Journal of Monetary Economics 48 (2001) 241–270

Journal of MONETARY ECONOMICS

www.elsevier.com/locate/econbase

Peso problem explanations for term structure anomalies ☆

Geert Bekaert[a,b], Robert J. Hodrick[a,b], David A. Marshall[c,*]

[a] *Columbia University, New York, NY 10027, USA*

[b] *National Bureau of Economic Research, Cambridge, MA 02138, USA*

[c] *Research Department, Federal Reserve Bank of Chicago, 230 South LaSalle Street, Chicago, IL 60604, USA*

Received 11 November 1998; received in revised form 28 June 2000; accepted 16 October 2000

Abstract

We investigate whether term structure anomalies in U.S. data may be due to a generalized peso problem, in which a high-interest-rate regime occurred less frequently in the U.S. sample than was rationally anticipated. We formalize this idea by estimating a regime-switching model of short-term interest rates with data from seven countries. Under the small-sample distributions generated by the model, the expectations hypothesis is rejected. When we allow moderate time variation in term premiums, the term-premium dynamics interact with peso-problem effects to generate small-sample distributions more consistent with the data. Nonetheless, our model cannot fully account for U.S. term structure anomalies. © 2001 Published by Elsevier Science B.V.

JEL classificaion: E43; C22; G15

Keywords: Term structure; Regime switching; Expectations hypothesis

☆ We are grateful to Christian Gilles, Karen Lewis, Chris Telmer, Paula Tkac, and the participants in the many seminars at which this article was presented for their suggestions, and to an anonymous referee, whose comments greatly improved the paper. Geert Bekaert and Robert Hodrick gratefully acknowledge financial support from NSF grants and Bekaert acknowledges additional support from the Financial Research Initiative at Stanford University. Andrew Ang and Denise Duffy provided excellent research assistance.

*Corresponding author. Tel.: +1-312-322-5102; fax: +1-312-322-2357.

E-mail address: dmarshall@frbchi.org (D.A. Marshall).

0304-3932/01/$ - see front matter © 2001 Published by Elsevier Science B.V.
PII: S0304-3932(01)00075-7

1. Introduction

When researchers test the expectations hypothesis (EH) of the term structure with U.S. dollar (USD) data, an interesting paradox emerges. Briefly, the change in the long-term interest rate does not behave as predicted by the EH, whereas future short rates do change in the direction predicted by the EH. Even so, at the short end of the term structure, future short rates do not move enough and the theory is still rejected (Campbell and Shiller, 1991). General equilibrium attempts to explain these observations with time-varying risk premiums generally fail.[1] The goal of our project is to see whether these problems may be driven, not by a failure of economic theory, but by a failure of the asymptotic distribution theory used to examine the Campbell–Shiller regression tests.

In this paper, we attempt to explain the anomalous patterns in the USD term structure by focusing on a particular issue in small-sample inference known as the peso problem. As in Evans (1996), we define a peso problem as arising whenever the ex post frequencies of states within the sample differ substantially from their ex ante probabilities, and where these deviations distort econometric inference. When a peso problem is present, the sample moments calculated from the available data do not coincide with the population moments that rational agents would have used when making their decisions.

Why do we think such peso problems may provide an explanation of the Campbell and Shiller (1991) anomalies in USD data? Consider the bond market in the early 1980s, when the five-year USD interest rate reached 15.9%, its maximum during our sample period. Under the EH, the long interest rate is the average of expected future short rates, so investors in the early 1980s would have expected future short rates to be drawn from a distribution centered around 15.9%. Let us conservatively assume that the standard deviation of this distribution equaled 3.66%, the unconditional standard deviation of short rates in our USD data.[2] Under this assumption, short rates as high as $23.22\% = 15.9\% + 2(3.66\%)$ would not have been unusual in the mid-1980s. But, the maximum value of the USD short rate during our sample occurs six months *before* the maximum of the USD long rate, and its value is only 16.3%.

There are two alternative interpretations of this example that are consistent with agent rationality. The first is to make the usual assumption of rational expectations econometric practice that the population distribution used by

[1] Recent examples of general equilibrium macroeconomic models of the term structure include Backus et al. (1989), den Haan (1995), and Bekaert et al. (1997b). Fisher and Gilles (1996), Roberds and Whiteman (1999), and Backus et al. (1997) develop the implications of the affine class of general equilibrium financial models for the Campbell and Shiller (1991) results.

[2] This assumption is conservative since the conditional standard deviation of short rates tends to be larger when interest rates are high.

agents corresponds to the empirical distribution estimated from post-war USD data. Under this interpretation, one would conclude that the data reject the EH. The second interpretation is that the population distribution of USD rates includes states in which the short rate is well above 16.3%, but there were no realizations of these states in the post-war USD data. This second interpretation corresponds to the peso-problem intuition.

Evans (1996) surveys the substantial literature suggesting peso-problem explanations for economic anomalies. However, he also notes a fundamental econometric problem in empirically implementing the peso problem intuition: the small sample of data available. How can one estimate the population distribution underlying a peso-problem model when, *by definition*, a peso problem only exists when there are insufficient data to estimate that population distribution? In this paper we overcome this small-sample problem in USD data by utilizing short-rate data from several different countries simultaneously. To do so, we assume that these data are all drawn from the same unconditional distribution. We regard this seemingly strong assumption as a reasonable starting point. The developed countries of the world face a similar set of technological shocks and have similar political economies, yet their experiences with inflation and real interest rates are quite different in small samples. If countries' rates of inflation and real interest rates vary over time for similar reasons, the short-term interest rates observed in any given country represent potential realizations that could occur in any of the other countries.

As an example of how our approach might change one's inference from the data, let us return to the example discussed above. While no interest rates approaching 23.22% are observed in the USD data, values close to this level are observed in other developed countries, such as Italy, Japan, and Australia.[3] If one believes that data from other developed countries contain information about possible realizations of USD interest rates, then it is not clear that the EH is incompatible with observed data. Rather, one might conclude that all possible realizations of the short rate are not in the particular small sample contained in the USD data set.

A contribution of this paper is to formalize and test this approach to inference. We estimate a regime-switching model using data from seven developed countries, including the U.S.[4] By including these other countries' data in our estimation of the data generating process, we create a hypothetical economy where the distribution of future interest rates resembles that which

[3] Three-month interest rates above 19% were observed in all of these countries during the post-war period.

[4] Evans (1996) suggests using a regime-switching model to formalize peso problems. Lewis (1991) models multiple policy regimes to explore potential peso problems in U.S. interest rates during 1979–1982. Other papers that explore regime-switching models of interest rates include Hamilton (1988), Sola and Driffill (1994), Evans and Lewis (1994, 1995), and Gray (1996).

agents confronted in real time, but which did not occur in actual U.S. data. We then incorporate the estimated short-rate process into two term structure models. The first imposes the EH; that is, term premiums are assumed to be constant over time. In the second, we allow for time-varying term premiums by assuming that long-term interest rates are generated from a one-factor pricing kernel as in Backus (1993) and Duffie and Kan (1996). For each of these models, we compute the small-sample distributions of the two Campbell–Shiller tests, and we use these small-sample distributions to re-evaluate these statistics.

Our first result is that our model of the peso problem intuition is unable to salvage the EH of the term structure. While evidence against the EH is considerably weakened when peso problems are taken into consideration, the hypothesis is still rejected in joint tests of the Campbell–Shiller statistics. Our second result is that peso problems interact with time-varying risk premiums in ways that are important for economic inference. In particular, when we allow for small time variation in the term premium (representing a slight departure from the EH in population), peso effects become magnified. We are unable to reject the model for the 12- and 36-month maturities. However, this model is still rejected for the 60-month maturity. We conclude that, while our single-factor term structure model represents an incomplete explanation for the Campbell–Shiller anomalies, it is potentially misleading to test term structure models by comparing the data to the population distribution implied by the model. Peso effects also should be taken into account.

The structure of the paper is the following. Section 2 briefly reviews the EH, and discusses why it is a useful starting point for our investigations. Section 3 provides evidence on the EH using data from the currencies of the United States (USD), the United Kingdom (GBP), and Germany (DEM). Section 4 develops our regime-switching model, and Section 5 provides estimates of the model's parameters and discusses the model's implications. Section 6 presents the small-sample distributions of term-structure test statistics implied by our estimates when the EH is imposed and evaluates whether our peso problem intuition can explain the empirical behavior of these statistics. Section 7 provides the analysis of the time-varying term-premium model. Section 8 concludes and outlines some directions for future research.

2. The expectations hypothesis of the term structure

In economies that do not admit arbitrage opportunities, the term structure of interest rates follows the relation

$$\exp(-r_{t,n}n) = \mathrm{E}_t\left[\prod_{i=1}^{n} M_{t+i}\right], \tag{1}$$

where $r_{t,n}$ denotes the continuously compounded yield to maturity on a zero-coupon bond purchased at date t and paying one dollar at $t+n$, and M_t denotes a positive pricing kernel for dollar assets purchased at date $t-1$ that pay off at date t.[5] Following Campbell and Shiller (1991), we define the EH as the hypothesis that continuously compounded zero-coupon bond yields are the averages of expected future continuously compounded short interest rates, plus a time-invariant term premium. Formally,

$$r_{t,n} = \frac{1}{n}\sum_{i=0}^{n-1} \mathrm{E}_t(r_{t+i,1}) + c_n. \tag{2}$$

Eq. (2) can be derived from the basic asset-pricing equation (1) under particular distributional assumptions. In particular, Bekaert et al. (1997c) demonstrate that the term premium c_n is a function of the second and higher-order conditional moments of the pricing kernel. If these moments are time-invariant, as in the Vasicek (1977) model, Eq. (2) holds. If the moments vary over time, term premiums are variable. Nevertheless, as Backus and Zin (1994) note, economists have encountered difficulty constructing a reasonable economic model in which the pricing kernel displays sufficient conditional heteroskedasticity to generate the term structure patterns documented below. We therefore ask whether small-sample econometric problems can account for these patterns.[6] The EH is a natural starting point, since it imposes the extreme assumption that *no* portion of the expectations hypothesis's failure can be attributed to time variation in term premiums. We impose this assumption in Section 6, but relax it in Section 7.

