Nonlinear Mean Reversion in the Term Structure of Interest Rates

Byeongseon Seo *

Department of Economics, Soongsil University †

May 2000

Abstract

The expectations hypothesis implies that the yield curve provides information on the future change in the short-term interest rate. However, transaction costs exist in the financial market, which prevents investors from realizing the arbitrage opportunity, if the arbitrage does not fully cover the transaction costs. This paper wants to assess the effect of transaction costs on the predictability of the term structure by using the threshold vector error correction model, which allows for the nonlinear adjustment to the long-run equilibrium relationship. A significant amount of threshold effect is found, and the adjustment coefficients are regime-dependent. The empirical result supports the nonlinear mean reversion in the term structure of interest rates.

Key words: Expectations hypothesis; Transaction costs; Threshold cointegration

JEL classification: C12; C22

*I would like to thank Nathan Balke for useful comments and suggestions. All errors are my responsibility.

†Sangdo-5 Dong, Seoul, 156-743, Korea. E-mail: seo@saint.soongsil.ac.kr. Phone: 82-2-820-0552. Fax: 82-2-824-4384.

1 Introduction

According to the expectations hypothesis, the long-term interest rate is an average of the current and future short-term rates, and hence it implies that the term structure or the yield curve provides information on the future change in the short-term interest rate. An upward-sloping yield curve causes the rise of the short rate in the future, and a downward-sloping curve causes the fall of the short rate. Thus, the expectations hypothesis implies that the term spread is mean-reverting. Because the expected returns on roll-over and maturity investment strategies are equalized, the arbitrage opportunity should be captured and realized by investors immediately if it exists. However, the transaction costs exist in the financial market, and the arbitrage opportunity cannot be realized and persist as long as the arbitrage does not fully cover the transaction costs and produce net gain.

This paper wants to assess the effect of the transaction costs on the adjustment process of the term structure of interest rates by using the threshold cointegration model, which has been proposed by Balke and Fomby (1997). As noted by Mankiw and Summers (1984), the term structure plays a role in the transmission mechanism between the real sector and the money market. Thus, this paper is useful and necessary because we investigate the econometric model which explains the empirical regularity of the term structure.

The expectations hypothesis provides the analytical framework which simplifies the rational behavior in the financial market. Particularly, the term structure provides information on the future changes in the short-term interest rates. However, the predictability of the term structure often fails to be consistent with the empirical findings. For example, Shiller et al. (1983) has shown that the term structure or the yield curve does not provide information on the future change in the short rate. Mankiw and Miron (1986) shows the empirical results that the predictability disappears after the foundation of the Federal Reserve. Rudebusch (1995) and Balduzzi et al. (1997) also find that the changes in the interest rate are due to the unexpected changes in the Fed targeting. However, these empirical results are based on the linear regression model, which does not consider the transaction costs, and thus the predictability of the term structure cannot be correctly identified.

The main objective of this paper is to find an econometric model that is consistent with the theory and evidence. Although Mankiw and Summers (1984) has considered time-varying liquidity or risk premium, the nonlinear mean reversion in the term structure has not been discussed. The predictability of the term structure cannot be assessed appropriately in the linear regression model. This paper assesses the predictability of the term structure by using the threshold vector error correction model, which allows for the nonlinear adjustment to the long-run equilibrium relationship.

As Friedman (1977, 1979) has shown, the transaction costs such as bid-ask spreads and brokerage fees prevent investors from reallocating asset portfolios immediately in the short run. The transaction costs reduce substitutability in the financial market, which in the short run support Culbertson's (1957) segmented markets theory or 'preferred habitat' named by Modigliani and Sutch (1966, 1967). The arbitrage opportunity cannot be utilized if it does not exceed the transaction costs, making the term spread persist in the short-run. However, investors will take advantage of the arbitrage opportunity in the long-run or if the opportunity exceeds the transaction costs. Thus, the optimal adjustment model provides an asymmetric adjustment of the term structure, and it is necessary to assess the nonlinear adjustment of the term spread by using an appropriate econometric model.

This paper analyzes the term structure of the U.S. bond market with an econometric model of threshold cointegration, which allows for nonlinear adjustment and cointegration. Threshold cointegration has been proposed by Balke and Fomby (1997), and since then it has attracted much attention in the recent literature of econometrics and its applications. Particularly, the model of threshold cointegration can be applied to economic models with transaction costs. Since the expectations hypothesis predicts that the short-term interest rate and the long-term interest rate have a long-run relationship, threshold cointegration can be useful in the analysis of the term structure.

This paper finds a significant threshold effect in the term structure of interest rates by using the method of Hansen and Seo (2000). The term premium is persistent if the market disequilibrium does not exceed the threshold values that may signify the transaction costs. On the other hand, the term spread is mean-reverting if the equilibrium error exceeds the threshold values. Therefore, our results supports the nonlinear mean reversion in the term structure of interest rates.

Section 2 begins with the expectations hypothesis and its implications, and we modify the expectations model to allow for nonlinear mean reversion. Section 3 deals with econometric methods to assess the nonlinear mean reversion in the term structure. Main empirical results are provided in Section 4. Estimation and testing results of the threshold vector error correction model with an unknown cointegrating vector are discussed in Section 5.

