ANNALS OF ECONOMICS AND FINANCE **15-2**, 411–431 (2014)

Yield Factors, the Expectations Hypothesis and Regime Shifts*

Rui Cao

Central University of Finance and Economics

and

Xiaoneng Zhu

The Chinese Academy of Finance and Development
Central University of Finance and Economics, Beijing, China
E-mail: xiaonengz@gmail.com

and

Shahidur Rahman

KIMEP University

This paper reexamines the expectations hypothesis of interest rates using US monthly data for bond yields ranging in maturity from 1 month to 10 years. We extend the Campbell-Shiller (1987) method to increase the test power: (a) by introducing economic variables as conditioning information; and (b) by explicitly taking regime shifts into consideration. We apply the new method to Treasury bond yields and find that the expectations hypothesis is rarely rejected by the US term structure data. We also find that two regimes are related to the business cycle and monetary policy.

Key Words: Expectations hypothesis; Regime shift; Risk premiums; Term structure; Yield factors.
JEL Classification Numbers: C12, C32, E43, G12.

* We thank Dennis Alba, Jun Yu, Yangru Wu, Lancaster Robert, Victoria Long, Longzheng Fan, Choy Keen Meng, Jianjun Li, Fei Wu, for helpful comments. Xiaoneng Zhu acknowledges the financial support from the Nature Science Foundation of China (Grant No. 71473281).

1529-7373/2014

1. INTRODUCTION

The expectations hypothesis (EH) is a central view of the term structure of interest rates. According to the EH, the long-term yield is a weighted average of the expected future short-term yields plus a maturity-specific constant risk premium.[1] Empirically, the EH has been intensively examined using a variety of tests and data since it has long been recognized as a basic workhorse model of the term structure of interest rates. Some examples include Bekaert and Hodrick (2001), Sarno, Thornton, Valente (2007), Bulkley and Giordani (2011), Bulkley, Harris, and Nawosah (2011), Frankel and Froot (1987), Froot (1989), Bekaert, Hodrick,and Marshall (1997), Zhu (2014). Contrary to prior belief, most empirical studies reject the expectations hypothesis.[2] In addition to the statistical rejection, some recent studies (see, for example, Campbell and Shiller, 1991; Cochrane and Piazzesi, 2005; Fama and Bliss, 1987) find that yield spreads or forward rates have predictive power on future excess bond returns. This appears to be the critical evidence on the empirical failure of the EH since the EH implies that excess bond returns are unpredictable.

However, the standard tests of the EH have missing motivations. Bekaert and Hodrick (2001) provide three potential reasons for understanding the empirical failure of the EH. They are respectively the failure of the rational expectations assumption, the presence of the time-varying risk premium, and the poor properties of the statistical tests in finite samples caused by regime shifts. In the spirit of Bekaert and Hodrick (2001), we attempt to extend the Campbell and Shiller (1987) testing procedure and reexamine the expectations hypothesis of interest rates.

The first extension is to allow for regime shifts in the testing of the EH. This extension is inspired by a large literature that highlights the importance of regime shifts in describing term structure dynamics such as Ang and Bekaert (2002), Bansal and Zhou (2002), Bansal, Tauchen, and Zhou (2004), Gray (1996), Fuhrer (1996), Hamilton (1988), Xiang and Zhu (2013), Zhu (2012). In particular, Dai, Singleton, and Yang (2007) show that regime shifts are a risk factor and regime uncertainty will be reflected in the risk premium. There are also some well-accepted economic

[1] This general form of the expectations hypothesis is typically examined in the empirical literature. In this paper, we focus on this form. Indeed, many different forms have appeared in the literature. Cox, Ingersoll, and Ross (1981) show that some forms are inconsistent with each other and argue that some forms imply the existence of arbitrage opportunities. However, these inconsistencies seldom occur with this general form (see Campbell (1986) and Longstaff (2000a) for further discussion).

[2] There are some exceptions. For example, Longstaff (2000) found that the EH is valid at the very short end of the yield curve. After controlling for year-end rate increases in commercial paper rates, Downing and Oliner (2007) find considerable support for a generalized form of the expectations hypothesis.

reasons for understanding regime shifts in yield curve movements such as the business cycle (e.g., Bansal and Zhou 2002), shifts in monetary policy regimes (e.g., Ang, Bovin, Dong, and Loo-Kung, 2009; Fuhrer, 1996; Li, Li, and Yu, 2011), inflation (e.g., Ang, Bekaert, and Wei, 2008), the risk premium (e.g., Dai, Singleton, and Yang, 2007), and the real interest rate (e.g., Garcia and Perron, 1996). Given the importance of regime shifts in term structure dynamics, it is therefore important to take regime shifts into consideration in the testing of the EH.

The second extension is to use a larger information set in the testing of the EH. In so doing, we move beyond the bivariate comparisons of short- and long-term yields which dominate the existing literature. This is motivated by the growing literature linking yield factors and macro factors to the dynamics of the term structure of interest rates (e.g., Sarno, Thornton, and Valente, 2007; Ang and Piazzesi, 2003; Carriero, Favero, and Kaminska, 2006; Kozicki and Tinsley, 2001; Evans and Marshall, 2002; Bekaert, Cho, and Moreno, 2004; Rudebusch and Wu, 2008). In particular, Sarno, Thornton, and Valente (2007) suggest that a larger information set can improve the power of the EH testing procedure. In the spirit of Sarno, Thornton, and Valente (2007), we hence use a larger information set.

Our method for testing the EH is developed in a matrix tractable way. We apply the method to the US data. We find that the EH is rarely rejected by the US term structure data. This contrasts to a vast body of previous results (see, for example, Campbell and Shiller, 1991). Our empirical results highlight the importance of including yield factors in the conditional information set and taking into account regime shifts. The results have two implications for risk premiums. On the one hand, the empirical findings may suggest that yield factors can partially capture time-varying risk premiums. On the other hand, regime-switching models can better describe the term structure of interest rates.

