

Expectations and the Term Structure of Interest Rates: Evidence and Implications

Robert G. King and André Kurmann

Interest rates on long-term bonds are widely viewed as important for many economic decisions, notably business plant and equipment investment expenditures and household purchases of homes and automobiles. Consequently, macroeconomists have extensively studied the *term structure of interest rates*. For monetary policy analysis this is a crucial topic, as it concerns the link between short-term interest rates, which are heavily affected by central bank decisions, and long-term rates.

The dominant explanation of the relationship between short- and long-term interest rates is the *expectations theory*, which suggests that long rates are entirely governed by the expected future path of short-term interest rates. While this theory has strong implications that have been rejected in many studies, it nonetheless seems to contain important elements of truth. Therefore, many central bankers and other practitioners of monetary policy continue to apply it as an admittedly imperfect yet useful benchmark. In this article, we work to quantify both the dimensions along which the expectations theory succeeds in describing the link between expectations and the term structure and those along which it does not, thus providing a better sense of the utility of this benchmark.

Following Sargent (1979) and Campbell and Shiller (1987), we focus on linear versions of the expectations theory and linear forecasting models of

■ The authors would like to thank Michael Dotsey, Huberto Ennis, Pierre-Daniel Sarte, and Mark Watson for helpful comments. The views expressed in this article are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System. Robert G. King: Professor of Economics, Boston University, and consultant to the Research Department of the Federal Reserve Bank of Richmond. André Kurmann: Department of Economics, University of Québec at Montréal.

future interest rate expectations. In this context, we reach five notable conclusions for the period since the Federal Reserve-Treasury Accord of March 1951.¹

First, cointegration tests confirm that the levels of both long and short interest rates are driven by a common stochastic trend. In other words, there is a permanent component that affects long and short rates equally, which accords with one of the basic predictions of the expectations theory.

Second, while changes in this stochastic trend dominate the month-to-month changes in long-term interest rates, the same changes affect the short-term rate to a much less important degree. We summarize our detailed econometric analysis with a useful rule of thumb for applied researchers: it is optimal to infer that the stochastic trend in interest rates has varied by 97 percent of any change in the long-term interest rate.² In this sense, the long-term interest rate is a good indicator of the stochastic trend in interest rates in general.³

Third, according to cointegration tests, the spread between long and short rates is not affected by the stochastic trend, which is consistent with the expectations theory. Rather, the spread is a reasonably good indicator of changes in the temporary component of short-term interest rates. Developing a similar rule of thumb, we compute that on average, a 1 percent increase in the spread indicates a 0.71 percent decrease in the temporary component of the short rate, i.e., in the difference between the current short rate and the stochastic trend.

Fourth, the expectations theory imposes important rational expectations restrictions on linear time series models in the spread and short-rate changes. Like Campbell and Shiller (1987), who pioneered testing of the expectations theory in a cointegration framework, we find that these restrictions are decisively rejected by the data. But our work strengthens this conclusion by using a longer sample period and a better testing methodology.⁴ We interpret the rejection as arising from predictable time-variations in term premia. Under the strongest form of the expectations theory, term premia should be constant and fluctuations in the spread should be entirely determined by expectations about future short-rate changes. However, our calculations indicate that—as another

¹ See Hetzel and Leach (2001) for an interesting recent account of the events surrounding the Accord.

² The sense in which this measure is optimal is discussed in more detail below, but it is based on minimizing the variance of prediction errors over our sample period of 1951 to 2001.

³ By contrast, a similar calculation indicates that changes in short-term interest rates are a much less strong indicator of changes in the stochastic trend: the comparable adjustment coefficient is 0.17 rather than 0.97. This finding is consistent with other evidence of important temporary variations in short-term interest rates, presented in this article and other studies.

⁴ We impose the cross-equation restrictions on the VAR and calculate a likelihood ratio test that compares the fit of the constrained and unconstrained VAR, while Campbell and Shiller (1987) use a Wald-type test of the restrictions on an estimated unrestricted VAR. It is now understood that Wald tests of nonlinear restrictions are sensitive to the details of how such tests are set up and suffer from much more severe small-sample bias than the method we employ here (see Bekaert and Hodrick [2001]).

rule of thumb—a 1 percent deviation of the spread from its mean signals a 0.69 percent fluctuation of the expectations component with the remainder viewed as arising from shifts in the term premia.

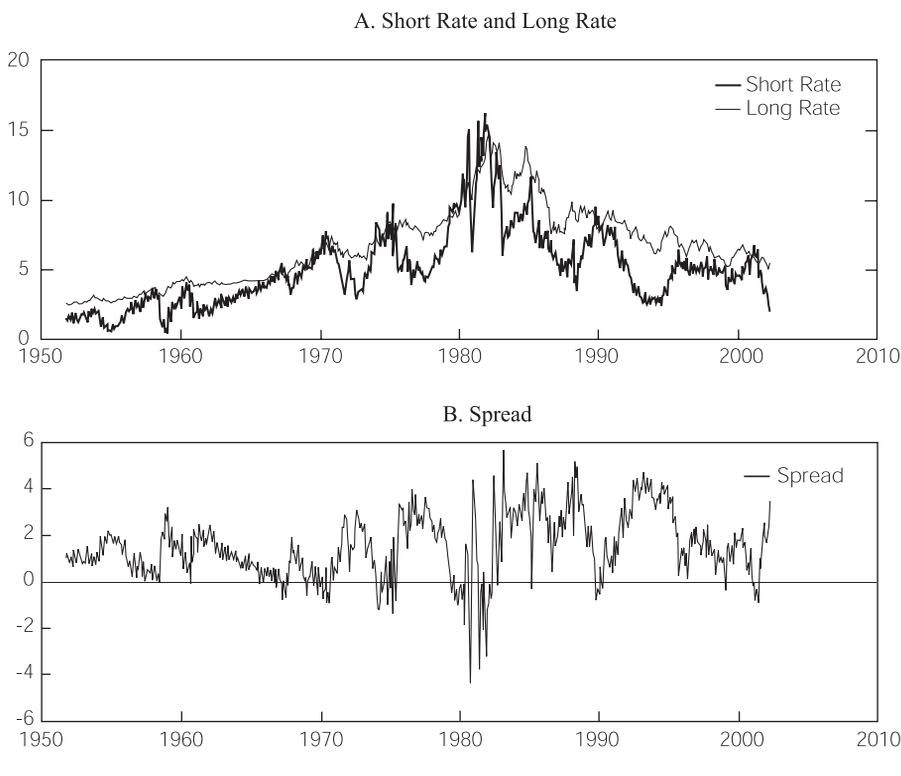
Fifth, based on the work by Sargent (1979), we show how to adapt the restrictions implied by the expectations theory to a situation where term premia are time-varying but unpredictable over some forecasting horizons. Our tests indicate that these modified restrictions continue to be rejected with forecasting horizons of up to a year. Thus, departures from the expectations theory in the form of time-varying term premia are not simply of a high frequency form, although the cointegration results indicate that the term premia are stationary.

Our empirical findings should provide some guidance for macroeconomic modeling, including work on small-scale econometric models and on monetary policy rules. In particular, our results suggest that the presence of a common stochastic trend in short and long nominal rates is a feature of post-Accord history that deserves greater attention. Furthermore, the detailed empirical results and the summary rules-of-thumb can be considered as a useful guide for monetary policy discussions. As an example, we ask whether the general patterns in the 50-year sample hold up over the period 1986–2001. Interestingly, we find a reduced variability in the interest rate stochastic trend: it is only about half as volatile as during the entire sample period. Nevertheless, the appropriate rule of thumb is still to view 85 percent of any change in the long rate as reflecting a shift in the stochastic trend. Our analysis also indicates that the expectations component of the spread (the discounted sum of expected short-rate changes) is of larger importance in the more recent sample, justifying an increase of the relevant rule-of-thumb coefficient from 69 percent to 77 percent. One interpretation of these different results is that they indicate increased credibility of the Federal Reserve System over the last decade and a half, which Goodfriend (1993) describes as the Golden Age of monetary policy because of enhanced credibility.

1. HISTORICAL BEHAVIOR OF INTEREST RATES

The historical behavior of short-term and long-term interest rates during the period April 1951 to November 2001 is shown in Figure 1. The two specific series that we employ have been compiled by Ibbotson (2002) and pertain to the 30-day T-bill yield for the short rate and the long-term yield on a bond of roughly twenty years to maturity for the long rate. One motivation for our use of this sample period is that the research of Mankiw and Miron (1986) suggests that the expectations theory encounters particular difficulties after the founding of the Federal Reserve System, particularly during the post-Accord period, because of the nonstationarity of short-term interest rates.

In this section, we start by discussing some key stylized facts that have previously attracted the attention of many researchers. We then conduct some

Figure 1 The Post-Accord History of Interest Rates

basic statistical tests on these series that provide important background to our subsequent analysis.

Basic Stylized Facts

We begin by discussing three important facts about the levels and comovement of short-term and long-term interest rates and then discuss two additional important facts about the predictability of these series.

Wandering levels: The levels of short-term and long-term interest rates vary substantially through time, as shown in Figure 1. Table 1 reports the very different average values over subsamples: in the 1950s, the short rate averaged 1.85 percent and the long rate averaged 3.02 percent; in the 1970s, the short rate averaged 6.13 percent and the long rate averaged 7.57 percent; and in the 1990s, the short rate averaged 4.80 percent and the long rate averaged 7.10 percent. These varying averages suggest that there are highly persistent factors that affect interest rates.

Table 1 Decade Averages

	Short Rate	Long Rate	Spread
1950s	1.85	3.02	1.17
1960s	3.81	4.63	0.82
1970s	6.13	7.57	1.45
1980s	8.54	10.69	2.15
1990s	4.80	7.10	2.30
Full Sample	5.13	6.67	1.57

Notes: All values are in percent per annum.

Comovement: While the levels of interest rates wander through time, subperiods of high average short rates are also periods of high average long rates. Symmetrically, short-term and long-term interest rates have a tendency to simultaneously display low average values within subperiods. This suggests that there may be common factors affecting long and short rates.

Relative stability of the spread: The spread between long- and short-term interest rates is much more stable over time, with average values of 1.17 percent, 1.45 percent, and 2.30 percent over the three decades discussed above. This again suggests that there is a common source of persistent variation in the two rates.

Predictability of the spread: While apparently returning to a more or less constant value, the spread between long and short rates appears relatively forecastable, even from its own past, because it displays substantial autocorrelation. This predictability has made the spread the focus of many empirical investigations of interest rates.

Changes in short-term and long-term interest rates: Figure 2 shows that changes in short and long rates are much less auto correlated. The two plots also highlight the changing volatility of short-term and long-term interest rates, which has been the subject of a number of recent investigations, including that of Watson (1999).

Basic Statistical Tests

The behavior of short-term and long-term interest rates displayed in Figures 1 and 2 has led many researchers to model the two series as stationary in first differences rather than in levels.

Unit root tests for interest rates: Accordingly, we begin by investigating whether there is evidence against the assumption that each series is stationary

in differences rather than in levels. For this purpose, the first two columns of Table 2 report regressions of the augmented Dickey-Fuller (ADF) form. Specifically, the regression for the short rate R_t takes the form

$$\Delta R_t = a_0 + a_1 \Delta R_{t-1} + a_2 \Delta R_{t-2} + \dots + a_p \Delta R_{t-p} + f R_{t-1} + e_{Rt}.$$

Our null hypothesis is that the short-term interest rate is difference stationary and that there is no deterministic trend in the level of the rate. In particular, stationarity in first differences implies that $f = 0$; if a deterministic trend is also absent, then $a_0 = 0$ as well. The alternative hypothesis is that the interest rate is stationary in levels ($f < 0$); in this case, a constant term is not generally zero because there is a non-zero mean to the level of the interest rate. The relevant test is reported in Table 2 for a lag length of $p = 4$.⁵ It involves a comparison of fit of the constrained regression in the first column and the unconstrained regression in the second column, with the former appropriate under the null hypothesis of a unit root and the latter appropriate under the alternative of stationarity. There is no strong evidence against the null, since the Dickey-Fuller F-statistic of 2.94 is less than the 10 percent critical value of 3.78.⁶ Looking at comparable results for the long rate R_t^L , we find even less evidence against the null hypothesis.⁷ The value of the Dickey-Fuller F-statistic is even smaller.⁸ We therefore model both interest rates as first difference stationary throughout our analysis.

In these regressions, we also find the first evidence of different predictability of short-term and long-term interest rates, a topic that will be a focus of much discussion below. Foreshadowing this discussion, we will find in every case that long-rate changes are less predictable than short-rate changes. In Table 2, the unconstrained regression for changes in the long rate accounts for about 3.5 percent of its variance, and the unconstrained regression for changes in the short rate accounts for about 8 percent of its variance.⁹

A simple cointegration test: Since we take the long-term and short-term rate as containing unit roots, the spread $S_t = R_t^L - R_t$ may either be

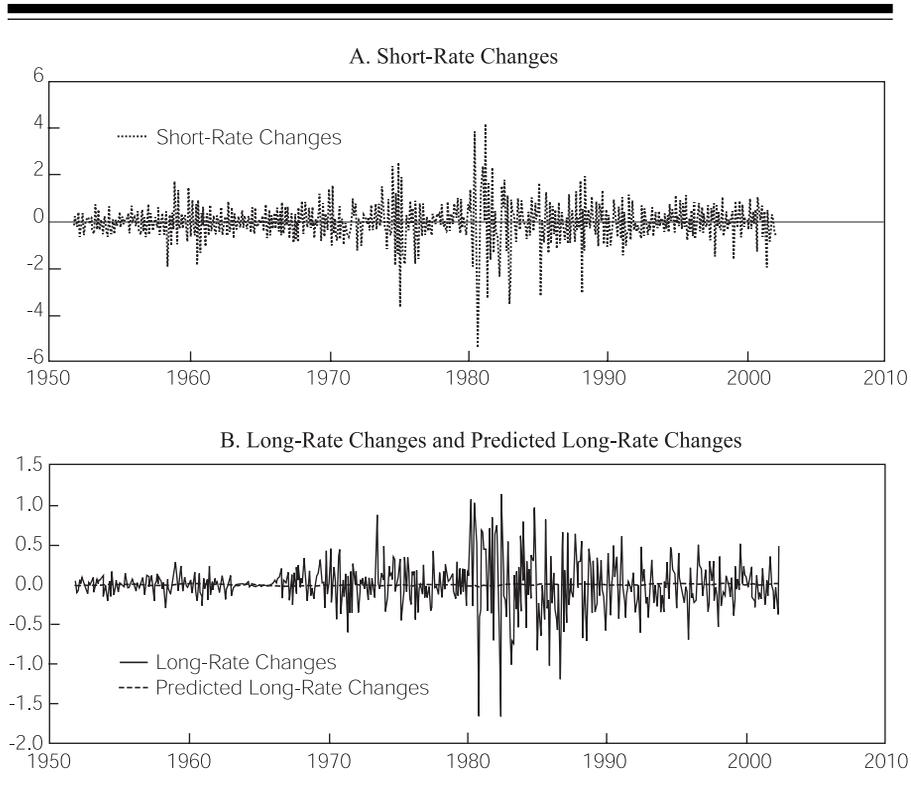
⁵ For the sake of simplicity, we use the same lag length of four months throughout the article. However, we also performed the different econometric tests with a higher lag length of $p = 6$ (as used for example by Watson [1999]) and found our results to be robust to this change.