3. Evidence on the expectations hypothesis of the term structure

3.1. The Campbell–Shiller regressions

Campbell and Shiller (1991) propose the following tests of Eq. (2) that involve current term spreads, $r_{t,n} - r_{t,m}$, where $n > m$ and $k \equiv n/m$ is an integer. First, Eq. (2) implies that a maturity-specific multiple of the term spread predicts the m-period change in the longer term bond yield. In particular, the slope coefficient α_1 should equal unity in the following regression equation:

$$r_{t+m,n-m} - r_{t,n} = \alpha_0 + \alpha_1\left(\frac{m}{n-m}\right)[r_{t,n} - r_{t,m}] + u_{t+m}. \tag{3}$$

[5] If markets are incomplete, there may be multiple pricing kernels consistent with no arbitrage. Eq. (1) and the subsequent discussion hold for any valid kernel.

[6] The peso-problem intuition is but one reason why rational expectations econometrics may be inapplicable. An alternative explanation is that agents take time to learn about the underlying structure of the economy (see, for example, Lewis 1989).

Second, Eq. (2) implies that the current term spread between the n-period yield and the m-period yield forecasts the average of future m-period interest rates minus the current m-period rate. In particular, the slope coefficient δ_1 should equal unity in the following regression equation:[7]

$$\frac{1}{k}\left[\sum_{i=0}^{k-1} r_{t+im,m}\right] - r_{t,m} = \delta_0 + \delta_1[r_{t,n} - r_{t,m}] + v_{t+n-m}. \tag{4}$$

Panel A of Table 1 displays results from regressions (3) and (4) using data from government bonds denominated in the USD, the GBP, and the DEM, with $m = 3$.[8] Appendix A describes the construction of the data. Two things are noteworthy in Panel A, although the reader is cautioned that the interpretation uses the asymptotic distributions, which are suspect. First, there appears to be strong evidence against the EH using the USD, especially from Eq. (3). Second, the evidence against the EH is much weaker using GBP and DEM data. If one uses the asymptotic distributions of the OLS slope estimators, one concludes that the GBP regressions show only slight evidence against the hypothesis that the slope coefficients equal unity for both regressions. The regressions using DEM data reject the hypothesis of a unit slope coefficient for all three horizons in Eq. (3), but the point estimates are closer to 1.0 than in the USD regressions. The results for Eq. (4) using DEM data show evidence against the EH only for the 12- and 60-month bonds.

3.2. Interpretations of the evidence

The results for the USD in Panel A of Table 1 confirm the findings of Campbell and Shiller (1991, p. 505) that "the slope of the term structure almost always gives a forecast in the wrong direction for the short-term change in the yield on the longer bond, but gives a forecast in the right direction for long-term changes in short rates". It is possible that these results may be driven by small-sample anomalies due to peso problems in the data analysis. Suppose that short interest rates can evolve in three different regimes, with the mean and volatility of interest rates increasing together as we move across regimes.

[7] A third way to investigate the expectations hypothesis, which is closely related to the second Campbell and Shiller (1991) specification, is to examine the forward interest rates implicit in the term structure as predictors of future spot interest rates as in Fama (1984), Fama and Bliss (1987), Stambaugh (1988), and Backus et al. (1997). Campbell and Shiller (1991) propose tests based on vector autoregressions (VAR) of short rates and spreads. Bekaert et al. (1997c) also examine various VAR statistics, which are not reported here to conserve space.

[8] We estimate Eq. (3), as does much of the literature, with the approximation $r_{t+3,n-3} = r_{t+3,n}$ because, for most commonly used values of n (such as 12, 24, etc.), data on yields with maturities of $n-3$ months are not available. Bekaert et al. (1997a) note that this approximation produces an inconsistent estimator. Because estimates for the three currencies would be subject to the same inconsistency, we do not adjust the results.

Table 1
Slope coefficient estimates for Eqs. (3) and (4)[a]

Panel A: OLS slope coefficients for Eqs. (3) and (4)

Horizon n (months)	USD Eq. (3)	USD Eq. (4)	GBP Eq. (3)	GBP Eq. (4)	DEM Eq. (3)	DEM Eq. (4)
12	−2.012	0.152	0.095	0.581	−0.054	0.388
	(0.311)	(0.120)	(0.395)	(0.186)	(0.161)	(0.125)
36	−3.098	0.448	0.894	1.300	−0.137	1.059
	(0.628)	(0.238)	(0.620)	(0.170)	(0.390)	(0.227)
60	−4.211	0.678	0.874	1.426	−0.325	1.285
	(0.988)	(0.244)	(0.891)	(0.155)	(0.543)	(0.108)

Panel B: The effects of large spreads on Eq. (3)

Horizon n (months)	USD Normal spreads	USD Large spreads	GBP Normal spreads	GBP Large spreads	DEM Normal spreads	DEM Large spreads
12	−1.186	−2.328	0.268	0.008	0.413	−0.222
	(0.504)	(0.318)	(0.503)	(0.432)	(0.291)	(0.145)
36	−1.164	−4.692	1.224	0.723	0.256	−0.371
	(0.949)	(0.687)	(1.150)	(0.696)	(0.786)	(0.346)
60	−2.177	−6.554	1.554	0.520	−0.034	−0.548
	(0.135)	(1.522)	(1.919)	(0.935)	(0.923)	(0.533)

[a]Note: The table reports estimates of Eqs. (3) and (4) using term-structure data from three currencies, the U.S. dollar (USD), the British pound (GBP), and the Deutsche mark (DEM). The sample contains monthly data from 1972:01 through 1996:09. The short rate is the 3-month rate for each currency. The long-rate maturity ("horizon") is indicated in column 1. The maximum number of observations is used in each regression. Hence, for Eq. (3) there are 294 observations, while for Eq. (4) there are $297-(n-3)$ observations. Hansen's (1982) GMM standard errors are in parentheses and are computed using the method of Newey and West (1987) to accommodate the overlapping error structure induced by using monthly observations with a multiperiod forecasting horizon. There are 3 Newey–West lags for Eq. (3) and $n-3$ lags for Eq. (4), which is one lag more than is necessary under the null hypothesis. For Panel B, the observations are split into two groups based on the size of the term spreads and Eq. (3) is re-estimated allowing for different slope coefficients for normal and large spreads. Spreads falling outside a band of 1.3 standard deviations of the mean spread are defined to be "large". The fraction of large spreads varies between 9.86% and 21.09% of the total sample, yielding a minimum of 29 observations to estimate the large spread slope coefficient.

Further, suppose that any shock that increases (decreases) the short rate also increases the probability of switching to a higher rate (lower rate) regime. Then, as short rates rise, the term spread may rise as agents rationally forecast transitions into a higher rate regime. But, if in a particular sample the higher rate regimes are observed less frequently than their unconditional probabilities, this increase in the spread will appear unjustified ex post. In such a sample, regression (3) will fail to deliver an estimated slope of unity and could produce negative coefficients if increases in the spread are subsequently followed by surprising transitions to lower rate regimes. The slope coefficient estimates in regression (4) will also be less than unity, but the estimation error here is likely to be smaller than in regression (3), since the short rates immediately following the shock will tend to be higher than their unconditional value even if rates stay within a regime because of the high serial correlation of short rates.

Suppose that this small-sample explanation of the USD evidence is true and that other currencies follow the same regime-switching model. Due to sampling variation, other currency interest rates need not resemble the USD experience. The differences in estimated coefficients in Table 1 for the different currencies could be due to different small-sample realizations from the same population distribution. One would expect there to be less evidence against the EH in currencies with a sample that is more representative of the population distribution. High and volatile interest rates were more common during the sample in the U.K. than in the U.S. and Germany. Strikingly, the U.K. data provide the weakest evidence against the EH.

Furthermore, if this peso explanation is true, the USD data and to a lesser extent the DEM data may contain observations in which spreads increase dramatically (because a shift to a high-rate regime is anticipated) but the shift does not actually occur. These observations may have a disproportionate effect on the slope coefficients of regression (3). Panel B of Table 1 investigates this possibility by allowing different coefficients on "normal" and "large" spreads in Eq. (3), where "large" spreads are more than 1.3 standard deviations from the mean. For all nine regressions the slope coefficients for the large spreads are lower than those for normal spreads. The DEM coefficients on large spreads are significantly different from one for all maturities using the asymptotic standard errors, whereas the "normal" coefficients are insignificantly different from one at (or close to) the 5% significance level. For the USD, both slope coefficients are significantly different from one for all maturities. Most GBP coefficients are close to one.

4. A regime-switching model of interest rates

This section presents a regime-switching model as a characterization of the process governing interest-rate data from several currencies.

4.1. The basic model

In our model, the short interest rate is determined by a regime-switching model in which the regimes follow a Markov process. In each regime the 3-month interest rate follows a first-order, autoregressive, conditionally heteroskedastic process. The parameters that determine the conditional mean and variance within a regime are all regime dependent, and we assume there are three regimes. Let s_t be an indicator variable such that $s_t = i$ if regime i prevails at date t. For convenience, we use the simpler notation r_t for the 3-month interest rate, $r_{t,3}$. The law of motion for r_{t+1} when $s_{t+1} = i$ is

$$r_{t+1} = \mu_i + \beta_i r_t + h_i(r_t)\varepsilon_{t+1}, \tag{5}$$

where $\{\varepsilon_t\}$ is a sequence of independent standard normal random variables, and the conditional standard deviation $h_i(r_t)$ is

$$h_i(r_t) = \sigma_i r_t^{\gamma_i} \tag{6}$$

for $i = 1, 2, 3$. Note that the realization of r_{t+1} is affected by two random shocks, the realization of the regime, s_{t+1}, and the innovation in ε_{t+1}.

