmetric time series; Threshold adjustment; Threshold autoregressive process.

It is widely acknowledged that many important economic variables display asymmetric adjustment paths. The observation that firms are more apt to raise than to lower prices is a key feature of many macroeconomic models. There is also a sizable literature concerning the asymmetric adjustment of real variables. Neftci (1984) began an important debate by showing that several measures of U.S. unemployment display asymmetric adjustment over the course of the business cycle. Falk (1986) found little evidence in favor of asymmetry when he applied Neftci's method to real U.S. gross national product (GNP), investment, and productivity and to industrial production in Canada, France, Italy, Germany, and the United Kingdom. Nevertheless, the recent consensus seems to be in favor of asymmetric adjustment. Teräsvirta and Anderson (1992) found that industrial production in 13 countries responds more sharply to negative shocks than to positive shocks. Similarly, Granger and Lee (1989) found that U.S. sales, production, and inventories display asymmetric adjustment toward their long-run equilibrium relationship. Potter (1995) modeled changes in real U.S. GNP as a threshold adjustment process, and Balke and Fomby (1996) showed that various short-term interest rates exhibit threshold cointegration.

In fact, the focus of the debate seems to have changed. Instead of trying to determine whether or not there is asymmetry, recent works attempt to ascertain the specific type of asymmetry. For example, Sichel (1993) discussed the distinction between "sharp" versus "deep" cycles. Sharpness occurs when contractions are steeper than expansions and deepness occurs when troughs are more pronounced than peaks. He found that U.S. unemployment, industrial production, and GNP display evidence in favor of deepness but that only unemployment displays evidence of sharpness. Ramsey and Rothman (1996) found both steepness and deepness in several of the Nelson and Plosser (1982) data series.

In spite of the interest in asymmetric adjustment models, standard unit-root tests assume a symmetric adjustment process. One aim of this article is to describe a class of models that can be used as the basis of unit-root tests in the presence of asymmetric adjustment. In particular, the threshold autoregressive (TAR) model developed by Tong (1983) allows the degree of autoregressive decay to depend on the state of the variable of interest. Such a model can capture the key aspects of any "deep" movements in a series. If autoregressive decay is fast when the variable is above trend and slow when the variable is below trend, troughs will be more persistent than peaks. A second aim of the article is to introduce the momentum threshold autoregressive (M-TAR) model. The M-TAR model allows a variable to display differing amounts of autoregressive decay depending on whether it is increasing or decreasing. The momentum model can capture the possibility of asymmetrically "sharp" movements in a series. The TAR and M-TAR models are described in Section 1. Section 2 contains the critical values that can be used to test the null hypothesis of a unit root against an alternative of stationarity with asymmetric adjustment. The power of the test for simple TAR adjustment is low relative to that of the usual Dickey–Fuller test. The power of the test is increased, however, in a multithreshold setting. Within a range of adjustment parameters relevant to many economic time series, the power of the test for M-TAR adjustment can be substantially greater than that of the corresponding Dickey–Fuller test. In Section 3 the asymmetric adjustment tests are used to analyze the relationship between long-term and short-term interest rates. The tests

Journal of Business & Economic Statistics
July 1998, Vol. 16, No. 3

are used to reconfirm the well-established result that the interest-rate differential is stationary. It is shown, however, that the movements toward the equilibrium relationship are asymmetric in a way characterized by M-TAR adjustment. The result is supported by the estimations of the implied error-correction model. Section 4 contains some concluding remarks and a discussion of an important limitation of the testing procedure.

1. THRESHOLD AND MOMENTUM THRESHOLD MODELS

Standard time series models assume linearity and symmetric adjustment. Consider the simple linear relationship used as the basis for the Dickey–Fuller test:

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

where ε_t is a white-noise disturbance.

The standard procedure is to estimate ρ and to ascertain whether $-2 < \rho < 0$ using the appropriate critical values. Equation (1) can be modified in many different ways including the introduction of deterministic regressors, the addition of lagged changes in Δy_t, allowing for structural breaks, and allowing $\{\varepsilon_t\}$ to be weakly dependent and heterogeneously distributed. Notice that the alternative hypothesis entails a symmetric adjustment process around $y_t = 0$. Formally, the homogeneous portion of (1) can be written as a first-order linear difference equation with constant coefficients. Convergence is assured if $-2 < \rho < 0$ because the homogeneous solution to (1) is $y_t = A(1+\rho)^t$, where A is an arbitrary constant.

The Dickey–Fuller test and its extensions are misspecified if adjustment is asymmetric, however. Consider an alternative specification—called the threshold autoregressive (TAR) model—such that

$$\Delta y_t = \begin{cases} \rho_1 y_{t-1} + \varepsilon_t & \text{if } y_{t-1} \geq 0 \\ \rho_2 y_{t-1} + \varepsilon_t & \text{if } y_{t-1} < 0. \end{cases}$$

A sufficient condition for the stationarity of $\{y_t\}$ is $-2 < (\rho_1, \rho_2) < 0$. Moreover, if the sequence is stationary, Tong (1983) proved that the least squares estimates of ρ_1 and ρ_2 have an asymptotic multivariate normal distribution. This result easily generalizes to higher-order autoregressive processes. Tong (1990) also developed many of the properties of the TAR model. A formal way to quantify the adjustment process is to write

$$\Delta y_t = I_t \rho_1 y_{t-1} + (1 - I_t)\rho_2 y_{t-1} + \varepsilon_t, \tag{2}$$

where I_t is the Heaviside indicator function such that

$$I_t = \begin{cases} 1 & \text{if } y_{t-1} \geq 0 \\ 0 & \text{if } y_{t-1} < 0. \end{cases} \tag{3}$$

If the system is convergent, $y_t = 0$ is the long-run equilibrium value of the sequence. If y_{t-1} is above its long-run equilibrium value, the adjustment is $\rho_1 y_{t-1}$, and if y_{t-1} is below long-run equilibrium, the adjustment is $\rho_2 y_{t-1}$. Because adjustment is symmetric if $\rho_1 = \rho_2$, (1) is a special case of (2) and (3). Notice that the TAR model can capture aspects of "deep" movements in a sequence. If, for example, $-1 < \rho_1 < \rho_2 < 0$, then the negative phase of the $\{y_t\}$ sequence will tend to be more persistent than the positive phase.

