No. 9209

THRESHOLD COINTEGRATION

by

Nathan S. Balke

and

Thomas B. Fomby

Research Paper

Federal Reserve Bank of Dallas

Threshold Cointegration*

Nathan S. Balke

and

Thomas B. Fomby

Department of Economics
Southern Methodist University

Research Associates
Federal Reserve Bank of Dallas

November 11, 1991
Revised February 27, 1992
This version, May 7, 1992

Abstract Cointegration has often been used to model long-run equilibrium relationship between nonstationary variables. However, there may be many instances in which the presence of fixed costs of adjustment prevent continuous adjustment towards the equilibrium. Thus, only when the system gets too far from the equilibrium, does the system move back towards the equilibrium. We model this discontinuous adjustment to a long-run equilibrium as threshold cointegration. Here, the equilbrium error follows a threshold autoregression that is mean reverting outside a given range and a unit root inside this range. This process, while displaying unit root behavior locally, is nonetheless asymptotically stationary. Traditional tests for unit roots such as the Dickey-Fuller test while capable of distinguishing between threshold stationarity and a unit root asymptotically will often have low power against the threshold alternative. We examine the long-run relationship between the Fed Funds rate and the Discount rate and find that this relationship can be characterized as threshold cointegration.

*A previous draft of this paper was presented at the Southern Economic Association Meetings, November 24-26, 1991. We wish to thank Herman Bierens, Margarida Genius, Robert Kunst, Essie Maasoumi, Baldev Raj, and participants of the SMU Econometrics Workshop for helpful comments. This paper in no way represents the views of the Federal Reserve Bank of Dallas or the Federal Reserve System. Of course, we are responsible for any errors.

I. Introduction

The concept of cointegration has been used to capture the notion that nonstationary variables may nonetheless possess long-run equilibrium relationships and, thus, have a tendency to move together in the long-run (see Granger (1986) and Engle and Granger (1987). Cointegration has been used to examine, among many others, the relationship between consumption and income (Campbell (1987)), stock prices and dividends (Campbell and Shiller (1987)), money demand (Johansen and Juselius (1990)), and purchasing power parity (Corbae and Ouliaris (1988)). Granger (1983) showed that systems in which variables are cointegrated can be characterized by an error correction model (ECM). This error correction model describes how the variables respond to deviations from the equilibrium. One can think of the ECM as the adjustment process through which the long-run equilibrium is maintained.

Implicit in much of the discussion of cointegration and its corresponding error correction model (ECM) is the assumption that such a tendency to move toward a long-run equilibrium is always present (for every time period). Yet, it is possible to think of situations in which movement toward the long run equilibrium does not occur in every period. For example, the presence of fixed costs of adjustment may prevent economic agents from adjusting continuously. Only when the deviation from the equilibrium exceeds a critical threshold, do the benefits of adjustment exceed the costs and, hence, economic agents act to move the system back towards the equilibrium.[1] This type of discrete adjustment process has been used to describe many

[1] This type of threshold behavior can be generated from control problems where there are fixed and/or linear adjustment costs of control (see Dixit (1991)). The (S,s), target zone, and reflecting barrier problems are examples of control problems that can generate threshold-type behavior.

economic phenomena including the behavior of inventories, money balances, consumer durables, prices, and employment.[2] Even in efficient financial markets, the presence of transaction costs may create a band in which asset returns are free to diverge and in which arbitrage possibilities exist.

Discrete adjustment may equally apply to policy interventions. For example, exchange rate management and commodity price stabilization are often characterized by discrete interventions. For exchange rate target zones, exchange rates are allowed to fluctuate freely within a given band, yet, when exchange rates exceed the target band, central banks intervene in the foreign exchange market. Similarly, for commodity price stabilization programs, only when the market price gets too far from the target price does the government intervene by buying or selling stocks or by changing the target price. Another example might include the Federal Reserve control of the Fed Funds rate and the Discount rate. If the spread between the two rates gets too large, the Fed intervenes to change the Fed Funds rate or Discount rate or both to prevent sending conflicting signals about monetary policy.

In this paper, we attempt to characterize this discrete adjustment in terms of threshold cointegration. In particular, we examine the case where the cointegrating relationship is inactive inside a given range and then becomes active once the system gets too far from the "equilibrium". That is, once the system exceeds a certain threshold, cointegration becomes active. The concept of threshold cointegration captures the essence of the nonlinear adjustment process envisioned to hold for many economic phenomena, yet, as we show below, allows one to use many of the tools developed for more

---

[2] See, for example, Scarf (1959), Miller and Orr (1966), Bertola and Caballero (1990), Sheshinski and Weiss (1983), and Bentolila and Bertola (1990).

traditional models of cointegration.

The remainder of this paper is organized as follows. In Section II we formally describe two types of threshold cointegration models. One corresponds to a threshold adjustment process that tends towards an equilibrium point while the other corresponds to an adjustment process that reverts to an equilibrium band or target zone. We discuss the properties of the threshold models in Section III. While these processes behave like a random walk inside the threshold range, they are, nonetheless, stationary stochastic processes. In Section IV we examine the asymptotic and finite-sample performance of standard time series methods such as the Dickey-Fuller unit root test. In Section V we examine and describe what appears to be a threshold cointegration relationship between the Fed Funds rate and the Discount rate. In Section VI we suggest additional topics for further research.

II. A Model of Threshold Cointegration

The Basic Threshold Model

To be precise about what we mean by threshold cointegration, consider a simple bivariate system (y_t, x_t) similar to that in Engle and Granger (1987) with:

(1) $y_t + \alpha x_t = z_t$, where $z_t = \rho^{(i)} z_{t-1} + \epsilon_{1t}$

(2) $y_t + \beta x_t = B_t$, where $B_t = B_{t-1} + \epsilon_{2t}$.

For simplicity let ϵ_{1t} and ϵ_{2t} be iid, mean zero random variables with variances σ_1^2 and σ_2^2 respectively.[3] Equation (1) represents the equilibrium relationship between y_t and x_t, where z_t is the deviation from equilibrium and

[3] In general, ϵ_{1t} and ϵ_{2t} could be serially correlated.

the cointegrating vector is given by $(1,\alpha)$. B_t, in equation (2), represents the common stochastic trend of y_t and x_t.

Rather than a linear autoregression with constant parameters as in Engle and Granger (1987), in our case the value of $\rho^{(i)}$ depends on past realizations of z_t. In particular,

$$\begin{aligned} \rho^{(i)} &= 1 && \text{if } |z_{t-d}| \leq \theta \\ &= \rho, \text{ with } |\rho| < 1 && \text{if } |z_{t-d}| \geq \theta \end{aligned}$$

where d is a positive integer. That is, departures from the equilibrium, z_t, follow a threshold autoregression (see Tong (1983)) where the threshold is given by θ.[4] As long as $|z_{t-d}| \leq \theta$, z_t acts as if it had an unit root and, consequently, there is no tendency for the system to drift back towards the equilibrium relationship. Once $|z_{t-d}| > \theta$, z_t becomes a stationary autoregression that has a tendency to revert back to a constant mean (in the example above, zero). Thus, if the equilibrium error is less than the threshold value, then y_t and x_t do not have a tendency to revert to some equilibrium (i.e. are not cointegrated); if the equilibrium error is greater than the threshold then y_t and x_t do tend to move towards some equilibrium (i.e. are cointegrated). The integer d represents the delay in the error correction process and reflects the possibility that economic agents or controllers may react to deviations from the equilibrium with a lag.

An alternative way to represent the threshold cointegration idea is in terms of an error correction model. We can rewrite the system given by equations (1) and (2) as

$$(3) \quad \Delta y_t = \gamma_1^{(i)} z_{t-1} + v_{1t}$$

---

[4] In general, there is no reason to restrict the threshold (or ρ) to be symmetric. We do so here for notational ease.

(4) $\Delta x_t = \gamma_2^{(i)} z_{t-1} + v_{2t}$

where $\gamma_1^{(i)} = -(1-\rho^{(i)})\beta/(\beta-\alpha)$, $\gamma_2^{(i)} = (1-\rho^{(i)})/(\beta-\alpha)$, $v_{1t} = [\beta/(\beta-\alpha)]\epsilon_{1t} - [\alpha/(\beta-\alpha)]\epsilon_{2t}$, $v_{2t} = [1/(\beta-\alpha)](\epsilon_{2t} - \epsilon_{1t})$, and $z_{t-1} = y_{t-1} + \alpha x_{t-1}$. The error correction term, z_{t-1}, represents the error in or deviation from the equilibrium condition while the parameters $\gamma_1^{(i)}$ and $\gamma_2^{(i)}$ capture how y_t and x_t respond to deviations from the equilibrium relationship. As long as deviations from the equilibrium condition are not greater than the threshold, the error correction parameters $\gamma_1^{(i)}$ and $\gamma_2^{(i)}$ are zero and y_t and x_t do not respond to deviations from the equilibrium condition. Only if the deviations exceed the threshold are $\gamma_1^{(i)}$ and $\gamma_2^{(i)}$ nonzero and y_t and x_t respond to deviations from the equilibrium.

In addition to the basic threshold model described above, there is a related threshold model that is also of interest. In the basic threshold model, the error correction model responds to the deviations from the equilibrium relationship; the strength of the error correction effect depends, in part, on how far the variable is away from the equilibrium relationship. In a control context, it is as if the controller is trying to return the controlled variable back to its equilibrium value. However, one can conceive of situations where the controller is satisfied if the process is within a band centered around the equilibrium. An example would include exchange rate target zones. In this case, the controller tries to return the variable to within the target band and not necessarily back to an equilibrium point.

In this case the error correction model responds to deviations from the target band; the strength of the error correction term then depends on how far the variable is from the equilibrium <u>band</u>. For this type of target zone threshold model, deviations from equilibrium are described by

$$(5)\quad z_t = z_{t-1} + \epsilon_{1t} \qquad \text{if } |z_{t-d}| \le \theta$$
$$= z_{t-1} + (\rho-1)(z_{t-1} - \theta^{(i)}) + \epsilon_{1t} \qquad \text{if } |z_{t-d}| > \theta$$

where $\theta^{(i)} = \theta$ if $z_{t-d} > \theta$ and $\theta^{(i)} = -\theta$ if $z_{t-d} < -\theta$. Thus, the series has a tendency to revert back to the range $[-\theta,\theta]$ and not to zero as in the previous threshold model.