⁶ See Dickey and Fuller (1981) for a discussion of the nonstandard distribution of this test statistic and a table of critical values.

⁷ A weaker null hypothesis, advocated for example by Hamilton (1994, 511–12), does not require $a_0 = 0$. This allows there to be a deterministic trend in the level of nominal rates, which seems implausible to us. But the second column of Table 2 also shows that there is no strong evidence against this null hypothesis, since $\hat{f} = -0.0283$ with a standard error of 0.0116. More specifically, the value of the Dickey-Fuller t-statistic is -2.43 , which is less than the 10 percent critical level of -2.57 .

⁸ The estimated level coefficient is also smaller and the associated Dickey-Fuller t-statistic takes on a value of -1.62 .

⁹ The constrained regressions display a similar pattern, although there are the familiar difficulties with interpreting R^2 when no constant term is present (see, for example, Judge et al. [1985, 30–31]).

Figure 2 The History of Interest Rate Changes

nonstationary or stationary. If the spread is stationary, then the long-term and short-term interest rates are cointegrated in the terminology of Engle and Granger (1987), since a linear combination of the variables is stationary. One simple test for cointegration when the cointegrating vector is known, discussed for example in Hamilton (1994, 582–86), is based on a Dickey-Fuller regression. In our context, we run the regression

$$\Delta S_t = a_0 + a_1 \Delta S_{t-1} + a_2 \Delta S_{t-2} + \dots + a_p \Delta S_{t-p} + f S_{t-1} + e_{St}.$$

As above, we take the null hypothesis to be that the spread is nonstationary, but that there is no deterministic trend in the level of the spread. The alternative of stationarity (cointegration) is a negative value of f ; the value of a_0 then captures the non-zero mean of the spread. The results in Table 2 show that we can reject the null at a high critical level: the value of the Dickey-Fuller F-statistic is 9.67, which exceeds the 5 percent critical level of 4.59.

Thus, we tentatively take the short-term and long-term interest rate to be cointegrated, but we will later conduct a more powerful test of cointegration. The regression results in Table 2 also highlight the fact that the spread is

Table 2 Unit Root Tests

Full Sample Estimates (1951.4–2001.11)						
	ΔR_t		ΔR_t^L		S_{t-1}	
	con.	uncon.	con.	uncon.	con.	uncon.
constant	0	0.0123 (0.0057)	0	0.0043 (0.0025)	0	0.0154 (0.0043)
lagged level	0	-0.0283 (0.0116)	0	-0.0068 (0.0042)	0	-0.1149 (0.0261)
lag 1	-0.2151 (0.0406)	-0.0198 (0.0411)	0.0896 (0.0409)	0.0918 (0.0409)	-0.3256 (0.0404)	-0.2471 (0.0437)
lag 2	-0.1649 (0.0415)	-0.1499 (0.0419)	-0.0441 (0.0407)	-0.0418 (0.0407)	-0.2610 (0.0425)	-0.1954 (0.0444)
lag 3	-0.0082 (0.0416)	0.0037 (0.0417)	-0.1390 (0.0407)	-0.1369 (0.0407)	-0.0759 (0.0425)	-0.0268 (0.0433)
lag 4	-0.1193 (0.0406)	-0.1094 (0.0407)	0.0384 (0.0409)	0.0398 (0.0409)	-0.1521 (0.0404)	-0.1157 (0.0407)
R-square	0.0721	0.0811	0.0301	0.0348	0.1322	0.1594
F-value	2.9352		1.4415		9.6688	

Notes: Numbers in parentheses represent standard errors. The critical 5 percent (10 percent) value for the Adjusted Dickey-Fuller F-test is 4.59 (3.78).

more predictable from its own past than are either of its components. In the unconstrained regression, 16 percent of month-to-month changes in the spread can be forecast from past values.

Cointegration of short-term and long-term interest rates is a formal version of the second stylized fact above: there is comovement of short and long rates despite their shifting levels. It is based on the third stylized fact: the spread appears relatively stationary although it is variable through time.

2. THE EXPECTATIONS THEORY

The dominant economic theory of the term structure of interest rates is called the expectations theory, as it stresses the role of expectations of future short-term interest rates in the determination of the prices and yields on longer-term bonds. There are a variety of statements of this theory in the literature that differ in terms of the nature of the bond which is priced and the factors that enter into pricing. We make use of a basic version of the theory developed in

Shiller (1972) and used in many subsequent studies.¹⁰ This version is suitable for empirical analyses of yields on long-term coupon bonds such as those that we study, since it delivers a simple linear formula for long-term yields. The derivation of this formula, which is reviewed in Appendix A, is based on the assumption that investors equate the expected holding period yield on long-term bonds to the short-term interest rate R_t , plus a time-varying excess holding period return k_t , which is not described or restricted by the model but could represent variation in risk premia, liquidity premia and so forth. It is based on a linear approximation to this expected holding period condition that neglects higher order terms. More specifically, the theory indicates that

$$R_t^L = \beta E_t R_{t+1}^L + (1 - \beta)(R_t + k_t), \quad (1)$$

where $\beta = 1/(1 + R^L)$ is a parameter based on the mean of the long-term interest rate around which the approximation is taken.¹¹

This expectational difference equation can be solved forward to relate the current long-term interest rate to a discounted value of current and future R and k :

$$R_t^L = (1 - \beta) \sum_{j=0}^{\infty} \beta^j [E_t R_{t+j} + E_t k_{t+j}]. \quad (2)$$

Various popular term-structure theories can be accommodated within this framework. The pure expectations theory occurs when there are no k terms, so that $R_t^L = (1 - \beta) \sum_{j=0}^{\infty} \beta^j E_t R_{t+j}$. This is a useful form for discussing various propositions about long-term and short-term interest rates that also arise in richer theories.

Implication for permanent changes in interest rates: Notably, the pure expectations theory predicts that if interest rates increase at date t in a manner which agents expect to be permanent, then there is a one-for-one effect of such a permanent increase on the level of the long rate because the weights sum to one, i.e., $(1 - \beta) \sum_{j=0}^{\infty} \beta^j = (1 - \beta)/(1 - \beta) = 1$. This is a basic and important implication of the expectations theory long stressed by analysts of the term structure and that appears capable of potentially explaining the comovement of short-term and long-term interest rates that we discussed above.

Implications for temporary changes in interest rates: Temporary changes in interest rates have a smaller effect under the pure expectations theory, with the extent of this effect depending on how sustained the temporary changes are assumed to be. Supposing that the short-term interest rate is governed by the simple autoregressive process $R_t = \rho R_{t-1} + e_{Rt}$ with the error term being

¹⁰ See, for example, Campbell, Shiller, and Schoenholz (1983) or Campbell and Shiller (1987).

¹¹ For our full sample, the average of the long rate equals 6.67 percent, or expressed as a monthly fraction: $R^L = 0.0667/12 = 0.00556$.

unforecastable, it is easy to see that $E R_{t+j} = \rho^j R_t$. It follows that a rational expectations solution for the long-term rate is

$$\begin{aligned} R_t^L &= (1 - \beta) \sum_{j=0}^{\infty} \beta^j E_t R_{t+j} \\ &= (1 - \beta) \sum_{j=0}^{\infty} \beta^j \rho^j R_t = \frac{1 - \beta}{1 - \beta \rho} R_t = \theta R_t. \end{aligned}$$

This solution can be used to derive implications for temporary changes in short rates. If these are completely transitory, so that $\rho = 0$, there is a minimal effect on the long rate, since $\theta = 1 - \beta \approx 0.005$. On the other hand, as the changes become more permanent (ρ approaches one) the θ coefficient approaches the one-for-one response previously discussed as the implication for fully permanent changes in the level of rates. Accordingly, the response of the long rate under the expectations theory depends on the degree of persistence that agents perceive in short-term interest rates, a property that Mankiw and Miron (1986) and Watson (1999) have exploited to derive interesting implications of the term structure theory that accord with various changes in the patterns of short-term and long-term interest rates in different periods of U.S. history.

The spread as an indicator of future changes: There has been much interest in the idea that the expectations theory implies that the long-short spread is an indicator of future changes in short-term interest rates. With a little bit of algebra, as in Campbell and Shiller (1987), we can rewrite (2) as

$$R_t^L - R_t = (1 - \beta) \sum_{j=0}^{\infty} \beta^j [(E_t R_{t+j} - R_t)] = \sum_{j=1}^{\infty} \beta^j E_t \Delta R_{t+j},$$

when there are no term premia.¹² Hence, the spread is high when short-term interest rates are expected to increase in the future, and it is low when they are expected to decrease. Further, permanent changes in the level of short-term interest rates, such as those considered above, have no effect on the spread because they do not imply any expected future changes in interest rates.

While these three implications can easily be derived under the pure expectations theory, they carry over to other more general theories so long as the changes in interest rates do not affect $(1 - \beta) \sum_{j=0}^{\infty} \beta^j E_t k_{t+j}$ in (2). Further, while the pure expectations theory is a useful expository device, it is simply rejected: one of the stylized facts is that long rates are generally higher than short rates (there is a positive average value to the term spread). For this reason, all empirical studies of the effects of expectations on the long rate

¹² To undertake this derivation, note that $R_{t+j} - R_t = R_{t+j} - R_{t+j-1} + \dots + (R_{t+1} - R_t)$. Hence, each expected change enters many times in the sum, with a total effect of $\sum_{h=j}^{\infty} \beta^h E_t (R_{t+h} - R_{t+h-1}) = \frac{\beta^j}{1-\beta} E_t (R_{t+j} - R_{t+j-1})$.

minimally use a modified form

$$R_t^L = (1 - \beta) \sum_{j=0}^{\infty} \beta^j E_t R_{t+j} + K,$$

where K is a parameter capturing the average value of the term spread that comes from assuming that k_t is constant.¹³

The Efficient Markets Test

As exemplified by the work of Roll (1969), one strategy is to derive testable implications of the expectations theory that (i) do not require making assumptions about the nature of the information set that market participants use to forecast future interest rates and that (ii) impose restrictions on a single linear equation. In the current setting, such an efficient markets test is based on manipulating (1) so as to isolate a pure expectations error, $R_t^L = \frac{1}{\beta} R_{t-1}^L - (\frac{1-\beta}{\beta})(R_{t-1} + K) + \xi_t$, where $\xi_t = R_t^L - E_{t-1} R_t^L$. As in Campbell and Shiller (1987, 1991), this condition may be usefully reorganized to indicate that the long-short spread (and only the spread) should forecast long-rate changes,

$$R_t^L - R_{t-1}^L = (\frac{1}{\beta} - 1)(R_{t-1}^L - R_{t-1} - K) + \xi_t,$$

which is a form that is robust to nonstationarity in the interest rate.

The essence of efficient markets tests is to determine whether any variables that are plausibly in the information set of agents at time $t - 1$ can be used to predict $\xi_t = R_t^L - R_{t-1}^L - (\frac{1}{\beta} - 1)(R_{t-1}^L - R_{t-1} - K)$. The forecasting relevance of any stationary variable can be tested with a standard t-statistic and the relevance of any group of p stationary variables can be tested by a likelihood ratio test, which has an asymptotic χ_p^2 distribution. Table 3 reports a battery of such efficient markets tests. The first regression simply is a benchmark, relating $R_t^L - R_{t-1}^L$ to a constant and to $(\frac{1}{\beta} - 1)S_{t-1}$ in the manner suggested by the efficient markets theory. The second regression frees up the coefficient on S_{t-1} and finds its estimated value to be negative rather than positive. The t-statistic for testing the hypothesis that the coefficient equals $(\frac{1}{\beta} - 1) = 0.005$ takes on a value of 2.345, which exceeds the standard 95 percent critical level. This finding has been much discussed in the context of long-term bonds and some other financial assets, in that financial markets spreads have a “wrong-way” influence on future changes relative to the predictions of basic theory.¹⁴ At the same time, the low R^2 of 0.0051 indicates that the prediction performance of the regression is very modest.

¹³ Below, we use the notation $K_t = (1 - \beta) \sum_{j=0}^{\infty} \beta^j E_t k_{t+j}$. But if $k_t = k$, then $K = k$.

¹⁴ See, for example, Campbell and Shiller (1991) for the term structure of interest rates or Bekaert and Hodrick (2001) for foreign exchange rates.

Table 3 Efficient Markets Tests

Full Sample Estimates (1951.4–2001.11)					
	test 1	test 2	ΔR_t^L test 3	test 4	test 5
constant	0.0002 (0.0010)	0.0023 (0.0015)	0.0033 (0.0015)	0.0032 (0.0016)	0.0034 (0.0016)
S_{t-1}	0.0050	-0.0147 (0.0084)	-0.0215 (0.0085)	-0.0214 (0.0095)	-0.0229 (0.0094)
ΔR_{t-1}				-0.0284 (0.0159)	-0.0147 (0.0171)
ΔR_{t-2}				-0.0321 (0.0160)	-0.0164 (0.0175)
ΔR_{t-3}				-0.0258 (0.0157)	-0.0250 (0.0168)
ΔR_{t-4}				0.0295 (0.0152)	0.0301 (0.0152)
ΔR_{t-1}^L			0.1002 (0.0410)		0.1048 (0.0409)
ΔR_{t-2}^L			-0.0496 (0.0406)		-0.0335 (0.0429)
ΔR_{t-3}^L			-0.1523 (0.0409)		-0.1328 (0.0440)
ΔR_{t-4}^L			0.0248 (0.0411)		0.0507 (0.0443)
R-square	-0.0040	0.0051	0.0408	0.0272	0.0550

Notes: Numbers in parentheses represent standard errors. F-stat (Regression 3 vs. Regression 2) = 5.598. F-stat (Regression 4 vs. Regression 2) = 3.436. F-stat (Regression 5 vs. Regression 3) = 2.280. F-stat (Regression 5 vs. Regression 4) = 4.411. The critical 5 percent (1 percent) F(4,400) value is 2.39 (3.36).