There are three important ways to modify Equations (2) and (3):

1. *Alternative Linear Attractors:* Equation (2) assumes a long-run equilibrium point around $y_t = 0$. Equations (4) and (5) use two other important attractors:

$$\Delta y_t = I_t \rho_1 [y_{t-1} - a_0] + (1 - I_t)\rho_2 [y_{t-1} - a_0] + \varepsilon_t, \tag{4}$$

where

$$I_t = \begin{cases} 1 & \text{if } y_{t-1} \geq a_0 \\ 0 & \text{if } y_{t-1} < a_0, \end{cases}$$

and

$$\Delta y_t = I_t \rho_1 [y_{t-1} - a_0 - a_1(t-1)] + (1 - I_t)\rho_2 [y_{t-1} - a_0 - a_1(t-1)] + \varepsilon_t, \tag{5}$$

where

$$I_t = \begin{cases} 1 & \text{if } y_{t-1} \geq a_0 + a_1(t-1) \\ 0 & \text{if } y_{t-1} < a_0 + a_1(t-1). \end{cases}$$

In (4), if $-2 < (\rho_1, \rho_2) < 0$, the attractor is such that long-run equilibrium occurs at the point $y_t = a_0$. Clearly, if $\rho_1 = \rho_2 = 0$, the sequence is a pure random walk. Notice that symmetric adjustment emerges as a special case. If $\rho_1 = \rho_2 \neq 0$, it is possible to write (4) as the AR(1) model: $\Delta y_t = a_0 + \rho y_{t-1} + \varepsilon_t$.

In (5), if $-2 < (\rho_1, \rho_2) < 0$, the trend line $y_t = a_0 + a_1 t$ is an attractor such that the $\{y_t\}$ sequence is trend stationary. The sequence tends to decay at the rate ρ_1 if y_{t-1} is above the trend and at the rate ρ_2 if y_{t-1} is below the trend. If either ρ_1 or ρ_2 lies outside the interval $(-2, 0)$, however, the $\{y_t\}$ sequence may not be trend stationary. For example, if $\rho_1 = 0$, the sequence will exhibit random-walk behavior whenever $y_t > a_0 + a_1 t$. Again, symmetric adjustment emerges as a special case. If $\rho_1 = \rho_2 \neq 0$, (5) can be written as $\Delta y_t = \rho(a_1 - a_0) - \rho a_1 t + \rho y_{t-1} + \varepsilon_t$.

2. *Higher-order Processes:* Equations (2), (4), and (5) can be augmented with lagged changes in the $\{y_t\}$ sequence. For example, (2) can be augmented such that it becomes the pth-order process

$$\Delta y_t = I_t \rho_1 y_{t-1} + (1 - I_t)\rho_2 y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta y_{t-i} + \varepsilon_t. \tag{6}$$

In working with specifications such as (6), it is possible to use diagnostic checks of the residuals, such as the autocorrelogram of the residuals and Ljung–Box tests, and various model-selection criteria to determine the appropriate lag length (Tong 1983). To ensure that there is no more than a single unit root, all the values of r satisfying the inverse characteristic equation $1 - \beta_1 r - \beta_2 r^2 + \cdots + \beta_{p-1} r^{p-1} = 0$ must lie outside the unit circle.

3. *Alternative Adjustment Specifications:* In (3), the Heaviside indicator depends on the *level* of y_{t-1}. A useful

alternative is to allow the decay to depend on the previous period's *change* in y_{t-1}. Consider setting the Heaviside indicator according to the following rule:

$$I_t = \begin{cases} 1 & \text{if } \Delta y_{t-1} \geq 0 \\ 0 & \text{if } \Delta y_{t-1} < 0. \end{cases} \tag{7}$$

Replacing (3) by (7) is especially valuable when adjustment is asymmetric such that the series exhibits more "momentum" in one direction than the other. Models constructed using (2) and (7) can be called momentum threshold autoregressive (M-TAR) models. If, for example, $|\rho_1| < |\rho_2|$, the M-TAR model exhibits little decay for positive Δy_{t-1} but substantial decay for negative Δy_{t-1}. In a sense, increases tend to persist but decreases tend to revert quickly toward the attractor. As such, the momentum model can be used to represent Sichel's (1993) notion of "sharpness."

Combining various aspects of the models is possible. For example, it is possible to allow for an attractor with long-run equilibrium at $y_t = a_0$ using the Heaviside indicator of (7) or the trend attractor in a model augmented by lagged changes in Δy_t.

2. TESTING FOR UNIT ROOTS VERSUS TAR AND M-TAR ADJUSTMENT

To conduct a Monte Carlo experiment that can be used to test the null hypothesis of a unit root against the alternative of a TAR or an M-TAR model, 100,000 random-walk processes of the following form were generated:

$$y_t = y_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, T. \tag{8}$$

For $T = 50, 100, 250$, and $1,000$, a total of $T + 100$ normally distributed and uncorrelated pseudorandom numbers with standard deviation equal to unity were drawn to represent the $\{\varepsilon_t\}$ sequence. Setting the initial value of the sequence (i.e., y_0) equal to 0, the remaining values of $\{y_t\}$ were generated using (8). For a test of asymmetric adjustment to be sensible, however, the generated sequence must cross the attractor at least once. Hence, if a generated sequence did not cross the line $y_t = 0$, it was discarded and replaced by another randomly generated sequence. Note that this issue does not arise for the other attractors to be considered later (i.e., $y_t = a_0$ and $y_t = a_0 + a_1 t$, where a_0 and a_1 are estimated from the data).