A special case of this target zone threshold model is a two-sided barrier model in which the series is not allowed to exit the interval $[-\theta,\theta]$. That is

$$(6)\quad z_t = z_{t-1} + \epsilon_{1t} \qquad \text{if } |z_{t-1} + \epsilon_{1t}| \le \theta$$
$$= \theta^{(i)} \qquad \text{if } |z_{t-1} + \epsilon_{1t}| > \theta$$

where $\theta^{(i)} = \theta$ if $(z_{t-1} + \epsilon_{1t}) > \theta$ and $\theta^{(i)} = -\theta$ if $(z_{t-1} + \epsilon_{1t}) < -\theta$. Thus, the series has a unit root as long as it is within the band, but it is not allowed to exit from the band. This is similar to the modified threshold model when $\rho = 0$. A continuous time analog of the Barrier model is a Brownian Motion on the interval $[-\theta,\theta]$ where $\pm\theta$ are reflecting boundaries.

Generalization of the threshold cointegration model.

We can describe the above threshold models in a somewhat more general context. Consider the system:

$$(7)\quad A^{(i)}(L)X_t = \epsilon_t$$

where $A^{(i)}(L)$ is a nxn polynomial lag matrix and X_t is a nx1 vector of nonstationary variables. Suppose there are k cointegrating vectors. The presence of cointegration reduces the rank of the $A^{(i)}(1)$ matrix. Threshold cointegration in the general setting implies that $A^{(i)}(1)$ is of reduced rank when $|\alpha' X_{t-d}|_j > \theta_j$ for the j-th cointegrating vector $(1 \le j \le k)$ and $A^{(i)}(1)$

is of full rank when $|\alpha' X_{t-d}|_j \leq \theta_j$ for all j. Here α' denotes the kxn matrix of cointegrating vectors and θ_j is the threshold for the j-th cointegrating relationship.

In terms of the error correction model, we can rewrite the general system

(8) $\Delta X_t = C(L)\Delta X_t + \delta^{(i)}\alpha' X_{t-1} + v_t$

where $\delta^{(i)}$ is a nxk matrix containing the error correction parameters. Like the simple example described above, the matrix of error correction parameters has the following property:

$$\delta^{(i)}{}_j = 0, \text{ if } |\alpha' X_{t-d}|_j \leq \theta_j$$

$$\delta^{(i)}{}_j \neq 0, \text{ if } |\alpha' X_{t-d}|_j > \theta_j,$$

where $\delta^{(i)}{}_j$ is the j-th column of error correction parameter matrix, $\delta^{(i)}$, that corresponds to the j-th cointegrating vector.

III. Properties of the Threshold Cointegration Equilibrium Error

In this section we examine some of the stochastic properties of threshold cointegrated variables. In particular, we focus on the behavior of the equilibrium error, z_t.

As suggested above, the behavior of the equilibrium error, z_t, depends on which region the equilibrium error is in; if z_{t-d} is in the interval $[-\theta,\theta]$ then z_t is a unit root process, if z_{t-d} is outside the interval then z_t is a stationary, mean reverting process. This threshold model displays very different "local" behavior as compared to its global behavior. During substantial portions of the sample (i.e., when the series is inside the range $[-\theta,\theta]$) the series behaves much like a random walk; yet, asymptotically, the series has a stationary distribution.

To see the distinction between the local and asymptotic (or global behavior), consider the target zone threshold model given by equation (5). Without loss of generality, we set the delay parameter, d, equal to one. Thus, z_t is given by

$$(9)\quad z_t = z_{t-1} + \epsilon_{1t} \qquad \text{if } |z_{t-1}| \le \theta$$

$$= (1-\rho)\theta^{(i)} + \rho z_{t-1} + \epsilon_{1t} \qquad \text{if } |z_{t-1}| > \theta$$

where ϵ_{1t} is iid $(0,\sigma_1^2)$, and $\theta^{(i)} = \theta$ if $z_{t-1} > \theta$ and $\theta^{(i)} = -\theta$ if $z_{t-1} < -\theta$. Clearly as long as z_{t-1} is in the interval $[-\theta,\theta]$, z_t behaves like a random walk.

However, asymptotically, z_t is a stationary stochastic process. Starting at $t = 0$ with $|z_0| \le \theta$ and by recursively substituting, we obtain the following equation for z_t when $|z_{t-1}| \le \theta$:

$$(10)\quad z_t = \rho^{N(t)} z_0 + \sum_{i=1}^{NE(t)} \{ \rho^{[N(t)-N_i]} (1-\rho^{[\kappa(i)-\tau(i)]})\theta^{(i)} \}$$

$$+ \sum_{i=1}^{NE(t)} \{ \rho^{[N(t)-N_{i-1}]} \sum_{j=\kappa(i-1)+1}^{\tau(i)} \epsilon_{1j}$$

$$+ \rho^{[N(t)-N_i]} \sum_{j=\tau(i)+1}^{\kappa(i)} \rho^{[j-\tau(i)-1]} \epsilon_{1j} \}$$

$$+ \sum_{j=\kappa(NE(t))+1}^{t} \epsilon_{1j}.$$

$NE(t)$ is the number of times the process exited from $[-\theta,\theta]$ in the time interval $[0,t]$. $\tau(i)$ is the time of the i-th exit from $[-\theta,\theta]$ while $\kappa(i)$ is the time of the i-th entrance into $[-\theta,\theta]$ with $\kappa(0) = 0$. The term $\kappa(i)-\tau(i)$ is the time outside the boundary for the i-th exit, and, since $|z_0| \le \theta$, $\kappa(i)-\tau(i) > 0$. The term $N(t) = \sum_{i=1}^{NE(t)} [\kappa(i)-\tau(i)]$ is the total time the process has spent outside $[-\theta,\theta]$ during the interval $[0,t]$ while N_i is the time the process has spent outside $[-\theta,\theta]$ during the first i exits.[5]

[5] A similar description for z_t can be derived for the case where z_{t-1} is outside $[-\theta,\theta]$ and for the case where $|z_0| > \theta$. Likewise, to obtain z_t for the basic threshold model, all one must do is set the $\theta^{(\cdot)}$ terms in equation (10) equal to zero.

The terms $\tau(i)$, $\kappa(i)$, N(t), and NE(t) are themselves random variables. In general, these hitting times will be a function of the boundaries, the degree of mean reversion outside of the boundaries, as well as realizations of the random variable ϵ_{1t}. The expected first passage time to the boundaries starting from z_k inside the interval $[-\theta,\theta]$ is approximately equal to $(\theta^2 - z_k^2)/\sigma_1^2$. Thus, the size of the boundaries relative to the variance of ϵ_{1t} plays an important role in the number of times the process crosses the boundaries in the time interval [0,t].

For the z_t process described in equation (10), note that the effect on z_t of z_0 (as well as hits on the boundary and ϵ_{1i} innovations early in the sample) diminishes the more times the process exceeds the boundaries. Because the boundaries are finite, the expected hitting time to the boundaries from anywhere inside the region $[-\theta,\theta]$ is finite. Furthermore, since z_t is mean reverting outside of the boundaries, the expected hitting time to the boundaries from outside the region $[-\theta,\theta]$ is also finite. Consequently, as $t \to \infty$, $NE(t) \xrightarrow{a.s} \infty$. Similarly, as $t \to \infty$, $N(T) \xrightarrow{a.s} \infty$. This implies that $corr(z_t,z_0)$ approaches zero as t goes to infinity. Essentially, each time the process exceeds the boundaries, some of the memory of the process is eliminated. Because the number of times the process will exit $[-\theta,\theta]$ in a given time interval goes to infinity as the time interval goes to infinity, events that are separated by large time intervals are almost independent. This suggests that since none of the other parameters are time dependent, z_t will be a stationary stochastic process.

In principle, it is possible to solve for the stationary distributions of z_t. Unfortunately, because the time that the process spends outside of the boundaries is stochastic, solving for the stationary asymptotic distributions of the above threshold processes is not in general practicable. However, in Appendix A, we solve for the stationary distributions of continuous time analogs of both the basic and the target zone threshold models. In general, the unconditional variance of the process depends on σ_1^2 and on the value of the boundaries.

IV. Unit root tests and threshold cointegration

In this section we examine how standard time series methods would work in the presence of threshold cointegration (stationarity). We show that the standard time series analyses used for linear cointegration are likely to be valid asymptotically for the threshold cointegration case. However, in finite samples, traditional linear methods such as the Dickey-Fuller test will have lower power against the threshold alternative.

Proposition Consider the threshold cointegration system given by equations (1) and (2).

(i) Consider the Dickey-Fuller regression given by

$$\hat{\rho}_{DFz} = {}_{t=2}\Sigma^{T}\, z_{t-1}z_t \; / \; {}_{t=2}\Sigma^{T}\, z_{t-1}^2.$$

As long as z_t is "α-mixing", then $\rho < \text{plim}\ \hat{\rho}_{DFz} < 1$.

The term ρ is the autoregressive parameter when $|z_{t-d}| > \theta$.

(ii) As long as the boundaries, ρ, and $\text{Var}(\epsilon_{1t}) = \sigma_1^2$ are such that the α-mixing conditions in Phillips (1987) (Assumption 2.1) are satisfied for z_t, then

(a) for $\hat{\rho}$ from Dickey-Fuller regressions for x_t and y_t,

$$\text{plim}\ T^{1-\delta}(\hat{\rho}-1) = 0, \quad \delta > 0,$$

(b) and the least squares estimate of the cointegrating vector, $\hat{\alpha}$,

$\operatorname{plim} T^{1-\delta}(\hat{\alpha}-\alpha) = 0, \quad \delta > 0.$ ■

Part (i) implies that the Dickey-Fuller regression is capable of distinguishing between threshold stationarity and a unit root, asymptotically. However, the estimated value of ρ understates the degree of mean reversion outside of the boundaries. Parts (ii) and (iii) imply that the super-consistency of least squares estimates of the Dickey-Fuller coefficient for the unit root processes x_t and y_t (Phillips (1987)) as well as the super-consistency of the cointegrating vector, α, (Stock (1987)) hold in the threshold cointegration case. In summary, Proposition 1 suggests that techniques designed to detect unit roots and cointegration in the linear case should work, asymptotically, for the threshold case.