2 Expectations Hypothesis

There are two important theories of the term structure: the expectations hypothesis and the segmented markets theory, and these two theories provide contrasting predictions on the mean-reverting behavior of the term structure. According to the expectations hypothesis, the interest rate on a long-term bond is an average of short-term interest rates over the life of the long-term bond. If the short-term interest rate is expected to rise, then the long-term interest rate tends to be higher than the current short-term rate. Therefore, the term spread shows a behavior of mean reversion, and the expectations hypothesis implies the predictability of the term structure that the term structure or the yield curve can be used to predict the future changes in the short-term interest rate.

On the other hand, the segmented markets theory is based on the proposition that markets for different-maturity bonds are completely separate and segmented. Investors have strong preference for bonds of one maturity over another, and hence bonds of different maturities are not substitutes. Because the market for bonds of different maturities are completely segmented, there is no reason for a rise in interest rates on a bond of one maturity to affect the interest rate on a bond of another maturity. Therefore, the segmented markets theory predicts that the term spread is persistent.

The predictability of the term structure has been denied by many empirical results such as Shiller et al. (1983) and Mankiw and Miron (1986). However, these results cannot reject the expectations hypothesis in favor of the segmented markets theory. As Campbell and

Shiller (1987) has shown, the long-term and the short-term interest rates are cointegrated, which implies a long-run equilibrium relationship. The data reveals that the term spread is persistent in the short-run although it is eventually mean-reverting. Thus, this paper postulates a hypothesis of nonlinear mean reversion and provides a formal assessment by using threshold cointegration which allows for a nonlinear adjustment to the long-run relationship.

First, we consider the expectations hypothesis and its implications, and then we develop a term structure model which allows for nonlinear mean reversion. Suppose r_t is the one-period interest rate and R_t is the interest rate on the bond with a maturity of m. According to the expectations hypothesis, the long-term interest rate R_t is the average of current and future expected returns on the bond with one-period maturity with a constant liquidity premium as follows:

$$R_t = \frac{1}{m} \sum_{i=1}^{m} E_t(r_{t+i-1}) + \kappa, \tag{1}$$

where κ is the liquidity premium.

We also obtain the following equation:

$$s_t = R_t - r_t = \frac{1}{m} \sum_{i=1}^{m-1} \sum_{j=1}^{i} E_t \Delta r_{t+j} + \kappa. \tag{2}$$

If we assume that the future change in the short rate is stationary, the term spread s_t is stationary and thus mean-reverting. According to the definition of Engle and Granger (1987), the long rate forms a long-run relationship with the short rate.

Although the expectations hypothesis predicts a long-run relationship, the relationship does not tell about the nonlinear mean reversion. Also, the empirical results that rejected the predictability of the term structure are based on the linear regression model, and thus the rejection of the expectations hypothesis may be overstated if we allow the nonlinear mean reversion.

The stylized fact found in the U.S. bond market is that the term spread is persistent as long as the term spread is not so large as to deviate an interval which is bounded by the shadow cost of mean reversion. Figure 1 depicts the response of the current term spread and the short and long rates to the past term spread for the period 1960:1-1999:12. The short rate corresponds to the yield on the 3-month Treasury bill (TB), and the long rate is the yield on the 10-year Treasury note (TN). As in Figure 1, the current change in the term spread responds to the past term spread and its response is different from the linear relationship. If the term spread is small, the response is not significant and the response function is flat. The term spread shows a significant mean reversion only if the term spread departs from the interval of flat response. Also, the response of the short rate to the term spread cannot be specified as a linear relationship, and as a result the predictability of the term spread cannot be correctly identified in the linear regression model.

In Figures 2-3, the yield on the 5-year Treasury note and the 1-year Treasury bill is used as the long rate, respectively. The short rate is the 3-month TB rate. In the same way as Figure 1, the response of the short rate switches depending on the magnitude of the term spread.

Friedman (1977, 1979) has shown that financial flow variables play a crucial role in the demand-for-bonds equations. The transaction costs of the flow adjustment are cheaper than those of the stock adjustment in the short run. As suggested in Friedman (1977), the main transaction costs in the U.S. bond markets are financial charges such as bid-ask spreads for institutional investors and brokerage fees for individual investors. The central banks regulation such as Fed targeting also affects the bond market and it provides frictional costs as suggested by Mankiw and Summers (1984), Makiw and Miron (1986), and Rudebusch (1995). The transaction costs keep investors from adjusting their portfolios fully and immediately.

To explain the persistence of the term spread, we modify the expectations hypothesis as follows:

$$R_t = \frac{1}{m}\sum_{i=1}^{m} E_t(r_{t+i-1}) + \kappa_t, \tag{3}$$

where κ_t is the liquidity premium.

Mankiw and Summers (1984) has shown that the future changes in the short rate can be explained by the time-varying risk premium. This paper also assumes that the liquidity premium is time-varying. Furthermore, we assume that the liquidity premium follows a threshold autoregressive process as follows:

$$\kappa_t = \begin{array}{l} \phi_1\kappa_{t-1} + \eta_t \text{ for } s_{t-1} \leq \gamma_1 \\ \phi_2\kappa_{t-1} + \eta_t \text{ for } \gamma_1 < s_{t-1} \leq \gamma_2 \\ \phi_3\kappa_{t-1} + \eta_t \text{ for } s_{t-1} > \gamma_2, \end{array} \tag{4}$$

where $E_{t-1}(\eta_t) = 0$.

The threshold autoregressive (TAR) model has been proposed by Tong (1978), and here, we assume the three-regime TAR model with the regime-specific autoregressive parameters. As the adjustment parameter is close to 1, the liquidity premium is persistent. Thus, our model allows for nonlinear mean reversion in the term structure.