For each of the 100,000 series, the first 100 realizations were discarded and the TAR model given by (2) and (3) was estimated and three different test statistics were tabulated. The t statistics for the null hypotheses $\rho_1 = 0$ and $\rho_2 = 0$ were recorded along with the F statistic for the null hypothesis $\rho_1 = \rho_2 = 0$. The most significant of the t statistics is called T-Max; the least significant of the t statistics is called T-Min, and the F statistic is called ϕ. Only the ϕ statistic is reported here because it was found to have better power than the T-Max and T-Min statistics. In Panel A of Table 1, critical values for the ϕ statistic are reported at the 10%, 5%, and 1% significance levels for each sample size. For example, for $T = 100$, the F statistic for the null $\rho_1 = \rho_2 = 0$ exceeded 3.95 in approximately 5% of the 100,000 trials.

The Monte Carlo experiment was repeated for an M-TAR model using the indicator function given by (7). The corresponding F test statistics, called ϕ^*, are reported in Panel B of Table 1. It is interesting to note that the critical values for the ϕ^* statistics are smaller than the corresponding values for the ϕ statistics.

The attractor $y_t = 0$ is especially convenient because it contains no coefficients that need to be estimated from the data. In fact, whenever the attractor is known, the data can be suitably adjusted so that the attractor $y_t = 0$ is applicable to the transformed data. In most instances, however, if the $y_t = a_0$ or $y_t = a_0 + a_1 t$ attractors are used, the values of a_0 and a_1 will need to be estimated from the data. Four additional sets of critical values are reported in Panels C–F of Table 1. It is straightforward to develop critical values for the attractor $y_t = a_0$, where a_0 is the sample mean of the $\{y_t\}$ sequence. These critical values can be used to test the null hypothesis of a random-walk process with a nonzero sample mean against the alternative of the TAR model given by (4). In essence, for values of $T = 50, 100, 250$, and $1,000$, 100,000 random walks were generated after initializing the first value of the sequence (i.e., y_0) equal to a constant. Each resulting series was regressed on a constant and the residuals called $\{\hat{y}_t\}$. For each of the $\{\hat{y}_t\}$ series, the following regression equation was estimated:

$$\Delta \hat{y}_t = I_t \rho_1 \hat{y}_{t-1} + (1 - I_t)\rho_2 \hat{y}_{t-1} + \hat{\varepsilon}_t. \tag{9}$$

Now, the F statistic for the null hypothesis $\rho_1 = \rho_2 = 0$ is called ϕ_μ. Because $\{\hat{y}_t\}$ is the "demeaned" value of the sequence, under the alternative hypothesis, (9) is equivalent to $\Delta y_t = I_t \rho_1 (y_{t-1} - \hat{a}_0) + (1 - I_t)\rho_2 (y_{t-1} - \hat{a}_0) + \varepsilon_t$, where $\hat{a}_0$ is the estimated sample mean of the $\{y_t\}$ sequence.

Table 1. The Critical Values for Rejecting the Null Hypothesis of a Unit Root

Sample size	Probability of a smaller value					
	90%	95%	99%	90%	95%	99%
	No estimated deterministic components					
	Panel A: The Φ statistic			Panel B: The Φ^* statistic		
50	3.30	4.12	6.09	2.98	3.81	5.79
100	3.18	3.95	5.69	2.83	3.60	5.38
250	3.10	3.82	5.53	2.68	3.41	5.10
1,000	3.04	3.75	5.36	2.51	3.21	4.85
	Estimated constant attractor					
	Panel C: The Φ_μ statistic			Panel D: The Φ^*_μ statistic		
50	3.84	4.73	6.85	4.17	5.14	7.43
100	3.79	4.64	6.57	4.11	5.02	7.10
250	3.74	4.56	6.47	4.05	4.95	6.99
1,000	3.74	4.56	6.41	4.05	4.95	6.91
	Estimated trend attractor					
	Panel E: The Φ_T statistic			Panel F: The Φ^*_T statistic		
50	5.41	6.52	9.14	5.89	7.07	9.77
100	5.27	6.30	8.58	5.74	6.83	9.21
250	5.18	6.12	8.23	5.64	6.65	8.85
1,000	5.15	6.08	8.12	5.60	6.57	8.74

The Panel C of Table 1 reports critical values of the ϕ_μ statistics. The same procedure was used to develop a test for the null hypothesis of a random-walk process against the alternative hypothesis of an M-TAR model obtained by replacing the indicator function in (4) with that in (7). The associated test statistics, denoted by ϕ_μ^*, are reported in Panel D.

A slight modification of the procedure was used to develop the critical values to test the null hypothesis of a random walk plus drift against the alternative that the data-generating process is given by (5). After the appropriate initialization (i.e., selecting values for a_0 and y_0), each set of 100,000 series was generated using

$$y_t = a_0 + y_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, T. \tag{10}$$

Thus, under the null hypothesis, $\{y_t\}$ is a random walk plus drift such that the initial condition may differ from 0. Each resulting series was regressed on a constant and a trend. The residuals—again called $\{\hat{y}_t\}$—were estimated in the form of (9). Thus, under the alternative hypothesis, $\{y_t\}$ is generated according to (5). Panel E reports the F statistic, now called ϕ_T, for the null hypothesis $\rho_1 = \rho_2 = 0$. Repeating the procedure for the M-TAR model yielded the values of the ϕ_T^* statistics shown in Panel F.

To use the statistics, perform the following four steps:

Step 1: As in the usual Dickey–Fuller test, critical values depend on the presence of the deterministic regressors (i.e., a_0 and a_1). There is no simple way to jointly test for the presence of the deterministic elements and a unit root, however, because a_0 and/or a_1 are estimated before setting the Heaviside indicator. In some instances, such as when regression residuals are used, it might be clear that the appropriate attractor to consider is $y_t = 0$. Other times, the data may have a clear trend so that the appropriate null is that of a random walk plus drift against an alternative of trend stationarity. If the form of the deterministic regressors is in doubt, we suggest the ad hoc procedure of fitting the model with a constant (i.e., a nonzero value of a_0). The ϕ and ϕ_μ statistics and the ϕ^* and ϕ_μ^* statistics are reasonably close so that an inappropriately included constant should not substantially affect the unit-root tests. An analysis of the residuals can indicate the possible omission of a trend.