The memory condition for z_t (α-mixing) in the Proposition is needed in order to apply the Law of Large Numbers and Central Limit Theorems for a serially correlated z_t. As it turns out, this memory condition is very likely to hold for threshold cointegration models in which θ is bounded.[6] As we pointed out above, each time the process exceeds the boundaries, some of the memory of the process is eliminated. Indeed, for the basic threshold model with ρ=0, the process essentially starts over again at zero every time it exceeds the boundary. Whether the decay in the memory is fast enough for the Central Limit Theorem to hold will depend on the value of θ relative to σ_1 and on the parameter ρ. All the threshold processes considered in the Monte

---

[6] See Appendix C for a more detailed discussion of the memory properties of the threshold stationary models examined here.

Carlo experiment conducted below appear to satisfy the Law of Large Numbers and the Central Limit Theorem.

Given that the threshold models described above are in fact stationary, perhaps it is not too surprising that standard time series methods such as Dickey-Fuller tests are capable of detecting threshold cointegration (stationarity). However, as we suggested above, since these threshold processes sometimes behave locally as if they have a unit root, in finite samples, traditional methods may have difficulty uncovering the presence of threshold cointegration. In particular, traditional tests for no cointegration (or a unit root in the deviation from equilibrium) may have low power against the alternative of threshold cointegration (or threshold stationarity).

Finite Sample Performance of Dickey-Fuller Test.

To evaluate the finite sample performance of standard tests of no cointegration/nonstationarity, we abstract from the problem of estimating the cointegrating vector and assume that the cointegrating vector is known. Thus, we need only consider how effective traditional methods such as the Dickey-Fuller test are in distinguishing the threshold models from unit root processes.[7]

Three threshold models are generated: one corresponding to the basic threshold model described in section II; one corresponding to the target zone threshold model in which there is a tendency for the series to return to a

[7] In a previous draft of this paper, we examined Cochrane (1988) variance ratio as well as Bieren's (1992) test for stationarity. Like the Dickey-Fuller test, both tests had difficulty distinguishing between the unit-root process and threshold stationary models.

target band but not necessarily to an equilibrium point; and one corresponding to a barrier process in which the process is not allowed to exit the interval $[-\theta,\theta]$. In the Monte Carlo experiment all three models have the same random innovations (i.e., they have the same ϵ_{1t}'s) but differ with respect to the threshold behavior. The shocks, ϵ_{1t}, were drawn from a N(0,1) distribution.

Three values of ρ were considered: (0.0, 0.4, 0.8). The value of ρ captures how strong the attraction of the equilibrium is, with $\rho = 0.0$ representing the case of the strongest attraction and $\rho = 0.8$ the least attraction. We also considered three threshold values, $\theta = 5$, 10, and 20. Given the random walk behavior when the series is within the threshold boundary, the expected first hitting time given that $z_t = 0$ at $t=0$ is approximately θ^2. Thus, starting from the equilibrium point (zero), the expected hitting times for the thresholds $\theta = 5$, 10, 20 are 25, 100, and 400 respectively. Sample sizes of 100, 250, 500, and 1000 are considered. Each experiment consists of 1000 replications.

Table 1 displays the power of the Dickey-Fuller t-test against the alternative hypothesis of the various threshold models. The Dickey-Fuller test has low power for small samples or for large threshold values. Indeed, the sample size relative to the value of θ^2/σ_1^2 (which is the expected hitting time starting from zero) seems to be the key determinant of the power of the Dickey-Fuller statistic.[8] Only when the sample size is substantially greater than the expected hitting time does the Dickey-Fuller test have high power

[8] In appendix A, we derive the asymptotic Dickey–Fuller coefficients for continuous time versions of the above threshold models. From the continuous time analysis, we show that these coefficient estimates are explicitly a function of ratio θ^2/σ_1^2.

against the threshold alternatives. Also, the Dickey-Fuller test have substantially less power against the target zone threshold and the barrier process alternatives than they do against the basic threshold alternative.

V. An Example of Threshold Cointegration

In this section, we examine whether the relationship between the Fed Funds rate and the Discount rate can be characterized by threshold cointegration. The Discount rate is the interest rate at which member banks can borrow from the Federal Reserve and is set by the Fed. The Fed Funds rate is a market determined interest rate for overnight loans between banks. While the Fed does not set the Fed Funds rate directly, it nonetheless can influence this rate through open market operations. There are several reasons why the Fed would not want the spread between the Fed Funds rate and the discount rate to get too large. Too large a spread may cause substantial swings in discount window borrowing--if the Fed Funds rate is too high relative to the discount rate, banks may attempt to take advantage of the interest rate spread by borrowing at the discount window--undermining the window's lender-of-last-resort role. Furthermore, both interest rates reflect the stance of monetary policy. The Fed does not want the spread between these rates to get too large since this would send conflicting signals about the conduct of monetary policy.

Because the spread between the Fed Funds rate and the Discount rate is the control variable of interest for the Federal Reserve in this example, we look for stationarity in the spread as evidence of an equilibrium relationship between the Fed Funds rate and the Discount rate. Therefore, the cointegrating vector is taken as known and equal to (1, -1). The data are monthly and the sample spans from January 1955 to December 1990.

Before examining the spread between the Fed Funds and the Discount rates, we need to determine the univariate time series properties of these series. Both the Fed Funds rate and the Discount rate individually show evidence consistent with unit roots; the augmented Dickey-Fuller t-statistics (with 12 lags) for the Fed Funds and the Discount rates are -1.99 and -1.93, respectively.[9] However, when the spread is examined, we can reject the null hypothesis of a unit-root--the Dickey-Fuller t-statistic is -6.08. Thus, the full sample suggests that Fed Funds and the Discount rate are cointegrated. We next consider the possibility that the Fed Funds and the Discount rate are threshold cointegrated.

We use the methodology suggested by Tsay (1989) to test for and model the threshold autoregression for the spread. The Tsay threshold autoregression procedure consists of several steps. First, a tentative AR model of order p and a set of possible threshold variables (the spread at t-d) is selected. An autoregression with 2 lags for the spread is sufficient to reduce the residuals to white noise, so p is tentatively set equal to 2. Because the Federal Reserve Open Market Committee typically meets about every six weeks, we set the possible range for the threshold lag, d, from 1 to 4 months.

For each possible threshold lag variable (d = 1 to 4), the Tsay (1989) test for threshold nonlinearity is conducted based on an arranged autoregression. An arranged autoregression orders the data according to the

[9] We also considered the possibility that the Fed Funds and Discount rates were themselves stationary threshold autoregressions. Using the procedure outlined below, we found a single threshold for the Discount rate and two thresholds for the Fed Funds rate. However, for both interest rate series, unit roots appear to be present in all of the threshold regimes.

value of the threshold variable. The Tsay statistic for threshold nonlinearity is the F-statistic from the regression of recursive residuals from the arranged autoregression on lagged values of the series. Under the null hypothesis of linearity, coefficients of lagged values of the series should be zero. Thus, high F statistics are evidence against the null of a linear autoregression. Because several observations (here 20 observations) are used to start up the recursive analysis, we consider arranging the autoregression from both low to high values of the threshold variable and from high to low values. The threshold lag is determined by choosing the lag d that yields the largest F statistic. In this application, d = 1 is the chosen threshold lag. The Tsay tests for nonlinearity (with d = 1) are:

autoregression arranged from low to high-- $F(3,409) = 13.47$ ($p = 0.0000$)

autoregression arranged from high to low-- $F(3,409) = 29.28$ ($p = 0.0000$).

Therefore, the null of linearity is strongly rejected.

Once the threshold lag is chosen, we use the arranged autoregression based on the chosen threshold variable (the spread at t-1) to identify the possible threshold values (the θ's). This is done by examining scatterplots of the recursive t-statistics of the autoregressive coefficients and/or recursive residuals. Since our focus is on the stationarity of the spread, we examine the recursive Dickey-Fuller t-statistics implied by the arranged autoregression. Figures 1 and 2 present scatterplots of the recursive Dickey-Fuller t-statistics for arranged autoregressions against possible threshold values. Breaks or changes in direction in the plots suggest candidate threshold values. Figures 1 and 2 suggest a clear break in the recursive t-statistics between values of 1.6 and 2.0 of the threshold variable. A second change of direction appears to be present at or around the threshold value of zero. Thus, there does indeed seem to be evidence of two thresholds--the upper threshold in the range 1.6 to 2.0 and the lower threshold around zero.

To determine the actual threshold values, we estimate several threshold autoregressions with possible lower threshold values of (-0.2, -0.1, 0.0, 0.1, 0.2) and upper threshold values of (1.6, 1.7, 1.8, 1.9, 2.0). The final threshold values were those that minimized the sum of squared errors from the threshold autoregression for the spread. Of the possible combinations of lower and upper threshold values, an upper threshold of 1.6 and a lower threshold -0.2 minimized the sum of squared errors. Figure 3 plots the spread between the Fed Funds and Discount rates and the identified threshold values.

Table 2 presents estimates of the threshold autoregression for the spread between Fed Fund rate and the Discount rate.[10] The Augmented Dickey-Fuller t-statistics implied by the threshold autoregression are also presented. From Table 2, it is clear that the relationship between the Fed Funds and Discount rates can be characterized by threshold cointegration. As long as the spread between the Fed Funds rate and the Discount rate is in the range [-0.2, 1.6] there does not appear to be any mean reversion--the spread has a unit root. However, when the spread is greater than 1.6 percentage points or less than -0.2 percentage points there is strong evidence of mean reversion. The estimated constant terms in the lower regime and upper regime

[10] In all the autoregressions in Table 2, a lag length of 2 was sufficient to eliminate any serial correlation in the residuals. Similarly, the residuals of the error correction models in Table 2 were also white noise.

suggest that the threshold model is more like the target zone threshold model given by equation (5) than the basic threshold model--that is, the spread tends to return to an equilibrium range rather than to an equilibrium point.

The estimated threshold error correction models for the Fed Funds rate and the Discount rate are also consistent with threshold cointegration between the Fed Funds and the Discount rates. For the Fed Funds rate, the error correction term (the coefficient on spr_{t-1}) is significant when the spread at t-1 is outside the range [-0.2, 1.6]; thus when the spread is too large (either positively or negatively) the Fed Funds rate adjusts to narrow that spread. Inside this range, the Fed Funds rate does not respond to the spread. The Discount rate does not appear to respond to the spread between the Fed Funds and the Discount rate. This suggests that the cointegration between the Fed Funds rate and the Discount rate appears to be in large part due to adjustments in the Fed Funds rate.[11]

VI. Summary and Further Research

Thus far, we have presented a model in which the cointegrating relationship between variables turns on and off. We modeled this on and off behavior explicitly as a threshold model in which the series are cointegrated if they get too far away from the equilibrium relationship but are not cointegrated as long as they are relatively close to the equilibrium. While standard time series methods should be able to detect threshold cointegration asymptotically, in finite samples and for relatively large threshold values these same methods may have trouble detecting threshold cointegration.