We now specify how the interest-rate process shifts among the three possible regimes. We identify higher numbered regimes with higher mean levels of interest rates. Because observed short rates move up and down gradually, we allow next period's regime to be either the same regime as today's or an adjacent regime. We do not allow jumps from regime 1 to 3 or 3 to 1.

We assume that regime transition probabilities depend on the current state of the economy. We parameterize these transition probabilities as follows:

$$\mathrm{Prob}(s_{t+1} = i | s_t = i, r_t) = \frac{\exp(a_{ii} + b_{ii} r_t)}{1 + \exp(a_{ii} + b_{ii} r_t)}, \quad i = 1, 3, \tag{7}$$

$$\mathrm{Prob}(s_{t+1} = 2 | s_t = 2, r_t) = \frac{\exp(a_{22} + b_{22} r_t)}{1 + \exp(a_{22} + b_{22} r_t) + \exp(a_{23} + b_{23} r_t)}, \tag{8}$$

$$\mathrm{Prob}(s_{t+1} = 3 | s_t = 2, r_t) = \frac{\exp(a_{23} + b_{23} r_t)}{1 + \exp(a_{22} + b_{22} r_t) + \exp(a_{23} + b_{23} r_t)}, \tag{9}$$

where $\{a_{ii}, b_{ii}, i = 1, 2, 3\}$ and $\{a_{23}, b_{23}\}$ are parameters of the model. Under Eqs. (5)–(9), the conditional distribution of r_{t+1} given r_t and s_t is a mixture of normals with state-dependent mixing probabilities. Gray (1996) examines a similar model with two regimes and finds that it fits the USD data better than alternative models.

4.2. Exploiting cross-currency data

We use interest-rate data from seven different currencies to estimate the parameters of the regime-switching model. Our hypothesis is that these data represent different draws from the same regime-switching process. Consequently, the parameters of the model are assumed to be the same for all currencies. The model allows three reasons why different countries' samples have different small-sample statistics. First, countries spend different amounts of time in the regimes. Second, some countries switch between regimes more frequently than others. Third, the shocks within a regime are idiosyncratic.

We use short interest rates denominated in the currencies of Australia, Germany, Italy, Japan, Sweden, the United Kingdom, and the United States. Fundamentally, each of these countries is an industrial democracy, and the people in these countries and their policy-makers would all like to have low inflation, low unemployment, and high real growth. Nevertheless, while these countries share common goals, their economic experiences have been quite different, which has resulted in a fairly wide range of interest-rate patterns.

Computational considerations require us to assume that the realizations across countries are independent observations. Consequently, we do not work with the G7 countries. We omit France because its interest rates are closely linked to German rates through attempts to fix exchange rates in the European Monetary System. We similarly omit Canada because Canadian monetary policy causes Canadian interest rates to be highly correlated with USD rates to prevent major changes in the value of the Canadian dollar relative to the U.S. dollar.

For the seven countries, the estimated bivariate interest-rate correlations range from -0.021 (Japan and Sweden) to 0.663 (Sweden and Italy). While the average estimated correlations are substantially positive, these estimates do not constitute strong evidence of cross-sectional correlation. Because short rates are highly persistent stochastic processes, substantial positive correlation can arise spuriously. For example, we conducted a Monte Carlo experiment drawing 2,000 pairs of independent interest-rate series of length 297 from our estimated model. The 95% quantile for these correlation coefficients was 0.675, and the 99% quantile was 0.795. Furthermore, we cannot reject the hypothesis that the correlations are jointly zero in a GMM test based on the small-sample distribution from our model.

5. Estimates of the regime-switching model

The basic regime-switching model has 20 parameters: $\{\mu_i, \beta_i, \sigma_i, \gamma_i, a_{ii}, b_{ii}, a_{23}, b_{23}\}$, $i = 1, 2, 3$. We estimate the model using a cross-sectional extension of Gray's (1995) recursive maximum-likelihood procedure.[9] Because we identify higher numbered regimes with higher within-regime interest rate means, we estimate $\mu_i/(1-\beta_i)$, rather than estimating μ_i directly, and we constrain $\mu_i/(1-\beta_i)$ to be increasing in i. This approach only makes sense if we constrain $\beta_i \leqslant 1$. We impose the stronger constraint that β_i lie in the interval $[-1, 1]$. Ang and Bekaert (1998) show that this constraint, along with the condition that

[9] The appendix of Bekaert et al. (1997c) contains a detailed, self-contained, description of the estimation method.

$\beta_i \in (-1, 1)$ for at least one i, is sufficient for stationarity of a regime-switching process. The parameters μ_i are then recovered from the $\mu_i/(1-\beta_i)$ estimates along with the estimates of β_i. Note that μ_i must equal zero if $\beta_i = 1$.

Table 2 displays our parameter estimates. Note first that $\beta_1 = 1$. This presents two problems. First, $\beta_1 = 1$ implies that μ_1 is an exact function of β_1, so the variance–covariance matrix of the parameter estimates is singular. Furthermore, our estimated parameter vector is on the boundary of the parameter space, so the usual procedure for computing maximum likelihood standard errors (which uses first-order conditions of the likelihood maximization) is inapplicable. The standard errors reported in Table 2 are computed by fixing $\mu_1 = 0$ and $\beta_1 = 1$. These standard errors ignore parameter uncertainty about μ_1 and β_1. This is not a major problem for our purposes, since our primary use of the estimates is to calibrate the regime-switching model. In

Table 2
Estimates of parameters of the regime-switching model[a]

Parameter	Estimate (SE)	Parameter	Estimate (SE)
μ_1	0.0000	a_{11}	0.6828
	(—)		(0.4713)
μ_2	0.0299	b_{11}	0.0475
	(0.0401)		(0.0484)
μ_3	2.0066	a_{22}	0.9239
	(0.7924)		(0.7509)
β_1	1.0000	b_{22}	0.1337
	(—)		(0.0911)
β_2	0.9891	a_{23}	−2.1740
	(0.0050)		(0.8424)
β_3	0.8506	b_{23}	0.2025
	(0.0618)		(0.0897)
σ_1	0.0361	a_{33}	−1.0386
	(0.0139)		(1.1167)
σ_2	0.1890	b_{33}	0.1806
	(0.0507)		(0.1035)
γ_3	1.1315		
	(0.4689)		
γ_1	0.3869		
	(0.1427)		
γ_2	0.3882		
	(0.1112)		
γ_3	0.1274		
	(0.1529)		

[a] Note: The table reports estimation of Eqs. (5)–(9) using monthly data on the 3-month interest rate for seven countries, Australia, Germany, Italy, Japan, Sweden, the United Kingdom, and the United States, from 1972:01 through 1995:12. Maximum likelihood robust standard errors (see White, 1982) with μ_1 and β_1 fixed at 0 and 1, respectively, are in parentheses.

particular, we do not use the standard errors to conduct formal hypothesis tests of parameter values. We include these standard errors in the table as a guide to how much information the data contain about the model's parameters.

The parameter estimates in Table 2 capture a number of appealing features. The estimates of $\mu_i/(1-\beta_i)$ are 0.00%, 2.74%, and 13.43% for regimes 1, 2, and 3, respectively.[10] The degree of mean reversion is also increasing in i. Regime 1 is a random walk, while the regime-specific autocorrelation for regime 3 (β_3) is around 0.85.[11] Although a random walk is nonstationary, our estimated interest-rate process is stationary. The estimate of b_{23} is positive, so persistent high interest rates eventually lead to a switch into the high-mean regime, which exhibits substantial mean reversion. In fact, in a simulation of 100,000 observations from our model, the maximum interest rate observed is only 23.96%. In an analogous experiment with a unit-root process, the maximum interest rate is above 200%.

The conditional volatility of interest rates within a regime also is increasing in the regime. In particular, when we compute the mean value of the conditional volatility function $h_i(r_t) \equiv \sigma_i r_t^{\gamma_i}$ conditional on regime i prevailing at date t, we find that the average conditional volatilities are 0.08, 0.42, and 1.54 for regimes 1, 2, and 3, respectively. Note also that γ_1 and γ_2 (but not γ_3) are within two standard errors of 0.5, the "square root" process assumed by Cox et al. (1985).

The point estimates for b_{22}, b_{23}, and b_{33} are all positive. These signs imply that the transition probabilities depend on the levels of the interest rates in intuitively plausible ways. In particular, $b_{33} > 0$ indicates that if the current regime is regime 3, the probability of remaining in this regime is increasing in r_t. Also, the positive point estimate of b_{23} indicates that if the economy is currently in regime 2, the probability of switching into regime 3, the higher mean, higher volatility regime is increasing in r_t. The positive point estimate of b_{22} also indicates that if the economy is currently in regime 2, the probability of staying in this regime is increasing in r_t. These latter two features imply that the probability of switching from regime 2 to regime 1 is decreasing in the interest rate.

Table 3 provides some diagnostic statistics on the three regime model. We first report the mean values of the conditional standard deviations of interest-rate innovations within each regime and for each of the seven countries. The values across countries are reasonably uniform as would be expected if the same model is appropriate. Table 3 also reports estimates of the fraction of

[10] For $\beta_i < 1$, $\mu_i/(1-\beta_i)$ can be interpreted as the within-regime mean. This interpretation is not valid for regime 1, since the estimated $\beta_1 = 1$.

[11] Mankiw and Miron (1986) and McCallum (1994) argue that interest-rate smoothing by monetary authorities induces high persistence.