Additional evidence against the efficient markets view comes when lags of short-rate changes and lags of long-rate changes or both are added to the above equation. As regressions 3 through 5 in Table 3 show, the estimated coefficient on S_{t-1} remains significantly different from its predicted theoretical value. Furthermore, the prediction performance remains small (the R^2 is less than 10 percent for all the cases) and the F-tests reported at the bottom of the

table indicate that adding lagged variables does not significantly increase the explanatory power compared to the original efficient markets regression.¹⁵

The efficient markets regression again highlights that there is a substantial amount of unpredictable variation in changes in long bond yields, which makes it difficult to draw strong conclusions about the nature of predictable variations in these returns.¹⁶ One measure of the degree of this unpredictable variation is presented in panel B of Figure 2, where there is a very smooth and apparently quite flat line that is labelled as the “predicted changes in long rates.” Those predicted changes are $(\frac{1}{\beta} - 1)(R_{t-1}^L - R_{t-1})$ with a value of β suggested by the average level of long rates over our sample period. Panel B of Figure 2 highlights the fact that the expectations theory would explain only a tiny portion of interest rate variation if it were exactly true. Sargent (1979) refers to this as the “near-martingale property of long-term rates” under the expectations hypothesis. But it would not look very different if the fitted values of the other specifications in Table 3 were employed. Changes in the long rate are quite hard to predict and their predictable components are inconsistent with the efficient markets hypothesis.

Where Do We Go from Here?

Given that the efficient markets restriction is rejected, some academics simply conclude we know nothing about the term structure.¹⁷ However, central bankers and other practitioners actually do seem to employ the expectations theory as a useful yet admittedly imperfect device to interpret current and historical events (examples in this review are Dotsey [1998], Goodfriend [1993], and Owens and Webb [2001]). In the remainder of this analysis, we recognize that the expectations theory is not true but instead of simply rejecting it, we use modern time series methods to understand the dimensions along which it appears to succeed and those along which it does not. Section 3 develops and tests the common stochastic trend/cointegration restrictions that the expectations theory imposes. Consistent with earlier studies, we find that U.S. data

¹⁵ One potential explanation for the failure of the efficient markets tests—highlighted in Fama (1977)—is that there may be time-variation k_t in the equilibrium returns, which investors require to hold an asset. Then the theory predicts that

$$R_t^L - R_{t-1}^L = (\frac{1}{\beta} - 1)(R_{t-1}^L - R_{t-1} - k_{t-1}) + \xi_t.$$

But the researcher conducting the test does not observe time variation in k , which may give rise to a biased estimate on the spread. Fama stresses that efficient markets tests involve a joint hypothesis about the efficient use of information and a model of equilibrium returns, so that a rejection of the theory may arise from either element.

¹⁶ See the discussion of Nelson and Schwert (1977) on testing for a constant real rate.

¹⁷ For example, at a recent macroeconomics conference, one prominent monetary economist argued that the expectations theory of the term structure has been rejected so many times that it should never be built into any model.

do not allow us to reject these restrictions and, thus, that the theory appears to contain an important element of truth as far as the common stochastic trend implication is concerned. Section 4 then follows Sargent (1979) in developing and testing a variety of cross-equation restrictions that the expectations theory implies. These restrictions are rejected in the data. Finally, in Section 5, we build on the approach by Campbell and Shiller (1987) to extract estimates of changes in market expectations, which also allows us to extract estimates of time-variation in term premia.

3. COINTEGRATION AND COMMON TRENDS

A basic implication of the expectations theory is that an unexpected and permanent change in the level of short rates should have a one-for-one effect on the long rate. In other words, the theory implies that there is a common trend for the short and the long rate. This idea can be developed further using the concept of cointegration and related methods can be used to estimate the common trend.

The starting point is Campbell and Shiller's (1987) observation that present value models have cointegration implications, if the underlying series are nonstationary in levels, and that these implications survive the introduction of stationary deviations from the pure expectations theory such as time-varying term premia. In the context of the term structure, we can rewrite the long-rate equation (2) as

$$R_t^L - R_t = (1 - \beta) \sum_{j=0}^{\infty} \beta^j [(E_t R_{t+j} - R_t) + E_t k_{t+j}] \quad (3)$$

$$= \sum_{j=1}^{\infty} \beta^j E_t \Delta R_{t+j} + (1 - \beta) \sum_{j=0}^{\infty} \beta^j E_t k_{t+j} \quad (4)$$

so that the expectations theory stipulates that the spread is stationary so long as (i) first differences of short rates are stationary and (ii) the expected deviations from the pure expectations theory are stationary. Thus, cointegration tests are one way of assessing this implication of the theory.

In Section 1, we found evidence against the hypothesis that the spread contains a unit root and suggested that a stationary spread was a better description of the U.S. data. That is, we found some initial evidence consistent with modeling the short rate and the long rate as cointegrated. Here, in Section 3, we confirm that the spread also passes a more rigorous cointegration test. Given this result, we then define and estimate the common stochastic trend for the short rate and the long rate. We also present an easy-to-use rule of thumb that decomposes fluctuations of the short and the long rate into fluctuations in the common trend and fluctuations in the temporary components.

Testing for Cointegration

To develop the intuition behind the more rigorous cointegration tests, consider a vector autoregression (VAR) in the first difference of the short rate and the first difference of the long rate:

$$\Delta R_t = \sum_{j=1}^p a_j \Delta R_{t-j}^L + \sum_{j=1}^p b_j \Delta R_{t-j} + e_{Rt}, \quad (5)$$

$$\Delta R_t^L = \sum_{j=1}^p c_j \Delta R_{t-j}^L + \sum_{j=1}^p d_j \Delta R_{t-j} + e_{Lt}. \quad (6)$$

By virtue of the Wold decomposition theorem, we may be tempted to believe that such a VAR in first differences can approximate the dynamics of short- and long-rate changes arbitrarily well, so long as the vector $\Delta x_t = [\Delta R_t \ \Delta R_t^L]$ is a stationary stochastic process (this last condition being asserted by the Dickey-Fuller tests of the last section). However, if the two variables R_t and R_t^L are also cointegrated, then this argument breaks down. The above VAR represents a poor approximation in such circumstances because the short and long rate only contain one common stochastic trend and first differencing both variables thus deletes useful information.¹⁸

However, as Engle and Granger (1987) demonstrate, if first differences of x_t are stationary and there is cointegration among the variables of the form αx_t , then there always exists an empirical specification relating Δx_t , its lags Δx_{t-p} , and αx_{t-1} that describes the dynamics of Δx_t arbitrarily well. Such a system of equations is called a vector error correction model (VECM). In our context, if R_t and R_t^L are cointegrated, as under the weak form of the expectation theory discussed above, then the following VECM should provide a better description of the dynamics of Δx_t than the VAR in (5) and (6):

$$\Delta R_t = \sum_{j=1}^p a_j \Delta R_{t-j}^L + \sum_{j=1}^p b_j \Delta R_{t-j} + f[S_{t-1} - K] + e_{Rt}, \quad (7)$$

$$\Delta R_t^L = \sum_{j=1}^p c_j \Delta R_{t-j}^L + \sum_{j=1}^p d_j \Delta R_{t-j} + g[S_{t-1} - K] + e_{Lt}. \quad (8)$$

In these equations, f and g capture the effects of the lagged spread on forecastable variations in the short and long rates; K is the mean value of the spread.

¹⁸ In more technical terms, when R_t and R_t^L are cointegrated, then the vector moving average representation of $\Delta x_t = [\Delta R_t \ \Delta R_t^L]$ (which exists by definition of the Wold decomposition theorem) is noninvertible. As a result, no corresponding finite-order VAR approximation can exist. See Hamilton (1994, 574–75) for details.

To test for cointegration, we estimate both the VAR and the VECM and compare their respective fit. A substantial increase in the log likelihood of the VECM over the VAR signals that the cointegration terms aid in the prediction of interest rate changes. More specifically, a large likelihood ratio results in a rejection of the null hypothesis in favor of the alternative of cointegration. In particular, we follow the testing procedure by Horvath and Watson (1995) and assume *a priori* that the cointegrating relationship is given by the spread $S_t = R_t^L - R_t$ rather than estimating the cointegrating vector.¹⁹ Table 4 reports estimates of the VAR and VECM models for the lag length of $p = 4$, which we choose as the reference lag length throughout. Before discussing the cointegration test results in detail, it is worthwhile looking at a few elements that the VAR and VECM regressions have in common. First, changes in short rates are somewhat predictable from past changes in short rates, as was previously found with the Dickey-Fuller regression in Table 2. In addition, past changes in long rates are important for predicting changes in short rates in both the VAR and the VECM.²⁰ Finally, changes in short rates are predicted by the lagged spread: if the long rate is above the short rate, then short rates are predicted to rise. Second, changes in long rates are still fairly hard to predict with either the VAR or the VECM.

Moving to the cointegration test, the likelihood ratio between the VECM and the VAR equals $2 * (L_{VECM} - L_{VAR}) = 27.67$, which exceeds the 5 percent critical level of 6.28 calculated by the methods of Horvath and Watson (1995).²¹ In other words, we can comfortably reject the hypothesis of no cointegration between R_t and R_t^L , which is consistent with earlier studies

¹⁹ This type of test is somewhat more powerful than the unit root test on the spread reported in Table 2, which may be revealed by taking the difference between the two VECM equations and reorganizing the results slightly to obtain

$$\Delta S_t = \sum_{j=1}^p (c_j - a_j) \Delta R_{t-j}^L + \sum_{j=1}^p (d_j - b_j) \Delta R_{t-j} + (g - f) S_{t-1} + (e_{Lt} - e_{Rt}),$$

which can be further rewritten as

$$\Delta S_t = \sum_{j=1}^p (c_j - a_j) \Delta S_{t-j} + \sum_{j=1}^p (c_j + d_j - a_j - b_j) \Delta R_{t-j} + (g - f) S_{t-1} + (e_{Lt} - e_{Rt}).$$

That is, the Horvath-Watson test essentially introduces some additional stationary regressors to the forecasting equation for changes in the spread that was used in the DF test. Adding these regressors can improve the explanatory power of the regression, resulting in a more powerful test.

²⁰ That is, long rates Granger-cause short rates.

²¹ As Horvath and Watson (1995) stress, the relevant critical values for the likelihood ratio must take into account that the spread is nonstationary under the null. Thus, we cannot refer to a standard chi-square table. We estimate the VAR and VECM without constant terms, since we are assuming no deterministic trends in interest rates. However, we allow for a mean value of the spread, which is not zero as shown in (7) and (8). Unfortunately, this combination of assumptions means that we cannot use the tables in Horvath and Watson (1995), but must conduct the Monte Carlo simulations their method suggests to calculate the critical values reported in the text. Details are contained in replication materials available at <http://people.bu.edu/rking>.

Table 4 VAR/VECM Estimates

Full Sample Estimates (1951.4–2001.11)					
	ΔR_t		ΔR_t^L		
	VAR	VECM	VAR	VECM	
S_{t-1}		0.1101 (0.0237)		–0.0229 (0.0094)	
ΔR_{t-1}	–0.3095 (0.0408)	–0.2382 (0.0430)	0.0001 (0.0160)	–0.0147 (0.0171)	
ΔR_{t-2}	–0.1997 (0.0427)	–0.1393 (0.0440)	–0.0038 (0.0168)	–0.0164 (0.0175)	
ΔR_{t-3}	–0.0051 (0.0417)	0.0426 (0.0423)	–0.0151 (0.0164)	–0.0250 (0.0168)	
ΔR_{t-4}	–0.0879 (0.0377)	–0.0466 (0.0382)	0.0387 (0.0148)	0.0301 (0.0152)	
ΔR_{t-1}^L	0.8712 (0.1038)	0.8209 (0.1026)	0.0943 (0.0408)	0.1048 (0.0409)	
ΔR_{t-2}^L	0.6250 (0.1095)	0.5954 (0.1078)	–0.0397 (0.0430)	–0.0335 (0.0429)	
ΔR_{t-3}^L	0.1791 (0.1123)	0.1698 (0.1104)	–0.1347 (0.0441)	–0.1328 (0.0440)	
ΔR_{t-4}^L	0.0430 (0.1133)	0.0520 (0.1114)	0.0526 (0.0445)	0.0507 (0.0443)	
R-square	0.2220	0.2492	0.0459	0.0553	
F-statistic	21.2172	21.9030	3.5785	3.8610	

Notes: Numbers in parentheses represent standard errors. The likelihood ratio statistic of the VECM against the VAR is 27.6704. Comparing this value to the corresponding critical value in Horvath and Watson's tables leads to strong rejection of null of two unit roots (p-value higher than 0.01).

and reinforces the statistical support for the common trend implication of the expectations theory. Therefore, the data is consistent with the basic implication of cointegration of the expectations theory and we thus view the VECM as the preferred specification and assume cointegration for the remainder of our analysis.²²

²² An alternative approach in this section would be to estimate the cointegrating vector and use the well-known testing method of Johansen (1988). Horvath and Watson (1995) establish that their procedure is more powerful if the cointegrating vector is known.

Uncovering the Common Stochastic Trend

A key implication of cointegration in our context is that the short and long rates share a common stochastic trend, which we will now work to uncover.²³ Following Beveridge and Nelson (1981), the stochastic trend of a single series such as the short-term interest rate is defined as the limit forecast $\bar{R}_t = \lim_{k \rightarrow \infty} E_t R_{t+k}$, or equivalently

$$\bar{R}_t = R_{t-1} + \lim_{k \rightarrow \infty} \sum_{j=0}^k E_t \Delta R_{t+k}. \quad (9)$$

However, in order to obtain a series of \bar{R}_t , we need to take a stand on how to compute the $E_t \Delta R_{t+k}$ terms. The VECM suggests a straight-forward way to do so. Specifically, suppose that the system expressed by equations (7) and (8) is written in the form

$$z_t \equiv \begin{bmatrix} \Delta R_t \\ \Delta R_t^L \\ S_t \end{bmatrix} = H x_t$$

$$x_t = M x_{t-1} + G e_t,$$

where e_t is the vector of one-step-ahead forecast errors $e_t = [e_{Rt} \ e_{Lt}]'$ and $x_t = [\Delta R_t \ \Delta R_t^L \ \Delta R_{t-1} \ \Delta R_{t-1}^L \ \dots \ \Delta R_{t-(p-1)} \ \Delta R_{t-(p-1)}^L \ S_t]$ is the vector of information that the VECM identifies as useful for forecasting future spreads and interest rate changes. The matrix H simply selects the elements of x_t , and the elements of M and G depend on the parameter estimates $\{a, b, c, d, f, g\}$ in a manner spelled out in Appendix B.