Step 2: When using the attractors of (4) or (5), regress the data on the deterministic components (i.e., the constant a_0 or the constant and trend $a_0 + a_1 t$) and save the residuals in the sequence $\{\hat{y}_t\}$. Next, depending on the type of asymmetry under consideration, set the indicator function I_t according to (3) or (7). Estimate a regression equation in the form of (9) and obtain the sample F statistic for the null hypothesis $\rho_1 = \rho_2 = 0$. Compare this sample statistic with the appropriate critical values shown in Table 1 to determine whether the null hypothesis of a unit root can be rejected.

If the alternative hypothesis is accepted, it is possible to test for symmetric versus asymmetric adjustment because ρ_1 and ρ_2 converge to multivariate normal distributions. As such, the restriction that adjustment is symmetric (i.e., the null hypothesis $\rho_1 = \rho_2$) can be tested using the usual F statistic.

Step 3: Diagnostic checking of the residuals should be undertaken to ascertain whether the $\{\hat{\varepsilon}_t\}$ can reasonably be characterized by a white-noise process. If the residuals are correlated, return to Step 2 and reestimate the model in the form $\Delta\hat{y}_t = I_t\rho_1\hat{y}_{t-1} + (1 - I_t)\rho_2\hat{y}_{t-1} + \beta_1\Delta\hat{y}_{t-1} + \cdots + \beta_{p-1}\Delta\hat{y}_{t-p+1}$.

Lag lengths can be determined by an analysis of the regression residuals and/or using model-selection criteria such as the Akaike information criterion (AIC) or Bayesian information criterion (BIC).

Step 4: Once it has been determined whether or not the series is stationary, model building can occur along the lines suggested by Chan (1993) and Tsay (1989). In a model with asymmetric adjustment, Tong (1983) demonstrated that the sample mean is a biased estimate of the attractor. Using (2) and (3) as an example, if $-2 < \rho_1 < \rho_2 < 0$, the $\{y_t\}$ sequence will exhibit more persistence whenever $y_{t-1} < 0$. As such, the value of the attractor (i.e., 0) will exceed the expected value of the sequence. Fortunately, Chan (1993) showed that searching over all values of a_0 so as to minimize the sum of squared errors from the fitted model yields a super-consistent estimate of the threshold.

At this point, the dynamic specification in (6) can be generalized in several directions. It is possible to allow Δy_t to display asymmetric adjustment to its lagged changes. For example, the magnitude of each β_i could depend on whether Δy_{t-1} was positive or negative. Moreover, as discussed by Granger and Teräsvirta (1993), the values of ρ_1 and ρ_2 might be allowed to smoothly adjust over time.

Power Tests

Because unit-root tests suffer from low power, it is of interest to compare the power of the proposed test statistics to the power of the more traditional Dickey–Fuller test. Toward this end, for various values of ρ_1 and ρ_2, 2,500 series were generated using (4) for $T = 100$. Following Steps 1 and 2, each series was regressed on a constant and an equation in the form of (9) was estimated. For each of the 2,500 regressions, the sample ϕ_μ statistics were calculated and compared to the appropriate critical values. The percentage of times that the null hypothesis was correctly rejected is reported in the center portion of Table 2.

For comparison purposes, the following regression equation was estimated for each of the generated sequences: $\Delta y_t = a_0 + \rho y_{t-1} + \varepsilon_t$.

Table 2. Power Tests for the ϕ_μ Statistics

		ϕ_μ test			Dickey–Fuller test		
ρ_1	ρ_2	10%	5%	1%	10%	5%	1%
−.05	−.05	18.52	9.72	1.84	22.88	11.64	2.44
−.10	−.10	46.24	28.48	8.08	53.56	33.44	8.96
−.10	−.20	67.92	50.96	19.12	74.28	55.00	21.04
−.10	−.50	90.88	80.44	48.96	92.36	81.20	48.80
−.10	−.75	94.32	87.36	60.60	94.16	86.48	57.88
−.10	−1.50	98.44	96.04	82.60	98.40	94.96	77.24

NOTE: For each significance level, the entry is the percentage of instances for which the null hypothesis of a unit root is correctly rejected.

Table 3. Alternative Power Tests for the ϕ_μ Statistics

Thresholds		Autoregressive coefficients			Φ_μ test			Dickey–Fuller test		
k_1	k_2	ρ_1	ρ_2	ρ_3	10%	5%	1%	10%	5%	1%
−1	+1	−.025	−.25	−.50	37.32	25.76	12.28	38.20	25.64	11.12
−2	+2	−.025	−.125	−.25	35.80	24.08	8.72	38.68	26.24	9.52
		−.025	−.125	−.50	40.16	30.20	14.28	41.08	29.44	13.36
		−.025	−.25	−.50	53.36	42.44	26.88	50.36	40.56	24.56
		−.025	−.25	−1.00	57.84	49.76	35.52	53.68	44.48	29.36
−2	+1	−.025	−.25	−1.00	58.80	51.04	37.76	54.76	44.68	29.76
−1	+2	−.025	−.25	−1.00	41.16	30.84	15.60	41.48	29.44	12.48
−2	+.2	−.025	−.125	−.50	42.08	31.84	16.24	41.08	30.48	15.16
		−.025	−.25	−.50	56.52	46.00	28.12	52.40	42.20	25.24
		−.025	−.25	−1.00	57.96	49.68	34.64	54.44	44.04	27.64
		−.050	−.25	−1.00	80.68	72.40	54.56	78.12	67.00	44.80

NOTE: For each significance level, the entry is the percentage of instances for which the null hypothesis of a unit root is correctly rejected.