---

[11] The fact that discount rate changes occur relatively infrequently may also account for the lack of statistical significance of the error correction term in the discount rate equation.

In addition to presenting the general threshold model, we examine the behavior of the Fed Funds rate and the Discount rate in light of possible threshold cointegration. Using the threshold modeling strategy suggested by Tsay (1989), we find significant evidence of threshold cointegration between the Fed Funds rate and the Discount rate. As long as the spread between the two interest rates is within a given range, there is no cointegration. But when the spread is outside this range, the Fed Funds rate and the Discount rate are cointegrated.

Several tasks remain to be done. Examining the properties of multivariate procedures such as Stock and Watson (1988) or Johansen (1991) in the presence of threshold cointegration needs to be considered. As we suggested above, these methods are still likely to be asymptotically capable of finding threshold cointegration but with a loss of power (or incorrect size) relative to the basic linear model. Perhaps, the nonlinear attractors approach of Granger and Hallman (1990) may be more effective at distinguishing threshold cointegration from no cointegration than standard linear methods.

In the application above, we took the cointegrating vector to be known. However, in practice this assumption is rarely valid. Therefore, estimation and inference for cointegrating vectors in the threshold cointegration context needs to be examined. The super-consistency of least squares estimates of the cointegrating vector (Stock (1987)) will still hold. However, in finite samples, the estimated cointegrating vector is likely to be noisier for threshold cointegration than for conventional linear cointegration.

In the example above, we used the Tsay (1989) procedure to estimate a threshold cointegration model. However, there are circumstances in which the Tsay procedure may not be as useful or appropriate. While the Tsay threshold modeling procedure is able to model threshold cointegration, given that the cointegrating vector is already estimated, it may not be very useful if there are only a few observations outside of the thresholds. In particular, the Tsay threshold modeling procedure is unlikely to be able to detect and model the two-sided Barrier process. In addition, we would like to consider examining threshold cointegration in a systems context; univariate methods may not be as efficient as a systems approach would be.

Finally, we would like to consider some additional economic examples in which threshold cointegration might be present. There is a large literature that uses cointegration to examine purchasing power parity (for example, Corbae and Ouliaris (1988)). Perhaps, the rejections of purchasing power parity are due to the relatively lower power of tests of cointegration in the presence of threshold cointegration. Another example of interest would be whether consumption, especially consumer durables, and income are characterized by threshold cointegration. Recently, Bertola and Caballero (1990) have applied (S,s) techniques to model the purchases of consumer durables. That is, consumers wait until their stock of durable purchases reaches a given threshold (either upper and lower thresholds) before making a durables purchase. This implies that at an individual level, consumer durables and income are threshold cointegrated. Because of aggregation, aggregate consumer durables and income may be modeled as a smooth transition threshold model (Terasvirta (1990)).

Appendices.

Appendix A. Continuous Time Threshold Models.

In continuous time, the threshold cointegration model is given by

(A1) $y_t + \alpha x_t = z_t$

(A2) $y_t + \beta x_t = B_t,$

where B_t is a Brownian Motion and z_t is the continuous time version of the threshold autoregression. We consider continuous time versions of both the basic threshold model and the target zone threshold model.

The basic threshold model would correspond to a Brownian motion that "jumps" back towards the equilibrium level whenever it hits the boundaries $-\theta$ and θ. This process is described by a Brownian motion of the form

(A3) $dz_t = \sigma dW_t, \qquad -\theta < z_t < \theta$

where W_t is a Wiener Process with $(W_{t+k}-W_t) \sim N(0,k)$. When the process hits the boundaries $+(-)\theta$, the process jumps to $+(-)\rho\theta$ $(0 \le \rho < 1)$. The value of ρ determines how far the process jumps back towards the equilibrium; for $\rho = 0$, the process jumps all the way back to the equilibrium.

We can also construct a continuous time version of target zone threshold model in which the process drifts back to an equilibrium (or target) zone. This process has the form

$$\text{(A4)}\quad dz_t = \begin{cases} -\nu(z_t + \theta)dt + \sigma dW_t & \text{for } z_t < -\theta \\ \sigma dW_t & \text{for } -\theta \le z_t \le \theta. \\ -\nu(z_t - \theta)dt + \sigma dW_t & \text{for } z_t > \theta \end{cases}$$

Thus, the process is a Brownian Motion inside the region $[-\theta,\theta]$ and an mean reverting Ornstein-Uhlenbeck process outside that region. The parameter ν controls the strength (or speed) of the attraction to the boundaries; for ν small there is weak attraction while for ν large there is strong attraction.

In the absence of the boundaries, the distribution of a uncontrolled Brownian Motion $(B_t|B_0)$ is given by $P(B_t \leq B \mid B_0) = \Phi((B-B_0)/\sigma\surd t)$ where Φ is the cumulative distribution function for the standard normal distribution. Note that like the discrete time random walk, the variance of the uncontrolled Brownian Motion grows linearly with time. However, the presence of the boundaries causes both threshold process to have stationary distributions. The "jump" process z_t described in equation (A3) has a stationary distribution described by the density function

$$\text{(A5)}\quad \phi(z) = \begin{cases} 0 & \text{for } z \leq -\theta \\ (\theta + z)/[(1-\rho^2)\theta^2] & \text{for } -\theta < z < -\rho\theta \\ 1/[(1+\rho)\theta] & \text{for } -\rho\theta \leq z \leq \rho\theta, \\ (\theta - z)/[(1-\rho^2)\theta^2] & \text{for } \rho\theta < z < \theta \\ 0 & \text{for } z \geq \theta \end{cases}$$

The asymptotic distribution of target zone threshold process described by equation (A4) has the following density function:

$$\text{(A6)}\quad \phi(z) = \begin{cases} \{[2\pi\sigma^2/(2\nu)]^{1/2} + 2\theta\}^{-1} \exp[-(z+\theta)^2/(\sigma^2/\nu)] & \text{for } z < -\theta \\ \{[2\pi\sigma^2/(2\nu)]^{1/2} + 2\theta\}^{-1} & \text{for } -\theta \leq z \leq \theta. \\ \{[2\pi\sigma^2/(2\nu)]^{1/2} + 2\theta\}^{-1} \exp[-(z-\theta)^2/(\sigma^2/\nu)] & \text{for } z > \theta \end{cases}$$

This process has two interesting limiting cases. As $\nu \to \infty$ the process becomes a reflected Brownian Motion on $[-\theta,\theta]$ which is equivalent to the two-sided barrier process described above. The asymptotic distribution for this process is a uniform distribution. As $\theta \to 0$, the process becomes an Ornstein-Uhlenbeck process over the entire range of z and is the continuous time analog of a mean reverting autoregressive process. This process has a asymptotic distribution of $N(0,\sigma^2/(2\nu))$.

To find the stationary distribution, we use standard results from the diffusion processes literature (see Karlin and Taylor (1981)). For the basic threshold model, z_t is a Brownian motion in the interval $[-\theta,\theta]$ that jumps to $\pm\rho\theta$ when the process hits the boundaries. For this process the stationary limiting distribution $\phi(z)$ will satisfy the following Kolmogorov forward differential equation

$$\sigma^2\phi''(z)/2 = 0 \qquad \text{for } -\theta < z < \theta.$$

(In the steady state $\partial\phi/\partial t = 0$, which yields the above differential equation). On the boundaries and return points $\phi(z)$ must satisfy

$$\phi(\theta) = 0, \qquad \phi(-\theta) = 0,$$

$$\phi'_-(\rho\theta) = \phi'_+(\rho\theta) - \phi'_-(\theta), \qquad \phi'_-(-\rho\theta) = \phi'_+(-\rho\theta) + \phi'_+(-\theta),$$

$$\phi_-(\rho\theta) = \phi_+(\rho\theta), \text{ and } \qquad \phi_-(-\rho\theta) = \phi_+(-\rho\theta),$$

where (-) indicates evaluated from below and (+) indicates evaluated from above. Solving the above differential equation and imposing the boundary conditions along with the adding up constraint, $\int_{-\theta}^{\theta} \phi(z)\,dz = 1$, yields the limiting distribution given in equation (A5).

Similarly, the limiting distribution of the continuous time process whose behavior is given by equation (A4) will satisfy the following Kolmogorov forward differential equations:

$$0 = \sigma^2\phi''(z)/2 + \phi'(z)\nu(z-\theta) + \nu\phi(z) \qquad \text{for } z > \theta$$

$$0 = \sigma^2\phi''(z)/2 \qquad \text{for } -\theta < z < \theta$$

$$0 = \sigma^2\phi''(z)/2 + \phi'(z)\nu(z+\theta) + \nu\phi(z) \qquad \text{for } z < -\theta.$$

In addition, the following boundary conditions will need to be satisfied

$$\phi'_-(\theta) = \phi'_+(\theta), \qquad \phi'_+(-\theta) = \phi'_-(-\theta),$$

$$\phi_-(\theta) = \phi_+(\theta), \text{ and } \qquad \phi_+(-\theta) = \phi_-(-\theta).$$

Finally, we have the adding up constraint $\int_{-\infty}^{\infty} \phi(z)\,dz = 1$. Solving the three differential equations and imposing the various boundary and adding up constraints yields the limiting distribution given by equation (A6).

Appendix B Continuous Time Dickey Fuller Regressions

Consider the continuous time version of the Dickey-Fuller regression

$$-\hat{\nu} = {}_0\int^T z_t \, dz_t \, / \, {}_0\int^T (z_t)^2 \, dt$$

where z_t is a continuous time stochastic process and $-\nu$ represents the reversion or the strength of attraction to the equilibrium. If z_t is the continuous time version of a unit root process, i.e. an uncontrolled Brownian Motion, then $\nu = 0$.