Given this setup, forecasts of ΔR_{t+k} conditional information on x_t are easily computed as

$$E[\Delta R_{t+k} | x_t] = h_R E[z_{t+k} | x_t] = h_R H E[x_{t+k} | x_t] = h_R H M^k x_t,$$

where $h_R = [1 \ 0 \ 0]$ such that $\Delta R_t = h_R z_t$. Mapping these forecasts of ΔR_{t+k} into (9), we obtain a closed-form solution for the stochastic trend of the short rate:

$$\bar{R}_t = R_{t-1} + \sum_{k=0}^{\infty} h_R H M^k x_t = R_{t-1} + h_R H [I - M]^{-1} x_t.$$

The same procedure for computing multiperiod forecasts also provides a recipe for computing the stochastic trend in the long rate, that is,

$$\bar{R}_t^L = R_{t-1}^L + \lim_{k \rightarrow \infty} \sum_{j=0}^k E_t \Delta R_{t+k}^L$$

²³ The idea that cointegration implies common stochastic trends is developed in Stock and Watson (1988) and King, Plosser, Stock, and Watson (1991).

$$= R_{t-1}^L + \sum_{k=0}^{\infty} h_L H M^k x_t = R_{t-1}^L + h_L H [I - M]^{-1} x_t,$$

where $h_L = [0 \ 1 \ 0]$ such that $\Delta R_t^L = h_L z_t$. Finally, the difference between \bar{R}_t^L and \bar{R}_t is the limit forecast of the spread. By definition of cointegration, the spread is stationary and therefore its limit forecast must be a constant:²⁴

$$K = \lim_{k \rightarrow \infty} E_t S_{t+k} = \lim_{k \rightarrow \infty} E_t R_{t+k}^L - \lim_{k \rightarrow \infty} E_t R_{t+k} = \bar{R}_t^L - \bar{R}_t.$$

Thus, the trends for the long rate and the short rate differ only by the constant K : in other words, the long rate and the short rate have a common stochastic trend component. Since this is sometimes termed the permanent component, deviations from it are described as temporary components. Using this language, the temporary component of the short rate is $R_t - \bar{R}_t$ and that of the long rate is $R_t^L - \bar{R}_t^L$.

A Stochastic Trend Estimate: 1951–2001

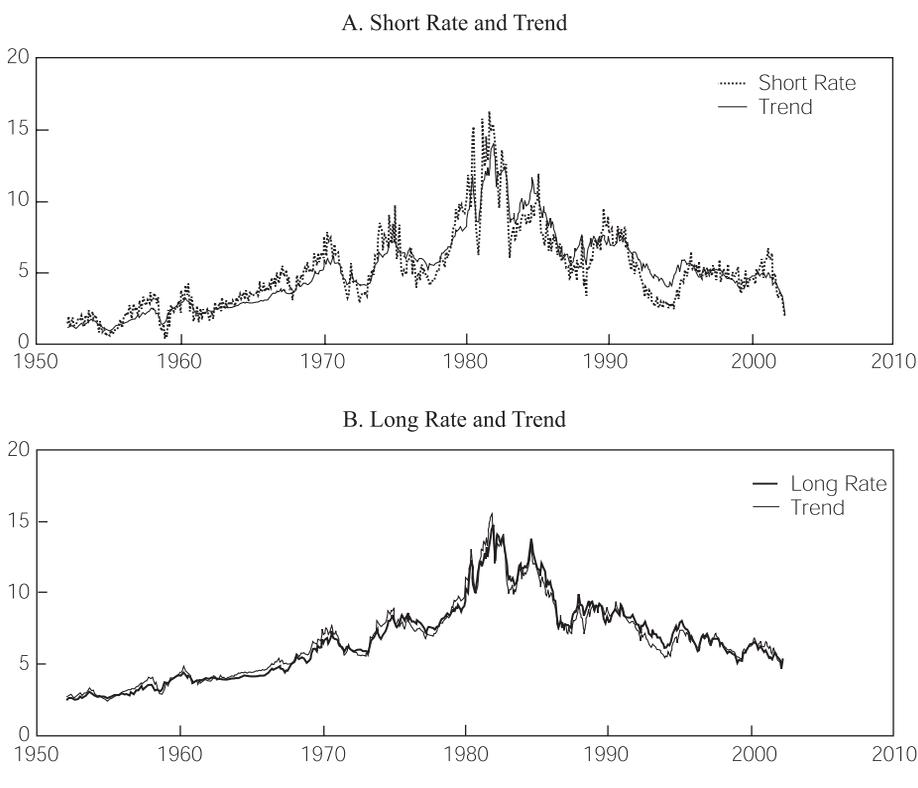
Figure 3 shows the common stochastic trend in long and short rates based on the VECM from Table 3, constructed using the method that we just discussed. In line with the expectations theory, we interpret this stochastic trend as describing permanent changes in the level of the short rate, which are reflected one-for-one in the long rate.

Short rates and the stochastic trend: In panel A, we see that the short rate fluctuates around its stochastic trend. There are some lengthy periods, such as the mid-1960s, where the short rate is above the stochastic trend for a lengthy period and others, such as the mid-1990s, where the short rate is below the stochastic trend. The vertical distance is a measure of the temporary component to short rates, which we will discuss in greater detail further below.

Long rates and the stochastic trend: In panel B, we see that the long rate and the stochastic trend correspond considerably more closely. This result accords with a very basic implication of the expectations theory: long rates should be highly responsive to permanent variations in the short-term interest rate.²⁵

²⁴ Under the expectations theory with a constant term premium, the average value of the spread must be the term premium K . So, to avoid proliferation of symbols, we use that notation here.

²⁵ To understand the sensitivity of the trend to the form of the estimated equation for the long rate, we compared three alternative measures of the trend. The first was the test measure based on the estimated VECM (i.e., the one reported in this section); the second was based on replacing the long-rate equation with the result of a simple regression of long-rate changes on the spread (i.e., the specification that we used for testing the efficient markets restriction above) so that there was a small negative weight on the spread in the long-rate equation; and the third

Figure 3 Interest Rates and the Common Stochastic Trend

Variance Decompositions

It is useful to consider a decomposition of the variance of short-rate and long-rate changes into contributions in terms of changes in the temporary and permanent components. For the short-rate changes, since $var(\Delta R_t) = var(\Delta \bar{R}_t + \Delta(R_t - \bar{R}_t))$, this decomposition takes the form

$$\begin{aligned} var(\Delta R_t) &= var(\Delta \bar{R}_t) + var(\Delta(R_t - \bar{R}_t)) \\ &\quad + 2 * cov(\Delta \bar{R}_t, \Delta(R_t - \bar{R}_t)) \\ 0.656 &= 0.105 + 0.544 + 2 * (0.004) \end{aligned}$$

was based on the efficient markets restriction (i.e., placed a small positive weight on the lagged spread). While there were some differences in these trend estimates on a period-by-period basis, they tell the same basic story in terms of the general pattern of rise and fall in the stochastic trend.

Table 5 Summary Statistics for Permanent-Temporary Decomposition

Full Sample Estimates (1951.4–2001.11)		
A. Short-Rate Changes		
Total	Permanent	Temporary
0.6559	0.1083	0.5477
0.4133	0.1046	0.0036
0.9168	0.0152	0.5440
B. Long-Rate Changes		
Total	Permanent	Temporary
0.0826	0.0802	0.0023
0.8631	0.1046	−0.0244
0.0499	−0.4614	0.0268
C. Long-Short Spread		
Total	Temporary Long Rate	Temporary Short Rate
1.9318	0.5649	−1.3668
−0.9920	−0.3841	0.9827
0.9559	0.1808	−0.9114

Notes: Table 5 is based on the VECM estimates in Table 4. Each panel contains a 3 by 3 matrix. On the diagonal, variances are reported (e.g., the variance of changes in long rates is 0.0826). Above the diagonal, covariances are listed (e.g., the covariance between changes in the long rate and changes in its permanent component is 0.0802). Below the diagonal, the corresponding correlation is reported (e.g., the correlation between changes in the long rate and changes in its permanent component is 0.8631).

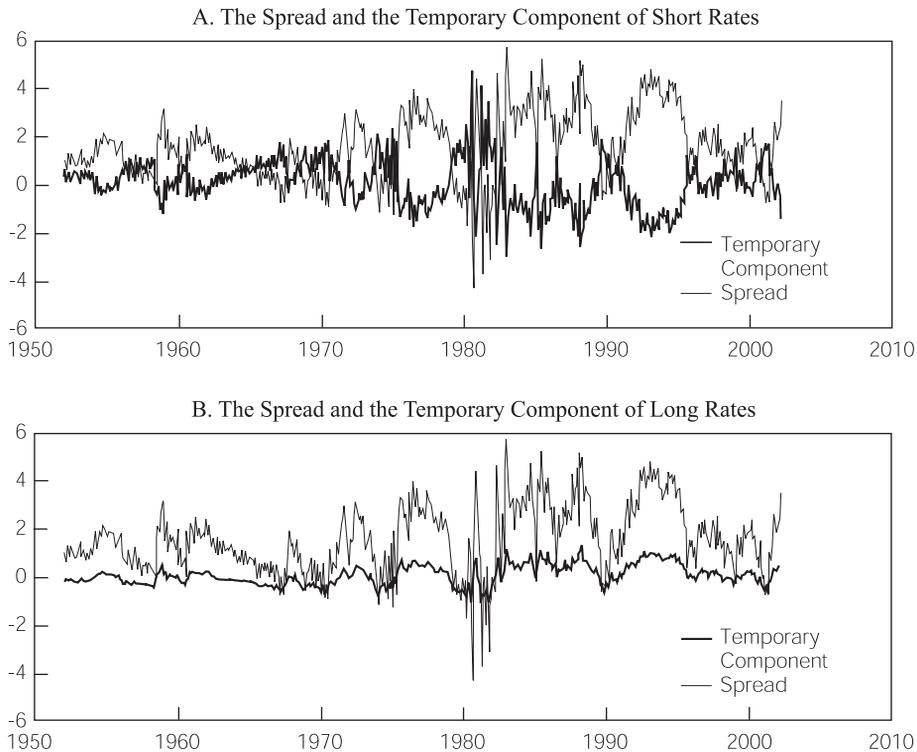
with the last line drawn from the first panel of Table 5.²⁶

The variance of month-to-month changes in interest rates is 0.66. Changes in the temporary component account for the great bulk (82.9 percent) of this variance, while the variance of changes in the permanent component contributes 15.9 percent and the covariance between the two components contributes only about 1.2 percent.

For the long rate, the decomposition takes conceptually the same form, but we find a very different result in terms of relative contributions:

$$\begin{aligned} \text{var}(\Delta R_t^L) &= \text{var}(\Delta \bar{R}_t^L) + \text{var}(\Delta(R_t^L - \bar{R}_t^L)) \\ &\quad + 2 * \text{cov}(\Delta \bar{R}_t^L, \Delta(R_t^L - \bar{R}_t^L)) \end{aligned}$$

²⁶ Here and below, our estimate of the stochastic trend allows us to calculate the variance decomposition, including the variance of changes in the trend and the covariance term. Note that due to rounding errors, the variance decompositions do not add up exactly.

Figure 4 The Spread and Temporary Components

$$0.083 = 0.104 + 0.027 + 2 * (-0.024)$$

First, the overall variance of month-to-month changes in the long rate is much smaller. In contrast to the short rate, this variance is dominated by the variance in its permanent component, which is actually somewhat larger because there is a negative correlation between the permanent and the transitory component.

The permanent-temporary decomposition also permits us to undertake a decomposition of the long-short spread, which is displayed in Figure 4. The spread and the two temporary components are connected via the identity

$$S_t - S = R_t^L - R_t - S = (R_t^L - \bar{R}_t^L) - (R_t - \bar{R}_t).$$

Hence, there is a mechanical inverse relationship between the spread and the temporary component of the short rate, which is clearly evident in panel A of Figure 4: everything else equal, whenever the short-term rate is high relative to its permanent component, the spread is low on this account. We can undertake

a similar decomposition of the variance of the spread to those used above,

$$\begin{aligned} \text{var}(S_t) &= \text{var}(R_t^L - \bar{R}_t^L) + \text{var}(R_t - \bar{R}) \\ &\quad - 2 * \text{cov}((R_t^L - \bar{R}_t^L), (R_t - \bar{R})) \\ 1.93 &= 0.18 + 0.98 - 2 * (-0.38). \end{aligned}$$

According to this expression, there is a variance of 1.93. Of this, 51 percent is attributable to the variability of the temporary component of the short rate, 9 percent is attributable to the temporary component of the long rate, and a substantial amount (39 percent) is attributable to the covariance between these two expressions.²⁷

Simple Rules of Thumb

Suppose that we observe just the change in the long rate and want to know how much of a change has taken place in the permanent component. Our variance decompositions let us provide an answer to this and related questions below. Specifically, we derive a simple rule of thumb as follows. First, define the change in the permanent component as an unobserved zero-mean variable Y_t . This variable is known to be connected to the observed zero-mean variables ΔR_t^L according to the identity $Y_t = \Delta R_t^L + U_t$, where U_t is an error. Then we can ask the question: What is the optimal linear estimate of Y_t given the observed series ΔR_t^L ? To calculate this measure, $\hat{Y}_t = b\Delta R_t^L$, we minimize the expected squared errors, $\text{var}(Y_t - \hat{Y}_t) = \text{var}(Y_t) + b^2\text{var}(\Delta R_t^L) - 2bcov(Y_t, \Delta R_t^L)$. The optimal value of b is the familiar OLS regression coefficient

$$b = \frac{\text{cov}(Y_t, \Delta R_t^L)}{\text{var}(\Delta R_t^L)}.$$

Using our estimates of the common stochastic trend, we compute that the variance of long-rate changes is 0.0826 and that the covariance of long-rate and permanent component changes is 0.0802 (see second panel of Table 5). Thus, the coefficient b takes on a value of 0.97, which leads to the following rule of thumb.

Long-rate rule of thumb: *If a 1 percent rise (fall) in the long rate occurs, then our calculations suggest that an observer should increase (decrease) his or her estimate of the permanent component by 97 percent of this rise (fall).*²⁸

²⁷ There is also substantial serial correlation in the spread, as well as in the temporary components of the short rate and the long rate. The first order autocorrelations of these series are, respectively, 0.81, 0.72, and 0.93.

²⁸ Of course, we could have devised a similar rule of thumb for the short rate by replacing ΔR_t^L by ΔR_t in the formula for the coefficient b . The result would have been a much more

A similar rule of thumb can be derived by linking changes in the unobserved temporary component of the short rate ($R_t - \bar{R}_t$) to the spread.²⁹

Spread rule of thumb #1: *If the spread exceeds its mean by 1 percent, then our estimates suggest that the temporary component of short-term interest rates is low by -0.71 percent ($-0.71 = (-1.37)/(1.93)$).*

Our two rules of thumb indicate that changes in the long rate are dominated by changes in the permanent component and the level of the spread (relative to its mean) is substantially influenced by the temporary component.

4. RATIONAL EXPECTATIONS TESTS

A hallmark of rational expectations models of the term structure, stressed by Sargent (1979), is that they impose testable cross-equation restrictions on linear time series models. In this section, we describe the strategy behind rational expectations tests along the lines of Sargent (1979) and Campbell and Shiller (1987); we also discuss how to extend the tests to accommodate time-varying term premia. We then implement these tests and find that there is a broad rejection of the rational expectations restrictions that we trace to divergent forecastability of the spread and changes in short-term interest rates.