The t statistic for the null hypothesis $\rho = 0$ was compared to the Dickey–Fuller τ_μ statistic (i.e., -2.89 at the 5% level and -3.51 at the 1% level). The percentage of times the Dickey–Fuller test correctly rejected the null hypothesis is shown in the right portion of Table 2.

To take a specific example, using the values $\rho_1 = -.10$ and $\rho_2 = -.20$, the ϕ_μ statistic correctly indicated stationarity in 67.92%, 50.96%, and 19.12% of the trials at the 10%, 5%, and 1% significance levels, respectively. The Dickey–Fuller performed substantially better. At the 10%, 5%, and 1% significance levels, the null of a unit root was correctly rejected in 74.28%, 55.00%, and 21.04% of the trials. Inspection of Table 2 reveals that, even in the presence of a substantial amount of asymmetry, the power of the Dickey–Fuller test generally exceeds the power of the ϕ_μ statistic. The poor performance of the ϕ_μ statistic is due to the use of a two-step procedure and to the estimation of one additional coefficient as compared to the Dickey–Fuller test. The resulting loss of power does not overcome the gain resulting from estimating a correctly specified model.

Table 2 reports results for the case in which the true data-generating process has only a single threshold. It is interesting to note that the relative power of the ϕ_μ statistic is improved in a multithreshold setting. Pippenger and Goering (1993) considered TAR processes that behave as random walks within a band but exhibit symmetric decay when outside the band. They showed that the power of the Dickey–Fuller test is substantially reduced in this three-regime setting. We shall examine the power of the Dickey–Fuller test when there is asymmetric decay outside of the band. To illustrate the point, for $T = 100$, 2,500 series were generated using the following modification of (4):

$$\Delta y_t = I(1)_t\rho_1[y_{t-1} - a_0] + I(2)_t\rho_2[y_{t-1} - a_0] + I(3)_t\rho_3[y_{t-1} - a_0] + \varepsilon_t, \quad (4')$$

where $I(1)_t = 1$ if $y_{t-1} - a_0 < k_1$ and 0 otherwise, $I(2)_t = 1$ if $k_1 \le y_{t-1} - a_0 \le k_2$ and 0 otherwise, and $I(3)_t = 1$ if $y_{t-1} - a_0 > k_2$ and 0 otherwise.

There are two distinct thresholds (k_1 and k_2), and the autoregressive nature of $\{y_t\}$ is dependent on each regime. As can be seen in Table 3, the ϕ_μ statistic can perform quite well as compared to the Dickey–Fuller statistic. For the threshold values $k_1 = -2$ and $k_2 = +2$ and for the autoregressive coefficients $\rho_1 = -.025, \rho_2 = -.25$, and $\rho_3 = -.5$, the ϕ_μ statistic correctly identified that the model was stationary in 53.36%, 42.44%, and 26.88% of the cases at the 10%, 5%, and 1% significance levels, respectively. At the same significance levels, the Dickey–Fuller τ_μ statistic identified that the model was stationary in 50.36%, 40.56%, and 24.56% of the cases. Inspection of Table 3 reveals that increasing the degree of asymmetry increases the relative power of the ϕ_μ test over the Dickey–Fuller test. For example, set $\rho_3 = -1$ but retain all of the other aforementioned values (i.e., $k_1 = -2, k_2 = +2, \rho_1 = -.025$, and $\rho_2 = -.25$). At the 1% level, the ϕ_μ statistic and the τ_μ statistic correctly identified a stationary series in 35.52% and 29.36% of the cases, respectively. The power of both tests increases in such a way that the power of the ϕ_μ statistic increases relative to that of the τ_μ statistic. As such, the relative power of the ϕ_μ test is greatest when one of the adjustment coefficients is very small.

Table 4 reports results using the two-regime version of the M-TAR model. Notice that the test for the M-TAR model has greater power than the test for the corresponding TAR model. Notice also that the power of the ϕ_μ^* statistic is often substantially larger than that of the corresponding τ_μ statistic. The power of the Dickey–Fuller test exceeds that of the ϕ_μ^* statistic when the true adjustment process is

Table 4. Power Tests for the Φ_μ^* Statistics

		Φ_μ^* test			Dickey–Fuller test		
ρ_1	ρ_2	10%	5%	1%	10%	5%	1%
−.025	−.05	17.76	9.08	2.48	18.04	9.92	2.04
−.025	−.10	32.00	18.88	4.72	28.48	15.12	3.56
−.025	−.20	76.64	60.68	26.36	60.60	38.60	11.12
−.05	−.05	20.84	11.20	2.72	22.88	11.64	2.44
−.05	−.10	35.92	21.52	5.80	36.20	20.92	5.60
−.05	−.20	76.28	58.96	25.08	68.64	46.52	13.96
−.10	−.10	45.40	28.40	7.60	53.56	33.44	8.96
−.10	−.20	79.04	63.48	25.96	80.72	62.60	24.92
−.10	−.50	100.00	99.92	98.88	99.92	99.28	90.72

NOTE: For each significance level, the entry is the percentage of instances for which the null hypothesis of a unit root is correctly rejected.