With a simple manipulation, we have the following result.

$$s_t = \begin{array}{l} \phi_1 s_{t-1} + v_{1t} \text{ for } s_{t-1} \leq \gamma_1 \\ \phi_2 s_{t-1} + v_{2t} \text{ for } \gamma_1 < s_{t-1} \leq \gamma_2 \\ \phi_3 s_{t-1} + v_{3t} \text{ for } s_{t-1} > \gamma_2, \end{array} \tag{5}$$

where $v_{kt} = \eta_t + \frac{1}{m}\sum_{i=1}^{m-1}\sum_{j=1}^{i} E_t\Delta r_{t+j} - \frac{\phi_k}{m}\sum_{i=1}^{m-1}\sum_{j=1}^{i} E_{t-1}\Delta r_{t+j-1}$ for $k = 1, 2, 3$.

We assume that the future change in the short-term interest rate is stationary. In general, the error term v_{kt} is dependent, and thus the term spread can be specified as a TAR(l) model, where l is the lag length.

3 Econometric Methods

This section develops econometric models that can be used to estimate the nonlinear mean-reverting behavior of the term structure of interest rates. We denote $x_t = (R_t,\ r_t)'$, and then the linear vector error correction model (VECM) can be defined as follows:

$$\Delta x_t = \mu + \alpha w_{t-1} + \sum_{i=1}^{l}\Gamma_i\Delta x_{t-i} + u_t, \tag{6}$$

where $E_{t-1}(u_t) = 0$.

The long-run relationship is defined as $w_t = (1\ \ -\beta)' x_t = R_t - \beta r_t$, which is stationary as discussed by Engle and Granger (1987). To estimate the response of the short rate to the term spread, we set $\beta = 1$ in this section. We allow for unknown β in section 5. If $\beta = 1$, the long-run relationship is the same as the term spread, that is, $w_t = s_t = (1\ -1)'x_t = R_t - r_t$.

The mean-reverting behavior depends on the adjustment vector α in the linear VECM. If it is close to 0, the equilibrium error or the term spread is likely to be persistent. The linear VECM assumes that the adjustment vector is constant and the response of the short rate to the past term spread is linear. As in Figures 1-3, the linear VECM cannot explain the nonlinear mean reversion in the term structure. Hansen and Seo (2000) extended the threshold autoregressive model to the threshold vector error correction model as a means of combining the long-run relationship and the nonlinear adjustment.

Suppose the time series of interest rates x_t follows a threshold vector error correction model as follows:

$$\Delta x_t = \begin{array}{l} (\mu_1 + \alpha_1 s_{t-1} + \sum_{i=1}^{l} \Gamma_{1i} \Delta x_{t-i}) 1(s_{t-1} \leq \gamma_1) + \\ (\mu_2 + \alpha_2 s_{t-1} + \sum_{i=1}^{l} \Gamma_{2i} \Delta x_{t-i}) 1(\gamma_1 < s_{t-1} \leq \gamma_2) + \\ (\mu_3 + \alpha_3 s_{t-1} + \sum_{i=1}^{l} \Gamma_{3i} \Delta x_{t-i}) 1(s_{t-1} > \gamma_2) + u_t, \end{array} \tag{7}$$

where $E_{t-1}(u_t) = 0$, and $1(\cdot)$ is the indicator function.

The long-run relationship $R_t - r_t = s_t$ is stationary and it determines three regimes. The threshold parameters γ_1 and γ_2 signify the shadow costs of mean reversion. Regime 1 is the period when the term spread satisfies $s_t \leq \gamma_1$. Regime 2 and Regime 3 correspond to the period satisfying $\gamma_1 < s_t \leq \gamma_2$ and $s_t > \gamma_2$, respectively. Accordingly, the adjustment vectors are α_1, α_2, and α_3 and they are regime-specific. Our model is different from the conventional regime switching model as the regime is determined by the equilibrium error and the threshold parameters. As the adjustment vector is close to 0, the interest rates do not respond to the term spread and then the term spread is persistent. On the other hand, if the adjustment vector is different from 0 to stationarity, then the interest rates react to the term spread and thus the term spread is mean-reverting.

As the TAR model considers the regime-specific autoregressive parameters, the threshold VECM allows the regime-specific adjustment vectors. There exists one-to-one correspondence between these two models. If the equilibrium error or the term spread does not exceed the transaction costs, the term spread is persistent and thus the autoregressive parameter is close to zero. In the threshold VECM, the adjustment vector is close to zero for the regime. On the other hand, the equilibrium error disappears quickly if the equilibrium error exceeds the transaction costs. In that case, the autoregressive parameter is less than zero, and the adjustment vector is far from zero to stationarity.

If we define the parameter vector $\theta = (\mu^{'}, \alpha^{'}, \Gamma_1^{'}, \cdots, \Gamma_l^{'})^{'}$, then the threshold VECM can be written compactly as follows:

$$\Delta x_t = z_{t-1}^{'}\theta_1 1(s_{t-1} \leq \gamma_1) + z_{t-1}^{'}\theta_2 1(\gamma_1 < s_{t-1} \leq \gamma_2) + z_{t-1}^{'}\theta_3 1(s_{t-1} > \gamma_2) + u_t, \tag{8}$$

where $z_{t-1} = (1, s_{t-1}, \Delta x_{t-1}^{'}, \cdots, \Delta x_{t-l}^{'})^{'}$.