To shed light on the economic interpretation of two regimes, we link two regimes to the business cycle and monetary policy. Our empirical analysis confirms the traditional wisdom that the high-volatility regime is more likely to be related to economic recessions. The empirical analysis also suggests that the Federal Reserve is more accommodative for growth in the high volatility regime and is more active in controlling for inflation in the low volatility regime.

The remaining paper is organized as follows. Section 2 discusses the expectations hypothesis in an arbitrage-free framework and derives the restrictions implied by the EH. The data descriptions and summary statistics of yields and yield factors are presented in Section 3. The arbitrage-free dynamic Nelson-Siegel model is also briefly introduced in this section. Section 4 applies the tests to the data and presents the empirical results. Section 5 concludes.

2. TESTING THE EXPECTATIONS HYPOTHESIS

2.1. The expectations hypothesis

The expectations hypothesis of the term structure of interest rates states that the n-period continuously compounded yield $i_{t,n}$ equals a weighted average of the current and expected short yields plus a maturity-specific constant risk premium[3]

$$i_{t,n} = (1-\delta)\sum_{j=0}^{\infty}\delta^j E_t i_{t+j,1} + c_n, \tag{1}$$

where the discount factor δ is a parameter reflecting the impatience of economic agents; c_n denotes a constant maturity-specific premium; and E_t is the conditional expectations based on the information set at time t. To be consistent with the economic theory, a class of modern asset pricing models impose no-arbitrage restrictions. It is straightforward to demonstrate that the EH in equation (1) is consistent with arbitrage-free conditions. Let M_{t+1} denote the pricing kernel, it is well-known that any gross return R_{t+1} in an economy that does not admit arbitrage opportunities can be correctly priced by

$$E_t(M_{t+1}R_{t+1}) = 1. \tag{2}$$

To be statistically tractable, it is assumed that returns and pricing kernels are conditionally log-normal. Following Bekaert and Hordrick (2001), equation (2) implies that

$$E_t(m_{t+1}) + 0.5V_t(m_{t+1}) + E_t(r_{t+1}) + 0.5V_t(r_{t+1}) + Cov_t(m_{t+1}, r_{t+1}) = 0, \tag{3}$$

where V_t and Cov_t respectively represent conditional variance and covariance, and the lower letters denote the logs of the corresponding uppercase letters, for example, $m_{t+1} = \log(M_{t+1})$. Since the return of one-period yield $i_{t,1}$ is observable at time point t, the last two items on the left-hand side of equations (3) disappear. Thus the expression for the one-period yield is

$$i_{t,1} = -[E_t(m_{t+1}) + 0.5V_t(m_{t+1})]. \tag{4}$$

[3]This version of the expectations hypothesis is a present value model of the term structure of interest rates. If we do not take this form, we cannot use yield factors to replace the long- and short-term yields in the testing of the EH. It is possibly that our results in support of the EH is due to this specific form of the EH.

Now let r_t in equations (3) represent the excess return of the holding period return ($h_{t+1,n}$) of a long-term bond over an one-period bond. Combined with equations (4), the expected excess return can be given by the following equation:

$$E_t(h_{t+1,n}) - i_{t,1} = -[Cov_t(m_{t+1}, h_{t+1,n}) + 0.5V_t(h_{t+1,n})]. \tag{5}$$

The right-hand side of equations (5) is a constant conditional on the time t information set. Let a_n denote this constant, then equation (5) can be expressed as:

$$E_t(h_{t+1,n}) = i_{t,1} + a_n. \tag{6}$$

The one-period holding return on a n-period bond can be approximated (see, for example, Shiller, 1979) by a linear function in the neighborhood of $i_{t,n} = i_{t+1,n-1}$.

$$h_{t+1,n} = \frac{i_{t,n} - \delta i_{t+1,n-1}}{1-\delta} \tag{7}$$

By taking expectations and using equations (4) and (5), after rearrangement, we have

$$i_{t,n} = (1-\delta)i_{t,1} + \delta E_t i_{t+1,n-1} + (1-\delta)a_n. \tag{8}$$

Using recursive substitution and letting $n \to \infty$, combined with the terminal condition $i_{t,0} = 0$, we have the present value version of the expectations hypothesis in equation (1) with $c_n = (1-\delta)\sum_{j=0}^{\infty} \delta^j a_n$. Hence, we show that the expectations hypothesis is consistent with no-arbitrage conditions.

2.2. Method

The nonstationarity of time series may invalid the statistical inference. The realization that yields are usually persistent and integrated of order one motivates Campbell and Shiller (1987, 1991) to test the EH using the yield spread $S_{t,n} = i_{t,n} - i_{t,1}$ and the first difference of short-term yields. The stationarity of the yield spread imposes a restriction on the long-run dynamics of yields. Specifically, by subtracting $i_{t,1}$ from both sides of equation (1) and rearranging, we have

$$S_{t,n} = \sum_{j=1}^{\infty} \delta^j E_t \Delta i_{t+j,1} + c_n. \tag{9}$$

The equation suggests that the yield spread is a weighted average of a stationary variable–the first difference of the short-term yield. However, this is a necessary but not sufficient condition of the EH since the validity of the EH also requires restrictions to be imposed on the short-run dynamics of yields.

Equations (9) provides the testable restrictions implied by the expectations hypothesis. To test the EH, we also need a data generating process. In the Campbell-Shiller approach, the data generating process is a vector autoregression (VAR) with the state variables $y_t = [S_{t,n}, \Delta i_{t,1}]'$. However, a large literature shows that a three-factor model is needed to accurately describe the movements of the yield curve. Thus, the Campbell-Shiller approach seems to have a missing term structure factor. To avoid potential mis-specification, we shall access the EH with an extended vector of state variables.