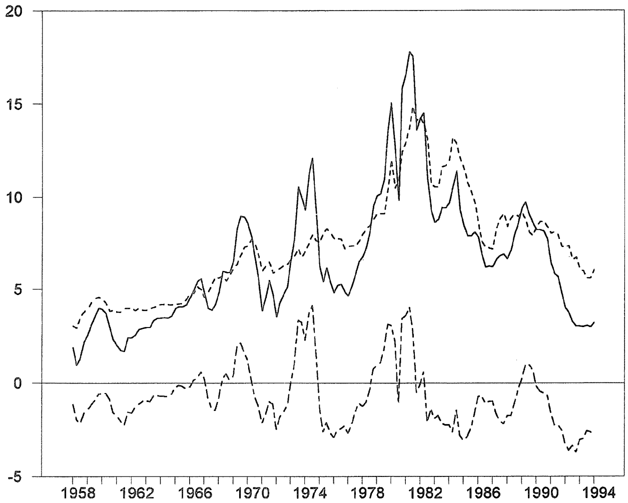

Figure 1. The Short-Term Rate, Long-Term Rate, and Interest-Rate Differential. The quarterly values of the short-term interest rate, the long-term interest rate, and the interest-rate differential are shown by the solid line, dotted line, and dashed line, respectively.

symmetric. When there is asymmetric adjustment, however, the relative power of the ϕ_μ^* statistic is enhanced. For example, if $\rho_1 = -.025$ and $\rho_2 = -.20$, the ϕ_μ^* test correctly indicated a stationary process 57% more often than the τ_μ statistic at the 5% level and more than twice as often at the 1% level.

3. ASYMMETRIC ADJUSTMENT OF INTEREST RATES

The appropriate use of the various test statistics just developed can be illustrated by considering the relationship between long-term and short-term interest rates. Toward this end, quarterly values of the federal funds rate and 10-year interest rate on U.S. government securities were obtained from the CD-ROM version of the *International Financial Statistics* over the 1958:Q1–1994:Q1 period. Figure 1 shows the short-term rate (r_S), the long-term rate (r_L), and the interest-rate differential (r_D). It is generally agreed (see Stock and Watson 1988) that short-term and long-term interest rates appear to be $I(1)$ variables that are cointegrated such that the interest-rate differential is stationary. To formally test the asymmetric-adjustment hypothesis, the interest-rate differential $r_S - r_L$ was regressed on a constant so that the equilibrium relationship has the form $r_S - r_L = a_0$.

As shown in the first two columns of Table 5, TAR models were estimated for various lag lengths. Notice that the AIC selects a model with one lagged change but the BIC selects a model without lags. In the model without any lagged changes, the sample value of ϕ_μ is 5.25, and in the model with one lagged change, the sample value of ϕ_μ is 6.77. Comparing these values to the 5% critical value of 4.64, it is possible to reject the null hypothesis $\rho_1 = \rho_2 = 0$. Note that the sample value of $\phi_\mu = 6.77$ is significant at the 1% level. Given that the interest-rate differential is stationary, it is possible to test the null $\rho_1 = \rho_2$ using the normal distribution. In each case, the value of the F statistic is small enough that it is not possible to reject the null of symmetric adjustment.

Next, the Heaviside indicator was set according to the momentum model given by (7). As shown in the third and fourth columns of Table 5, the M-TAR model was estimated for various lag lengths. Again the AIC selects a model with one lagged change, but the BIC selects a model without lags. In the model without any lagged changes, the sample value of ϕ_μ^* is 8.64, and in the model with one lagged change the sample value of ϕ_μ^* is 9.50. Comparing these values to the 1% critical value of 7.10, it is possible to reject the null hypothesis $\rho_1 = \rho_2 = 0$. Now, the test of the null hypothesis $\rho_1 = \rho_2$ is strongly rejected. With no lagged changes, the value of F is significant at the .009 level, and with one lagged change the significance level is .018. Hence, adjustment appears to be asymmetric such that the attractor is stronger for negative *changes* in the term structure.

In addition, as reported in the next two columns of Table 5, a standard Dickey–Fuller test was performed on the

Table 5. Estimates of the Interest-Rate Differential

	Linear attractor model		*Momentum model*		*Dickey–Fuller model*		*Momentum-consistent*[g]	
Lags	*0*	*1*	*0*	*1*	*0*	*1*	*0*	*1*
ρ_1	−.159	−.183	−.041	−.068	−.133	−.156	−.061	−.050
	(−3.01)[a]	(−3.42)	(−.746)	(−1.20)	(−.313)	(−3.58)	(−1.37)	(−.965)
ρ_2	−.085	−.106	−.264	−.271	NA	NA	−.293	−.299
	(−1.19)[b]	(−1.47)	(−4.09)	(−4.22)			(−5.19)	(−5.11)
AIC[c]	670.91	669.00	664.63	664.01	669.59	667.79	657.71	659.55
BIC	676.83	677.89	670.56	672.90	672.55	673.72	666.60	671.39
ϕ_μ[d]	5.25	6.77	8.64	9.50	NA	NA	13.64	13.08
$\rho_1 = \rho_2$[e]	.676	.778	7.02	5.77			16.44	12.42
	(.421)	(.379)	(.009)	(.018)			(.000)	(.000)
$Q(4)$[f]	2.81	.296	5.19	1.46	2.98	.377	1.75	2.33
	(.590)	(.990)	(.268)	(.834)	(.561)	(.984)	(.781)	(.675)

NOTES: [a] Entries in this row are the t statistic for the null hypothesis $\rho_1 = 0$. [b] Entries in this row are the t statistic for the null hypothesis $\rho_2 = 0$. [c] The AIC is calculated as $T^*\log(\text{SSR}) + 2^*n$, where T = number of usable observations, SSR = sum of squared residuals, and n = number of regressors. The BIC is calculated as $T^*\log(\text{SSR}) + n^*\log(T)$. Because the Dickey–Fuller tests were performed on the residuals of the interest differential regressed on a constant, the AIC and BIC for the Dickey–Fuller tests are directly comparable to the other values in the table. [d] Entries in this row are the sample values of ϕ_μ or ϕ_μ^*. [e] Entries in this row are the sample F statistics for the null hypothesis that the adjustment coefficients are equal. Significance levels are in parentheses below. [f] $Q(4)$ is the Ljung–Box statistic that the first four of the residual autocorrelations are jointly equal to 0. The significance level is in parentheses below. [g] An intercept was included in the momentum-consistent models because the mean differs from the attractor (i.e., the residuals do not have a zero mean).

interest-rate differential. The sample values of τ_μ (-3.13 with no lags and -3.58 with one lag) both indicate that the null hypothesis of nonstationarity can be rejected at conventional significance levels. Observe that for each lag length, the estimated adjustment coefficient is between the estimates of ρ_1 and ρ_2 from the corresponding asymmetric adjustment model. In comparing the various models, the diagnostic checks of the residuals indicated that each is adequate. The momentum models, however, fit substantially better than the TAR and the linear attractor models. The AIC and BIC both select the momentum model even though the M-TAR model entails estimating one more coefficient than does the linear adjustment model.