For computational convenience, we modify the threshold VECM as follows:

$$\Delta x_t = z_{t-1}^{'}\theta + z_{t-1}^{'}\delta_1 1(s_{t-1} \leq \gamma_1) + z_{t-1}^{'}\delta_3 1(s_{t-1} > \gamma_2) + u_t, \tag{9}$$

where $z_{t-1} = (1, s_{t-1}, \Delta x_{t-1}^{'}, \cdots, \Delta x_{t-l}^{'})^{'}$, $\theta = \theta_2 = (\mu^{'}, \alpha^{'}, \Gamma_1^{'}, \cdots, \Gamma_l^{'})^{'}$, $\delta_1 = \theta_1 - \theta$, and $\delta_3 = \theta_3 - \theta$ in equation (8).

Hence, the tests for nonlinear mean reversion can be based on the following hypotheses:

$$\mathcal{H}_0 : \delta = 0 \ \ \textit{against} \ \ \mathcal{H}_1 : \delta \neq 0,$$

where $\delta = (\delta_1^{'}, \delta_3^{'})^{'}$.

Under the null hypothesis, the threshold model is the same as the conventional error correction model, and then the term spread follows a linear mean-reverting behavior. If we fix the threshold parameters $\gamma = (\gamma_1, \gamma_2)$, the threshold VECM can be estimated by linear regression and the tests for nonlinear mean reversion can be based on the LM statistic as follows:

$$\mathrm{LM}_n(\gamma) = [\mathrm{vec}(\hat{\delta}(\gamma))]'[\mathrm{Est.Var}(\hat{\delta}(\gamma))]^{-1}[\mathrm{vec}(\hat{\delta}(\gamma))], \tag{10}$$

where $\mathrm{vec}(\cdot)$ is the column-stacking operator.

Because the conditional heteroskedasticity is general in the regression of interest rates, we use the White heteroskedasticity-consistent covariance estimator.

As discussed by Andrews and Ploberger (1994) and Hansen (1996), the threshold parameter γ cannot be identified under the null hypothesis, and as a result the standard methods cannot be applied. By using the optimality arguments of Andrews and Ploberger (1994), we use the Sup-LM statistic which does not depend on the nuisance parameter.

$$\mathrm{SupLM}_n = \mathrm{Sup}_{\gamma \in \Gamma^* \times \Gamma^*} \ \mathrm{LM}_n(\gamma),$$

where $\Gamma^* = [\gamma_L, \gamma_U]$, and γ_L and γ_U satisfy $P(s_t \leq \gamma_L) = p$ and $P(s_t \leq \gamma_U) = 1 - p$, respectively.

The threshold parameter γ_L is the p-th percentile of the term spread, and γ_U is the $(1-p)$-th percentile. Depending on the degrees of freedom, p can be set at .05, .10, .15.

The SupLM statistic has a nonstandard asymptotic distribution as shown by Hansen and Seo (2000). The bootstrapping p-values can be calculated and we reject the null hypothesis if the bootstrapping p-values are smaller than the size chosen.

To estimate the threshold VECM, we take a grid of $\Gamma^* = [\gamma_L, \gamma_U]$, which is equally-spaced. If we fix γ, the threshold VECM is linear in the parameter and the maximum likelihood estimator is the same as the least squares estimator. Thus, a sequential method of linear regression can be applied to the estimation of the threshold VECM as follows:

$$\mathrm{Min}_{\gamma \in \Gamma^* \times \Gamma^*} \ \mathrm{Min}_{\theta_1, \theta_2, \theta_3} \ -\mathcal{L}(\gamma, \theta_1, \theta_2, \theta_3)$$

where $\mathcal{L}(\gamma, \theta_1, \theta_2, \theta_3) = -n \log(2\pi) - \frac{n}{2} \log|\Sigma| - \frac{1}{2} \sum_{t=1}^{n} \mathrm{tr}(\Sigma^{-1} u_t u_t')$, where u_t is in equation (8).

4 Empirical Results

To assess the nonlinear adjustment in the U.S. term structure, this paper uses monthly interest rates on the 3-month Treasury bill as the short rate and the 1-year Treasury bill and the 5-year, 10-year Treasury notes as the long rate for the period 1960:1-1999:12.[1]

We first analyze the term spread defined as $s_t = R_t - r_t$, where r_t is the monthly interest rate on the 3-month Treasury bill (TB3M) and R_t is the rate on the 1-year Treasury bill

[1] The data is extracted from the St. Louis federal reserve databank at www.stls.frb.org/fred/.

(TB1Y) and the 5-year (TB5Y), 10-year (TB10Y) Treasury notes. The unit root hypothesis on each interest rate cannot be rejected by the ADF unit root test.

As Table 1 shows, the likelihood ratio statistics reject the null hypothesis of no cointegration at the 5% size. Thus, we find a long-run relationship between the long rates and the short rate by using Johansen's (1988) cointegration tests. For example, at the VAR lag-length 3 picked by Schwarz, the likelihood ratio statistic is 27.471 which exceeds the asymptotic critical value at the 5% size.

Table 1. Cointegration Tests

Model	(TB10Y, TB3M)	(TB5Y, TB3M)	(TB1Y, TB3M)
LR ($\mathcal{H}_0^* : rank = 0$)	27.471	27.986	37.314
LR ($\mathcal{H}_0^{**} : rank = 1$)	3.771	4.416	5.416

*: 5% critical value=19.96, **: 5% critical value=9.24.