Motivated by modern term structure modes, we include three yield factors in the EH testing. According to their effect on the yield curve, three factors in the term structure literature are usually labeled as the 'level', 'slope', and 'curvature' factors. Empirically, the level factor is a long-term factor because it is highly correlated with long-term yields. In contrast, the slope factor a short-term factor and the curvature is a medium-term factor. In order to shed light on the entire yield curve on the testing of the EH, instead of a pair of long- and short-term yields, we use the level, slope and curvature factors from the dynamic Nelson-Siegel model (DNS) (see, Diebold and Li, 2006) to test the EH. Indeed, the sum of the level and slope factors has a nice interpretation of short-term yield. In addition, minus slope factor can been taken as yield spread. By using the yield factors, we need not test each pair of long- and short-term yields.

An extended VAR approach has several advantages. First, the extension of information set may alleviate simultaneity bias in the estimation (see, for example, Carriero, Favero, and Kaminska, 2006). The improved power properties of the extended testing procedure provide the second reason (see, for example, Sarno, Thornton, and Valente, 2007). Third, yield factors may partially capture regime-dependent risk premiums.

Now Let $y_t = [MS_t, \Delta LS_t, C_t]'$ be the extended 3×1 vector of the state variables. In particular, MS_t is minus slope factor from the DNS model. In the testing of the EH, the slope factor represents the yield spread(9).[4] Δ is a first-difference operator. LS_t is the sum of the level and slope factor from the DNS model and represents the short-term yield. C_t is the curvature factor from the DNS model. We assume that the state vector y_t follows a

[4]The MS_t is a good proxy of the yield spreads between the long-term yield and the short-term yield. Panel D of Table 1 shows that the correlation between the slope factor and the spreads between 10-year yield and 3-month yield is 0.9928.

vector autoregressive process of finite order l

$$y_t = \mu + \sum_{j=1}^{l} \phi_j y_{t-j} + u_t. \tag{10}$$

For simplicity, the intercepts are removed from equation (10) since the EH does not impose any restriction on the constant risk premium. By expanding the state vector to the companion form $Y_t = [y_t', \ldots, y_{t-l}']'$, we can rewrite the state dynamics in a first-order representation:

$$Y_t = \Phi Y_{t-1} + U_t. \tag{11}$$

The information set $\Theta = [MS_{t-j},\ \Delta LS_{t-j},\ C_{t-j},\ j \geq 0]$ is observed by econometricians at time point t. Let $g' = [1, 0, \ldots, 0]$ and $h' = [0, 1, 0, \ldots, 0]$ be the selection vectors with $3l$ elements such that $MS_t = g'Y_t$ and $\Delta LS_t = h'Y_t$. It is easy to show that the m-period ahead optimal forecast is $\Delta L\hat{S}_{t+m} = h'\Phi^m Y_t$. Next, by projecting restrictions equation (9) onto the data generating process (11), we obtain[5]

$$g'Y_t = \delta h'\Phi(I - \delta\Phi)^{-1} Y_{t.} \tag{12}$$

Naturally, the testable cross-equation restrictions implied by the EH of the term structure of interest rate are

$$g' = \delta h'\Phi(I - \delta\Phi)^{-1}. \tag{13}$$

If the expectations hypothesis is true, equations (13) hold for sure and the selection of information set is not relevant. The intuitions is straightforward: given a constant term premium, all the relevant information of investors is embodied in the yield spread. However, any economic model is an approximation, in this sense, all economic models are false. The question is to what extent an economic model approximates the truth. As such, the selection of information set is important. First, selected variables should have enough information about the question being asked. Second, statistical tests based on selected variables should have good power and size properties. These two criteria account for the selection of state variables in y_t.

The CS test is also motivated by the unsatisfactory of pure econometric tests that may over-reject or under-reject the null hypothesis.[6] Furthermore, econometric tests do not tell us the economic significance of the

[5] Campbell and Shiller (1987) presents the details.

[6] Bekaert and Hordick (2001) provide the Monte Carlo simulation results for the widely used Wald and LM tests. The Wald test is found to grossly overreject the EH in a small sample. In constrast, the LM test slightly underrejects the null hypothesis.

EH. If the EH economically can explain most of the variations in actual yield spread, the EH is a good approximation of the term structure regardless of the statistical rejection or non-rejection. It is also possible that the observed sample contains little information about the expectations hypothesis. Hence, many economists are reluctant to see the statistical rejection or non-rejection of restrictions as a definitive answer. The economic significance should also play an important role in evaluating the EH.

To get some intuition on how well the model explains the economic significance, instead of resorting only to statistical significance, we can evaluate the theoretical spread

$$S^*_{t,n} = \delta h' \Phi (I - \delta \Phi)^{-1} Y_t. \tag{14}$$

This theoretical yield spread is the optimal forecast given the information set and the entertained model. If the present value model is a good enough approximation and our econometric model describes the joint dynamics of the yield factors well, the theoretical spread would be observed in financial markets. The difference between theoretical and actual spreads contains information about the adequacy of the EH. We can visually inspect the deviation by plotting both series in a diagram. The good fit of the $S_{t,n}$ and $S^*_{t,n}$ indicates a good economic significance. Another measure is to calculate the correlation between $S_{t,n}$ and $S^*_{t,n}$. The high degree of comovement suggests that economic agents accurately forecast the future changes of spread. Economic agents hence incorporate the optimal prediction into their investment decisions. As a result, no profitable arbitrage opportunity is available in the bond market.

2.3. The EH under regime shifts

We build our testing method upon a growing literature suggesting that regime-shifting models describe interest rate dynamics better than single-regime models (e.g., Ang and Bekaert 2002; Gray 1996). From an asset pricing perspective, term structure models with regime switches can account for some well-documented puzzles of the term structure (see, for example, Bansal and Zhou 2002; Dai, Singleton, and Yang 2007), for instance, the violation of the expectations hypothesis and the predictability of excess bond returns. From the central banks' perspective, it is important to understand the role of the yield curve in the monetary transmission mechanism in different regimes. Overall, regime-switching term structure models represent a parsimonious way to capture the nonlinear dynamics of interest rates.