Table 3
Model diagnostics[a]

	Australia	Germany	Italy	Japan	Sweden	U.K.	U.S.
Mean CV 1	0.008	0.006	0.009	0.005	0.007	0.008	0.006
Mean CV 2	0.211	0.156	0.261	0.141	0.202	0.211	0.160
Mean CV 3	2.271	2.055	2.442	1.968	2.239	2.278	2.073
Switches 1–2 or 2–1	25	41	21	30	24	25	26
Switches 2–3 or 3–2	24	26	22	6	32	32	8
% time 1	0.29	0.28	0.23	0.34	0.19	0.16	0.19
% time 2	0.34	0.62	0.57	0.53	0.69	0.67	0.70
% time 3	0.37	0.10	0.20	0.13	0.12	0.17	0.11

[a] Note: The table reports the mean conditional variance (CV i) in regime i, $i = 1, 2, 3$. Switches 1–2 or 2–1 (2–3 or 3–2) denote the estimated number of switches between regimes 1 and 2 (2 and 3). The estimated percentage of the time that each country spent in regime i is denoted % time i. Regimes are classified using the smoothed regime probabilities, prob$[s_t = i|I_T]$, $i = 1, 2, 3$, where I_T denotes the set of all available short interest rate data. The regime assigned to date t is that which has the highest smoothed regime probability.

time spent in each regime for each of the seven countries.[12] For comparison, the unconditional probabilities of the three regimes, computed by simulating the model for 200,000 time periods, are 26.0% for regime 1, 59.1% for regime 2, and 14.9% for regime 3. One sense in which there would be a peso problem for a particular currency is if the estimated fraction of time spent in the various regimes differs from these unconditional probabilities. The parameter estimates indicate that Germany and the United States spent too much time in regime 2 and too little time in regime 3, while the United Kingdom spent too much time in regime 3, compared to the unconditional probabilities of these regimes implied by our estimates. While these differences are suggestive that peso problems may be able to explain the differences across the countries that were documented above, the peso problem does not appear severe. The number of switches between regimes in a given time interval is another possible indicator of peso problems. For each currency, Table 3 also reports the number of switches between regimes 1 and 2, and between regimes 2 and 3. Here the USD looks somewhat different from the GBP and the DEM.

More generally, we find that the patterns of regime switches vary greatly among the seven countries. Italy, Sweden, and the U.K. move in and out of the high-rate regime intermittently throughout the sample. In contrast, Australia enters this regime around 1980, and remains there throughout the decade. Germany is in the high-rate regime very infrequently; Japan's experience with

[12] The true regime is, of course, unobservable. The recursive maximum likelihood algorithm delivers an estimate of the probability (given the information in the sample) that a given country is in a particular regime at each data point. We allocate each data point to the regime that is estimated to be most likely.

this regime largely coincides with the periods following the two big oil shocks; and, with the exception of one month in 1974, the U.S. is in the high-rate regime only during the 1979–1982 period of monetary targeting. These results show that the regime switches are not highly correlated across countries, which indicates that our assumption of cross-currency independence does not do substantial violence to the data.

6. Monte Carlo explorations of the peso problem under the expectations hypothesis

6.1. Monte Carlo methodology

To evaluate whether small-sample problems can explain the data, we must compute the long-maturity yields implied by the EH in the context of our regime-switching model. This, in turn, requires us to compute expected future short rates. Since the model is highly nonlinear, we compute the expected future short rates using a Markov chain approximation to the estimated regime-switching model with a grid of 550 points on the space of possible realizations of the short rate in each regime (implying 1,650 possible discrete states).[13] The resulting Markov chain approximation to the law of motion given by Eqs. (5)–(9) is highly accurate. In particular, when we re-estimate the model with 100,000 observations simulated using the Markov chain approximation, all point estimates are within two standard errors of the estimates given in Table 2, and all but two parameters (a_{23} and b_{23}) are within one-half standard error of their Table 2 values (where we use the asymptotic standard errors reported in Table 2).

We then derive the small-sample distributions of the slope coefficients in the Campbell and Shiller (1991) regressions under the assumption that the short rates are generated by the estimated regime-switching model. Specifically, we simulate the estimated model of the three-month short rate to create an artificial time series of 297 observations, we compute the long yields for 12-, 36-, and 60-month bonds implied by the EH under the estimated law of motion of our model, and we recompute the statistics reported in Table 1. This exercise is replicated 200,000 times to construct a small-sample distribution of the estimators for each of the statistics.

6.2. Monte Carlo results

A summary of the Monte Carlo exercise is reported in Table 4. In order to re-consider the evidence on the EH across the three currencies, Table 4 reports

[13] See Bekaert et al. (1997c) for details of our discretization method.

Table 4
Monte Carlo distributions of OLS slope coefficients under the expectations hypothesis using the regime-switching model as the data-generating process[a]

n	Mean	Median	σ	0.5%	2.5%	5%
Panel A: Eq. (3)						
12	1.965	1.976	1.544	−3.308	−1.741	−0.802
36	2.691	2.427	1.861	−4.188	0.087	0.641
60	2.880	2.483	1.868	−0.450	0.579	0.822
Panel B: Eq. (4)						
12	1.285	1.263	0.559	−0.244	0.200	0.416
36	1.603	1.532	0.650	0.104	0.539	0.693
60	1.677	1.804	0.707	0.035	0.548	0.723

[a] Note: The Monte Carlo evidence is based on 200,000 replications. The data-generating process is the discretized, regime-switching model based on the parameters reported in Table 2. There are 297 total observations in each experiment. The columns labelled Mean, Median, σ, 0.5%, 2.5%, and 5% are the sample mean, the median, the standard deviation, and the respective quantiles of the empirical distributions. Panels A and B report statistics from the empirical distributions of the OLS slope coefficients from Eqs. (3) and (4). As noted in footnote 8, the dependent variable $r_{t+3,n-3} - r_{t,n}$ in Eq. (3) is approximated by $r_{t+3,n} - r_{t,n}$.

the 0.5%, 2.5%, and 5% quantiles of the empirical distributions. These quantiles correspond to the relevant critical values for two-sided tests with sizes 1%, 5%, and 10%, respectively, although choice of an appropriate significance level is complicated by the skewness of the small-sample distributions.

When the estimates in Table 1 are evaluated using the small-sample distributions in Table 4, the evidence against the EH appears weaker than when standard asymptotic inference is used. The only rejection at the 1% level is for regression (3) at the 60-month horizon for the USD data. Although substantially negative values occur in the simulations, especially at shorter horizons, the severe positive bias in the estimators makes negative values very unlikely for longer horizons. For regression (4), we can reject the EH at the 5% level in only two cases (the 12- and 36-month maturities with USD data); there are no rejections at the 1% level.

While Table 4 summarizes the appropriate small-sample marginal distributions of the slope coefficients of the two Campbell–Shiller regressions, a more powerful test of the EH focuses on the joint distribution of these two statistics. Panels A–C of Fig. 1 display one-sided bivariate significance bounds for the two estimates (α_1, δ_1) for the 12-, 36-, and 60-month horizons. The marginal significance levels are 0.5% (solid line), 2.5% (dash–dot line), and 5% (dashed line). The solid line indicates the locus of points such that 0.5% of the 200,000 Monte Carlo experiments had estimates of the pair (α_1, δ_1) in the region to the southwest of the point. Note that these significance bounds asymptote to the

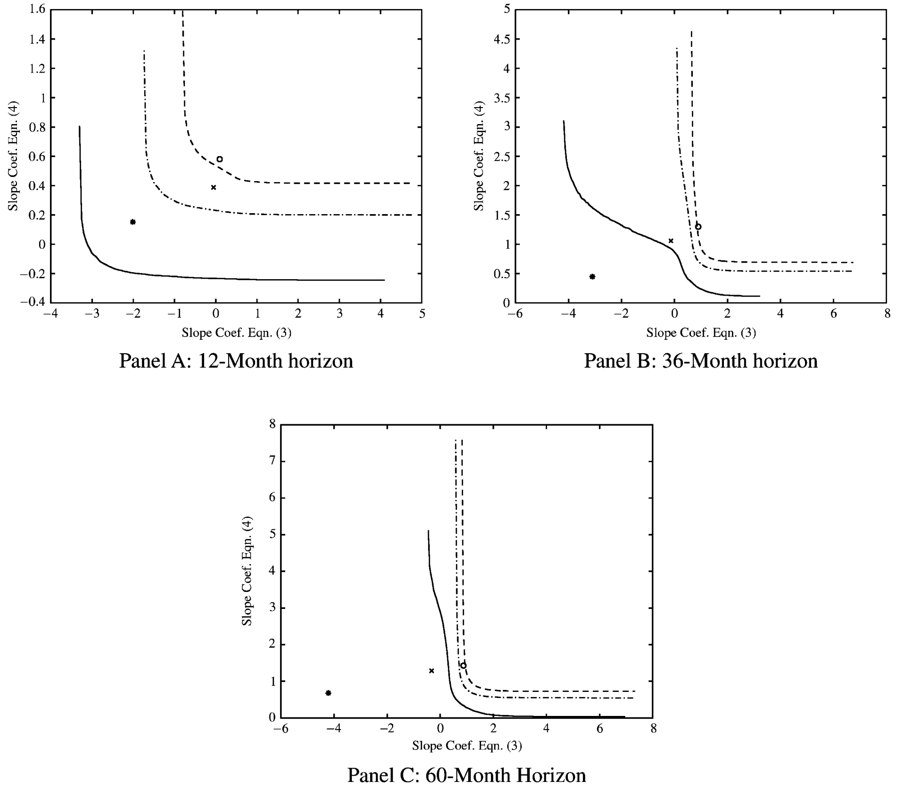

Fig. 1. Bivariate significance boundaries for the expectations hypothesis model. We simulate the model of Section 5, in which the short rate is generated by the three-regime switching process given by Eqs. (5)–(9), using the parameter estimates reported in Table 2, and the long rates are generated by the expectations hypothesis. We conduct 200,000 independent Monte Carlo simulations with 297 observations. For each simulation, we estimate the slope coefficient α_1 in Eq. (3) and the slope coefficient δ_1 in Eq. (4), with $m = 3$. The three panels report bivariate significance bounds corresponding to the 0.5% (solid line), 2.5% (dash–dot line) and 5% (dashed line). The lines give the locus of points for which the relevant percentage of the 200,000 Monte Carlo experiments had estimates of the pair (α_1, δ_1) to the southwest. Also displayed in these panels are the point estimates of (α_1, δ_1) reported in Table 1 for the USD ("*"), the GBP ("○"), and the DEM ("×").

marginal significance levels reported in Table 4. Also displayed in these figures are the point estimates of (α_1, δ_1) reported in Table 1 for the USD ("*"), the GBP ("○"), and the DEM ("×").