Proposition B1. For the Brownian Motion that returns to $\pm\, \rho\theta$ $(0 \leq \rho < 1)$ when it reaches the boundaries $\pm\, \theta$ as given by equation (A3), the

$$\text{plim } -\hat{\nu} = -\, 6\sigma^2/[\theta^2(1+\rho)(1+\rho^2)] < 0. \qquad \blacksquare$$

Proposition B1 suggests that the linear Dickey-Fuller test should be able to distinguish between the unit root process and the threshold-jump process. The threshold model looks more like the uncontrolled Brownian Motion the smaller process jumps back towards to the equilibrium (i.e. ρ is larger). In addition, the smaller the ratio σ^2/θ^2 (which is $1/E(\tau)$ where $E(\tau)$ is the expected hitting time of reaching the boundary starting from zero), the closer $\hat{\nu}$ is to zero. As σ^2 falls relative to θ^2, the thresholds are reached less often; hence, the threshold process looks more like an uncontrolled Brownian Motion.

Proposition B2. Consider the threshold process given by equation (A4) in which the process drifts back to the target range $[-\theta, \theta]$.

$$-\nu < \text{plim } -\hat{\nu} = -\ \nu\sigma^2/[\ \sigma^2 + 2\nu\theta^2/3 + 2\theta(\sigma^2 + 2\nu(\pi\sigma^2/\nu)^{.5})/IC\ ] < 0,$$

where $IC = [\pi\sigma^2/\nu]^{1/2} + 2\theta$. ∎

Thus, as in the case of the threshold jump process, the standard Dickey-Fuller regression can distinguish the target zone threshold process model from the random walk. However, the Dickey-Fuller regression understates the degree of mean reversion outside of the range $[-\theta,\theta]$. Note, that as σ^2/θ^2 gets small the process becomes more like an uncontrolled Brownian Motion (i.e. as $\sigma^2/\theta^2 \rightarrow 0$, plim $-\hat{\nu} \rightarrow 0$), while as σ^2/θ^2 gets large the process becomes more like a mean reverting Ornstein-Uhlenbeck process (as $\sigma^2/\theta^2 \rightarrow \infty$, plim $-\hat{\nu} \rightarrow -\nu$). For the special case of the reflecting Brownian Motion ($\nu = \infty$), plim $-\hat{\nu}$ approaches $-(3/2)(\sigma^2/\theta^2)$.

Proof of Proposition B1.

Note that for this process the unconditional mean and variance[12] equals

$$E(z) = 0 \text{ and } E(z^2) = (1+\rho^2)\theta^2/6.$$

Furthermore, because of symmetry of the boundaries and return points, the expected first passage time to the boundaries starting from either return point is $E(\tau) = (1-\rho^2)\theta^2/\sigma^2$.

Recall the continuous time analog of the Dickey-Fuller regression is

$$-\hat{\nu} = {}_0\!\int^T z_t dz_t \ /\ {}_0\!\int^T (z_t)^2\, dt.$$

Note that for the above jump process, when the process in the interior $E(z_t dz_t) = 0$. When the process hits the boundary and jumps bank into the interior, $z_t dz_t = [\pm\theta][(\rho-1)(\pm\theta)] = (\rho-1)\theta^2$. Thus,

[12] Robert Kunst pointed out an algebraic error in a previous version of the paper.

$$\text{plim} \, {}_0\!\int^T z_t dz_t/T = (\rho-1)\theta^2 \, \text{plim} \, N(T)/T,$$

where N(T) is the number times that the process hits the boundaries in the [0,T] time interval.

Because of the recursive nature of the above regulated Brownian Motion, the hitting times are independent and identically distributed with the exception of the first hitting time whose distribution depends on the starting value of z. As a consequence, from Renewal Theory,

$$\text{plim} \, N(T)/T = 1/E(\tau) = \sigma^2/[(1-\rho^2)\theta^2],$$

(see Ross (1983)). Note, $\lim_{T\to\infty} \text{Var}(N(T))/T = \text{Var}(\tau)/E(\tau)^3)$.

Furthermore, since z(t) has a stationary limiting distribution

$$\text{plim} \, {}_0\!\int^T (z_t)^2 \, dt/T = E(z^2) = (1+\rho^2)\theta^2/6.$$

Therefore,

$$\text{plim} \, -\hat{\nu} = - \, 6\sigma^2/[\theta^2(1+\rho)(1+\rho^2)].$$

<u>Proof of Proposition B2.</u>

For the threshold model given by equation (B4),

$$E[z_t dz_t \mid z_t = z] = \begin{cases} -\nu(z^2 + \theta z) \, dt & z < -\theta \\ 0 & -\theta \le z \le \theta. \\ -\nu(z^2 - \theta z) \, dt & \theta < z \end{cases}$$

Recall, that the limiting stationary distribution, $\phi(z)$, for this threshold is given by equation (12). Thus,

$$\text{plim} \, {}_0\!\int^T z_t dz_t/T = {}_{-\infty}\!\int^{\infty} E[z_t dz_t | z_t = z]\phi(z) \, dz$$

$$= {}_{-\infty}\!\int^{-\theta} -\nu(z^2+\theta z)\phi(z) \, dz + {}_{\theta}\!\int^{\infty} -\nu(z^2-\theta z)\phi(z) \, dz$$

$$= -\nu\sigma^2/(2\nu).$$

$$\text{plim} \, {}_0\!\int^T (z_t)^2 dt/T = E[z^2] = {}_{\infty}\!\int^{\infty} z^2 \, \phi(z) \, dz$$

$$= \sigma^2/(2\nu) + \theta^2/3 + 2\theta[\sigma^2/(2\nu) + \theta(\pi\sigma^2/\nu)/3]/IC,$$

where IC = $[\pi\sigma^2/\nu]^{1/2} + 2\theta$.

Therefore, after some rearranging, it can be shown that

$$\text{plim } -\hat{\nu} = \text{plim } {}_0\!\int^T z_t dz_t/T] \;/\; \text{plim } {}_0\!\int^T (z_t)^2 dt/T$$

$$= -\nu\sigma^2/[\sigma^2 + 2\nu\theta^2/3 + 2\theta(\sigma^2 + 2\nu(\pi\sigma^2/\nu)^{.5})/IC] < 0,$$

where IC = $[\pi\sigma^2/\nu]^{1/2} + 2\theta$.

Appendix C. Proof of Proposition.

(i) We prove the proposition for the case of d = 1 and the target zone threshold model. Furthermore, ϵ_{1t} is assumed to be iid with mean zero and variance σ_1^2. Recall that the modified threshold model is given by:

$$z_t = z_{t-1} + \epsilon_{1t} \qquad \text{if } |z_{t-1}| \le \theta$$

$$= z_{t-1} + (\rho-1)(z_{t-1} - \theta^{(i)}) + \epsilon_{1t} \qquad \text{if } |z_{t-1}| > \theta$$

where $\theta^{(i)} = \theta$ if $z_{t-1} > \theta$ and $\theta^{(i)} = -\theta$ if $z_{t-1} < -\theta$. The basic threshold model is just a special case where $\theta^{(i)} = 0$.

The least squares estimator of the autoregressive coefficient, ρ, from the Dickey-Fuller regression is given by

$$\hat{\rho}_{DFz} = \{ {}_{t=2}\Sigma^T z_t z_{t-1} \}/\{ {}_{t=2}\Sigma^T (z_{t-1})^2 \}.$$

Using the above threshold model, we can write

$${}_{t=2}\Sigma^T z_t z_{t-1} = {}_{t=2}\Sigma^T (z_{t-1})^2 + (\rho-1)\,{}_{j=1}\Sigma^{N(T)} (z_{\tau(i)-1} - \theta^{(i)}) z_{\tau(i)-1}$$

$$+ \quad {}_{t=2}\Sigma^T \epsilon_{1t} z_{t-1},$$

where N(T) is the number of times that the series is outside of the threshold range and $\tau(i)$ is the time period of the i-th exit from inside the threshold range. Since for starting values anywhere inside the finite interval $[-\theta,\theta]$ the expected hitting time to the boundaries is finite, as $T \to \infty$, $N(T) \overset{a.s}{\to} \infty$. N(T)/T is the proportion of the sample in which the process is outside of the threshold range. For finite boundaries, as $T \to \infty$ this proportion will be strictly positive and less than one. This can be seen from the continuous time results derived above. For the continuous time modified threshold model $\text{plim } N(T)/T = P(|z| > \theta) = 1 - 2\theta/[(\pi\sigma^2/\nu)^{1/2} + 2\theta]$ while for the basic threshold model or "jump" process $\text{plim } N(T)/T = \sigma^2/[(1-\rho^2)\theta^2]$.

Because z_t is assumed to satisfy the "α-mixing" conditions and ϵ_{1t} is independent of z_{t-1} then

$$\text{plim } {}_{t=2}\Sigma^T \epsilon_{1t}z_{t-1}/T = 0 \text{ and plim } {}_{t=2}\Sigma^T (z_{t-1})^2/T = \sigma_z^2 < \infty.$$

Furthermore, since

$${}_{i=1}\Sigma^{N(T)} (z_{\tau(i)-1} - \theta^{(j)})z_{\tau(i)-1} \leq {}_{t=2}\Sigma^T (z_{t-1})^2$$

and $\text{plim } N(T)/T < 1$, we have

$$\text{plim } [{}_{i=1}\Sigma^{N(T)} (z_{\tau(i)-1} - \theta^{(i)})z_{\tau(i)-1}/T] \;/\; \text{plim } [{}_{t=2}\Sigma^T (z_{t-1})^2/T] < 1.$$

Finally, since $|z_{\tau(i)-1}| > |\theta^{(i)}|$ and $0 < \text{plim } N(T)/T < 1$,

$$\text{plim } [\; ({}_{j=1}\Sigma^{N(T)} (z_{\tau(i)-1} - \theta^{(i)})z_{\tau(i)-1}/T \;]$$

$$= \text{plim } [\; \{{}_{i=1}\Sigma^{N(T)} (z_{\tau(i)-1} - \theta^{(i)})z_{\tau(i)-1}/N(T)\}\; N(T)/T \;] > 0.$$

Hence,

$$\text{plim } \hat{\rho}_{DFz} = 1 + (\rho-1)\text{ plim } [{}_{i=1}\Sigma^{N(T)} (z_{\tau(i)-1} - \theta^{(i)})z_{\tau(i)-1}/T] \;/\; \text{plim } {}_{t=2}\Sigma^T (z_{t-1})^2/T.$$

which in turn implies that

$$\rho < \text{plim } \hat{\rho}_{DFz} < 1.$$

For parts (ii) and (iii) of the Proposition, given that z_t is stationary and is assumed to satisfy the mixing conditions, then the results of Phillips (1987) and Stock (1987) will hold for the threshold case.