Since it has been well-accepted that term structure dynamics are subject to discrete shifts between distinct regimes, we assume that the data generating process is a regime-switching process. Specifically, we assume that the dynamics of state variables can be described by a Markov-switching (MS) VAR model

$$Y_t = \Phi_{k_t} Y_{t-1} + U_{k_t}, \tag{15}$$

where the subscript $k_t = \{1, 0\}$ denotes the unobservable state variable, which is assumed to be governed by a discrete-time Markov chain.[7] The specification of a first-order Markov chain is not as restrictive as it seems. The first order chain offers a good approximation to higher order Markov chain regime shift (see, Hamilton, 1994; chapter 22). In addition, it also provides an approximation to some continuous regime shifts. The nonlinear process with regime shift may better characterize the yield dynamics of the selected sample than a linear process. Thus, the regime switching specification provides a parsimonious way to express complicated dynamics, which might otherwise require an ARIMA model with long lags. Alternatively, equations (15) can be rewritten as

$$Y_t = k_t \Phi_1 Y_{t-1} + (1 - k_t)\Phi_0 Y_{t-1} + k_t U_{1t} + (1 - k_t) U_{0t} \tag{16}$$

With an MS-VAR data generating process, the restrictions implied by the EH can be projected onto the MS-VAR model. As usual, the Markov chain that governs the state variable is

$$P = \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix}. \tag{17}$$

To simplify the projection, we defined a matrix M as

$$M = \begin{bmatrix} p\Phi_1 & (1-q)\Phi_1 \\ (1-p)\Phi_0 & q\Phi_0 \end{bmatrix}. \tag{18}$$

Now it is straightforward to produce the optimal forecasts given the Markov-switching data generating process. If the prevailing regime is 0, then the optimal forecast is

$$\hat{Y}_{t+m} = J M^m Q_0 Y_t. \tag{19}$$

Alternatively, if we start from the regime 1,

$$\hat{Y}_{t+m} = J M^m Q_1 Y_t, \tag{20}$$

where

[7]This is a general specification. It allows regime-dependent coefficients and heteroscedasticity. When testing the EH against data, the model selection is an empirical issue.

$$J = \iota' \otimes I_{3l}, \tag{21}$$
$$\iota = (1,1). \tag{22}$$

In addition, we have

$$Q_j = e_j \otimes I_{3l}, \tag{23}$$

where e_j is the jth column of 2×2 identity matrix in association with the state we are in, and I_{3l} is the $3l$-dimension identity matrix.

With the above well-defined notions and conditional on the prevailing regime 0, we can project equation (9) onto the data generating process equation (16) to have

$$g\prime Y_t = \delta h' JM(I - \delta M)^{-1} Q_0 Y_t. \tag{24}$$

Starting from regime 1, the project generates the following forecasts

$$g\prime Y_t = \delta h' JM(I - \delta M)^{-1} Q_1 Y_t. \tag{25}$$

Equations (24) and (25) are counterparts of equations (12) in a MS-VAR model. Thus, conditional on regime 0, the cross-equation restrictions implied by the present value model of the expectations hypothesis in equations (1) are:

$$g\prime = \delta h' JM(I - \delta M)^{-1} Q_0. \tag{26}$$

Starting from regime 1, the restrictions are

$$g\prime = \delta h' JM(I - \delta M)^{-1} Q_1. \tag{27}$$

The maximum likelihood estimation of MS-VAR with highly nonlinear restrictions is complicated. In particular, the nonlinear restrictions lead to a numerical trustworthiness problem of the maximum likelihood method. Instead of the likelihood ratio test, the Wald test is proposed to serve as an alternative.[8] Because all restricted parameters are presented in matrix M, in regime k, the first-order derivatives of the restrictions in equations

[8]Since the Monte Carlo simulation (see, Bekaert and Hodrick, 2001) results show that the Wald test usually overrejects the EH in a small sample. We suppose that an LR test will lead to empirical results more favorable to the EH.

(26) and (27) with respective to the parameters are

$$\frac{\partial C_{(\hat{\theta})}}{\partial \hat{M}_{ij}} = \delta h' JM \frac{dM}{dM_{ij}} (I - \delta M)^{-1} Q_k + \delta^2 h' JM (I - \delta M)^{-1} \frac{dM}{dM_{ij}} (I - \delta M)^{-1} Q_k \quad (28)$$

where $\frac{dM}{dM_{ij}}$ is the derivative of matrix M with respect to parameter M_{ij}. With the above derivatives, the Wald test statistics can be calculated easily.

3. DATA ISSUES

3.1. Bond Yields

The yield curve consists of the end-of-month observations of 1, 3, 6, 12, 24, 36, 60, 84, 120 months zero-coupon yields on treasury securities. The sample covers a period from January 1983 to May 2009. The data source is the Federal Reserve Bank of St. Louis. Figure **??** plots the U.S. yield curves. One stylized fact of yields is that they tend to exhibit considerable persistence and are thus believed to be nonstationary or better approximated by an integrated process. This feature has profound implications for the estimation and the statistical inference. Panel A of Table 1 provides the evidence of persistence of bond yields. The classic Campbell-Shiller regression requires that yield spreads are stationary. Panel B of Table 1 presents the Johansen cointegration analysis results.[9] The cointegration results suggest the presence of the cointegration relationship between bond yields of different maturities. The cointegration relationship thus help to explain another important stylized fact of the yield curve: spreads are less persistent than yields.