The tests for cointegration show that a long-run equilibrium relationship exists in the term structure, which implies that the term spread is mean-reverting. The long-run relationship supports the expectations hypothesis in the sense that the term spread has a long-run predictability. This result is consistent with that of Campbell and Shiller (1987).

Table 2. Linear Error Correction Model

Model	TB10Y	TB3M	TB5Y	TB3M	TB1Y	TB3M
β		1.111		1.083		1.068
s.e.		0.125		0.100		0.029
α	-0.017	0.049	-0.023	0.053	0.010	0.169
s.e.	0.013	0.030	0.020	0.036	0.087	0.093
μ	0.015	-0.031	0.019	-0.033	0.001	-0.032
s.e.	0.018	0.037	0.023	0.040	0.030	0.032
Likelihood	578.207		546.280		604.676	

Table 2 shows the estimates of the linear error correction model that are estimated by reduced rank regression. The VAR lag length chosen is 3 by Schwarz criterion. Standard errors are based on the heteroskedasticity-consistent covariance estimator. The cointegrating vectors are estimated close to 1, but the cointegrating estimate of (TB1Y, TB3M) is statistically different from 1. The adjustment coefficient of the short rate is positive in each model and its magnitude increases as the maturity of the long rate decreases. However, the estimate of the adjustment coefficient is not significant at the 5% size, and thus, the long-run relationship does not provide significant information on the future change in the short rate. This result is consistent with the empirical results of Shiller et al. (1983) and Mankiw and Miron (1986). The adjustment coefficient of the long rate is not statistically significant in each model.

The persistence of a shock in the long-run relationship can be measured by its half life. The half life of a shock is estimated at 9.43, 8.32, and 3.72 for the model (TB10Y, TB3M), (TB5Y, TB3M), and (TB1Y, TB3M), respectively. Thus, the term spread becomes more persistent as the maturity of the long rate increases.

To allow for the transaction costs in the U.S. bond market, we use the threshold vector error correction model proposed by Hansen and Seo (2000). The tests for threshold effects supports the hypothesis of nonlinear mean reversion in the term structure as Table 3. For example, the SupLM statistic for the tests of nonlinear mean reversion in the model (TB10Y, TB3M) is calculated at 51.027 with a bootstrapping p-value of 0.007. The tests are based on the threshold VECM with a VAR lag length of 3 and the trimming parameter $p = 0.10$. The bootstrapping p-values are calculated on the linear error correction model with 1,000 bootstrapping replications.

Table 3. Tests for Nonlinear Mean Reversion

Model	(TB10Y, TB3M)	(TB5Y, TB3M)	(TB1Y, TB3M)
SupLM	51.027	43.860	46.367
5% c.v.	44.583	44.163	44.528
p-value	0.007	0.054	0.029

In general, the bond markets of different maturities may be segregated by the transaction costs such as the bid-ask spread and the price impact. Also, different taxes and investment regulations prevent investors from allocating assets immediately even if an arbitrage opportunity exists.

Because the monetary authority controls interest rates, the persistent behavior of the term spread may depend on the monetary policy. For instance, the monetary authority maintained a high level of short rate and a low level of long-term rate in the episode of 'the Operational Twist.' Makiw and Miron (1986) provides the empirical results that the predictability of the yield curve began to disappear after the foundation of the Federal Reserve. Also, Rudebusch (1995) and Balduzzi et al. (1997) show that the unexpected changes in the Fed targeting explain a large part of the future change in the interest rates. If the monetary authority controls interest rates as a means of economic policy, the term spread is likely to be persistent.

Table 4 shows the estimation result of the threshold VECM, which is estimated by maximum likelihood estimation at the VAR lag-length 3. Standard errors are calculated from the heteroskedasticity-robust covariance estimator. The trimming parameter p is set at 0.10.

In the model (TB10Y, TB3M), the threshold estimates are 0.070, 2.730, and $P(s_{t-1} \leq \gamma_1)$, $P(\gamma_1 < s_{t-1} \leq \gamma_2)$, and $P(s_{t-1} > \gamma_2)$ are estimated at 0.130, 0.717, and 0.153, respectively. The adjustment coefficient of the short rate is positive in each regime, but its magnitude is significant only in Regime 1. The adjustment coefficient of the long rate is not significant in Regimes 1 and 2, but the long rate responds to the term spread in Regime 3.

Figure 4 shows the response function of the short and long rates to the term spread estimated on the model (TB10Y, TB3M). The response function is based on the estimates of the intercept and the adjustment vector in each regime given the other short-run dynamics. In Regime 1, the short rate increases as the lagged term spread increases, and thus the term spread provides information on the future change in the short rate. In Regime 2, the term spread is persistent because the short and long rates do not respond to the term spread. Thus, the predictability of the term spread vanishes in Regime 2. In Regime 3, the adjustment coefficient of the short rate is not significant, but the term spread is mean-reverting because the long rate responds to the term spread.