A Simple Reference Model

To illustrate the nature of the cross-equation restrictions that the expectations theory imposes and to motivate the ensuing discussion of rational expectations tests, consider the following simple model. Suppose that the short-term interest rate is governed by

$$R_t = \tau_t + x_t,$$

where τ_t is a relatively persistent permanent component that we model as a unit root process and x_t is a relatively less persistent temporary component. In addition, suppose that agents observe τ_t and x_t separately and also understand that these evolve according to

$$\begin{aligned}\tau_t &= \tau_{t-1} + e_{\tau,t} \\ x_t &= \rho x_{t-1} + e_{x,t},\end{aligned}$$

with $-1 < \rho < 1$ and with $e_{\tau,t}, e_{x,t}$ being white noises. Suppose also that the expectations theory holds true. Using equation (2) and setting (1 –

modest rule of thumb coefficient ($0.1651 = 0.1083/0.6559$). This smaller coefficient reflects the fact that temporary variations are much more important for the short rate.

²⁹ For this purpose, we interpret Y_t as the change in the temporary component of the short rate and replace ΔR_t^L with the spread (less its mean) in the above formula for b . Based on the third panel of Table 5, the covariance between changes in the temporary component of the short rate and the spread equals -1.37 and the variance for changes in the spread is 1.93 .

$\beta) \sum_{j=0}^{\infty} \beta^j E_t k_{t+j} = K = 0$ for all t , the dynamics of the long rate can thus be described as³⁰

$$\begin{aligned} R_t^L &= (1 - \beta) \sum_{j=0}^{\infty} \beta^j E_t R_{t+j} \\ &= (1 - \beta) \sum_{j=0}^{\infty} \beta^j E_t [\tau_{t+j} + x_{t+j}] = \tau_t + \theta x_t, \end{aligned} \tag{10}$$

where $\theta = (1 - \beta)/(1 - \beta\rho) < 1$ since $\rho < 1$ as in Section 2 above. Finally, notice that the spread by definition takes the form

$$S_t = R_t^L - R_t = (\theta - 1)x_t,$$

which implies that under the expectations theory, the spread is a perfect negative indicator of the temporary component of short-term interest rates.

Cross-equation restrictions on a stationary VAR system: By assuming a unit root component τ_t in the short rate and the expectations theory being true, we determined above that both the short rate and the long rate in our reference model are stationary in first differences rather than levels. We therefore follow Campbell and Shiller (1987) and study the bivariate system in short-rate changes,

$$\begin{aligned} \Delta R_t &= \Delta \tau_t + \Delta x_t = e_{\tau,t} + e_{x,t} + (\rho - 1)x_{t-1} \\ &= e_{\tau,t} + e_{x,t} + \frac{\rho - 1}{\theta - 1} S_{t-1} = e_{\tau,t} + e_{x,t} + \frac{1 - \beta\rho}{\beta} S_{t-1}, \end{aligned}$$

and in the spread,

$$S_t = (\theta - 1)x_t = \rho S_{t-1} + (\theta - 1)e_{x,t}.$$

Both of these variables are stationary, which has the advantage that testable restrictions are easier to develop in the presence of time-varying, but stationary, term premia.³¹

As stressed by Sargent (1979), the expectations theory imposes cross-equation restrictions. In the case of ΔR_t and S_t , these restrictions become immediately apparent when we compare the two model equations above to an unrestricted bivariate, first order vector autoregression:

$$\begin{aligned} \Delta R_t &= a \Delta R_{t-1} + b S_{t-1} + e_{\Delta R,t}. \\ S_t &= c \Delta R_{t-1} + d S_{t-1} + e_{S,t}. \end{aligned}$$

³⁰ According to the expectations theory, K does not have to equal zero. For the sake of convenience, we set $K = 0$, which can be reconciled with the data if we consider all variables as deviations from their respective means.

³¹ Such a stationary system is sometimes called a VECM in Phillips's triangular form. See Hamilton (1994, 576-78) and Appendix C.

In particular, we see that the expectations theory imposes $a = c = 0$, $b = (1 - \beta\rho)/\beta$, $d = \rho$, and $e_{\Delta R,t} = e_{\tau,t} + e_{x,t}$, $e_{S,t} = (\theta - 1)e_{x,t}$.³² In our econometric analysis below, we will focus on deriving and testing similar restrictions for a more general rational expectations framework that contains the assumption of agents having more information than the econometrician.³³

Restrictions on a VAR Model

For the purpose of testing the cross-equation restrictions in the data, we adopt a general strategy initially put forth by Sargent (1979). Following Campbell and Shiller (1987), we consider a bivariate VAR in the short-rate change and the spread:³⁴

$$\Delta R_t = \sum_{i=1}^p a_i \Delta R_{t-i} + \sum_{i=1}^p b_i S_{t-i} + e_{\Delta R,t}. \quad (11)$$

$$S_t = \sum_{i=1}^p c_i \Delta R_{t-i} + \sum_{i=1}^p d_i S_{t-i} + e_{S,t}. \quad (12)$$

In this section, we work under the assumption that the expectations theory is exactly true, which we relax later. Under this condition, term premia are

³² VECM regressions like (7) and (8) in the previous section are also restricted by the expectations theory. According to our simple model, the dynamics of short- and long-rate changes take the form

$$\begin{aligned} \Delta R_t &= e_{P_t} + e_{T_t} + \frac{1 - \beta\rho}{\beta} S_{t-1}, \\ \Delta R_{L,t} &= \Delta \tau_t + \theta \Delta x_t = e_{\tau,t} + \theta e_{x,t} + \frac{1 - \beta}{\beta} S_{t-1}. \end{aligned}$$

The second equation for the long-rate change is simply the efficient markets restriction.

³³ In our simple model, the VECM approach (discussed in the previous footnote) helps to correctly uncover some features of the data that are not known *a priori* by the econometrician. First, the temporary component x_t of the short rate is reflected in a temporary component of the long rate, but with a much dampened magnitude for plausible values of β and ρ . For example, if $1/\beta = 1.005$ and $\rho = 0.8$, then the composite coefficient θ takes on a value of $0.005/0.025 = 0.2$. Second, the spread is predicted to be a significant predictive variable for interest rates in the VECM, but especially for the temporary component of interest rates. These features of the model appear broadly in accord with the estimated VECM and its outputs, particularly in terms of the implication that there is a much smaller volatility of the temporary component of the long rate than the temporary component of the short rate. In addition, the generally poor predictive performance for changes in the long rate seems consistent with the importance of permanent shocks in that equation, relative to the small effect of the spread. Finally, the spread and the temporary component of the short-term interest rate are negatively associated in the example as in the outputs of the VECM. But other features of the model are at variance with the results obtained via estimating a VECM. In particular, the temporary component of the long rate has a strong positive association with the temporary component of the short rate in the model, while there is a negative correlation in the estimates discussed in the preceding section.

³⁴ The example we discussed above used one lag for analytical convenience, but in this empirical context we use multiple lags to capture the dynamic interactions between the variables more completely.

constant and the expression for the spread in (4) reduces to³⁵

$$S_t = \sum_{j=1}^{\infty} \beta^j E_t \Delta R_{t+j}, \quad (13)$$

as we saw in Section 3 above. This expression is important for two reasons. First, it says that according to the expectations theory the spread is simply the discounted sum of future expected short-rate changes. Second, in terms of econometrics, it reveals that as long as short rates are stationary in first differences, the spread must be stationary as well.

The derivation of testable restrictions that (13) imposes on (11) and (12) has four key ingredients. First, the *law of iterated expectations* implies that for any information set ω_t which is a subset of the market's information set Ω_t ,

$$E[E_t \Delta R_{t+j} | \omega_t] = E[E \Delta R_{t+j} | \Omega_t] | \omega_t = E[\Delta R_{t+j} | \omega_t].$$

Practically, this says that an econometrician's best estimate of market expectations of future short-rate changes, given a data set ω_t , is equal to the econometrician's forecast of these short-rate changes given his or her data. Thus, under the assumption that the expectations theory is exactly true and using the fact that the current spread is in the information set, we can rewrite (13) as

$$S_t = \sum_{j=1}^{\infty} \beta^j E[\Delta R_{t+j} | \omega_t]$$

so that the spread formula is unchanged when the information set is reduced.³⁶

Second, the *Wold decomposition theorem* guarantees that if ΔR_t is stationary, it can be well described by a vector autoregression (possibly of infinite order p) where the explanatory variables are composed of information Ω_{t-1} available to the market at date $t - 1$.

³⁵ Note that we have dropped the constant K from the equation for the sake of notational simplicity. In econometric terms, this simply means that, without a loss of generality, we have to test the expectations theory with demeaned data.

³⁶ As Campbell and Shiller (1987) stress, the explanation for this result is subtle: the expectations theory says that the spread is simply the discounted sum of future expected short-rate changes. Under the null that the theory is true, all the relevant information that market participants use to forecast future short-rate changes must by definition be embodied in the actual spread. As long as S_t is part of the econometrician's information set ω_t , it must thus be the case that $E[\sum_{j=1}^{\infty} \beta^j \Delta R_{t+j} | \Omega_t] = E[\sum_{j=1}^{\infty} \beta^j \Delta R_{t+j} | \omega_t]$. It is important to note that this result is conditional on the expectations theory holding exactly. If we relax the null to allow for time-varying term premia or even a simple error term, S_t no longer embodies all necessary information about expected future short-rate changes.

Third, since we want to derive restrictions on the bivariate system composed of (11) and (12), we define the data set ω_t as p lags of ΔR and S each.³⁷ The econometrician's best linear one-period forecast of short-rate changes thus becomes $E[\Delta R_{t+1}|\omega_t] = h_{\Delta R}E[\omega_{t+1}|\omega_t] = h_{\Delta R}M\omega_t$, where $h_{\Delta R}$ is a selection vector equaling $[1\ 0\ \dots\ 0]$ and where M is the companion matrix corresponding to (11) and (12), written in first order form as $\omega_t = M\omega_{t-1} + e_t$; i.e.:

$$\begin{bmatrix} \Delta R_t \\ \dots \\ \Delta R_{t-p+1} \\ S_t \\ \dots \\ S_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & \dots & a_p & b_1 & \dots & b_p \\ 1 & & & & & \\ & & & & & \\ c_1 & \dots & c_p & d_1 & \dots & d_p \\ \dots & & & 1 & & \\ & & & & & \dots \end{bmatrix} \begin{bmatrix} \Delta R_{t-1} \\ \dots \\ \Delta R_{t-p} \\ S_{t-1} \\ \dots \\ S_{t-p} \end{bmatrix} + \begin{bmatrix} e_{\Delta R,t} \\ \dots \\ 0 \\ e_{S,t} \\ \dots \\ 0 \end{bmatrix}, \quad (14)$$

Fourth, given $\omega_t = M\omega_{t-1} + e_t$, *multi-period linear predictions* of short-rate changes are easy to form:

$$E[\Delta R_{t+j}|\omega_t] = h_{\Delta R}M^j\omega_t.$$

Mapping these forecasts into $S_t = \sum_{j=1}^{\infty} \beta^j E[\Delta R_{t+j}|\omega_t]$ and expressing $S_t = h_S\omega_t$ where h_S is a selection vector with a one in the position corresponding to the spread and zeros elsewhere, we finally derive:

$$h_S\omega_t = \sum_{j=1}^{\infty} \beta^j h_{\Delta R}M^j\omega_t = h_{\Delta R}M[I - \beta M]^{-1}\omega_t,$$

or equivalently:

$$h_S = h_{\Delta R}\beta M[I - \beta M]^{-1}. \quad (15)$$

Expression (15) represents a set of $2p$ cross-equation restrictions that the expectations theory imposes on the bivariate VAR system and that are sometimes called the hallmark of rational expectations models. Specifically, (11) and (12) contain $4p$ parameters $\{a_i\}_{i=1}^p$, $\{b_i\}_{i=1}^p$, $\{c_i\}_{i=1}^p$ and $\{d_i\}_{i=1}^p$. However, under the null that the expectations theory holds true, only $2p$ of these parameters are free while the remaining half is constrained by the cross-equation restrictions in (15).³⁸

Working with the same vector autoregression in short-rate changes and the spread, Campbell and Shiller (1987) test such rational expectations restrictions on U.S. data between 1959 and 1983 by means of a Wald test and conclude

³⁷ This restriction to the past history of interest rates follows Sargent (1979) and Campbell and Shiller (1987). It would be of some interest to explore the implications of adding other macroeconomic variables.

³⁸ As Campbell and Shiller (1987) note, the cross-equation (15) can be simplified to a linear set of restrictions. Specifically, we can rewrite them as $h_S[I - \beta M] = h_{\Delta R}\beta M$, which implies that $a_i = -c_i$ for $i = 1, \dots, p$; $d_1 = 1/\beta - b_1$; and $b_i = -d_i$.

Table 6 VAR Tests of the Expectations Hypothesis

Full Sample Estimates (1951.4–2001.11)				
	ΔR_t		S_t	
	unconstrained VAR	VAR consistent with ET	unconstrained VAR	VAR consistent with ET
ΔR_{t-1}	0.5782 (0.1095)	0.5739 (0.1088)	-0.4927 (0.1171)	-0.5739 (0.1088)
ΔR_{t-2}	0.4580 (0.1124)	0.4604 (0.1116)	-0.5059 (0.1201)	-0.4604 (0.1116)
ΔR_{t-3}	0.2192 (0.1125)	0.2268 (0.1117)	-0.3701 (0.1202)	-0.2268 (0.1117)
ΔR_{t-4}	-0.0447 (0.0379)	-0.0464 (0.0377)	0.0767 (0.0405)	0.0464 (0.0377)
S_{t-1}	0.9254 (0.1021)	0.9218 (0.1014)	0.1507 (0.1091)	0.0838 (0.1014)
S_{t-2}	-0.2228 (0.1542)	-0.2159 (0.1532)	0.0875 (0.1649)	0.2159 (0.1532)
S_{t-3}	-0.4233 (0.1552)	-0.4184 (0.1541)	0.3263 (0.1659)	0.4184 (0.1541)
S_{t-4}	-0.1693 (0.1104)	-0.1761 (0.1096)	0.3023 (0.1180)	0.1761 (0.1096)

Notes: All variables represent deviations from their respective means. Numbers in parentheses represent standard errors. The likelihood ratio test of the unconstrained VAR against the VAR consistent with the expectations theory (ET) is 35.7131. Since the corresponding critical 0.1 percent χ^2 value for 8 degrees of freedom is only 26.1, the restrictions imposed by the ET are strongly rejected.

that the expectations theory is strongly rejected. Alternatively, Sargent (1979) advocates assessing the expectations theory by means of a likelihood ratio test with an asymptotic chi-square distribution, which is the approach that we follow here. The likelihood ratio is $2[L_{UVAR} - L_{ETVAR}]$, that is, the difference between the log likelihood values of the unrestricted VAR and the VAR subject to the restriction in (15), respectively. For a given significance level, the restrictions are then rejected if the likelihood ratio is larger than the critical chi-square value for $2p$ degrees of freedom.