This bivariate test reveals more evidence against the EH than the univariate statistics of Table 4. The EH fares fairly well at the 12-month horizon. However, the significance bounds for the 36- and 60-month horizons display nonconvexities that lead to rejection of the hypothesis. Consider Panel B of

Fig. 1. In USD data, the EH would not be rejected by a one-sided test at the 0.5% significance level when the univariate distributions are used separately. However, the joint test rejects the hypothesis at this significance level. Similarly, the DEM data at the 60-month horizon would not reject this hypothesis at the 0.5% significance level with the univariate tests in Table 4. However, the joint test displayed in Panel C of Fig. 1 rejects the hypothesis in DEM data at this significance level. Not surprisingly, when we test the EH using the joint small-sample distribution of all six test statistics (two coefficients at three different horizons), the EH is rejected at the 0.005%, 0.424%, and 0.073% marginal significance levels for the USD, GBP, and DEM respectively.

A natural question is whether the results in Table 4 are driven by peso problems, modelled as multiple interest-rate regimes, or whether they are driven by other aspects of the model such as persistence or heteroskedasticity. In particular, Bekaert et al. (1997a) show that extreme persistence can induce upward bias and extreme dispersion in the small-sample distributions for the slope coefficients in Eqs. (3) and (4), without introducing multiple regimes. We do find that persistence is essential for the results of Table 4. When we replicate the Table 4 analysis using a three-regime data generating process in which the within-regime process for the short rate is serially uncorrelated, we find that the small-sample distributions display little bias or dispersion.[14] However, persistence alone cannot deliver the results of Table 4. When we repeat the analysis using a single-regime model whose mean, persistence, volatility, and conditional heteroskedasticity mimic that implied by our estimated three-regime model,[15] the distribution is rather different. Unlike Table 4, there is virtually no probability mass in the negative region for either slope coefficient, and the means and standard deviations are decreasing (rather than increasing) in the maturity of the long bond. Furthermore, the standard deviations are lower than in Table 4 for the longer horizons. We conclude that persistence and heteroskedasticity alone cannot account for the key features displayed in Table 4. Multiple regimes play an essential role.[16]

[14] This result obtains whether or not the within-regime process is conditionally heteroskedastic, and whether or not the regime-switching probabilities are state dependent.

[15] Specifically, we simulate our three-regime model at the parameter estimates in Table 2 (sample size 100,000), and use these simulated data to estimate a single-regime AR(1) process with constant elasticity of variance: $r_{t+1,3} = \mu + \beta r_{t,3} + \sigma r_{t,3}^{\gamma} \varepsilon_{t+1}$. The resulting parameter estimates are: $\mu = 0.1683$; $\beta = 0.9809$; $\sigma = 0.2438$; $\gamma = 0.4947$. Detailed results from this experiment can be obtained from the authors.

[16] We also conduct tests for sensitivity of our small-sample distribution to parameter uncertainty. When we draw parameter vectors from the joint parameter distribution and rank the draws according to the value of parameter b_{23}, we find that the draws in the lowest 10th percentile reject the expectations hypothesis slightly more often than do the parameter estimates of Table 2. The effect, however, is not dramatic.

6.3. Can peso problems explain the regression patterns in U.S. data?

To further isolate the importance of peso problems as an explanation of the poor small-sample behavior of the test statistics, we consider situations where the high-mean, high-variance regime is under-represented in the sample relative to the population.[17] In Fig. 2, we order the 200,000 slope coefficients of our experiments on regression (3) in bins of 500 ranked according to size and graph the average slope coefficients from these bins relative to the average frequency of regime 3 from the relevant 500 samples. Fig. 2 illustrates the importance of peso effects in generating the tails of the small-sample distributions. Negative slope coefficients are associated with a frequency of regime 3 below 7%, considerably lower than the average regime 3 frequency of 14.9%.

Fig. 2 suggests that very low slope coefficients for regression (3) are associated with peso-type events. Is it the case, then, that the negative slope coefficients estimated using USD data are due to peso effects? To answer this question, we look at the simulations with slope coefficients near those found in USD data, and we compare the statistical properties of these simulations with the properties of U.S. interest rates. Results for the model of this section are in the first five columns of Table 5. For each long-rate maturity we select the simulations whose estimated slope coefficient is within one Monte Carlo standard error (reported in column 4 of Table 4, Panel A) of the slope coefficient estimated from USD data (reported in column 2 of Table 1, Panel A). Note first that the USD data differ from the implications of the estimated regime-switching model. In particular, the mean and standard deviation of the USD interest rate are lower than implied by the model, and the skewness, kurtosis, and first-order autoregressive coefficient are larger than implied by the model. In addition (as noted above), the estimated frequency of regime 3 and of shifts between regimes 2 and 3 are lower in USD data than in the model. When we look only at the subset of simulations where the slope coefficient of regression (3) is near that estimated in USD data, all of these statistics move in the direction needed to explain USD data: The mean, standard deviation, frequency of regime 3, and frequency of shifts between regimes 2 and 3 all decline, while the skewness, kurtosis, and AR(1) coefficient all rise. For many of these statistics, however, the magnitude of these changes far overshoots what is needed to fit USD data. Consider especially the large declines in the mean and standard deviation of the interest rate and in the frequency of regime 3. We conclude that the peso effect as modelled in this section does not constitute a complete explanation for the low estimates of regression (3) slope coefficients found in USD data.

[17] There are many other types of "peso problems". For example, one could also consider the number of switches from regime 2 to 3 (see Table 3). However, our frequency characterization is highly positively correlated with the "switches" definition.

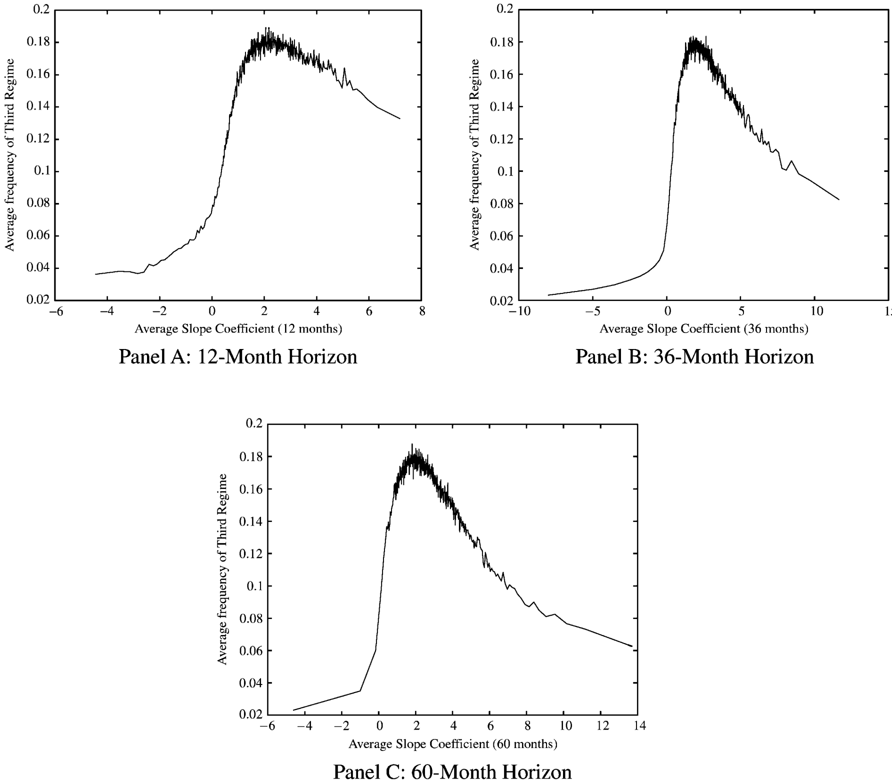

Fig. 2. Peso effects in the expectations hypothesis model. We simulate the model of Section 6, in which the short rate is generated by the three-regime switching process given by Eqs. (5)–(9), using the parameter estimates reported in Table 2, and the long rates are generated by the expectations hypothesis. We conduct 200,000 independent Monte Carlo simulations with 297 observations. For each simulation, we estimate the slope coefficient α_1 in Eq. (3) with $m = 3$. The 200,000 simulations are sorted into bins of 500 each, ranked according to the size of α_1. The graphs in this figure plot the average frequency of regime 3 in each bin against the average value of α_1 in that bin.