Given that the term structure displays M-TAR adjustment, Chan's (1993) method finds the consistent estimate of the threshold to be 2.64. As shown on the last portion of Table 5, the best-fitting M-TAR model using the consistent estimate of the attractor is given by (t statistics are in parentheses)

$$\begin{aligned}\Delta r_{Dt} = \underset{(2.57)}{.265} &- \underset{(-1.38)}{.061 I_t (r_{Dt-1} + 2.64)} \\ &- \underset{(-5.19)}{.293(1 - I_t)(r_{Dt-1} + 2.64)}. \quad (11)\end{aligned}$$

The residuals of (11) show no evidence of serial correlation, and introducing lagged changes of Δr_D reduces the fit. For example, with one lagged change, the t statistic of Δr_{Dt-1} is only $-.404$. The prominent feature of Table 5 is that the M-TAR models using the consistent estimator fit substantially better than the other models. Moreover, the evidence for asymmetric adjustment is enhanced. Using (11), the F statistic for the null hypothesis $\rho_1 = \rho_2 = 0$ is 13.64 and the F statistic for the null hypothesis $\rho_1 = \rho_2$ is 16.44.

The positive finding of cointegration with M-TAR adjustment justifies the estimation of the asymmetric error-correction model shown in Table 6. Using the consistent estimate of the threshold, fitted equations have the form (with t statistics in parentheses)

$$\begin{aligned}\Delta r_{L,t} = &\underset{F_{11} = 2.89}{A_{11}(L)\Delta r_{L,t-1}} + \underset{F_{12} = 1.12}{A_{12}(L)\Delta r_{S,t-1}} \\ &+ \underset{(2.59)}{.088 z\_\text{plus}_{t-1}} + \underset{(.959)}{.037 z\_\text{minus}_{t-1}}, \quad (12)\end{aligned}$$

and

$$\begin{aligned}\Delta r_{S,t} = &\underset{F_{21} = 9.98}{A_{21}(L)\Delta r_{L,t-1}} + \underset{F_{22} = .259}{A_{22}(L)\Delta r_{S,t-1}} \\ &+ \underset{(.445)}{.030 z\_\text{plus}_{t-1}} - \underset{(-3.10)}{.123 z\_\text{minus}_{t-1}}, \quad (13)\end{aligned}$$

where $z\_\text{plus}_{t-1} = I_t(r_{S,t-1} - r_{L,t-1} + 2.64)$, $z\_\text{minus}_{t-1} = (1 - I_t)(r_{S,t-1} - r_{L,t-1} + 2.64)$, I_t = momentum Heaviside indicator function, $A_{ij}(L)$ is a polynomial in the lag operator L, and F_{ij} is the F statistic for the null hypothesis that all coefficients of $A_{ij}(L) = 0$.

The key feature in (12) and (13) is the pattern of the estimated coefficients for z_plus and z_minus. In (12), the t statistics imply that long-term rate responds to a positive, but not a negative, discrepancy in the term structure. In (13), the t statistics indicate that the short-term rate responds to a negative, but not a positive, discrepancy. Thus, in response to a one-unit *positive* gap, the long-term rate is estimated to rise by .088 units, and in response to a one-unit *negative* gap, the short-term rate is estimated to rise by .123 units. Moreover, the F statistics concerning causality indicate that, at conventional significance levels, both the long-term and short-term rates respond to lagged changes in the long-term rate but not to lagged changes in the short-term rate.

Although not shown in Table 6, (12) and (13) were also estimated using lags of 8, 4, and 1 quarters and various

Table 6. Estimates of the Error-Correction Models

1. The consistent M-TAR model with two lags							
$\Delta r_{L,t}$ =	$-.119$	$+ .255\Delta r_{L,t-1}$	$- .014\Delta r_{L,t-2}$	$- .045\Delta r_{S,t-1}$	$- .067\Delta r_{S,t-2}$	$+ .088z\_\text{plus}_{t-1}$	$+ .037z\_\text{minus}_{t-1}$
	(-1.72)	(2.40)	$(-.130)$	$(-.739)$	(-1.35)	(2.59)	(.959)
	F_{11} = 2.89 (.059)		F_{12} = 1.12 (.329)				
$\Delta r_{S,t}$ =	.129	$+ .793\Delta r_{L,t-1}$	$- .504\Delta r_{L,t-2}$	$- .057\Delta r_{S,t-1}$	$- .055\Delta r_{S,t-2}$	$+ .030z\_\text{plus}_{t-1}$	$- .123z\_\text{minus}_{t-1}$
	(.956)	(3.80)	(-2.44)	$(-.481)$	$(-.565)$	(.445)	(-3.10)
	F_{21} = 9.86 (.000)	F_{22} = .259 (.773)					AIC = -259.26; BIC = -217.88; X^2 = 10.1; Significance = .006
2. The symmetric error-correction model with two lags							
$\Delta r_{L,t}$ =	.013	$+ .218\Delta r_{L,t-1}$	$- .036\Delta r_{L,t-2}$	$- .002\Delta r_{S,t-1}$	$- .060\Delta r_{S,t-2}$	$+ .067(r_{St-1} - r_{Lt-1} + .761)$	
	(.319)	(2.14)	$(-.344)$	$(-.043)$	(-1.22)	(2.31)	
	F_{11} = 2.30 (.104)	F_{12} = .748 (.475)					
$\Delta r_{S,t}$ =	.130	$+ .601\Delta r_{L,t-1}$	$- .617\Delta r_{L,t-2}$	$+ .160\Delta r_{S,t-1}$	$- .022\Delta r_{S,t-2}$	$- .081(r_{St-1} - r_{Lt-1} + .761)$	
	(.155)	(2.92)	(-2.94)	(-1.62)	$(-.226)$	(1.38)	
	F_{21} = .785 (.001)		F_{22} = 1.33 (.267)				AIC = -252.64; BIC = -217.17