3.2. Yield factors

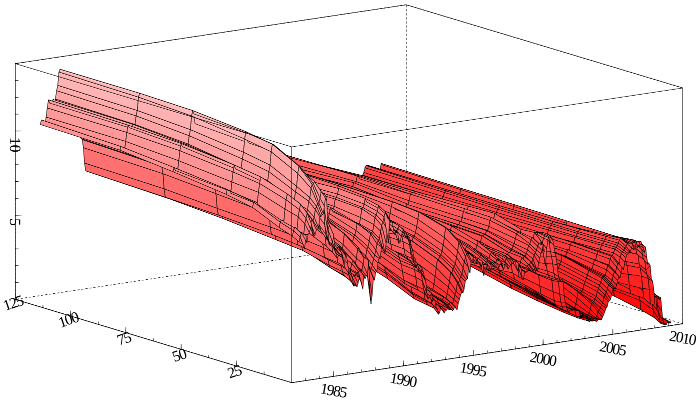
Yield factors are obtained by estimating the dynamic Nelson-Siegel model. Most term structure models use three factors to capture stylized facts of yields in cross-section and time series. By properly restricting the factor loadings in the statistical factor model, Diebold and Li (2006) propose the dynamic Nelson-Siegel model for the τ-period yield

$$i_{t(\tau)} = L_t + S_t\left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau}\right) + C_t\left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right) + \varepsilon_t,$$

where L_t is the level factor, S_t denotes the slope factor, and C_t represents the curvature factor. The parameter λ is the rate of change of factor

[9]The results for the pairwise cointegration tests are available upon request.

FIG. 1. U.S. Bond Yields.



The sample covers the period from January 1983 to May 2009. The yields plotted in this graph include, from the lowest to the highest line (with occasional crossovers), 1-, 3-, 6-, 12-, 23-, 36-, 60-, 84-, 120-month treasury zero-coupon bond yields. The yields are reported in annualized percentage terms.

loadings along maturity horizons and also determines the maturity at which the curvature loading achieves its maximum. Empirically, the level factor corresponds to the long-term interest rate, so the level factor is a long-term factor. By construction, the slope factor has a maximal impact at short maturities and a minimal effect on the longer maturity yields, so the slope factor is a short-term factor. In addition, the curvature is a medium-term factor since the factor loading of the curvature achieves its maximum at medium maturity.

The dynamic Nelson-Siegel model is flexible enough to match the changing shape of the yield curve, but it is still parsimonious and easy to estimate. We stick to the tradition of Diebold and Li (2006) and estimate the DNS model by OLS with fixed $\lambda = 0.0603$. In so doing, the yield factors at time point t only depend on the observable yields at time t. Thus, adding the yield factors in the information set will not lead to the use of posterior information. The summary statistics for the extracted factors are presented in panel C of Table 1. Panel D of Table is the pairwise correlations among the yield factors and empirical counterparts. Figure **??** plots three yield factors extracted from the DNS model and their empirical counterparts. The proxy for the empirical level factor is 10-year interest rate. The proxy for the empirical slope factor is the yield spreads between 10-year and 3-month yields. The empirical curvature factor is the average of 10-year, 2-year and 3-month yields. It is evident that the two sets of factors follow

TABLE 1.

Summary Statistics of Yields and Yield Factors

Panel A: Yields						
Maturity (months)	Mean	Std. Dev.	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$	ADF
1	4.7161	1.5176	0.9785	0.6623	0.1708	-1.0878
3	4.9747	1.5576	0.9857	0.6807	0.2130	-1.2846
6	5.7164	1.5752	0.9857	0.6870	0.2234	-1.3507
12	5.3656	1.5877	0.9854	0.7002	0.2622	-1.4357
24	5.7606	1.6064	0.9848	0.7202	0.3340	-1.1759
36	5.9690	1.5950	0.9846	0.7296	0.3745	-1.2097
60	6.2964	1.5682	0.9848	0.7427	0.4268	-1.2274
84	6.5403	1.5522	0.9856	0.7557	0.4595	-1.2260
120	6.6908	1.5277	0.9859	0.7663	0.4812	-1.3084

Panel B: Johansen Cointegration Analysis			
	Rank	Test Statistics	p_value
Trace Test:	0	505.85	[0.000]
	1	319.90	[0.000]
	2	212.12	[0.000]
	3	139.38	[0.000]
	4	86.55	[0.001]
	5	44.476	[0.100]
	6	18.937	[0.508]
	7	8.942	[0.378]
	8	1.309	[0.253]

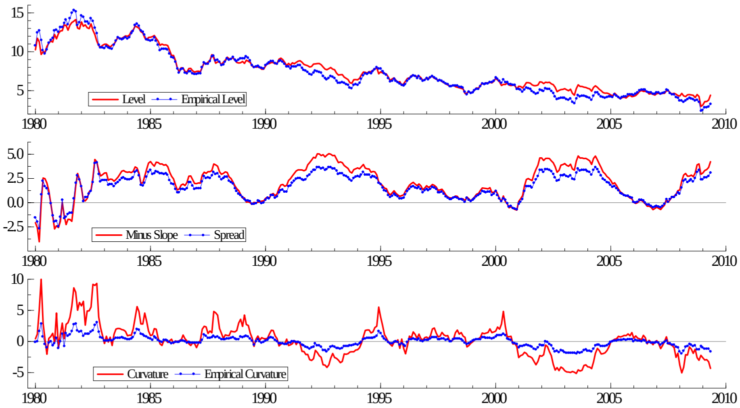
Panel C: Yield Factors						
	Mean	Std. Dev.	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$	ADF
Level	7.0027	1.4924	0.9854	0.7793	0.5039	-1.4655
Slope	-2.2428	1.2430	0.9760	0.4289	-0.2388	-2.5387
Curvature	-0.2290	1.4623	0.9250	0.4340	-0.0494	-3.2828

Panel D: Correlations	
Level and Empirical Level	0.9817
Slope and Empirical Slope	0.9928
Curvature and Empirical Curvature	0.9743

Note: $\rho(i)$ is autocorrelation with lag length i; ADF is the augmented Dickey-Fuller test with lag length selected by AIC.

each other very closely, with correlation coefficients of 0.96, 0.95 and 0.92 for the level, slope, and curvature factors respectively.