Table 6 reports the unrestricted and the restricted VAR estimates for our 1951–2001 sample using our reference lag length of $p = 4$.³⁹ Remarkably, none of the restricted point estimates differ by more than two standard errors from their unrestricted counterparts.⁴⁰ However, the computed likelihood ratio of 35.71 is larger than the critical 0.1 percent chi-square value of 26.1. Our data set thus comfortably rejects the restrictions imposed by the expectations theory, confirming Campbell and Shiller’s result over a substantially longer time period and using a more appropriate testing procedure.⁴¹

Time-Varying Term Premia

The restrictions in (15) are derived from the strong assumption that the expectations theory is exactly true up to term premia that are constant through time, which precludes even measurement error in the spread. Alternatively, we can adapt the testing approach discussed above and derive testable restrictions that allow for certain forms of time-variation in the term premia. To this end, reconsider the general formula (4) that links the long rate to the present value of future expected short rates and the expected term premia. Without imposing any restrictions, the spread can thus be expressed as the sum of two unobserved components:

$$S_t = F_t + K_t, \quad (16)$$

where $F_t = \sum_{j=1}^{\infty} \beta^j E[\Delta R_{t+j} | \Omega_t]$ and $K_t = (1 - \beta) \sum_{j=0}^{\infty} \beta^j E[k_{t+j} | \Omega_t]$ denote the present value of the market’s expectations about future short-rate changes and term premia, respectively. Combining this expression with the VAR framework $\omega_t = M\omega_{t-1} + e_t$, we can rewrite (16) as

$$S_t = E[F_t | \omega_t] + K_t + \xi_t,$$

where $\xi_t = F_t - E[F_t | \omega_t] = \sum_{j=1}^{\infty} \beta^j \{E[\Delta R_{t+j} | \Omega_t] - E[\Delta R_{t+j} | \omega_t]\}$ is the error arising from the fact that the econometrician is using a smaller data set than the market to forecast future short-rate changes.⁴² Equivalently, we can

³⁹ The reported results hold true for alternative lag lengths as well.

⁴⁰ Because of the specific linear nature of the cross-equation restrictions noted above, the constraint estimates and the standard errors for different pairs of VAR coefficients are identical.

⁴¹ Bekaert and Hodrick (2001) show that in the context of cross-equation restrictions tests of present-value models such as the expectations theory, Wald tests suffer from substantially larger sample biases than likelihood ratio tests or Lagrangean multiplier tests.

⁴² As noted in a previous footnote, under the null that the expectations theory holds, S_t embodies all necessary information about future short-rate changes, and thus $E[\Delta R_{t+j} | \Omega_t] = E[\Delta R_{t+j} | \omega_t]$ as long as S_t is part of ω_t . However, since now we have relaxed the assumption of constant term premia (i.e., the expectations theory does not hold), we can no longer assume that S_t contains all necessary information about future short-rate changes. This means that replacing the market’s information set Ω_t with the econometrician’s information set $\omega_t \subset \Omega_t$ (potentially) introduces a forecasting error.

Table 7 VAR Tests Based on Lagged Information

Full Sample Estimates (1951.4–2001.11)	
Information Lag	Likelihood Ratio (between unconstrained and constrained VAR)
0	35.7131
1	32.8594
3	33.6881
6	33.6300
12	35.6203

form expectations conditional on data ω_{t-l} :

$$E[S_t|\omega_{t-l}] = E[F_t|\omega_{t-l}] + E[K_t|\omega_{t-l}], \quad (17)$$

where we recognize that $E[\xi_t|\omega_{t-l}] = 0$ since ξ_t is uncorrelated by construction with any information in ω_{t-l} .

Finally, we impose that the term premia K_t is unforecastable from information ω_{t-l} , that is, $E[K_t|\omega_{t-l}] = 0$. Under this assumption, which is weaker than the assumption $K_t = 0$ employed in the tests of the expectations theory discussed earlier, we obtain the following testable restrictions:

$$h_S M^l = h_{\Delta R} \beta M [I - \beta M]^{-1} M^l, \quad (18)$$

where we used the same arguments as above to rewrite $E[S_t|\omega_{t-l}] = h_S M^l \omega_{t-l}$ and $E[F_t|\omega_{t-l}] = h_{\Delta R} \beta M [I - \beta M]^{-1} M^l \omega_{t-l}$.⁴³ This strategy is suggested by the fact that Sargent (1979) actually tests the expectations theory by considering such a relaxed form of the cross-equation restrictions with $l = 1$ (i.e., a one-period lag in the information set).

The restrictions in (18) can be evaluated using a likelihood ratio test similar to that used above, which compares the fit of the constrained and unconstrained vector autoregressions. Because of the assumed stationarity of the joint process for spreads and short-rate changes, the eigenvalues of the companion matrix M are all smaller than one in absolute value. It must be the case, then, that the restrictions are satisfied as l becomes very large, since both sides of the equation contain only zeros in the limit. However, restrictions of the form of (18) are valid and interesting so long as the researcher is willing to assume that term premia are unforecastable at some intermediate horizon.

⁴³ It might appear that one could “divide out” the terms M^l from both sides of (18), restoring the restrictions (15). However, the matrix M can be shown to be singular if $E[K_t|\omega_{t-l}] = 0$ is true (Kurmann [2002a]).

Table 7 reports likelihood ratios of the unrestricted VAR against the VAR subject to the restrictions in (18) for the forecasting horizons $l = 1, 3, 6,$ and 12 .⁴⁴ Notably, the restrictions are rejected for all of these lags. Thus, while the cointegration tests of Section 3 indicate that variations in the term premia are stationary, the results of Table 7 show that departures from the expectations theory are not only due to high-frequency deviations but also occur at intermediate, business cycle frequencies.

5. EXPECTATIONS AND THE SPREAD

The preceding section illustrates that the cross-equation restrictions implied by the expectations theory are soundly rejected, even when we allow for some limited time-variation in the term premia. However, as Campbell and Shiller (1987) argue, statistical tests of the cross-equation restrictions may be “highly sensitive to deviations from the expectations theory—so sensitive, in fact, that they may obscure some of the merits.”⁴⁵ In other words, even if the theory is not strictly true, it may contain important elements of the truth. This section builds on the ingenious approach of Campbell and Shiller (1987) in computing an estimate of the expectations component of the spread—which they call a “theoretical spread”—in order to shed more light on this issue. This approach also permits us to (i) extract an estimate of the term premium and (ii) to derive a rule of thumb linking the observed spread to unobserved expectations concerning temporary variations in the short-term interest rate.

Decomposing the Spread in Theory

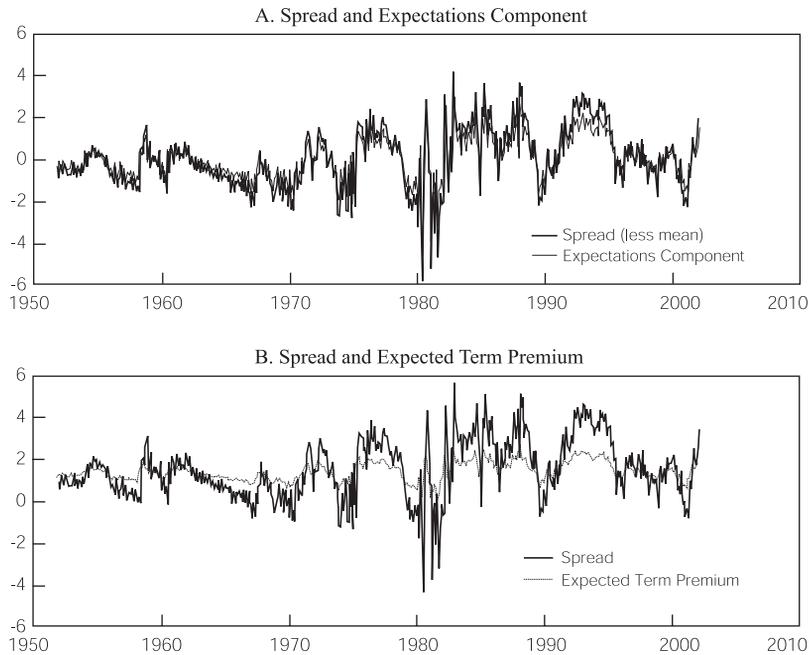
Our discussion above stresses that the observed spread is the sum of two unobserved components, $S_t = F_t + K_t$, which we call the expectations and term premium components. From (17) above, we know that the spread conditional on the econometrician’s information set ω_{t-l} can be written as:

$$E[S_t|\omega_{t-l}] = E[F_t|\omega_{t-l}] + E[K_t|\omega_{t-l}].$$

Under the expectations theory, we assumed that $E[K_t|\omega_{t-l}]$ is constant (or zero in deviations from the mean). In this section, we alternatively calculate an estimate of the expectations component given an information set and compare it to the prediction of the spread conditional on that same information set. From our results above, we know that the expectations component can be formed as $E[F_t|\omega_{t-l}] = \sum_{j=1}^{\infty} \beta^j E[\Delta R_{t+j}|\omega_{t-l}] = h_{\Delta R} \beta M [I - \beta M]^{-1} M^l \omega_{t-l}$, and

⁴⁴ The variables in the information set ω_{t-l} remain the same as for the cross-equation restriction tests above (i.e., ω consists of lags of ΔR and S). However, it would be interesting to assess the robustness of the reported results if we included additional variables that are likely to help forecast changes in the short rate.

⁴⁵ Campbell and Shiller (1987, 1080).

Figure 5 Spread, Expectations, and Term Premia

we also know that the predicted spread can be calculated as $E[S_t|\omega_{t-1}] = h_S M^1 \omega_{t-1}$. In these formulas, the coefficients from an *unrestricted* VAR are used to provide the elements of the matrix M that are relevant to forecasting. The difference between the two expressions, $E[K_t|\omega_{t-1}] = E[S_t|\omega_{t-1}] - E[F_t|\omega_{t-1}]$, is an implied variation in the term premium.

Decomposing the Spread in Practice

In view of the results from the prior section, we calculate two decompositions of the spread, based on different information sets.

Current information: We begin by calculating an estimate of the expectations component and the residual term premium using current information ω_t . In this setting, which corresponds to the analysis of Campbell and Shiller (1987), $E[S_t|\omega_t]$ simply equals the actual spread and $E[F_t|\omega_t] = h_{\Delta R} \beta M [I - \beta M]^{-1} \omega_t$.

Panel A of Figure 5 shows that the expectations component (the spread under the expectations theory) is strongly positively correlated with the actual spread (correlation coefficient = 0.99) and displays substantial variability. Panel B of Figure 5 shows the spread and the term premium (the gap between

Table 8 Summary Statistics for Expectations Component/Term Premium Decomposition

Full Sample Estimates (1951.4–2001.11)		
A. Based on Current Information		
Spread	Expectations	Term Premium
1.9318	1.3339	0.5979
0.9923	0.9355	0.3984
0.9633	0.9225	0.1994
B. Based on 6-months Forecasts		
Spread	Expectations	Term Premium
0.6495	0.4264	0.2231
0.9998	0.2800	0.1464
0.9995	0.9987	0.0767

Notes: Statistics correspond to Figures 5 and 6. Each panel contains a 3 by 3 matrix. On the diagonal, variances are reported (e.g., the variance of 6-months forecasts of the spread is 0.6495). Above the diagonal, covariances are listed (e.g., the covariance between the spread and expectations in the current information case is 1.3339). Below the diagonal, the corresponding correlation is reported (e.g., the correlation between the spread and expectations in the current information case is 0.9923).

the spread and the expectations component). The residual term premium is much less variable.

It is useful to consider a decomposition of variance for the spread, similar to that which we used for permanent and temporary components in Section 3:

$$\begin{aligned} \text{var}(S_t) &= \text{var}(F_t|\omega_t) + \text{var}(K_t|\omega_t) + 2 * \text{cov}(F_t|\omega_t, K_t|\omega_t) \\ 1.93 &= 0.94 + 0.20 + 2 * (0.40) \end{aligned}$$

Panel A of Table 8 reports second moments of the spread, the expectations component and the term premia. The variance of the spread is 1.93 (as noted in the derivation of the first spread rule of thumb), while the variance of the expectations component is 0.94. Since their respective standard deviations are not too different (1.39 and 0.97, respectively) and since they are virtually perfectly correlated, it is not surprising that a glance at the first panel of Figure 5 leads one to think that the expectations component explains most of the spread. By contrast, the standard deviation of the estimated term premium is much smaller (0.45), so it is natural to downplay its contribution after glancing at the second panel. But as panel A of Table 8 indicates, there is a very high estimated correlation of changes in the term premium and changes in the expectations component (0.94), so there is a substantial contribution to variability in the spread that arises from the covariance term (0.80 of a total of 1.93).

Economically, the spread appears excessively volatile relative to the estimated expectations component because there is a tendency for periods of high expectations components to occur when the term premium is also high.⁴⁶ Looking back to the first test of rational expectations restrictions, Figure 5 provides insight into why the cross-equation restrictions are rejected, since it highlights the distinct behavior of the spread and the expectations component. The spread contains information about the temporary component of interest rates highlighted by the expectations theory, but there are important departures as well.