Appendix D. Threshold models and α-mixing.

The α-mixing condition is needed so that Law of Large Numbers and Central Limit Theorems hold for z_t which is serially correlated. Essentially, α-mixing implies that the serial dependence dies out as the time interval between observations increases. In this section, we argue heuristically that the LLN and CLT are likely to apply to the threshold models considered in this paper.

Consider the definition of α-mixing (Bierens (1992)). Define F as the Borel Field generated by $\epsilon_{1t}, \epsilon_{1t-1}, \epsilon_{1t-2}, \ldots$, and F_{t+m} as the Borel Field generated by $\epsilon_{1t+m}, \epsilon_{1t+m+1}, \epsilon_{1t+m+2}, \ldots$. Define for $m \geq 0$

$$\alpha(m) = \sup_t \sup_{A \in F^t_{-\infty},\, B \in F^{\infty}_{t+m}} |P(A \cap B) - P(A)P(B)|$$

If $\lim_{m \to \infty} \alpha(m) = 0$, then z_t is called a strong (or α-) mixing process. Thus, $\alpha(m)$ measures the dependence between events separated by m time periods.

Recall that from equation (10) in the text, the effect on z_t of starting at z_0 and of hits on the boundary and ϵ_{1t} innovations in the distant past diminishes the more times the process exceeds the boundaries. As $t \to \infty$, the number of time the process exceeds the boundaries almost surely approaches infinity. This suggests that as $t \to \infty$, z_t and z_0 become independent and, hence, z_t is likely to be α-mixing.

To see this more clearly, consider the case where the process jumps back towards zero (i.e. the basic threshold model). For this model, $\rho = 0$ and $\theta(i) = 0$ for all i; hence, from equation (10)

(C1) $z_t = \sum_{j=\kappa(NE(t))+1}^{t} \epsilon_{1j}$.

Thus, once the boundary is hit, z_t does not depend on z_0. In fact, z_t is like a finite moving average, but one in which the order of the moving average is determined randomly by the probability of hitting the boundary.

For an informal demonstration that α-mixing is likely to hold for the threshold process, consider the case where the process hits the boundary and jumps back towards zero. Define the random variable $a = I[(z_t, z_{t-1}, z_{t-2}, \ldots) \in A]$ and $b_m = I[(z_{t+m}, z_{t+m+1}, z_{t+m+2}, \ldots) \in B]$, where $I[\,]$ is an indicator function. Define the random variable G_m to be 1 if the process hits the boundary between t and t+m and zero otherwise. Consider the joint probability distribution of a and b_m, $P(a,b_m)$. For the case of the basic threshold model with $\rho = 0$, once the boundary is hit the random variables a and b_m become independent; thus, $P[a,b_m|G_m=1] = P[a|G_m=1]P[b_m|G_m=1]$.

Now,

$$
\begin{aligned}
P[a,b_m] &= P[a,b_m|G_m=1]\ P[G_m=1] + P[a,b_m|G_m=0]\ P[G_m=0] \\
&= P[a|G_m=1]\ P[b_m|G_m=1]\ P[G_m=1] + P[a,b_m|G_m=0]\ P[G_m=0] \\
&= (P[a] - P[a|G_m=0]\ P[G_m=0])\ (P[b_m] - P[b_m|G_m=0]\ P[G_m=0])\ /\ P[G_m=1] \\
&\qquad + P[a,b_m|G_m=0]\ P[G_m=0] \\
&= P[a]P[b_m] + (P[a]P[b_m] - P[a|G_m=0]P[b_m] - P[a]P[b_m|G_m=0])P[G_m=0]/P[G_m=1] \\
&\qquad + P[a|G_m=0]\ P[b_m|G_m=0]\ P[G_m=0]^2/P[G_m=1] + P[a,b_m|G_m=0]\ P[G_m=0] \\
&= P[a]P[b_m] + O(\ -P[G_m=0]/P[G_m=1] + P[G_m=0]^2/P[G_m=1] + P[G_m=0]\ ) \\
&= P[a]P[b_m] + O(P[G_m=0])
\end{aligned}
$$

Thus, $\alpha(m) = O(P[G_m=0])$. Since $P[G_m=0] \to 0$ as $m \to \infty$, z_t is α-mixing. The rate of decay in $\alpha(m)$ depends on how fast $P[G_m=0]$ goes to zero. For the general model, the degree to which $\alpha(m)$ decays will depend on the relative values of θ and σ_1 which, in part, determine the frequency with which the boundary is hit, on whether the process jumps back towards an equilibrium value or whether the process drifts back to the $[-\theta,\theta]$, and on the value of ρ.

As an additional check that the z_t process satisfied conditions needed for the Law of Large Numbers (LLN) and Central Limit Theorem (CLT), we verified whether the LLN and the CLT held for the threshold models used in Table 1. We generated one thousand replications of the threshold series and calculated for each replication the following statistics:

$_{t=1}\Sigma^T z_t/T$, $_{t=1}\Sigma^T z_t^2/T$, and $_{t=1}\Sigma^T z_t/(T^{1/2})$.

If LLN holds for z_t, the first two statistics tend towards a constant as $T \rightarrow \infty$, while if the CLT holds for z_t the distribution of the last statistic converges to a normal distribution as $T \rightarrow \infty$. We consider sample sizes of T = 100, 250, 500, 1000, and 5000. For all the combinations considered in Table 1, the threshold model appears to satisfy both the LLN and the CLT; however, for cases where the threshold boundaries are large, sample sizes must be quite large before the CLT is approximated.

REFERENCES

Bentolila, S. and G. Bertola (1990), "Firing Costs and Labor Demand: How Bad is Eurosclerosis?" Review of Economic Studies, 57, 381-402.

Bertola, G. and R. J. Caballero (1990), "Kinked Adjustment Costs and Aggregate Dynamics," Brookings Papers on Economic Activity, 237-295.

Bierens (1992), Topics in Advanced Econometrics, manuscript in preparation.

Bierens (199*),.

Billingsley, P. (1968) Convergence of Probability Measures, (New York: John Wiley and Sons).

Campbell, J.Y. (1987), "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," Econometrica, 55, 1249-1274.

Campbell, J.Y. and R. J. Shiller (1987) "Cointegration and Tests of Present Value Models," Journal of Political Economy, 95, 1062-1088.

Corbae D., and S. Ouliaris (1988) "Cointegration and Tests of Purchasing Power Parity," Review of Economics and Statistics, 70, 508-511.

Dixit, A. (1991) "A Simplified Treatment of the Theory of Optimal Regulation of Brownian Motion," Journal of Economic Dynamics and Control, 15, 657-673.

Engle, R.F. and C.W.J. Granger (1987), "Co-integration and Error Correction: Representation, Estimation, and Testing," Econometrica, 55, 251-276.

Granger, C.W.J. (1986), "Developments in the Study of Cointegrated Economic Variables," Oxford Bulletin of Economics and Statistics, 46, 213-228.

Granger, C.W.J. and J. Hallman (1991), "Long Memory Series with Attractors," Oxford Bulletin of Economics and Statistics, 53, 11-26.

Harrison, J. Michael (1985), Brownian Motion and Stochastic Flow Systems (New York: Wiley).

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models," Econometrica, 59, 1551-1580.

Johansen, S. and K. Juselius (1990), "Maximum Likelihood Estimation and Inference on Cointegration--with Applications to the Demand for Money," Oxford Bulletin of Economics and Statistics, 52, 169-210.

Karlin, S. and H. Taylor (1975), A First Course in Stochastic Processes, (New York: Academic Press).

Karlin S. and H. Taylor (1981), A Second Course in Stochastic Processes, (New York: Academic Press).

Miller, M. and D. Orr (1966), "A Model of the Demand for Money by Firms," Quarterly Journal of Economics, 100, 529-539.

Phillips, P.C.B. (1987) "Time Series Regression with a Unit Root," Econometrica, 55, 277-302.

Phillips, P.C.B. (1991) "Error Correction and Long-Run Equilibrium in Continuous Time," Econometrica, 59, 967-980.

Ross, S.M. (1983), Stochastic Processes, (New York: Wiley & Sons).

Scarf, H. (1959) "The Optimality of (s,S) Policies in the Dynamic Inventory Method," in K. Arrow, S. Karlin, and P. Suppes (eds.), Mathematical Methods in the Social Sciences, Stanford University Press, Ch. 13.

Sheshinski, E. and Y. Weiss (1983) "Optimum Pricing Policy under Stochastic Inflation," Review of Economic Studies, 43, 513-529.

Stock, J. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," Econometrica, 55, 1035-1056.

Terasvirta T. (1990), "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," mimeo, University of California, San Diego, October.

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

Tsay, R. S. (1989), "Testing and Modeling Threshold Autoregressive Processes," Journal of the American Statistical Association, 82, 590-604.