FIG. 2. Level, slope, and curvature.



The figure plots the level, slope, and curvature factors extracted from the DNS model (solid lines) and their empirical counterpartsl (the dotted lines). The upper panel plots the level factors. The middle panel plots the slope factors. The lower panel plots the curvature factors. All factors are reported in annualized percentage terms.

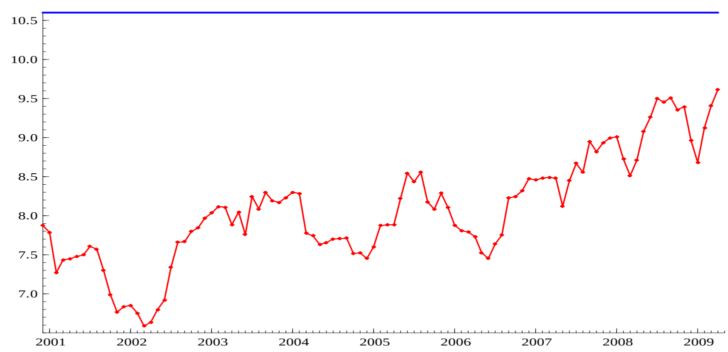
4. EMPIRICAL ANALYSIS

4.1. Testing results

The testing framework of the expectations hypothesis in this article circumvents the pairwise investigation of yields, but still sheds light on the accuracy of the EH on the entire yield curve. The curvature factor is a missing factor in the Campbell-Shiller method. It contains information about the middle section of the yield curve that are not captured by the level and slope factors. The omission of the curvature factor means that we do not use market information efficiently. The inclusion of the curvature factor also help identify two regimes.

The standard Hamilton (1989, 1994) algorithm can be used to estimate the MS-VAR model in equation (15). The smoothed probabilities that are usually parameters of interest can be calculated from the Kim (1996) filter. The entertained model is a MSH-VAR(2) model, which means a Markov-switching heteroscadasticity model with a lag length of 2. The model is selected by specification analysis on the error terms. The recursive tests are conducted since January 2001 because this is the point after which the transition probabilities p and q are usually significant. The recursive Wald statistics of testing the present value model are plotted in Figure 3. The asymptotic distribution of the test statistics is a $\chi^2(6)$ distribution. The straight line represents the ten percent critical value. It is evident from the figure that we generally cannot reject the expectations hypothesis. We therefore provide supporting evidence on the EH.

FIG. 3. The Wald statistics.



The figure plots the recursive Wald statistics for testing the validity of the cross-equation restrictions implied by the expectations hypothesis. The bold straight line represents the 10 percent critical value.

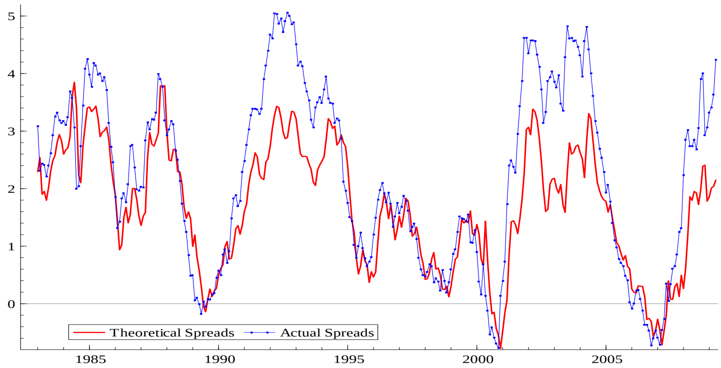
The evidence in support of the EH implies that both the conditional information set and regime shifts play an important role in understanding the empirical failure of the expectations hypothesis. Bekaert and Hodrick (2001) summarize three potential reasons for the rejection of the EH. The second interpretation is omitted variables that might capture time-varying risk premiums. The extension of the information set for testing the EH might provide an insight on this issue if time-varying risk premiums can be captured by yield factors. Our testing framework also explicitly takes regime shifts into consideration, thus it also sheds light on the peso problems, a second explanation raised by Bekaert and Hodrick for explaining the empirical failure of the EH.

Having conducted the statistical tests on the EH, it is important to investigate the economic significance of the EH. Figure 4 is a plot of the theoretical and actual yield spreads. It plots the slope factor from the DNS model and the theoretical slope factor computed from equations (24) and (25). The DNS and theoretical slope factors are highly correlated with a correlation coefficient 0.91. The result seems to suggest that in an economic sense, the expectations hypothesis is a good approximation of the term structure of interest rates. Hence, the EH is supported by the term structure data not only in a statistical sense, but also in an economic sense.

4.2. Understanding regimes

Figure 4 plots the smoothed probabilities of being in a high volatility regime. In the two-regime MS-VAR model, the most straightforward inter-

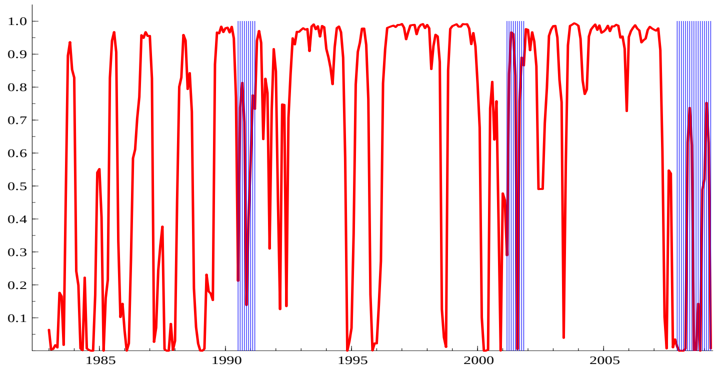
FIG. 4. Theoretical and actual spreads.