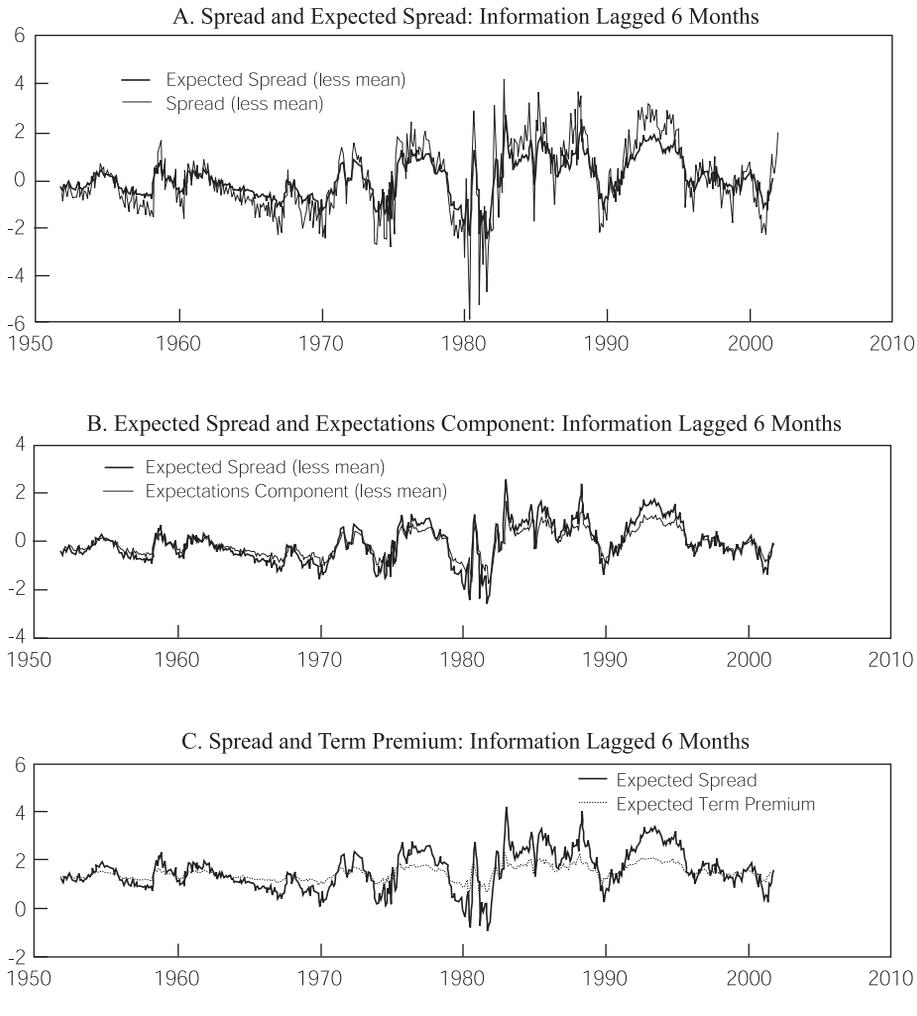
Results based on lagged information: Figure 6 and panel B of Table 8 use forecasts from the vector autoregression, using information six months previous. In panel A of Figure 6, the actual spread S_t and the forecast $ES_t|\omega_{t-6}$ are plotted. While these series move together, the forecasted spread is much less volatile than the actual spread (the variance of the forecasted spread is 0.65, which is about one-third of the actual spread's variance of 1.95). In panel B, the forecasted spread $ES_t|\omega_{t-6}$ and the forecasted expectations component $EF_t|\omega_{t-6}$ are plotted. While the forecasted expectations component is highly correlated with the forecasted spread, it is clearly less volatile as well. In panel C, the forecasted spread $ES_t|\omega_{t-6}$ and the forecasted term premium component $EK_t|\omega_{t-6} = ES_t|\omega_{t-6} - EF_t|\omega_{t-6}$ are plotted. This residual is positively associated with $ES_t|\omega_{t-6}$, with a near-perfect correlation. Its variance (0.076) is also somewhat more than one-third of the variance of the term premium measure $EK_t|\omega_t$ that is shown in Figure 5.

This figure illustrates, we conjecture, why the rational expectations restrictions are rejected when the information set is lagged, as reported previously in Table 7 and discussed in detail above. The deviations of the forecastable part of the spread $ES_t|\omega_{t-6}$ from the forecastable part of the expectations component $EF_t|\omega_{t-6}$ appear important. Indeed, there is some evidence that $EK_t|\omega_{t-6}$ are more serially correlated than either $ES_t|\omega_{t-6}$ or $EF_t|\omega_{t-6}$, as opposed to being unforecastable in the manner required for the rational expectations restrictions to be satisfied.

A Second Rule of Thumb for the Spread

If the spread rises by 1 percent, then how great a rise in the expectations component should an observer infer has occurred? This is a natural question, analogous to one earlier posed for the temporary component of the nominal interest rate, identified via the VECM. Since the variance of the spread is 1.93 and the covariance between the spread and the expectations component is 1.33, the rule

⁴⁶ While these are point estimates and do not take into account uncertainty implied by the fact that the unrestricted VAR is estimated rather than known, preliminary results in Kurmann (2002b) suggest that there may not be too much uncertainty in our context.

Figure 6 Conditioning S, F, and K on Lagged Information

of thumb coefficient is $b = 0.69 = 1.33/1.93$. Hence, we have the following.

Spread rule of thumb #2: *If the spread exceeds its mean by 1 percent, then our estimates suggest that the expectations component is high by 0.69 percent.*

Earlier, we derived a very similar implication—a coefficient of 0.71 but with an opposite rule sign—for the link between the temporary component of the short-term interest rate and the spread. It is not an accident that these two measures are very closely associated. The temporary component of the short-term rate is defined as $R_t - \bar{R}_t$, with $\bar{R}_t = R_{t-1} + \lim_{k \rightarrow \infty} \sum_{j=0}^k E_t \Delta R_{t+k}$. It

is accordingly given by $R_t - \bar{R}_t = -[\lim_{k \rightarrow \infty} \sum_{j=1}^k E_t \Delta R_{t+k}]$. The expectations component studied in this section is $E[F_t | \omega_{t-l}] = \sum_{j=1}^{\infty} \beta^j E_t[\Delta R_{t+j}]$. In each case, the expectations terms are made operational by use of very similar linear forecasting models; there are small differences because β is slightly smaller than one, but the essential theoretical and empirical properties are very similar except for the change in sign.

6. FOCUSING ON RECENT HISTORY

Many studies of recent macroeconomic history document changes in the pace and pattern of macroeconomic activity that have occurred over the past two decades.⁴⁷ Other studies suggest that a major reason for these changes is that the Federal Reserve System has altered its behavior in important ways. For example, Goodfriend (1993) argues that the U.S. monetary policy decision-making came of age—gaining important recognition and credibility—during this period, after having earlier traveled on a wide-ranging odyssey. Accordingly, in this section, we explore how some key features of our previous analysis change if we restrict attention to 1986.7–2001.11. The start date of this period was selected as descriptive of recent U.S. monetary policy with increased credibility, following the narrative history of Goodfriend (2002): it includes the last few years of the Volcker period and the bulk of the Greenspan period. We focus our attention on two sets of issues. First, how did the estimated variability in the permanent component of interest rates change during this period? Second, how did the estimated importance of the expectations effects on the long-short spread change during this period?

The Stochastic Trend in Interest Rates

One important conclusion from our earlier analysis is that there is a common stochastic trend in interest rates, which is closely associated with the long rate. To conduct the analysis for the recent period, we start by reestimating the VECM discussed in Section 3 and reported in Table 4. Then, we calculate the permanent component suggested by this specification, producing the results reported in Figure 7 and Table 9.

We focus on two main results. First, as Figure 7 shows, the stochastic trend continues to be an important contributor to the behavior of both the long-term and short-term interest rates. As in the full sample period, it is closely associated with the long rate. Further, it is much less closely associated with the short rate.

⁴⁷ For example, see Blanchard and Simon (2001) or Stock and Watson (2002).

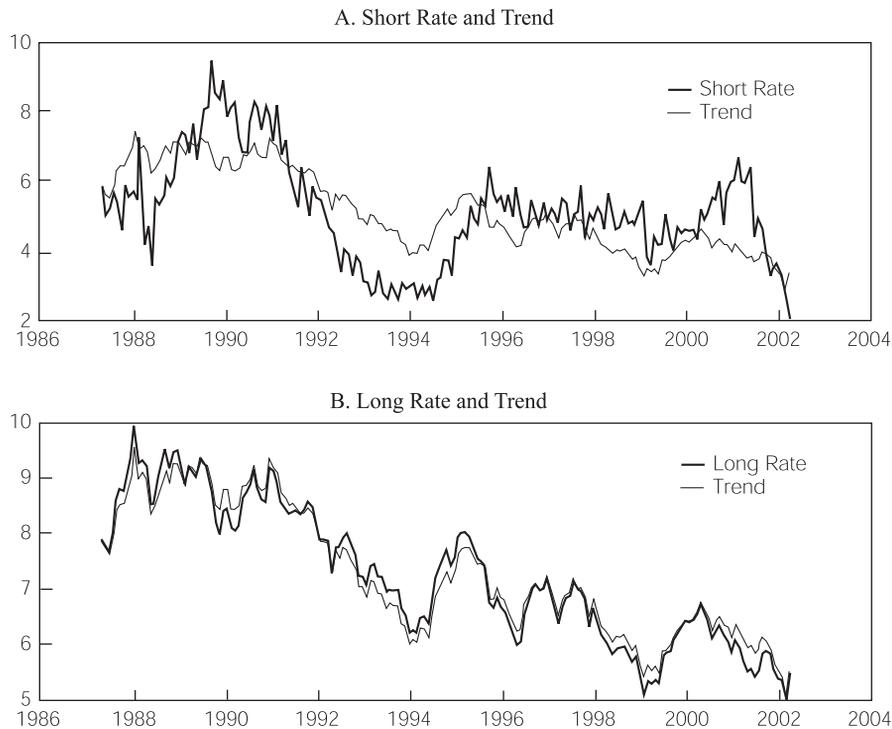
Table 9 Summary Statistics for Two Decompositions

Subsample Estimates (1951.4–2001.11)		
A. Short-Rate Changes		
Total	Permanent	Temporary
0.4409	−0.0003	0.4411
−0.0019	0.0476	−0.0478
0.9501	−0.3137	0.4890
B. Long-Rate Changes		
Total	Permanent	Temporary
0.0636	0.0536	0.0100
0.9747	0.0476	0.0061
0.6314	0.4422	0.0039
C. Spread		
Spread	Expectations	Term Premium
1.8917	1.4574	0.4343
0.9950	1.1341	0.3232
0.9475	0.9108	0.1111

Notes: Each panel contains a 3 by 3 matrix in a manner similar to Tables 5 and 8. On the diagonal, variances are reported. Above the diagonal, covariances are listed. Below the diagonal, the corresponding correlation is reported.

Panel B of Table 9 provides more detail. It shows that changes in the common stochastic trend (permanent component) have a variance of 0.048, which is less than one-half of the comparable variance reported in Table 5. Thus, there is evidence that the stochastic trend is less important for both short-term and long-term interest rates. We can measure this reduced influence on our rule of thumb. Based on the full sample, we calculate that a 1 percent rise in the long rate should bring about a 97 percent rise in the predicted permanent component. On the recent sample, this rule-of-thumb coefficient is smaller: a 1 percent rise should bring about only a 84 percent increase in the predicted permanent component.⁴⁸ Yet, while the effect is smaller, changes in long rate still strongly signal changes in the stochastic trend.

⁴⁸ In terms of elements of Table 9, the rule-of-thumb coefficient is calculated as $b = 0.0536/0.0636 = 0.84$.

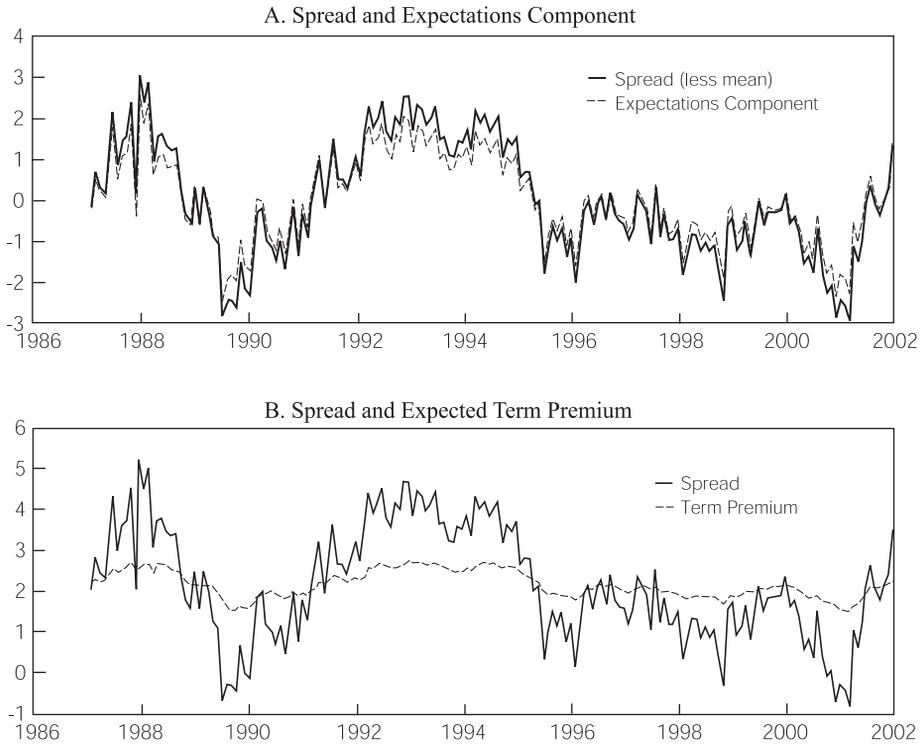
Figure 7 Interest Rates and Stochastic Trend, Once Again

Expectations and the Spread

Another important conclusion of our analysis above is that the spread is an indicator of forecastable temporary variation in short-term interest rates and, in particular, of market expectations of these variations. Figure 8 and panel C of Table 9 show that this relationship has been maintained and, indeed, has apparently gained strength during the recent period. In particular, if we look at rule of thumb #2 for the spread, which indicates the extent to which a high spread should be interpreted as reflecting a high expectations component, then the rule-of-thumb coefficient is $0.77 = 1.46/1.89$ for the recent period, whereas it was only 0.69 for the entire sample period.⁴⁹

In sum, the two reported differences for this more recent period are intriguing, and it is natural to think about possible sources of the change in

⁴⁹ We think that a natural next stage of research involves a more systematic inquiry into the evolving nature of the links between short-term rates and long-term rates. For example, Watson (1999) argues that increased persistence in short-term interest rates—which in our case would involve evolving VAR coefficients—helps explain the increased variability of long-term rates from the 1965–1978 period to the 1985–1998 period. This section, by contrast, argues that the changes in the persistent component in interest rates (the stochastic trend) were less important during 1986–

Figure 8 Spread, Expectations, and Term Premia, Once Again

stochastic properties of the term structure. For example, we might conjecture that the reduced importance of the permanent component is the result of a more credible, inflation-stabilizing monetary policy. Given the lack of structure in our present analysis, however, it is impossible to support such a claim with statistical evidence or to quantify its importance compared to other potential explanations. Rather, we consider that these findings highlight a topic that warrants further investigation.

7. SUMMARY AND CONCLUSIONS

We conclude that expectations about the level of interest rates are very important for the behavior of long-term interest rates on two dimensions. First,

2001 than over the 1951–2001 sample that includes the volatile 1979–1984 period not studied by Watson. A recent attempt to take into account time variations in the VAR parameters is Favero (2001), who computes the long rate under the expectations theory using a rolling regression VAR approach.

changes in the long-term interest rate substantially reflect changes in the permanent component (stochastic trend) in the level of the short-term rate. Second, the spread between long-term and short-term rates depends heavily on a temporary component (deviations from stochastic trend) of the level of short-term rates. Although the strong form of the expectations theory is rejected by a battery of statistical tests, it remains a workable approximation for many applied purposes. Changes in the long rate are largely a signal that the common trend in rates has shifted; a high spread is an important signal that future short rates will rise. More specifically, we provide rules of thumb for interpreting the expectations component of changes in long rates and the level of the long-short spread.