Let us return briefly to Fig. 2. Note that peso effects of the kind that we thought would generate negative coefficients, also seem responsible for highly positive slope coefficients. (This is especially apparent in Panel C of Fig. 2, which displays results for the 60-month maturity.) That is, peso problems contribute to the dispersion of the small-sample distributions in both tails. Why are samples with infrequent realizations of the third regime sometimes associated with highly positive slope coefficients in regression (3)? First, when the third regime is under-represented in the sample, short rates are likely to be more persistent (since the third regime displays less persistence than the other two regimes). This effect is apparent from Table 5. This increased persistence

Table 5
Characteristics of simulations with low slope coefficients for Eq. (3)[a]

	USD data	Model	Subsample with low slope for Eq. (3): expectations hypothesis			Subsample with low slope for Eq. (3): one-factor pricing Kernel		
			12 months	36 months	60 months	12 months	36 months	60 months
Mean	7.03	8.92	5.25	3.98	2.84	6.20	4.62	3.40
Std. Dev.	2.78	2.93	1.95	1.64	1.29	2.27	1.98	1.64
Skewness	1.02	0.41	0.45	0.53	0.65	0.46	0.61	0.83
Ex. Kurtosis	1.03	−0.16	−0.09	0.17	0.87	−0.09	0.37	1.22
AR(1) coef.	0.975	0.964	0.970	0.973	0.976	0.970	0.975	0.979
Freq. regime 3	0.10	0.16	0.05	0.03	0.02	0.07	0.04	0.03
Trans. to regime 3	0.027	0.041	0.026	0.020	0.015	0.029	0.021	0.017
Bounds on slope in Eq. (3) for subsample			−3.556	−4.959	−6.079	−4.359	−6.108	−7.178
			−0.468	−1.237	−2.343	0.335	−0.088	−1.244
Fract. simulations in subsample			0.058	0.010	0.0015	0.108	0.017	0.0026

[a] Note: Column 1 reports the mean, standard deviation (both in percent per annum), skewness, excess kurtosis, and first-order autoregression coefficient estimated for the 3-month U.S. Treasury Bill yield (monthly data 1972:01–1996:09), as well as the frequency of regime 3 and the empirical probability of a transition from regime 2 to 3 estimated for these data (as reported in Table 3). Column 2 reports the means of these statistics computed across 200,000 Monte Carlo simulations of the regime-switching model (evaluated at the estimated parameters reported in Table 2), each of length 297. Columns 3–8 report the means of these statistics for the subsample of Monte Carlo simulations for which the slope coefficient of Eq. (3) is within one Monte Carlo standard error of the point estimate reported for USD data in Table 1. (The lower and upper bounds on this slope coefficient used to select the subsample are given in rows 8 and 9; the fraction of Monte Carlo simulations contained in the subsample is given in row 10.) In columns 3–5, the long rates are generated by the expectations hypothesis (as in Table 4), with long bond maturity of 12, 36, and 60 months respectively. In columns 6–8, the long rates are generated by the one-factor pricing kernel with $\lambda = -0.8$ (as in Table 6).

exacerbates the upward bias in small-sample estimators of this slope coefficient. Second, peso effects are associated with an unusually low frequency of the third regime only if there are a substantial number of observations where short-rate shocks significantly drive up the term spread. If these increases in the spread do not occur or they are not of sufficient magnitude, we are simply left with a sample of very persistent short rates and low-variance term spreads. Finally, the derivative of the regime transition probabilities with respect to the short rate is increasing in the level of the short rate. At low interest rates (which are more likely when the high-rate regime occurs infrequently), interest-rate shocks are less variable (see again Table 5) and consequently less likely to generate large changes in term spreads.

This last point is related to a more general problem with this model. A rise in the short rate has two effects on the current spread. The direct effect is to narrow the spread since the short rate enters with a negative sign. The indirect effect is to increase the long rate through an increase in expected future short rates. Peso problems, if present, are generated by this indirect effect. This effect depends on the sensitivity of the transition probability to short-rate changes, and is likely to dissipate for longer horizons. This is why, in Table 4, longer maturity is associated with greater upward bias and fewer negative observations for regression (3) slope coefficients. Unfortunately, this pattern is opposite to that found in USD data, where longer maturities are associated with more negative slope coefficients.

One potential solution to this problem is to assume that agents observe a variable that affects the transition probability but is imperfectly correlated with the current short rate. This extension of the model raises too many technical complexities to be explored within the context of the current paper. A second possibility is to consider alternatives to the EH by allowing for time variation and regime dependence in term premiums. We explore this possibility in the next section.

7. Regime switching and term premiums

7.1. Model and calibration

This section presents an alternative model in which time-varying, regime-dependent term premiums are derived from an extension of the discrete-time affine class of models (see Backus et al., 1997; Campbell et al., 1997). We postulate the following law of motion for the natural logarithm of M_t, the pricing kernel for nominal assets:

$$m_{t+1} = -\left(1 + \frac{\lambda^2}{2}\right) r_t - \lambda \sqrt{r_t} \varepsilon_{t+1}, \tag{10}$$

where $m_t \equiv \ln(M_t)$, λ is a parameter, and $\{\varepsilon_t\}$ is a sequence of independent standard normal random variables. To avoid introducing additional sources of noise into the model, we assume that the disturbance term ε_t in Eq. (10) is identical to the disturbance to r_t in Eq. (5). Note that because $\exp(-r_t) = \mathrm{E}_t[\exp(m_{t+1})]$, the $\{m_t\}$ process correctly prices the short asset.

If there were but a single regime, r_t would be a Gaussian process, and the model given by Eqs. (5) and (10) would fall into the discrete-time affine class. In the multi-regime case, r_t is not Gaussian, so this model does not inherit the analytic tractability of an affine model. Nonetheless, its implications for long-term bond yields can be computed using our discrete Markov chain approximation to the short-rate process given by Eqs. (5)–(9). From Eq. (1)

$$\begin{aligned}\exp(-r_{t,n}n) = \mathrm{E}_t\{&\exp(m_{t+1})\mathrm{E}_{t+1}[\exp(m_{t+2})\mathrm{E}_{t+2}(\exp(m_{t+3})\cdots \\ &\mathrm{E}_{t+n-1}[\exp(m_{t+n})]\cdots)]\}.\end{aligned} \tag{11}$$

The right-hand side of Eq. (11) is straightforward to evaluate because in the discrete state-space approximation $\mathrm{E}_{t+i}[\exp(m_{t+i+1})]$ can only take a finite number of values (one for each state). Once these values have been computed, the right-hand side of Eq. (11) can be evaluated by recursively applying the state transition matrix.

The only additional parameter in this model is λ, which determines the innovation variance of m_t. This parameter can be interpreted as the market price of risk. It determines the average slope of the term structure. (A positively sloped average term structure requires a negative value for λ.)[18] In our data, the average spread between the 3- and 60-month interest rates equally weighted across the three currencies is 0.64%.[19] In simulations of our model at the parameter estimates of Table 2, setting $\lambda = -0.8$ implies an average spread of 0.67%. Since this is quite close to the average spread in the data, we use this value in our simulations.

An additional technical issue involves the timing interval in Eq. (10), which is one month because the ε_t in Eq. (5) is a monthly process. Therefore, the m_t process satisfying Eq. (10) is consistent with the one-month rate. However, because one-month rates for all currencies are unavailable in the 1970s, we estimate Eq. (5) using data for three-month interest rates. To accommodate this slight inconsistency between the timing interval in the model and that in the data, we treat the discretized r_t process generated by our estimates of the model in Eqs. (5)–(9) as if this was the process for the one-month rate. In practice, the distortions thus induced are small, since one-month rates track three-month rates very closely in all countries for which we have data on both

[18] A negative λ implies that a positive innovation to ε_{t+1}, which increases short-term interest rates, also increases m_{t+1}. Since long-term bond prices fall when r_{t+1} rises, long-term bonds are risky assets when $\lambda < 0$. Hence, the average term spread must be positive.

[19] The average spreads are 1.1% for USD, 0.14% for GBP, and 0.68% for DEM.

maturities. We then compute the (discretized) one-month nominal pricing kernel m_t according to Eq. (10), and generate (discretized) n-period rates according to Eq. (11) for n ranging from 3 to 60 months. In our subsequent Monte Carlo analysis of Eqs. (3) and (4), we use the three-month rate generated by Eq. (11) as the short rate. This insures that all interest rates are generated in an internally consistent arbitrage-free manner. It also makes the results of this model comparable to those reported in Section 6, above, since the three-month rate is the short rate in that analysis.

7.2. Monte Carlo results for the time-varying term-premium model

General equilibrium models, and affine models in particular, have not generated population distributions for the term premiums that resolve the Campbell–Shiller puzzles. Our model, given by Eqs. (5)–(11), shares this feature. The second and third columns in Panel A of Table 6 show that our model does not deliver slope coefficients for Eqs. (3) and (4) that are substantially below one in population. Panel A also reports the population moments of the term premiums, the difference between the actual long rate and its value predicted by the EH with term premiums set to zero. The term-premium means are positive and increase with maturity, and the term premium is positively correlated with the level of the short rate. Note that the standard deviation of the term premiums is rather small (varying from 8 basis points at 12 months to 19 basis points at 60 months), so this model represents a quantitatively small departure from the EH.