Table 4. Estimation of Threshold VECM

Model	TB10Y	TB3M	TB5Y	TB3M	TB1Y	TB3M
α_1	0.031	0.603	0.054	0.498	0.470	1.110
s.e.	0.111	0.242	0.143	0.263	0.425	0.505
μ_1	0.122	0.255	0.137	0.105	-0.074	-0.227
s.e.	0.046	0.099	0.048	0.073	0.105	0.120
α_2	-0.017	0.000	-0.025	-0.040	0.205	0.250
s.e.	0.016	0.024	0.027	0.037	0.072	0.072
μ_2	0.019	-0.002	0.027	0.057	-0.092	-0.114
s.e.	0.022	0.035	0.027	0.046	0.045	0.043
α_3	-0.311	0.040	-0.253	0.122	-0.467	0.073
s.e.	0.136	0.210	0.128	0.194	0.246	0.324
μ_3	0.958	-0.107	0.606	-0.297	0.513	-0.155
s.e.	0.427	0.660	0.322	0.485	0.324	0.412
$\gamma = (\gamma_1, \gamma_2)$	0.070, 2.730		0.170, 2.090		0.300, 1.090	
P_1, P_2	0.130	0.717	0.130	0.648	0.233	0.631
Likelihood	672.270		623.315		685.329	

$P_1 = P(s_{t-1} \leq \gamma_1)$, $P_2 = P(\gamma_1 < s_{t-1} \leq \gamma_2)$.

The response of the model (TB5Y, TB3M) is estimated in Figure 5, which is similar to Figure 4 although the probability of staying in Regime 2 decreases. Figure 6 shows the response function of the model (TB1Y, TB3M). The short rate responds to the term spread in Regime 1, and its response is also significant in Regime 2. The adjustment coefficient of the short rate is positive in Regime 3, but its magnitude is not statistically significant. Instead, the adjustment coefficient of the long rate is significant in Regime 3.

The term spread and threshold estimates are depicted in Figure 7. Particularly, the term spread is temporarily persistent in Regime 2 and it reverts to the equilibrium when its deviation exceeds the thresholds. The estimation results reveal an asymmetric adjustment behavior in the term structure; the adjustment speed varies and the adjustment instrument switches depending on the regime which is determined by the current state of market disequilibrium.

We estimated the linear vector error correction model in Table 2. Obviously, the asymmetric adjustment behavior cannot be explained by the linear cointegration model. The mean-reversion depends only on the short rate with a constant adjustment speed. The difference in the likelihood between these two models is not negligible, and hence the threshold VECM explains the behavior of the term spreads in the U.S. bond market better than the linear VECM.

Table 5. Transition Probability and Regime Duration

	Transition Probability P_{ij}			Duration	
	$S_{t-1} = 1$	$S_{t-1} = 2$	$S_{t-1} = 3$	Mean	Median
	TB10Y, TB3M				
$S_t = 1$	0.839	0.029	0.000	6.200	3.500
$S_t = 2$	0.161	0.930	0.192	13.680	7.000
$S_t = 3$	0.000	0.041	0.808	5.214	2.000
	TB5Y, TB3M				
$S_t = 1$	0.774	0.045	0.000	4.429	2.000
$S_t = 2$	0.226	0.906	0.142	10.300	4.000
$S_t = 3$	0.000	0.049	0.858	7.067	5.000
	TB1Y, TB3M				
$S_t = 1$	0.730	0.097	0.015	3.700	2.000
$S_t = 2$	0.252	0.833	0.338	5.902	3.000
$S_t = 3$	0.018	0.070	0.646	2.826	1.000

$P_{ij} = P(S_t = i | S_{t-1} = j)$, for $i, j = 1, 2, 3$.

Table 5 computes the transition probability from one regime to another and the regime duration in each model. In the model (TB10Y, TB3M), the probability of staying in Regime 2 is estimated at 0.930, which is larger than that of staying in Regime 1 or 3. In the model (TB10Y, TB3M), the duration of Regime 1 has a mean of 6.2 while that of Regime 2 has a mean of 13.68. The median of duration is 3.5, 7, and 2 for Regime 1, 2, and 3, respectively. Thus, Regime 2 has a longer duration than the other regimes, which implies that the persistence of the term spread depends on the regime. If the current term spread belongs to Regime 2, it is more likely to stay in Regime 2.

Table 6. Predictive Accuracy

	Random Walk (A)	Linear VECM (B, B/A)	Threshold VECM (C, C/A)
	TB10Y, TB3M		
RMSE	0.536	0.481 (0.897)	0.423 (0.789)
MAE	0.293	0.292 (1.000)	0.260 (0.887)
	TB5Y, TB3M		
RMSE	0.536	0.475 (0.885)	0.430 (0.802)
MAE	0.293	0.288 (0.984)	0.259 (0.885)
	TB1Y, TB3M		
RMSE	0.536	0.447 (0.833)	0.418 (0.780)
MAE	0.293	0.277 (0.946)	0.272 (0.930)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{E}_{t-1}r_t - r_t)^2},\ \text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|\hat{E}_{t-1}r_t - r_t|.$$

Table 6 compares the predictive ability of the threshold VECM with that of the random walk and the linear cointegration models. In the model (TB10Y, TB3M), the root mean squared error (RMSE) of the threshold VECM is about 20% lower than the random walk model while the linear VECM reduces 10% of the RMSE. In the same model, the mean absolute error (MAE) of the threshold VECM is about 10% lower than the random walk model while that of the linear cointegration model is the same as that of the random walk model. The root mean squared error (RMSE) and the mean absolute error (MAE) are defined for one-step ahead forecast errors. In other models, the RMSE and the MAE of the threshold VECM are 10%-20% lower than the random walk model. Thus, the threshold VECM improves the predictive accuracy.

5 Extensions

Previous results are based on the known cointegrating vector, and thus they are different from the full-information maximum likelihood estimation. This section provides the estimation and testing results on the threshold VECM with an unknown cointegrating vector.