Table 1
Power of Dickey-Fuller t-statistic for Various Threshold Models

$\rho = 0.0$

Confidence level	Basic Threshold θ = 5	10	20	Target Zone Thresh. θ = 5	10	20	Barrier Process θ = 5	10	20
T=100									
5%	71.0	9.0	5.0	18.0	8.0	5.0	15.0	8.0	5.0
10%	88.0	20.0	11.0	27.0	15.0	11.0	25.0	15.0	11.0
T=250									
5%	100.0	39.0	7.0	37.0	12.0	6.0	42.0	12.0	6.0
10%	100.0	60.0	14.0	64.0	24.0	12.0	68.0	22.0	12.0
T=500									
5%	100.0	97.0	15.0	99.0	19.0	10.0	100.0	19.0	10.0
10%	100.0	99.0	28.0	100.0	30.0	17.0	100.0	28.0	17.0
T=1000									
5%	100.0	100.0	48.0	100.0	47.0	15.0	100.0	50.0	14.0
10%	100.0	100.0	68.0	100.0	79.0	24.0	100.0	80.0	23.0

$\rho = 0.4$

Confidence level	Basic Threshold θ = 5	10	20	Target Zone Thresh. θ = 5	10	20	Barrier Process θ = 5	10	20
T=100									
5%	33.0	11.0	5.0	15.0	7.0	5.0	15.0	8.0	5.0
10%	51.0	20.0	11.0	24.0	14.0	11.0	25.0	15.0	11.0
T=250									
5%	99.0	22.0	8.0	26.0	12.0	6.0	42.0	12.0	6.0
10%	100.0	35.0	16.0	47.0	22.0	12.0	68.0	22.0	12.0
T=500									
5%	100.0	65.0	17.0	97.0	19.0	10.0	100.0	19.0	10.0
10%	100.0	91.0	25.0	100.0	28.0	17.0	100.0	28.0	17.0
T=1000									
5%	100.0	100.0	25.0	100.0	36.0	14.0	100.0	50.0	14.0
10%	100.0	100.0	38.0	100.0	67.0	24.0	100.0	80.0	23.0

$\rho = 0.8$

Confidence Level	Basic Threshold θ = 5	10	20	Target Zone Thresh. θ = 5	10	20	Barrier Process θ = 5	10	20
T=100									
5%	17.0	8.0	5.0	11.0	6.0	5.0	15.0	8.0	5.0
10%	28.0	16.0	11.0	19.0	13.0	11.0	25.0	15.0	11.0
T=250									
5%	50.0	16.0	6.0	17.0	10.0	6.0	42.0	12.0	6.0
10%	79.0	27.0	13.0	30.0	19.0	11.0	68.0	22.0	12.0
T=500									
5%	100.0	23.0	12.0	53.0	17.0	9.0	100.0	19.0	10.0
10%	100.0	36.0	20.0	82.0	25.0	16.0	100.0	28.0	17.0
T=1000									
5%	100.0	83.0	18.0	100.0	24.0	13.0	100.0	50.0	14.0
10%	100.0	99.0	29.0	100.0	45.0	22.0	100.0	80.0	23.0

Notes: 1000 replications.

Table 2
Threshold Autoregressions for the Spread Between the Fed Funds Rate and the Discount Rate

Threshold autoregression for the spread ($spr_t = ff_t - dr_t$)

$$
spr_t = \begin{cases}
\underset{(0.087)}{-0.193} + \underset{(0.101)}{0.696}\, spr_{t-1} - \underset{(0.056)}{0.102}\, spr_{t-2} + \underset{(0.38)}{e_t} & \text{if } spr_{t-1} < -0.2 \\
\underset{(0.020)}{0.013} + \underset{(0.068)}{1.411}\, spr_{t-1} - \underset{(0.059)}{0.406}\, spr_{t-2} + \underset{(0.26)}{e_t} & \text{if } -0.2 \le spr_{t-1} \le 1.6 \\
\underset{(0.346)}{1.358} + \underset{(0.151)}{0.985}\, spr_{t-1} - \underset{(0.140)}{0.547}\, spr_{t-2} + \underset{(0.99)}{e_t} & \text{if } spr_{t-1} > 1.6
\end{cases}
$$

Dickey-Fuller T-Stats in the Different Threshold Regimes

Threshold Regime	T-Stat	Sample Size
$spr_{t-1} < -0.2$	-3.702	T = 78
$-0.2 \le spr_{t-1} \le 1.6$	0.166	T = 278
$spr_{t-1} > 1.6$	-4.720	T = 74

Threshold Error Correction Models:

$$
\Delta ff_t = \begin{cases}
\underset{(0.09)}{-0.21} + \underset{(0.06)}{0.30}\, \Delta ff_{t-1} - \underset{(0.19)}{0.02}\, \Delta dr_{t-1} - \underset{(0.12)}{0.37}\, spr_{t-1} + \underset{(0.41)}{e_t} & \text{if } spr_{t-1} < -0.2 \\
\underset{(0.03)}{0.02} + \underset{(0.08)}{0.73}\, \Delta ff_{t-1} - \underset{(0.13)}{0.07}\, \Delta dr_{t-1} + \underset{(0.04)}{0.01}\, spr_{t-1} + \underset{(0.33)}{e_t} & \text{if } -0.2 \le spr_{t-1} \le 1.6 \\
\underset{(0.37)}{1.54} + \underset{(0.17)}{0.50}\, \Delta ff_{t-1} + \underset{(0.62)}{0.71}\, \Delta dr_{t-1} - \underset{(0.13)}{0.64}\, spr_{t-1} + \underset{(1.06)}{e_t} & \text{if } spr_{t-1} > 1.6
\end{cases}
$$

$$
\Delta dr_t = \begin{cases}
\underset{(0.03)}{-0.01} + \underset{(0.02)}{0.20}\, \Delta ff_{t-1} + \underset{(0.06)}{0.22}\, \Delta dr_{t-1} + \underset{(0.04)}{0.01}\, spr_{t-1} + \underset{(0.13)}{e_t} & \text{if } spr_{t-1} < -0.2 \\
\underset{(0.01)}{0.01} + \underset{(0.04)}{0.33}\, \Delta ff_{t-1} + \underset{(0.07)}{0.20}\, \Delta dr_{t-1} + \underset{(0.02)}{0.01}\, spr_{t-1} + \underset{(0.17)}{e_t} & \text{if } -0.2 \le spr_{t-1} \le 1.6 \\
\underset{(0.08)}{0.19} + \underset{(0.03)}{0.09}\, \Delta ff_{t-1} + \underset{(0.13)}{0.13}\, \Delta dr_{t-1} - \underset{(0.03)}{0.05}\, spr_{t-1} + \underset{(0.22)}{e_t} & \text{if } spr_{t-1} > 1.6
\end{cases}
$$

Notes: Standard errors are in parentheses.