Panel B of Table 6 reports the results from a Monte Carlo experiment mimicking the experiment underlying the results in Table 4, but for the new data generating process. Introducing rather small variation in term premiums dramatically increases the dispersion of the small-sample distributions for all statistics. The standard deviations increase by over 40% compared to Table 4, and the left tails reach further into the negative region. In particular, the distributions accommodate the increasingly negative slope estimates for Eq. (3) for the USD. Panels A–C of Fig. 3 display one-sided bivariate significance bounds, analogous to those of Fig. 1. The evidence against the model is weak for the 12- and 36-month horizons. There is clear evidence against the model only for the 60-month horizon. Due to the pronounced nonconvexity of the significance bounds in Fig. 3, Panel C, the joint test rejects the model at the 1% level in both USD and DEM data for the 60-month horizon. The joint distribution of the Campbell–Shiller statistics confirms these results. The joint distribution for all six statistics (including the 60-month horizon) rejects the model at the 0.5% significance level for all countries. In contrast, when the 60-month horizon is excluded, the joint distribution of the remaining four

Table 6
Characteristics of time-varying term premiums and Monte Carlo distributions using the regime-switching model with a one-factor pricing kernel as the data-generating process, $\lambda = -0.8$[a]

Panel A: Characteristics of time-varying term premiums

n	Slope coef.		Mean term premium	Standard deviation term premium	Correlation between term premium and short interest rate
	Eq. (3)	Eq. (4)			
12	1.829	1.138	0.141	0.083	0.883
36	1.909	1.136	0.401	0.156	0.970
60	1.853	1.116	0.567	0.185	0.967

Panel B: Monte Carlo distributions of OLS slope coefficients

n	Mean	Median	σ	0.5%	2.5%	5%
Eq. (3)						
12	2.336	2.530	2.347	−6.509	−3.978	−2.390
36	3.472	3.100	3.010	−11.548	−0.439	0.827
60	3.787	3.081	2.967	−1.287	0.741	1.030
Eq. (4)						
12	1.485	1.505	0.818	−1.251	−0.446	0.012
36	1.947	1.797	0.944	−0.528	0.547	0.783
60	2.017	1.802	1.082	−0.054	0.614	0.818

[a] Note: Panel A reports statistics on the term premium, which is the difference between the actual long yield and the yield predicted by the expectations hypothesis. The statistics are population values, that is, they are computed directly from the Markov chain approximation to the regime-switching model and the corresponding pricing kernel. Panel B presents the Monte Carlo evidence for Eqs. (3) and (4) based on 200,000 replications. There are 297 total observations in each experiment. The columns labelled Mean, Median, σ, 0.5%, 2.5%, and 5% are the sample mean, the median, the standard deviation, and the respective quantiles of the empirical distributions. As noted in footnote 8, the dependent variable $r_{t+3,n-3} - r_{t,n}$ in Eq. (3) is approximated by $r_{t+3,n} - r_{t,n}$.

statistics implies marginal significance levels of 1.0%, 2.3%, and 1.7% for the USD, GBP, and DEM, respectively.[20]

How can the introduction of rather small term premiums substantially alter these small-sample distributions? The reason is that variable term premiums considerably exacerbate the peso effects described above through the

[20] Lower absolute values for λ cause a deterioration in this model's performance, while higher absolute values make it more difficult to reject the model. In particular, when we set λ to -0.6 (implying a spread between 3-month and 5-year yields of 0.49% per annum), the model is rejected at the 0.5% marginal significance level (one-tailed test) using USD data with a 36-month maturity for the long bond. (Inference is not substantially changed for the other countries and horizons.) When λ is set to -1.0 (implying a spread of 0.86% per annum), the model is no longer rejected at any conventional significance level using USD data at the 12-month maturity. (Inference is not substantially changed for other horizons or countries.)

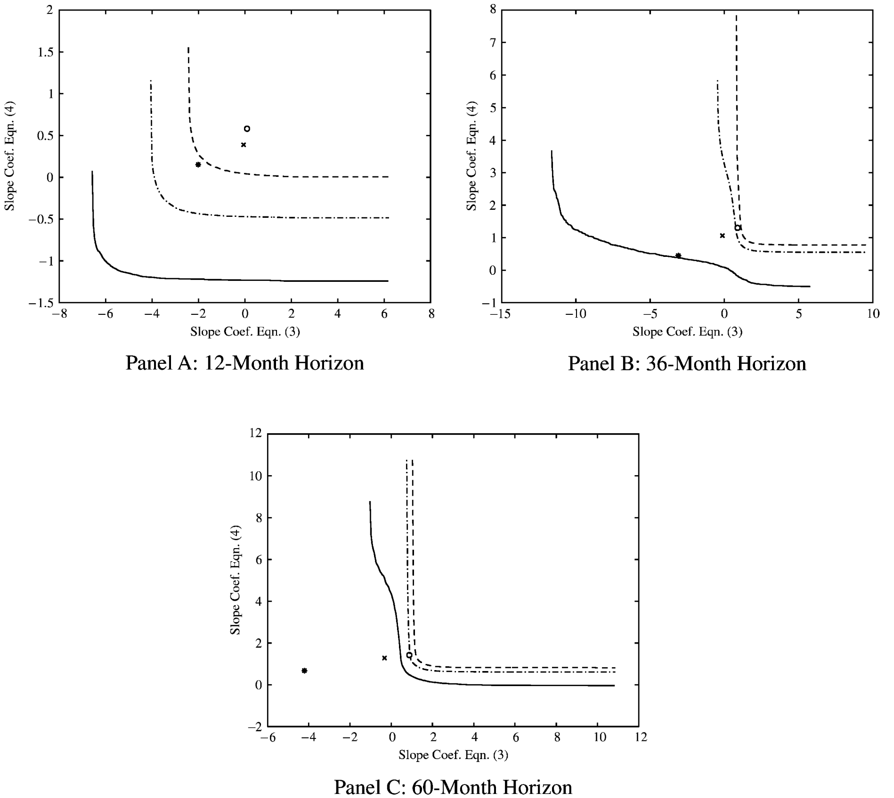

Fig. 3. Bivariate significance boundaries for the time-varying term-premium model. We simulate the model of Section 7, in which the short rate is generated by the three-regime switching process given by Eqs. (5)–(9), using the parameter estimates reported in Table 2, and the long rates are generated by the time-varying term-premium model, as given in Eqs. (10) and (11) with $\lambda = -0.8$. We conduct 200,000 independent Monte Carlo simulations with 297 observations. For each simulation, we estimate the slope coefficient α_1 in Eq. (3) and the slope coefficient δ_1 in Eq. (4), with $m = 3$. The three panels report bivariate significance bounds corresponding to the 0.5% (solid line), 2.5% (dash–dot line), and 5% (dashed line). The lines give the locus of points for which the relevant percentage of the 200,000 Monte Carlo experiments had estimates of the pair (α_1, δ_1) to the southwest. Also displayed in these panels are the point estimates of (α_1, δ_1) reported in Table 1 for the USD ("*"), the GBP ("○"), and the DEM ("×").

correlation of the short rate and the term premiums. As before, a positive shock to interest rates has a direct negative effect on the term spread, but, now, the indirect positive effect on spreads through expected regime changes is amplified by the positive correlation between the term premiums and the short rate presented in the last column in Panel A of Table 6. Hence, some samples may indeed experience dramatic peso effects, where large increases in spreads are not followed by corresponding increases in short rates. The effect on the distribution of the slope coefficients is most pronounced at the 36-month horizon, where the correlation between the short rate and the term premium is highest, the standard deviation of the term premium is quite high, and the indirect effect is not fully attenuated. To illustrate this phenomenon, Fig. 4 presents analyses for the model of this section analogous to those displayed in Fig. 2. The results show that extreme negative values of the estimated slope coefficient for Eq. (3) are unambiguously associated with those samples that have few realizations of the high-rate regime.

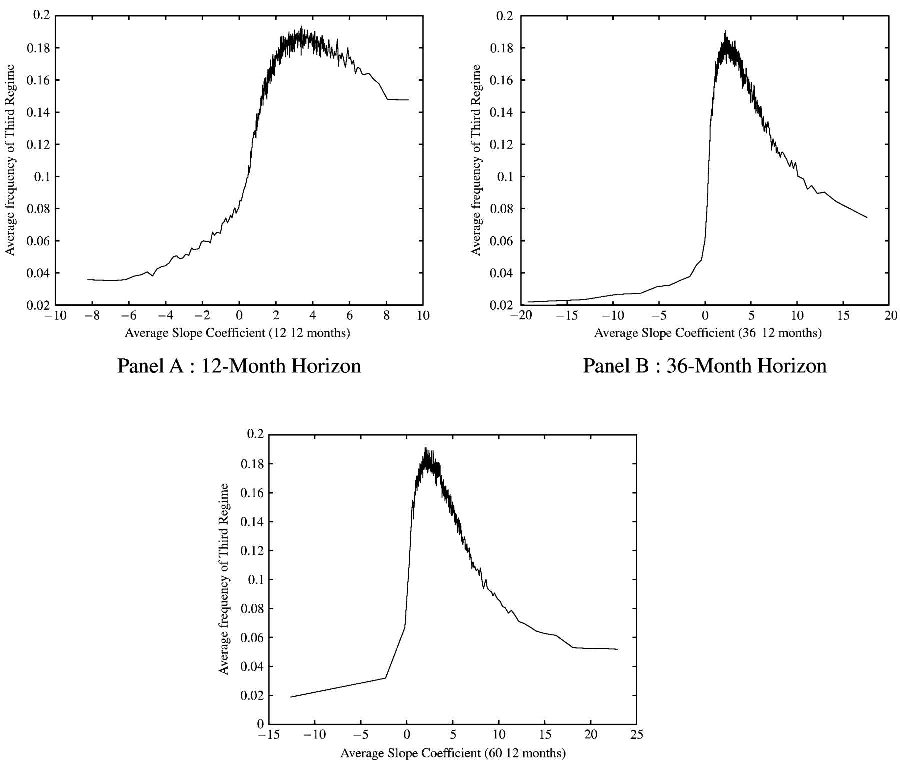

Fig. 4. Peso effects in the time-varying term-premium model. We simulate the model of Section 7, in which the short rate is generated by the three-regime switching process given by Eqs. (5)–(9), using the parameter estimates reported in Table 2, and the long rates are generated by the time-varying term-premium model, as given in Eqs. (10) and (11) with $\lambda = -0.8$. We conduct 200,000 independent Monte Carlo simulations with 297 observations. For each simulation, we estimate the slope coefficient α_1 in Eq. (3) with $m = 3$. The 200,000 simulations are sorted into bins of 500 each, ranked according to the size of α_1. The graphs in this figure plot the average frequency of regime 3 in each bin against the average value of α_1 in that bin.