While the expectations theory is rejected, our rational expectations statistical approach is constructive in highlighting the ways in which the linear expectations theory of the term structure fails. The nature of predictable departures from the expectations theory, which we interpreted as time-varying term premia, suggests to us the importance of studying linkages between these factors and the business cycle, since our analysis indicates that these were not simply high frequency deviations.

Finally, the econometric methods that we use are nonstructural, in that they do not take a stand on the specific economic model that determines short-term rates. Nevertheless, the results of our investigation do make some suggestions about the shape that structural models must take, since they indicate the presence of a stochastic trend in the level of the interest rate. Recent research on monetary policy rules, as exemplified by Clarida, Gali, and Gertler (1999), almost invariably assumes that the short-term interest rate is governed by a stable behavioral rule of the central bank, linking it simply to the level of inflation and the level of the output gap, a specification which would preclude such shifts in trend interest rates when incorporated into most macroeconomic models. Our results suggest that a crucial next step in the analysis of monetary policy rules must be the exploration of specifications that can give rise to a stochastic trend in interest rates. In addition, most current macroeconomic models would generally ascribe such shifts in interest rate trends to shifts in inflation trends. Our results thus suggest the importance of an analysis of the interplay between trend inflation, the long-term rate, and monetary policy.

APPENDIX A: THE SHILLER APPROXIMATION

The purpose of this appendix is to derive and exposit Shiller's approximate equation for the yield on a long-term bond. For a coupon bond of arbitrary maturity, N , the yield-to-maturity is the interest rate that makes the price

equal to the present discounted value of its future cash flows $\{C_{t+j}\}$, which may include both coupons and face value:

$$P_t^L = \sum_{j=1}^N \frac{C_{t+j}}{(1 + R_t^L)^j}.$$

In the particular case of a bond with infinite term, which is commonly called a consol, the relationship is

$$P_t^L = \sum_{j=1}^{\infty} \frac{C}{(1 + R_t^L)^j} = \frac{C}{R_t^L}.$$

Between t and $t + 1$, the holding period yield on any coupon bond is given by

$$H_{t+1} = \frac{P_{t+1} + C - P_t}{P_t}.$$

Accordingly, the holding-period yield on a consol is given by

$$H_{t+1} = \frac{(C/R_{t+1}^L) + C - (C/R_t^L)}{(C/R_t^L)} = \frac{(1/R_{t+1}^L) + 1}{(1/R_t^L)} - 1.$$

The ratio R_t^L/R_{t+1}^L is approximately $1 + \theta(R_t^L - E_t R_{t+1}^L)$ via a first order Taylor series approximation about the point $R_t^L = R_{t+1}^L = R^L$, $\theta = 1/R^L$. It then follows that the holding-period yield is approximately

$$H_{t+1} = \theta(R_t^L - R_{t+1}^L) + R_t^L.$$

Notice that small changes in the yield $R_t^L - R_{t+1}^L$ have large implications for the holding-period yield H_{t+1} because θ is a large number. For example, if the annual interest rate is 6 percent and the observation period is one month, then $\theta = 1/(0.005) = 200$. Defining $\beta = 1/(1 + R^L)$, this expression can be written as $H_{t+1} = \frac{1}{1-\beta} R_t^L - \frac{\beta}{1-\beta} R_{t+1}^L$, which is convenient for the discussion below.

Suppose next that this approximate holding-period yield is equated (in expected value) to the short-term interest rate R_t and a term premium k_t . Then, it follows that

$$E_t H_{t+1} = E_t \left[\frac{1}{1-\beta} R_t^L - \frac{\beta}{1-\beta} R_{t+1}^L \right] = R_t + k_t$$

or

$$R_t^L = \beta E_t R_{t+1}^L + (1 - \beta)(R_t + k_t),$$

which is the form used in the main text. This derivation highlights the fact that the linear coefficient β may “drift” over time if the average level of the long rate is very different. It also highlights the fact that this term structure formula is an approximation suitable for very long-term bonds.

APPENDIX B: FORECASTING WITH THE VECM

We estimate a VECM of the form

$$\begin{aligned}\Delta R_t &= a(B)\Delta R_{t-i}^L + b(B)\Delta R_{t-i} + fS_{t-1} + e_{Rt} \\ \Delta R_t^L &= c(B)\Delta R_{t-i}^L + d(B)\Delta R_{t-i} + gS_{t-1} + e_{Lt},\end{aligned}$$

where B is the backshift (lag) operator. We note that this difference between these two equations is

$$\begin{aligned}\Delta S_t &= [c(B) - a(B)]\Delta R_{t-i}^L + [d(B) - b(B)]\Delta R_{t-i} \\ &\quad + (g - f)S_{t-1} + e_{Lt} - e_{Rt},\end{aligned}$$

so that we can write

$$\begin{aligned}S_t &= [c(B) - a(B)]\Delta R_{t-1}^L + [d(B) - b(B)]\Delta R_{t-1} \\ &\quad + (1 + g - f)S_{t-1} + (e_{Lt} - e_{Rt}),\end{aligned}$$

so that it is easy to write the system in state space form defining $x_{t-1} = [\Delta R_{t-1} \ \Delta R_{t-2} \ \dots \ \Delta R_{t-p} \ \Delta R_{t-1}^L \ \Delta R_{t-2}^L \ \Delta R_{t-p}^L \ S_{t-1}]$, which captures all of the predictor variables in these three equations. The main state equation is of the form $x_t = Mx_{t-1} + Ge_t$, with the elements being

$$\begin{aligned}x_t &= \begin{bmatrix} \Delta R_t \\ \dots \\ \Delta R_{t-p+1} \\ \Delta R_t^L \\ \dots \\ \Delta R_{t-p+1}^L \\ S_t \end{bmatrix} \\ M &= \begin{bmatrix} a_1 & \dots & a_p & b_1 & \dots & b_p & 0 \\ 1 & & & \dots & & & 0 \\ & & & \dots & & & \\ c_1 & \dots & c_p & d_1 & \dots & d_p & 0 \\ \dots & & & 1 & & & 0 \\ c_1 - a_1 & \dots & c_p - a_p & d_1 - b_1 & \dots & d_p - a_p & 1 + g - f \end{bmatrix} \\ Ge_t &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ \dots & \dots \\ 0 & 1 \\ \dots & \\ -1 & 1 \end{bmatrix} \begin{bmatrix} e_{Rt} \\ e_{Lt} \end{bmatrix}.\end{aligned}$$

APPENDIX C: VARIOUS COINTEGRATED MODELS

In this appendix, we want to demonstrate that the vector autoregression system estimated by Campbell and Shiller (1987) implies a vector error correction model with the cointegrating vector $[1 \ -1]$. The discussion is a specific case of the existence of a Phillips triangular form for a cointegrated system (see Hamilton [1994, 576–78]).

We write the vector error correction model as

$$\begin{aligned}\Delta R_t^L &= a(B)\Delta R_{t-1}^L + b(B)\Delta R_{t-1} + f(R_{t-1}^L - R_{t-1}) + e_{Lt} \\ \Delta R_t &= c(B)\Delta R_{t-1}^L + d(B)\Delta R_{t-1} + g(R_{t-1}^L - R_{t-1}) + e_{Rt},\end{aligned}$$

where B is the back-shift (lag) operator.

We write the VAR system of the CS form as

$$\begin{aligned}S_t &= g(B)S_{t-1} + h(B)\Delta R_{t-1} + e_{St} \\ \Delta R_t &= i(B)S_{t-1} + j(B)\Delta R_{t-1} + e_{Rt}.\end{aligned}$$

Finding the first equation in the VECM: Add the second equation of the VAR to the first, resulting in

$$\begin{aligned}R_t^L - R_{t-1} &= [g(B) + i(B)]S_{t-1} \\ &\quad + [h(B) + i(B)]\Delta R_{t-1} + (e_{St} + e_{Rt}).\end{aligned}$$

Reorganize this as

$$\begin{aligned}R_t^L - R_{t-1}^L &= [1 + g(1) + i(1)]S_{t-1} \\ &\quad + [g(B) - g(1) + i(B) - i(1)]S_{t-1} \\ &\quad + [h(B) + i(B)]\Delta R_{t-1} + (e_{St} + e_{Rt}),\end{aligned}$$

where $g(1)$ is the sum of coefficients in the g polynomial (and similarly for i). Since the coefficients in $[g(B) - g(1)]$ sum to zero by construction, it is always possible to factor $[g(B) - g(1)] = \gamma(B)(1 - B)$ with $\gamma(B)$ having one less lag than $g(B)$. Further, we can similarly write $i(B) - i(1) = \phi(B)(1 - B)$. Hence, we can write the above equation as

$$\begin{aligned}R_t^L - R_{t-1}^L &= [1 + g(1) + i(1)]S_{t-1} \\ &\quad + [\gamma(B) + \phi(B)](\Delta R_{t-1}^L - \Delta R_{t-1}) \\ &\quad + [h(B) + i(B)]\Delta R_{t-1} + (e_{St} + e_{Rt}),\end{aligned}$$

which takes the general form of the VECM equation with suitable definitions of $a(B)$ and $b(B)$.

Finding the second equation in the VECM: Similarly, we can rearrange the second equation above as

$$\Delta R_t = [i(B) - i(1)]S_{t-1} + j(B)\Delta R_{t-1} + i(1)S_{t-1} + e_{Rt}.$$

Hence,

$$\Delta R_t = \phi(B)\Delta R_{t-1}^L + [j(B) - \gamma(B)]\Delta R_{t-1} + i(1)S_{t-1} + e_{Rt},$$

which is the same form as the second equation of the VECM system. Thus, the Campbell-Shiller VAR implies a VECM.

REFERENCES

- Bekaert, Geert, and Robert J. Hodrick. 2001. "Expectations Hypothesis Tests." *Journal of Finance* 56 (August): 1357–94.
- Beveridge, Stephen, and Charles R. Nelson. 1981. "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'." *Journal of Monetary Economics* 7 (March): 151–74.
- Blanchard, Olivier, and J. Simon. 2001. "The Long and Large Decline in U.S. Output Volatility." *Brookings Papers on Economic Activity* 1 (March): 135–64.
- Campbell, John Y., and Robert J. Shiller. 1987. "Cointegration and Tests of Present Value Models." *Journal of Political Economy* 95 (October): 1062–88.
- _____. 1991. "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *The Review of Economic Studies* 58 (May): 495–514.
- _____, and Kermit L. Schoenholtz. 1983. "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates." *Brookings Papers on Economic Activity* 1 (March): 173–217.
- Clarida, Richard, Jordi Gali, and Mark Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37 (December): 1661–707.
- Dickey, David N., and Wayne A. Fuller. 1981. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica* 49 (June): 1057–72.
- Dotsey, Michael. 1998. "The Predictive Content of the Interest Rate Term Spread for Future Economic Growth." Federal Reserve Bank of Richmond *Economic Quarterly* 84 (Summer): 31–51.
- Engle, Robert F., and Clive W. J. Granger. 1987. "Cointegration and Error Correction: Representation, Estimation and Testing." *Econometrica* 55

- (March): 251-76. Reprinted in *Long-Run Economic Relations: Readings in Cointegration*, ed. Robert F. Engle and Clive W. J. Granger. New York: Oxford University Press, 1991.
- Fama, Eugene F. 1977. *Foundations of finance: portfolio decisions and securities prices*. Blackwell.
- Favero, Carlo A. 2001. "Taylor Rules and the Term Structure." Working paper, IGIER Università L. Bocconi (December).
- Goodfriend, Marvin. 1993. "Monetary Policy Comes of Age: A 20th Century Odyssey." Federal Reserve Bank of Richmond *Economic Quarterly* 79 (Winter): 1–22.
- _____. 2002. "The Phases of U.S. Monetary Policy: 1987 to 2001." Manuscript. Prepared for the Charles Goodhart Festschrift, Bank of England, November 2001, revised January 2002.
- Hetzl, Robert L., and Ralph F. Leach. 2001. "The Treasury-Fed Accord: A New Narrative Account." Federal Reserve Bank of Richmond *Economic Quarterly* 87 (Winter): 57–64.
- Hamilton, James D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.
- Horvath, Michael T. K., and Mark W. Watson. 1995. "Testing for Cointegration When Some of the Cointegrating Vectors are Known." *Econometric Theory* 11 (December): 952–84.
- Ibbotson Associates. 2002. *Stocks, Bonds, Bills, and Inflation 2002 Yearbook*. Chicago: Ibbotson Associates.
- Johansen, Soren. 1988. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control* 12 (June/Sept.): 231–54. Reprinted in *Long-Run Economic Relations: Readings in Cointegration*, ed. Robert F. Engle and Clive W. J. Granger. New York: Oxford University Press, 1991.
- Judge, George G., W. E. Griffiths, R. Carter Hill, Helmut Lutkepohl, and Tsoung-Chao Lee. 1985. "The Theory and Practice of Econometrics." New York: Wiley.
- King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson. "Stochastic Trends and Economic Fluctuations." *American Economic Review* 81 (September): 819–40.
- Kurmann, André. 2002a. "Maximum Likelihood Estimation of Dynamic Stochastic Theories with An Application to New Keynesian Pricing." Chapter 2, *New Keynesian Price and Cost Dynamics: Theory and Evidence*, Ph.D. diss., University of Virginia.

- _____. 2002b. "Quantifying the Uncertainty about Theoretical Interest Rate Spreads." Working paper, University of Virginia.
- Mankiw, N. Gregory, and Jeffrey A. Miron. 1986. "The Changing Behavior of the Term Structure of Interest Rates." *Quarterly Journal of Economics* 101 (May): 211–28.
- Nelson, Charles R., and G. William Schwert. 1977. "Short-Term Interest Rates as Predictors of Inflation: On Testing the Hypothesis That the Real Rate of Interest Is Constant." *American Economic Review* 67 (June): 478–86.
- Owens, Raymond E., and Roy H. Webb. 2001. "Using the Federal Funds Futures Market to Predict Monetary Policy Actions." Federal Reserve Bank of Richmond *Economic Quarterly* 87 (Spring): 69–77.
- Roll, Richard. 1969. *The Behavior of Interest Rates: An Application of the Efficient Market Model to U.S. Treasury Bills*. New York: Basic Books.
- Sargent, Thomas J. 1979. "A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure." *Journal of Monetary Economics* 5 (January): 133–43.
- Shiller, Robert J. 1972. Rational Expectations and the Term Structure of Interest Rates. Ph.D. diss., M.I.T.
- Stock, James H., and Mark W. Watson. 1988. "Testing for Common Trends." *Journal of the American Statistical Association* (December): 1097–107.
- _____. 2002. "Has the Business Cycle Changed and Why?" Manuscript. Prepared for NBER Macroeconomics Annual (April).
- Watson, Mark W. 1999. "Explaining the Increased Variability in Long-Term Interest Rates." Federal Reserve Bank of Richmond *Economic Quarterly* 85 (Fall): 71–96.