The threshold VECM is based on the step function whose arguments include the cointegrating vector. The likelihood function is not smooth in the parameter vector, and the conventional methods of gradient hill-climbing cannot be applied. Hansen and Seo (2000) has proposed a grid-search algorithm for the threshold VECM, and this paper uses it.

The threshold VECM with an unknown cointegrating vector can be written as follows:

$$\Delta x_t = \begin{array}{l} (\mu_1 + \alpha_1 w_{t-1} + \sum_{i=1}^{l} \Gamma_{1i} \Delta x_{t-i}) 1(w_{t-1} \le \gamma_1) + \\ (\mu_2 + \alpha_2 w_{t-1} + \sum_{i=1}^{l} \Gamma_{2i} \Delta x_{t-i}) 1(\gamma_1 < w_{t-1} \le \gamma_2) + \\ (\mu_3 + \alpha_3 w_{t-1} + \sum_{i=1}^{l} \Gamma_{3i} \Delta x_{t-i}) 1(w_{t-1} > \gamma_2) + u_t, \end{array} \tag{11}$$

where $w_t = R_t - \beta r_t$, and $u_t \sim i.i.d.(0, \Sigma)$.

The log-likelihood function can be defined as follows:

$$\mathcal{L}(\beta, \gamma, \theta_1, \theta_2, \theta_3) = -n \log(2\pi) - \frac{n}{2} \log |\Sigma| - \frac{1}{2} \sum_{t=1}^{n} \mathrm{tr}(\Sigma^{-1} u_t u_t^{'}),$$

where u_t is in equation (11).

If we fix β and γ, the likelihood function is linear in the parameter and the threshold VECM can be estimated by linear regression. Thus, the grid-search estimation algorithm is based on the sequential linear regression of the threshold VECM.

First, we estimate the linear VECM, and then by using the estimate and the standard error of β we take a grid of $\beta^* = [\beta_L, \beta_U]$. We select β_L and β_U which satisfy $P(\tilde{\beta} \in \beta^*) = \phi$, where $\tilde{\beta}$ is the estimate of the linear VECM. Based on the normality of $\tilde{\beta}$, ϕ can be chosen close to 1. Also, the grid of $\Gamma^* = [\gamma_L, \gamma_U]$ is based on the estimate of the linear VECM. That is, $P(\tilde{w}_{t-1} \leq \gamma_L) = p$, and $P(\tilde{w}_{t-1} \leq \gamma_U) = 1 - p$, where $\tilde{w}_t = R_t - \tilde{\beta} r_t$ and p can be chosen at 0.05, 0.10, 0.15.

A sequential method of linear regression can be applied to the estimation of the threshold VECM as follows:

$$\mathrm{Min}_{\beta \in \beta^*, \gamma \in \Gamma^* \times \Gamma^*} \ \mathrm{Min}_{\theta_1, \theta_2, \theta_3} \ -\mathcal{L}(\beta, \gamma, \theta_1, \theta_2, \theta_3).$$

As Table 7 shows, the full-information maximum likelihood estimation provides similar results to those of the threshold VECM with a known cointegrating vector. The equilibrium error persists in Regime 2 because the adjustment coefficients are insignificant. The short rate responds to the equilibrium error in Regime 1, and the long rate responds in Regime 3.

Table 7. Estimation of Threshold VECM (β unknown)

Model	TB10Y	TB3M	TB5Y	TB3M	TB1Y	TB3M
β		1.247		1.071		1.037
α_1	0.003	0.193	0.055	0.366	0.673	1.261
s.e.	0.035	0.079	0.100	0.190	0.417	0.463
μ_1	0.031	0.314	0.167	0.307	0.168	0.174
s.e.	0.060	0.132	0.073	0.138	0.117	0.109
α_2	-0.013	0.019	-0.018	-0.021	0.021	0.079
s.e.	0.019	0.020	0.033	0.040	0.074	0.078
μ_2	0.003	0.001	0.003	0.012	0.009	-0.003
s.e.	0.015	0.018	0.024	0.027	0.026	0.029
α_3	-0.151	-0.180	-0.234	0.080	-0.438	0.116
s.e.	0.082	0.128	0.118	0.168	0.199	0.274
μ_3	0.361	0.457	0.472	-0.141	0.408	-0.079
s.e.	0.209	0.326	0.243	0.341	0.185	0.247
$\gamma = (\gamma_1, \gamma_2)$	-0.802, 2.012		-0.154, 1.597		-0.022, 0.686	
Likelihood	676.875		624.684		697.847	

As Table 8 shows, the bootstrap p-values of the SupLM statistic reject the null hypothesis of linear mean reversion at the 5% size. The p-values are computed on the linear VECM with 1,000 bootstrapping replications. Thus, the tests for threshold effect support the nonlinear mean reversion in the U.S. term structure of interest rates.

Table 8. Tests for Nonlinear Mean Reversion (β unknown)

Model	(TB3M, TB10Y)	(TB3M, TB5Y)	(TB3M, TB1Y)
SupLM	53.886	44.365	48.078
5% c.v.	44.911	43.975	43.572
p-value	0.000	0.047	0.009

6 Concluding Remarks

This paper assesses the effect of the transaction costs on the predictability of the term structure of interest rates. We find that the predictability is regime dependent, where the regimes are determined by the equilibrium error or the term spread. The short-term interest rate responds to the past term spread only if the term spread exceeds the threshold parameters. Our results are consistent with the optimal adjustment model which considers the transaction costs in the financial market. The transaction costs prevent investors from realizing an arbitrage opportunity of term spread if the opportunity does not exceed the transaction costs. Thus, our results provide one explanation regarding the predictability of the term spread, which has been rejected by many authors.