We conclude this section by asking whether peso effects coupled with the one-factor model can mimic USD data. The last three columns of Table 5 are analogous to columns 4–6, except that the one-factor model is assumed. The simulations whose regression (3) slope coefficient for the 12-month long rate is near that estimated from U.S. interest-rate data (column 6 of Table 5) more closely resemble USD data than when the EH is assumed. In particular, the mean, standard deviation, AR(1) coefficient, and frequency of shifts between regimes 2 and 3 are all tolerably close to the data. (The frequency of regime 3 is 7% for these simulations, compared with the USD estimate of 11%.) However, the subsamples whose slope coefficients are near the USD estimates using longer-maturity interest rates have short-rate means considerably lower than the value of 7.03% estimated from USD data. On the whole, the results of Table 5 argue against the peso problem as a complete explanation of the negative slope coefficients estimated for Eq. (3) in USD data.

8. Conclusions

In this paper, we ask whether the Campbell and Shiller (1991) term structure anomalies may be due to a generalized peso problem in which a high-interest-rate regime occurred less frequently in the sample of USD data than was rationally anticipated. We formalize this idea as a regime-switching model of short interest rates estimated with data from seven countries. Technically, this model extends recent research on regime-switching models with state-dependent transitions to a cross-sectional setting. The regime-switching model reveals the existence of a high-mean, high-variance regime in which short rates are much more mean reverting than in the two more "normal" regimes.

When we conduct inference with the small-sample distributions generated by the regime-switching model, the evidence against the EH weakens considerably. Notable features of the distributions are a substantial upward bias and much larger dispersion than the asymptotic distributions. Whereas Bekaert et al. (1997a) show that such features are also present when the data-generating process does not involve regime-switching induced peso effects, we demonstrate that peso effects contribute to the dispersion of the distribution. Nevertheless, the USD evidence remains somewhat anomalous, making it implausible that this is the correct data-generating process.

A better reconciliation of the data for all countries with our regime-switching model is achieved when we allow for a small, time-varying term premium. When we allow regime changes and other interest-rate movements to be priced as time-varying term premiums, the small-sample distributions become much more dispersed, skewed and biased in the direction of explaining the data. However, this model still cannot explain the regression results for the longest horizon, and simulations that match these regression results differ from USD data in a number of ways.

This last result has important implications for statistical inference. The population distribution implied by the model with time variation in the term premium is inconsistent with the data. Yet the implications of this model are far less at variance with the data because of the strong way the time-varying term premium interacts with peso effects in small samples. Thus, correct tests of this model require small-sample inference. This lesson applies well beyond the simple model examined in this paper. It is perhaps no surprise that a single-factor model is incapable of reconciling the behavior of short rates with long rates of all maturities; the message of this paper for econometric analysis of richer models of the term structure is that peso effects interact strongly with the model's dynamics and cannot be disregarded.

These results raise many questions for further research. The effects of combining peso problems with time-varying risk premiums are dramatic. While the model developed here is only illustrative, we believe it deserves further exploration. Future work should use additional information from the term structure and from macroeconomic processes such as inflation. For example, Evans (1998a, b) finds that the real term structure in the U.K. is well modeled by a single regime, while different inflation regimes cause the nominal term structure to be regime dependent. Introducing variables not perfectly correlated with the short rate (for example term spreads) in the transition probabilities also constitutes an important extension, but these extensions of our model raise many challenging technical issues.

Finally, the sensitivity of the Campbell–Shiller regressions to small-sample problems raises the econometric issue of developing estimators less prone to severe small-sample biases. This is all the more pressing since many important financial theories are tested with similar regressions. A good example is the unbiasedness hypothesis tests in foreign exchange, which suffer from similar problems (see Baillie and Bollerslev, 2000).

Appendix A. Data

Our data set for USD, GBP, and DEM zero-coupon bond yields updates the data originally used by Jorion and Mishkin (1991). We thank Philippe Jorion for generously providing us with his original data, which consist of monthly observations from 1972:01 through 1991:12 on implied zero-coupon, government bond yields for maturities of 3, 12, 24, 36, 48, and 60 months. Data from 1990:1 to 1996:9 for maturities of 3, 12, 36, and 60 months for the three currencies were obtained from a New York investment bank that wishes to remain anonymous.

The short rates used to estimate the regime-switching model are from a variety of sources. For the US, we use the three-month Treasury Bill data. For Germany, we use the three-month interbank rate from the Bank for International Settlements (BIS) database. A missing observation at 1995:12 for Germany was replaced by a linear interpolation between observations from 1995:11 to 1996:01. For the U.K., we use three-month Treasury Bill data from the BIS. A missing observation at 1995:12 for the U.K. was filled in with the observation for that date from the IFS database. The Australian short rate is the three-month Treasury bill rate taken from Datastream from September 1972 onwards; for the first 8 months of 1972 we use the commercial paper rate from the BIS. The Italian short rate is the three-month interbank rate from International Financial Statistics (IFS). The Japanese short rate is the three-month Gensaki rate from the BIS. The Swedish short rate is the three-month Treasury Bill rate from the IFS. Two outliers (at 1983:4 and 1983:5) were confirmed to be data errors by looking at interest series reported in the Sveriges Riksbank Quarterly Review (1983–1984) and were replaced by Eurocurrency rates drawn from Datastream.

References

Ang, A., Bekaert, G., 1998. Regime switches in interest rates. NBER Working Paper No. 6508.

Backus, D.K., 1993. Cox–Ingersoll–Ross in discrete time. Manuscript, New York University.

Backus, D.K., Zin, S.E., 1994. Reverse engineering the yield curve. NBER Working Paper No. 4676.

Backus, D.K., Gregory, A., Zin, S.E., 1989. Risk premiums in the term structure: evidence from artificial economies. Journal of Monetary Economics 24, 371–399.

Backus, D.K., Foresi, S., Mozumdar, A., Wu, L., 1997. Predictable changes in yields and forward rates. Manuscript, New York University.

Baillie, R.T., Bollerslev, T., 2000. The forward premium anomaly is not as bad as you think. Journal of International Money and Finance 19, 471–488.

Bekaert, G., Hodrick, R.J., Marshall, D., 1997a. On biases in tests of the expectations hypothesis of the term structure of interest rates. Journal of Financial Economics 44, 309–348.

Bekaert, G., Hodrick, R.J., Marshall, D., 1997b. The implications of first-order risk aversion for asset market risk premiums. Journal of Monetary Economics 40, 3–39.

Bekaert, G., Hodrick, R.J., Marshall, D., 1997c. Peso problem explanations for term structure anomalies. NBER Working Paper No. 6147.

Campbell, J.Y., Shiller, R.J., 1991. Yield spreads and interest rate movements: a bird's eye view. Review of Economic Studies 58, 495–514.

Campbell, J.Y., Lo, A., MacKinlay, C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton.

Cox, J., Ingersoll, J., Ross, S., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–408.

den Haan, W., 1995. The term structure of interest rates in real and monetary economics. Journal of Economic Dynamics and Control 19, 909–940.

Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Mathematical Finance 6, 379–406.

Evans, M.D.D., 1996. Peso problems: their theoretical and empirical implications. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics, Vol. 14. Elsevier Science, Amsterdam, pp. 613–646.
Evans, M.D.D., 1998a. Real rates, expected inflation, and inflation risk premia. Journal of Finance 53, 187–218.
Evans, M.D.D., 1998b. Regime shifts, risk, and the term structure. Manuscript, Georgetown University.
Evans, M.D.D., Lewis, K.K., 1994. Do risk premia explain it all? Evidence from the term structure. Journal of Monetary Economics 33, 285–318.
Evans, M.D.D., Lewis, K.K., 1995. Do expected shifts in inflation affect estimates of the long-run Fisher relation? Journal of Finance 50, 225–353.
Fama, E.F., 1984. The information in the term structure of interest rates. Journal of Financial Economics 13, 509–528.
Fama, E.F., Bliss, R., 1987. The information in long-maturity rates. American Economic Review 77, 680–692.
Fisher, M., Gilles, C., 1996. Term premia in exponential-affine models of the term structure. Manuscript, Board of Governors of the Federal Reserve System.
Gray, S.F., 1995. An analysis of conditional regime-switching models. Manuscript, Duke University.
Gray, S.F., 1996. Modelling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics 42, 27–62.
Hamilton, J.D., 1988. Rational expectations analysis of changes in regime: an investigation of the term structure of interest rates. Journal of Economic Dynamics and Control 12, 385–423.
Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.
Jorion, P., Mishkin, F., 1991. A multicountry comparison of term-structure forecasts at long horizons. Journal of Financial Economics 29, 59–80.
Lewis, K.K., 1989. Changing beliefs and systematic rational forecast errors with evidence from foreign exchange. American Economic Review 79, 621–636.
Lewis, K.K., 1991. Was there a "peso problem" in the U.S. term structure of interest rates: 1979–1982? International Economic Review 32, 159–173.
Mankiw, N.G., Miron, J.A., 1986. The changing behavior of the term structure of interest rates. Quarterly Journal of Economics 101, 211–228.
McCallum, B.T., 1994. Monetary policy and the term structure of interest rates. Manuscript, Carnegie-Mellon University.
Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Roberds, W., Whiteman, C.H., 1999. Endogenous term premia and anomalies in the term structure of interest rates: explaining the predictability smile. Journal of Monetary Economics 44, 555–580.
Sola, M., Driffill, J., 1994. Testing the term structure of interest rates using a vector autoregression with regime switching. Journal of Economic Dynamics and Control 18, 601–628.
Stambaugh, R., 1988. The information in the forward rates: implications for models of the term structure. Journal of Financial Economics 21, 41–69.
Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 339–341.
White, H., 1982. Maximum likelihood estimation of misspecified models. Econometrica 50, 1–25.