The figure plots the theoretical yield slope (solide line) and the yield slope from the DNS model estimation.

pretation for the two regimes is a low volatility regime and a high volatility regime. Though the two regimes can be labelled as low and high volatility regimes, this classification does not offer an insight on their economic interpretation. In this section, we attempt to relate the regimes to the business cycle and to a time-varying monetary policy.

FIG. 5. Regime classifications: the probabilities of being in a tranquil regime. The shaded bars indicate NBER recessions.



A visual check seems to suggest that the regime classification by and large coincides with the NBER business cycles. Specifically, the high volatility regime is more frequently observed in economic recessions. Having qualitatively related two regimes to business cycles, the next issue is to quantitatively investigate whether or not the probabilities of regimes are related to economic activity. Following Bansal and Zhou (2002) and Zhu (2013), we estimate a logit model to examine the relation between economic activity and regime classification.[10] The binary variable is defined to be one when the average monthly filtered probability of being in a low volatility regime is smaller than one-half and to be zero when the average monthly filtered probability of being in a low volatility regime is greater than or equal to one-half. Because the GDP data are not available at a monthly frequency, the explanatory proxy variable of real economic activity is the capacity utilization (U_t), which is retrieved from the Federal Reserve Bank of St. Louis. The logit model is as follows:

$$P_t(L) = \frac{\exp(\delta_0 + \delta_1 U_t)}{1 + \exp(\delta_0 + \delta_1 U_t)},$$

where $P_t(L)$ is the implied probability of being in a low volatility regime. The estimation shows that $\delta_1 = 0.07$ with a t-value 8.7. The pseudo-R^2 of the logit regression is 0.22. The result confirms the traditional wisdom that the probability of being in a low volatility regime is higher when an economy is in a boom. The logit regression thus provides supporting evidence on the business cycle interpretation of two regimes.

Two regimes are also likely to be related to monetary policy. A large body of narrative and empirical evidence (e.g., Ang, Bovin, Dong, and Loo-Kung, 2011; Li, Li, and Yu 2011) suggests that a time-varying monetary policy has important implications for the term structure of interest rates. The transmission channels include: (1) a time-varying monetary policy largely determines short-term interest rates through the Taylor (1993) rule; (2) a time-varying monetary policy affects the expected inflation and inflation risk premium. Eventually, it affects long-term interest rates; (3) a time-varying monetary policy affects expected future short-term interest rates, so it finally exerts an influence on long-term interest rates. Since a time-varying monetary policy affects the term structure of interest rates, two regimes identified from the term structure of interest rates should contain information about time-varying monetary policy.

For investigating the relation between time-varying monetary policy and regimes, we assume that the dynamic behavior of the short-term interest rate (the 1-month interest rate) follows the Taylor rule where monetary authority sets the short rate as a function of inflation (π_t) and the output gap (g_t).[11] We respectively estimate the Taylor rule in the high volatility regime and in the low volatility regime. Based on the observations in the

[10]Using the probit model does not significantly change the conclusion.

[11]Because the GDP data are now available in a monthly frequency, the output is measured by capacity utilization. The capacity utilization and consumer price index (for calculating the year-over-year inflation rate) are retrieved from the economic database, Federal Reserve Bank of St. Louis.

high volatility regime, the Taylor rule is:

$$i_{t(1)} = \underset{(0.47)}{0.14} + \underset{(11.84)}{1.06}\pi_t + \underset{(3.47)}{0.9}g_t + \epsilon_t,$$

where t-values are reported in parentheses. The Taylor rule in the low volatility regime is:

$$i_{t(2)} = \underset{(4.20)}{0.62} + \underset{(14.89)}{1.44}\pi_t + \underset{(1.32)}{0.09}g_t + \epsilon_t.$$

It is interesting to see that most coefficients are significantly different under the two regimes. The coefficient of inflation is 1.06 under the high volatility regime, while the coefficient is 1.44 in the low volatility regime. Therefore, the Federal Reserve is more (less) active in controlling inflation in the low (high) volatility regime. The coefficient of output gap is about 0.29 (0.09) under the high volatility regime (the low volatility regime). In particular, the coefficient of output gap is statistically insignificant in the low volatility regime, suggesting that the Fed is more accommodative for growth under the high volatility regime than the low volatility regime. Since the high volatility regime is related to economic recessions, the result is reasonable. The Taylor-rule estimates suggest that the Fed is proactive in controlling inflation in the low volatility regime but is accommodative for growth in the high volatility regime. The Taylor-rule analysis thus provides us supporting evidence on the monetary policy interpretation of regimes.

5. CONCLUDING REMARKS

This paper has reexamined the expectations hypothesis of the term structure of interest rates. The empirical results indicate that the expectations hypothesis cannot be rejected. The non-rejection of the expectations hypothesis is achieved through using the yield factors to capture time-varying risk premiums and taking into account regime switches. Furthermore, the regimes relate to business cycles. Our interpretation for the resurrection of the expectations hypothesis is the use of an appropriate information set and a Markov-switching VAR model as the data generating process for testing the EH.

The above results indicate a promising direction for future research. Since an appropriate information set and regime switches can account for the empirical failure of the expectations hypothesis on a posterior basis, it is interesting to investigated the role of predictable time-varying risk premiums on a prior basis. Using survey data on interest rate forecasts, Piazzesi and Schneider (2009) found that subject premiums are less volatile and not very cyclical. It seems interesting to further investigate the prior predictability of variable risk premiums.

REFERENCES

Ang, A. and G. Bekaert, 2002. Regime switches in interest rates. *Journal of Business and Economic Statistics* **20**, 163-182.

Ang, A., G. Bekaert, and M. Wei, 2008. The term structure of real rates and expected inflation. *Journal of Finance* **63**, 797-849.

Ang, A., J. Boivin, S. Dong, and R. Loo-Kung, 2011. Monetary policy shifts and the term structure. *Review of Economic Studies* **78**, 429-457.

Ang, A. and M. Piazzesi, 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* **50**, 745-787.