NOTE: AIC and BIC are the multiequation AIC and BIC, respectively. F_{ij} is the F statistic for the test that all coefficients of L in the polynomial $A_{ij}(L)$ are equal to 0, the significance level is shown in the parentheses below. The χ^2 statistic is the likelihood ratio test for the null hypothesis that $\rho_{11} = \rho_{12}$, $\rho_{21} = \rho_{22}$ allowing for the different estimated attractors; the significance level of the null hypothesis is also reported.

lag-length tests were performed. The multivariate AIC selected the model with eight lags, and the multivariate BIC selected the two-lag model. In the eight-lag model, aside from a single coefficient at lag 7 in the equation for the short-term interest rate, it was possible to set all coefficients past lag 2 equal to 0. It is plausible that the importance of this single coefficient is spurious. The key point is that the coefficients for the error-correction terms were quite robust to lag length.

By way of contrast, a standard error-correction model assuming symmetric adjustment is reported in the lower portion of Table 6. In spite of the extra two coefficient estimates, both the multivariate AIC and BIC select the asymmetric model over the symmetric model. Moreover, the likelihood ratio test for the restriction implied by symmetric error correction yields a χ^2 value of 10.1 (with a p value of .006). Notice that the coefficients on the error-correction terms in the symmetric model are each between those in the asymmetric model. Respectively, the long-term rate and the short-term rate are predicted to change by .067 units and $-.081$ units multiplied by *any* discrepancy (whether positive or negative) in the term structure. It is interesting that the t statistic for the error-correction term in the equation for the short-term rate is only 1.38. The implication is that only the long-term rate adjusts to restore the term structure.

Because the aim of this section is to illustrate the use of the various tests developed in the article, the impulse-response functions and variance decompositions are not reported here. Note, however, that the impulse responses and variance decompositions depend on the sign of the initial shocks. Given the asymmetric nature of the adjustment process, positive innovations will yield different time paths from negative innovations.

4. CONCLUDING REMARKS

A generalization of the Dickey–Fuller test was developed that can be used to test the null hypothesis of a unit root against the alternative hypothesis of stationarity with asymmetric adjustment. In addition to the standard TAR model, we introduced the momentum threshold autoregressive (M-TAR) model to capture possible "sharp" movements in a sequence. If adjustment is approximately symmetric, the Dickey–Fuller test is more powerful than any of the tests developed here. With several plausible types of asymmetry, however, the power of the proposed statistics—particularly those for the M-TAR model—exceed those of the corresponding Dickey–Fuller test. The use of the statistics was illustrated using the term structure of interest rates. An asymmetric error-correction model with M-TAR adjustment was estimated such that the long-term rate responds to the previous change in the long-term rate and to the deviation from the equilibrium relationship only when there is a *negative* discrepancy between the short-term and long-term rate. The short-term rate responds to the previous changes in both rates and to the deviation from the equilibrium relationship only when there is a *positive* discrepancy between the short-term and long-term rates.

ACKNOWLEDGMENTS

This article was written while the first author was a Visiting Scholar at the University of California at San Diego. We thank Frank J. Breidt, Graham Elliott, Wayne Fuller, four anonymous referees, and workshop members at the University of California at San Diego and the University of California at Riverside for helpful suggestions.

[Received March 1996. Revised August 1997.]

REFERENCES

Balke, N. S., and Fomby, T. B. (1996), "Threshold Cointegration," working paper, Southern Methodist University (submitted to *International Economic Review*).

Chan, K. S. (1993), "Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model," *The Annals of Statistics*, 21, 520–533.

Falk, B. (1986), "Further Evidence on the Asymmetric Behavior of Economic Time Series Over the Business Cycle," *Journal of Political Economy*, 94, 1096–1109.

Granger, C. W. J., and Lee, T. H. (1989), "Investigation of Production, Sales, and Inventory Relationships Using Multicointegration and Nonsymmetric Error-Correction Models," *Journal of Applied Econometrics*, 4, S145–S159.

Granger, C. W. J., and Teräsvirta, T. (1993), *Modelling Nonlinear Economic Relationships*, Oxford, U.K.: Oxford University Press.

Neftci, S. N. (1984), "Are Economic Time Series Asymmetric Over the Business Cycle?" *Journal of Political Economy*, 92, 307–328.

Nelson, C. R., and Plosser, C. I. (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139–162.

Pippenger, M. K., and Goering, G. E. (1993), "A Note on the Empirical Power of Unit Root Tests Under Threshold Processes," *Oxford Bulletin of Economics and Statistics*, 55, 473–481.

Potter, S. (1995), "A Nonlinear Approach to U.S. GNP," *Journal of Applied Econometrics*, 10, 109–125.

Ramsey, J. B., and Rothman, P. (1996), "Time Irreversibility and Business Cycle Asymmetry," *Journal of Money, Credit and Banking*, 28, 1–21.

Sichel, D. E. (1993), "Business Cycle Asymmetry: A Deeper Look," *Economic Inquiry*, 31, 224–236.

Stock, J., and Watson, M. (1988), "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097–1107.

Teräsvirta, T., and Anderson, H. M. (1992), "Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models," *Journal of Applied Econometrics*, 7, S119–S139.

Tong, H. (1983), *Threshold Models in Non-Linear Time Series Analysis*, New York: Springer-Verlag.

——— (1990), *Non-Linear Time-Series: A Dynamical Approach*, Oxford, U.K.: Oxford University Press.

Tsay, R. S. (1989), "Testing and Modeling Threshold Autoregressive Processes," *Journal of the American Statistical Association*, 84, 231–240.