Figure 1

RECURSIVE DICKEY-FULLER T-STATISTICS FROM ARRANGED AUTOREGRESSION VERSUS POSSIBLE THRESHOLDS

AUTOREGRESSION ARRANGED FROM LOW TO HIGH VALUES

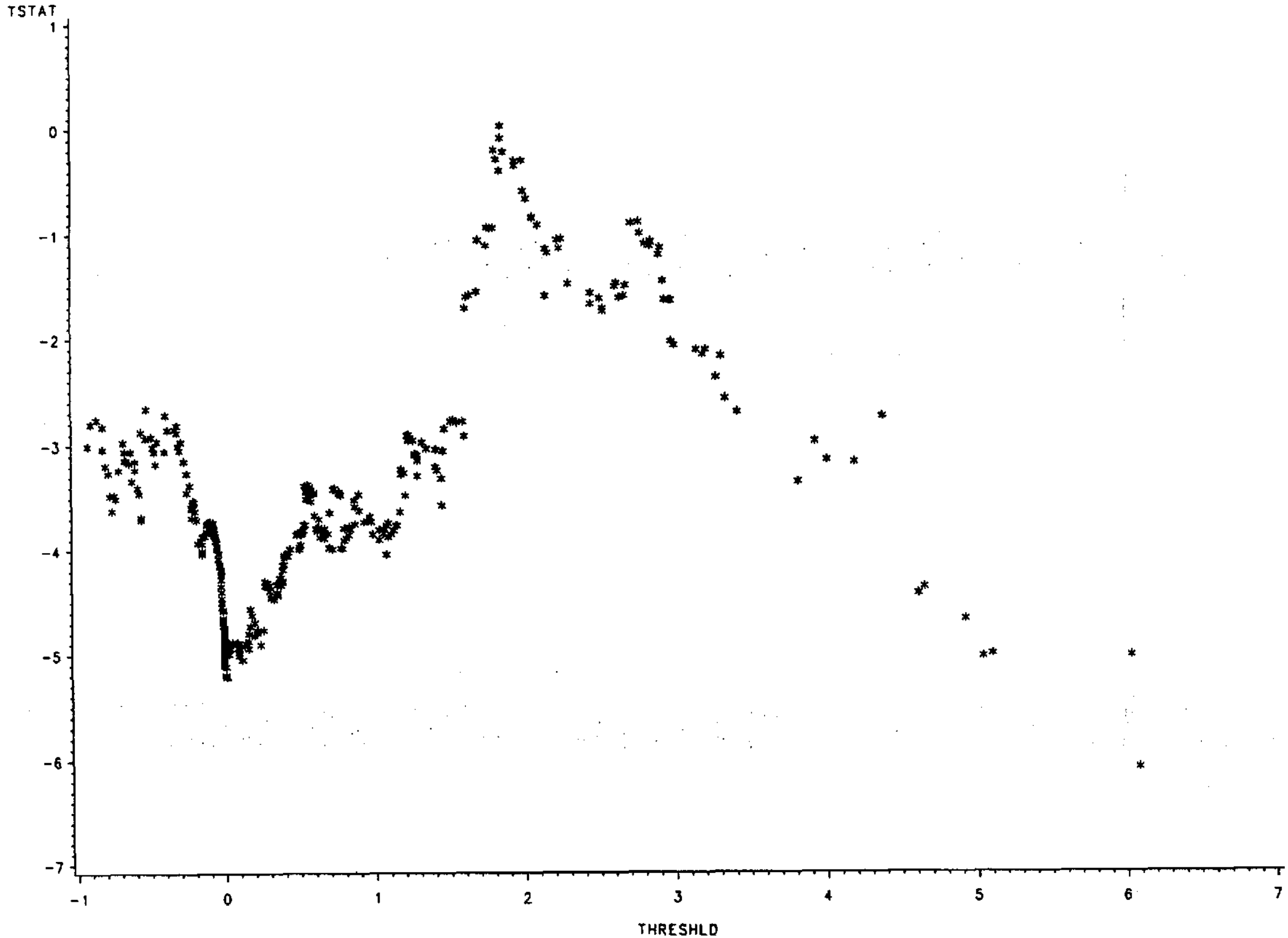


Figure 2

RECURSIVE DICKEY–FULLER T–STATISTICS FROM ARRANGED AUTOREGRESSION VERSUS POSSIBLE THRESHOLDS

AUTOREGRESSION ARRANGED FROM HIGH TO LOW VALUES

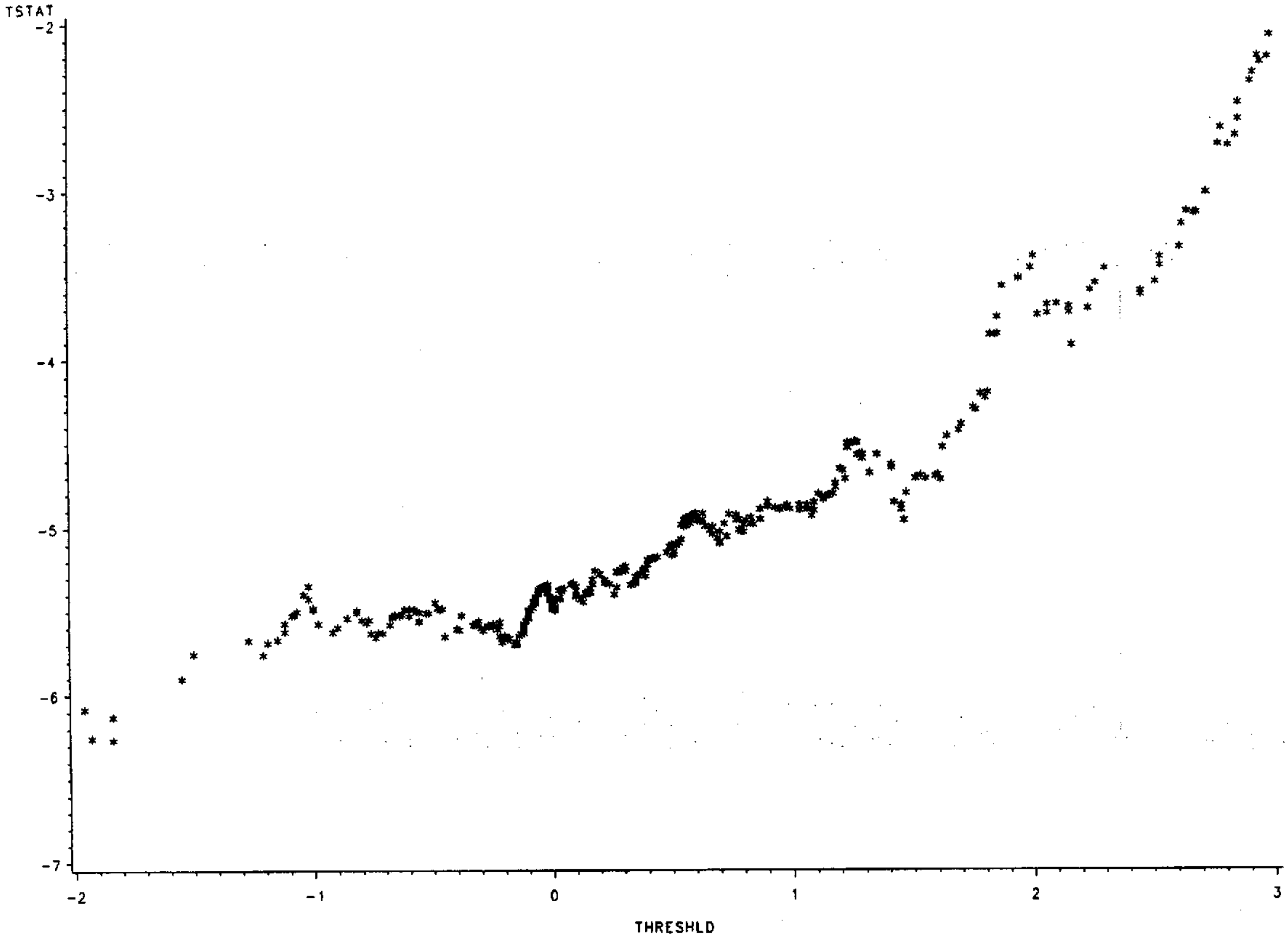
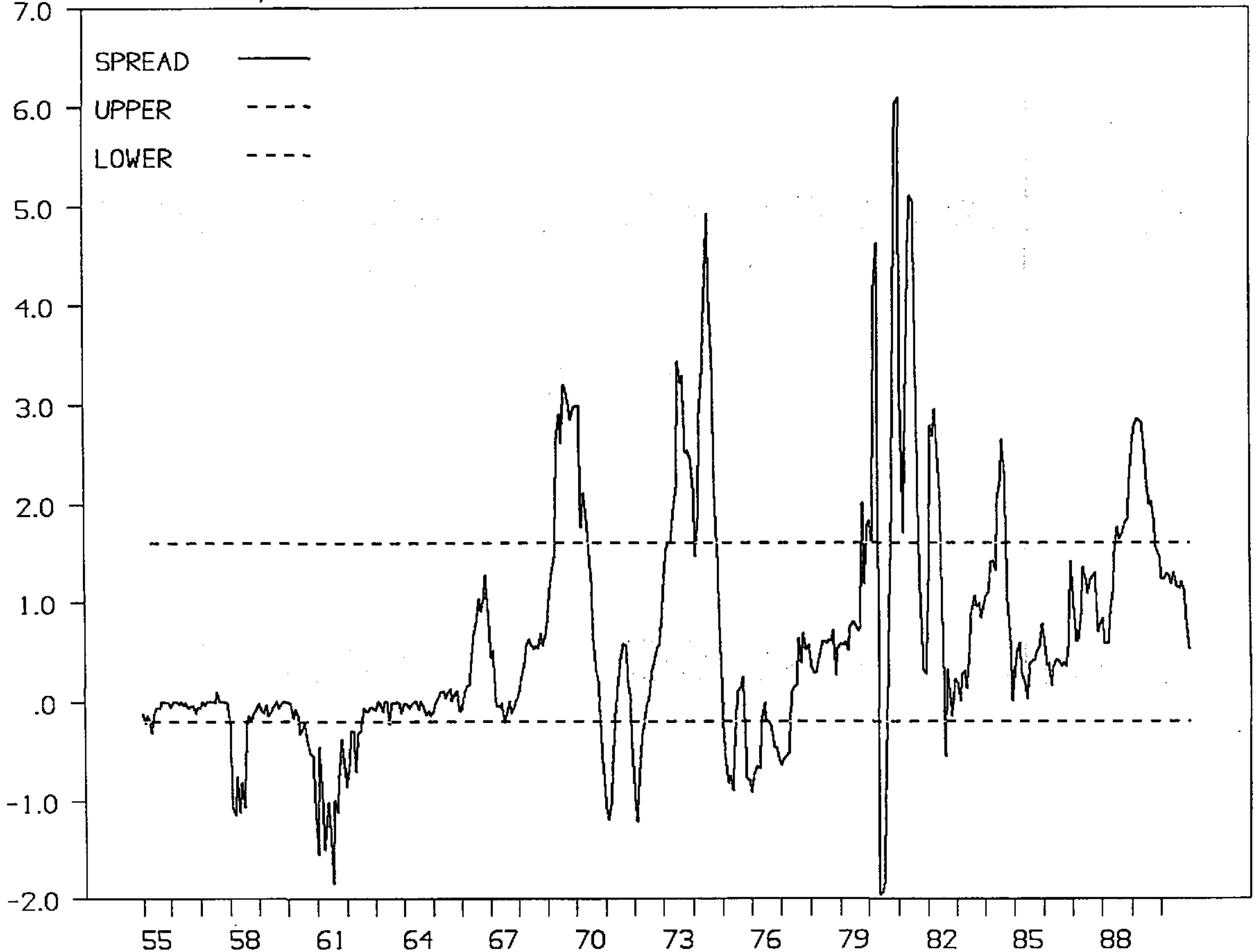


Figure 3

Spread Between the Federal Funds Rate and the Discount Rate



RESEARCH PAPERS OF THE RESEARCH DEPARTMENT
FEDERAL RESERVE BANK OF DALLAS

Available, at no charge, from the Research Department
Federal Reserve Bank of Dallas, Station K
Dallas, Texas 75222

9101 Large Shocks, Small Shocks, and Economic Fluctuations: Outliers in Macroeconomic Time Series (Nathan S. Balke and Thomas B. Fomby)

9102 Immigrant Links to the Home Country: Empirical Implications for U.S. and Canadian Bilateral Trade Flows (David M. Gould)

9103 Government Purchases and Real Wages (Mark Wynne)

9104 Evaluating Monetary Base Targeting Rules (R.W. Hafer, Joseph H. Haslag and Scott E. Hein)

9105 Variations in Texas School Quality (Lori L. Taylor and Beverly J. Fox)

9106 What Motivates Oil Producers?: Testing Alternative Hypotheses (Carol Dahl and Mine Yucel)

9107 Hyperinflation, and Internal Debt Repudiation in Argentina and Brazil: From Expectations Management to the "Bonex" and "Collor" Plans (John H. Welch)

9108 Learning From One Another: The U.S. and European Banking Experience (Robert T. Clair and Gerald P. O'Driscoll)

9109 Detecting Level Shifts in Time Series: Misspecification and a Proposed Solution (Nathan S. Balke)

9110 Underdevelopment and the Enforcement of Laws and Contracts (Scott Freeman)

9111 An Econometric Analysis of Borrowing Constraints and Household Debt (John V. Duca and Stuart S. Rosenthal)

9112 Credit Cards and Money Demand: A Cross-Sectional Study (John V. Duca and William C. Whitesell)

9113 Rational Inflation and Real Internal Debt Bubbles in Argentina and Brazil? (John H. Welch)

9114 The Optimality of Nominal Contracts (Scott Freeman and Guido Tabellini)

9115 North American Free Trade and the Peso: The Case for a North American Currency Area (Darryl McLeod and John H. Welch)

9116 Public Debts and Deficits in Mexico: A Comment (John H. Welch)

9117 The Algebra of Price Stability (Nathan S. Balke and Kenneth M. Emery)

9118 Allocative Inefficiency in Education (Shawna Grosskopf, Kathy Hayes, Lori Taylor, William Weber)

9119 Student Emigration and the Willingness to Pay for Public Schools: A Test of the Publicness of Public High Schools in the U.S. (Lori L. Taylor)

9201 Are Deep Recessions Followed by Strong Recoveries? (Mark A. Wynne and Nathan S. Balke)

9202 The Case of the "Missing M2" (John V. Duca)

9203 Immigrant Links to the Home Country: Implications for Trade, Welfare and Factor Rewards (David M. Gould)

9204 Does Aggregate Output Have a Unit Root? (Mark A. Wynne)

9205 Inflation and Its Variability: A Note (Kenneth M. Emery)

9206 Budget Constrained Frontier Measures of Fiscal Equality and Efficiency in Schooling (Shawna Grosskopf, Kathy Hayes, Lori Taylor, William Weber)

9207 The Effects of Credit Availability, Nonbank Competition, and Tax Reform on Bank Consumer Lending (John V. Duca and Bonnie Garrett)

9208 On the Future Erosion of the North American Free Trade Agreement (William C. Gruben)

9209 Threshold Cointegration (Nathan S. Balke and Thomas B. Fomby)