Also, this paper suggests the threshold error correction model, which is consistent with the stylized fact of the term structure. The threshold cointegration model outperforms the linear cointegration model and the random walk model, and thus it can be used in the evaluation of the monetary policy and economic forecasting.

References

[1] Andrews, D. and W. Ploberger, 1994, "Optimal Tests When a Nuisance Parameter is Present only under the Alternative," Econometrica 62, 1383-1414.

[2] Backus, D., A. Gregory, and S. Zin, 1989, "Risk Premiums in the Term Structure: Evidence from Artificial Economies," Journal of Monetary Economics 24, 371-399.

[3] Balduzzi, P., G. Bertola, and S. Foresi, 1997, "A Model of Target Changes and the Term Structure of Interest Rates," Journal of Monetary Economics 39, 223-249.

[4] Balke, N. and T. Fomby, 1997, "Threshold Cointegration," International Economic Review 38, 627-645.

[5] Campbell, J., 1995, "Some Lessons from the Yield Curve," Journal of Economic Perspectives 9, 129-152.

[6] Campbell, J. and R. Shiller, 1987, "Cointegration and Tests of Present Value Models," Journal of Political Economy, 95, 1062-1088.

[7] Campbell, J. and R. Shiller, 1991, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," Review of Economic Studies 58, 495-514.

[8] Cox, J., J. Ingersoll, and S. Ross, 1981, "A Re-examination of Traditional Hypotheses about the Term Structure of Interest Rates," Journal of Finance 36, 769-799.

[9] Cox, J., J. Ingersoll, and S. Ross, 1985, "A Theory of the Term Structure of Interest Rates," Econometrica 53, 385-407.

[10] Culbertson, J. M., 1957, "The Term-structure of Interest Rates," Quarterly Journal of Economics 71, 485-517.

[11] Engle, R. and C. Granger, 1987, "Co-integration and Error Correction: Representation, Estimation, and Testing," Econometrica 55, 251-276.

[12] Fama, E., 1984, "The Information in the Term Structure," Journal of Financial Economics 13, 509-528.

[13] Fama, E., 1984, "Term Premiums in Bond Returns," Journal of Financial Economics 13, 529-546.

[14] Friedman, B., 1977, "Financial Flow Variables and the Short-run Determination of Long-term Interest Rates," Journal of Political Economy 85, 661-689.

[15] Friedman, B., 1979, "Substitution and Expectation Effects on Long-term Borrowing Behavior and Long-term Interest Rates," Journal of Money, Credit, and Banking 11, 131-150.

[16] Hansen, B. and B. Seo, 2000, "Testing for Threshold Cointegration," unpublished manuscript.

[17] Jegadeesh, N. and G. Pennacchi, 1996, "The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures," Journal of Money, Credit, and Banking 28, 426-446.

[18] Johansen, S., 1988, "Statistical Analysis of Cointegration Vectors," Journal of Economic Dynamics and Control 12, 231-254.

[19] Lewis, K., 1991, "Was There a 'Peso Problem' in the U. S. Term Structure of Interest Rates: 1979-1982?" International Economic Review 32, 159-173.

[20] Mankiw, G. and J. Miron, 1986, "The Changing Behavior of the Term Structure of Interest Rates," Quarterly Journal of Economics 101, 211-228.

[21] Mankiw, G. and L. Summers, 1984, "Do Long-term Interest Rates Overreact to Short-term Interest Rates?" Brookings Papers on Economic Activity, 223-242.

[22] Modigliani, F. and R. Shiller, 1973, "Inflation, Rational Expectation and the Term Structure of Interest Rates," Economica 40, 12-43.

[23] Modigliani, F. and R. Sutch, 1966, "Innovations in Interest Rate Policy," American Economic Review 56, 178-197.

[24] Modigliani, F. and R. Sutch, 1967, "Debt Management and the Term Structure of Interest Rates: An Empirical Analysis of Recent Experience," Journal of Political Economy 75, 569-589.

[25] Rudebusch, G., 1995, "Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure," Journal of Monetary Economics 35, 245-274.

[26] Sargent, T., 1972, "Rational Expectations and the Term Structure of Interest Rates," Journal of Money, Credit and Banking 4, 74-97.

[27] Sargent, T., 1979, "A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure," Journal of Monetary Economics 5, 133-143.

[28] Seo, B., 1998, "Tests for Structural Change in Cointegrated Systems," Econometric Theory 14, 222-259.

[29] Shiller, R., 1979, "The Volatility of Long-term Interest Rates and Expectations Models of the Term Structure," Journal of Political Economy 87, 1120-1219.

[30] Shiller, R., J. Campbell, and K. Schoenholtz, 1983, "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates," Brookings Papers on Economic Activity, 173-217.

[31] Simon, D., 1990, "Expectations and the Treasury Bill-Fed Funds Rate Spread over Recent Monetary Policy Regimes," Journal of Finance 65, 567-577.

[32] Tong, H., 1978, On a Threshold Model, Pattern Recognition and Signal Processing, ed. C. H. Chen, Amsterdam: Sijhoff & Noordhoff.