Bansal, R., G. Tauchen, and H. Zhou, 2004. Regime shifts, risk premiums in the term structure, and the business cycle. *Journal of Business and Economic Statistics* **22**, 396-409.

Bansal R. and H. Zhou, 2002. Term structure of interest rates with regime shifts. *Journal of Finance* **57**, 1997-2042.

Bekaert, G., S. Cho, and A. Moreno, 2004. New-Keynesian macroeconomics and the term structure. Columbia Business School, mimeo.

Bekaert, G. and R. J. Hodrick, 2001. Expectations hypothesis tests. *Journal of Finance* **56**, 115-138.

Bekaert, G., R. J. Hodrick, and D. A. Marshall, 1997. Peso problem explanations for term structure anomalies. *Journal of Monetary Economics* **48**, 241-270.

Bulkley, G. and P. Giordani, 2011. Structural breaks, parameter uncertainty, and term structure puzzles. *Journal of Financial Economics* **102**, 222-232.

Bulkley, G., R. D. F. Harris, and V. Nawosah, 2011. Revisiting the expectations hypothesis of the term structure of interest rates. *Journal of Banking and Finance* **35**, 1202-1212.

Campbell, J.Y., 1986. A defense of the traditional hypotheses about the term structure of interest rates. *Journal of Fiance* **41**, 183-193.

Campbell, J.Y. and R. J. Shiller, 1987. Cointegration and tests of present value models. *Journal of Political Economy* **95**, 1062-1088.

Campbell, J.Y. and R. J. Shiller, 1991. Yields spreads and interest rate movements: a bird's eye view. *Review of Economics Studies* **58**, 494-514.

Carriero A., C. A. Favero, and I. Kaminska, 2006. Financial factors, macroeconomic information and the expectations theory of the term structure of interest rates. *Journal of Econometrics* **131**, 339-358.

Cochrane, J.H. and M. Piazzesi, 2005. Bond risk premia. *American Economic Review* **95**, 138-160.

Cox, J., J. Ingersoll, and S. Ross, 1981. A reexamination of traditional hypotheses about the term structure of interest rates. *Journal of Finance* **3**, 321-346.

Dai, Q. and K. J. Singleton, 2002. Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics* **63**, 415-441.

Dai, Q., K. J. Singleton, and W. Yang, 2006. Are regime shifts priced in U.S. treasury markets? *Review of Financial Studies* **20**, 1669-1702.

Della Corte, P., J. Sarno, and D. L. Thornton, 2008. The expectation hypothesis of the term structure of very short-term rates: Statistical tests and economic value. *Journal of Financial Economics* **89**, 158-174.

Diebold, F.X. and C. Li, 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* **130**, 337-364

Diebold, F.X., G. D. Rudebusch, and S. B. Aruoba, 2006. The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics* **131**, 309-338

Downing, C. and S. Oliner, 2007. The term structure of commercial paper rates. *Journal of Financial Economics* **83**, 59-86.

Evans, C. L. and D. A. Marshall, 2002. Economic determinants of the nominal treasury yield curve? Federal Reserve Bank of Chicago, mimeo.

Fama, E. F. and R. R. Bliss, 1987. The information in long-maturity forward rates. *American Economic Review* **65**, 269-282.

Frankel, J. A. and K. A. Froot, 1987. Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review* **77**, 133-153.

Froot, K. A., 1989. New hope for the expectations hypothesis of the term structure of interest rates? *Journal of Finance* **44**, 283-305.

Fuhrer, J. C. 1996. Monetary policy shifts and long-term interest rates. *Quaterly Journal of Economics* **111**, 1183-1209.

Garcia, R. and P. Perron, 1996. An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics* **78**, 111-125.

Gray, S. F., 1996. Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* **42**, 27-62

Hamilton, J. D., 1988. Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control* **12**, 385-423.

Hamilton, J. D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycles. *Econometric* **57**, 357-384.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press.

Kozicki, S. and P. A. Tinsley, 2001. Shifting endpoints in the term structure of interest rates. *Journal of Monetary Economics* **47**, 613-652.

Krolzig, H. M., 1997. Markov-switching vector autoregressions: modelling, statistical inference and application to business cycle analysis. Lecture Notes in Economics and Mathematical Systems, Volume 454; Berlin: Springer.

Li, H., T. Li, and C. Yu, 2011. No-arbitrage Taylor rules with switching regimes. Working paper.

Longstaff, F. A., 2000a. Arbitrage and the expectations hypotheses. *Journal of Finance* **55**, 989-994.

Longstaff, F. A. 2000b. The term structure of very short term rates: new evidence for the expectations hypothesis. *Journal of Financial Economics* **58**, 397-415.

Piazzesi, M. and M. Schneider, 2009. Trend and cycle in bond premia. Working paper, Standford University.

Rudebusch, G. D. and T. Wu, 2008. A macro-finance model of the term structure, monetary policy, and the economy. *Economic Journal* **118**, 908-926.

Sarno L., D. L. Thornton, and G. Valente, 2007. The empirical failure of the expectations hypothesis of the term structure of bond yields. *Journal of Financial and Quantitative Analysis* **42**, 81-100

Sola, M. and J. Driffill, 1994. Testing the term structure of interest rates using a stationary vector autoregression with regime switching. *Journal of Economic Dynamics and Control* **18**, 601-628.

Shiller, R. J., 1979. The volatility of long-term interest rates and expectations models of the term structure. *Journal of Political Economy* **87**, 1190-1219.

Taylor, J. B., 1993. Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* **39**, 195–214.

Xiang, J. and X. Zhu, 2013, A regime-switching Nelson-Siegel term structure model and interest rate forecasts. *Journal of Financial Econometrics*, **11**, 522–555.

Zhu, X., 2012. Interest rate forecasts: The role of no-arbitrage restrictions, macro factors, and regime shfits. Working paper.

Zhu, X., 2014. Peso problems and term structure anomalies of repo rates. *Review of Finance*, **18**, 1183–1215.