An Exploration of the Forward Premium Puzzle in Currency Markets

Ravi Bansal
Duke University

A standard empirical finding is that expected changes in exchange rates and interest rate differentials across countries are negatively related, implying that uncovered interest rate parity is violated in the data. This article provides new empirical evidence that suggests that violations of uncovered interest rate parity, and its economic implications, depend on the sign of the interest rate differential. A framework related to term structure models is developed to account for the puzzling relationship between expected changes in exchange rates and interest rate differentials. Estimation results suggest that a particular term structure model can account for the puzzling empirical evidence.

According to the hypothesis of uncovered interest rate parity, expected changes in the nominal exchange rate should be positively related to the difference in the nominal interest rate across countries. In particular, this hypothesis implies that the slope coefficient from the regression of the change in exchange rate on the interest rate differential should be one. The forward premium puzzle refers to the well-documented empirical finding that the slope coefficient from this regression is significantly negative [see Bilson (1981), Cumby and Obstfeld (1984), Fama (1984), Hansen

I thank Mark Carhart, John Coleman, Silverio Foresi, Steve Gray, Cam Harvey, Robert Hodrick, David Hsieh, Ravi Jagannathan (the editor), Nancy Keeshan, Bruce Lehmann, Ken Singleton, George Tauchen, S. Viswanathan, and the anonymous referees for many helpful comments. Address correspondence to Ravi Bansal, Fuqua School of Business, Duke University, Durham, NC 27708.

The Review of Financial Studies Summer 1997 Vol. 10, No. 2, pp. 369–403

and Hodrick (1983), Hodrick and Srivastava (1986), and Hsieh (1984)]. This empirical finding can be interpreted as evidence of time-varying forward risk premia.

The forward premium puzzle poses a formidable challenge since a wide array of economic models cannot explain this empirical finding. Mark (1988) explores the implications of the static CAPM, while Backus, Gregory, and Telmer (1993), Bansal et al. (1995), Bekaert (1996), and Hodrick (1989) explore the ability of consumption and money-based general equilibrium models developed in Lucas (1982) to explain the forward premium puzzle. In these models the forward risk premia is time varying and is determined by aggregate consumption and inflation risks across countries. Evidence contained in these articles suggests that the above models fail to explain the forward premium puzzle.

An alternative class of models that can be used to address the forward premium puzzle are term structure models. Backus, Foresi, and Telmer (1995) and Nielson and Saá-Requejo (1994) explore the implications of a two-country version of the term structure model of Cox, Ingersoll, and Ross (1985), for the forward premium puzzle. In general, nominal term structure models capture interest rate risks that are relevant in explaining nominal bond returns and movements of the nominal term structure. An implication of these models is that the forward risk premia and expected continuous change in exchange rates depend on the interest rates across countries. This property is quite attractive because, in the data, the forward risk premia and expected continuous change in exchange rate seem to be primarily related to interest rate movements across countries. Hence, *a priori*, it seems quite plausible that a particular term structure model may provide an explanation for the forward premium puzzle.

This article explores this possibility and makes a contribution to the literature related to the forward premium puzzle. First, it provides new empirical evidence that deepens the forward premium puzzle. This evidence shows that violations of uncovered interest rate parity systematically depend on the sign of the interest rate differential across countries. Second, the article develops a framework that provides sufficient conditions under which the forward risk premia and expected continuous change in expected exchange rate are primarily related to risks contained in default-free nominal interest rates movements.

The focus of this article is on countries that form the key currency blocks—the United States, Germany, and Japan—although additional evidence from other countries is also provided. The new empirical evidence documented is that the slope coefficient of the linear projection of changes in exchange rate on the interest rate differential (U.S. interest rate minus foreign interest rate) depends on the sign of the interest differential. When the interest differential is positive, the slope coefficient is negative and the hypothesis of uncovered interest rate parity is sharply rejected. However, when the interest differential is negative, the slope coefficient is positive, and for many currencies, insignificantly different from one.

One substantive economic implication of the negative slope coefficient, as shown by Fama (1984), is that the variance of forward risk premia must exceed the variance of expected depreciation. However, the empirical evidence provided in this article indicates that the opposite is often true when the interest rate differential is negative—the variance of expected depreciation must exceed the variance of the forward risk premia.

To provide an explanation of this puzzle, a conceptual framework is developed that relates the forward risk premia and expected exchange rate to risks contained in domestic bond markets. This framework provides sufficient conditions under which the forward risk premium is related to the differences in the market price of interest rate risk (mean excess return per unit of standard deviation) across countries. The framework also nests many existing term structure models, such as Brennan and Schwartz (1979), Cox, Ingersoll, and Ross (1985), (henceforth, CIR), and Vasicek (1977).

The empirical exercise jointly estimates the market price of risk in domestic bond markets and simultaneously imposes restrictions on the drift of the continuous change in exchange rates. The market price of risk is identified by relying on observed returns on default-free nominal bonds and changes in exchange rates. All estimation is conducted using the generalized method of moments (GMM) developed in Hansen (1982) and Hansen and Singleton (1982). Empirical results suggest that a model that incorporates sufficiently high interest rate elasticity of conditional variance of bond returns, especially for the United States, can account for the observed relationships between the expected changes in exchange rates and the interest rate differentials.

The rest of the article is divided into five sections. Section 1 provides details regarding the data and empirical evidence regarding the new forward premium puzzle. Section 2 discusses the conceptual framework. Section 3 provides details regarding specification. Section 4 discusses the estimation methods and the empirical results. Section 5 contains concluding remarks.

1. Data, the Puzzle, and Its Implications

1.1 Data

Weekly data for the United States, Germany, and Japan from January 1981–May 1995 is taken from Datastream. There are 750 observations for each series. All the data are simultaneously collected from London

Table 1
Sample Statistics

Variable	Mean	Std	Max	Min	Skew	Kurt
U.S.-7d	7.95	3.60	20.06	2.93	1.03	4.27
U.S.-6m	8.72	9.09	72.25	−44.88	1.89	13.71
U.S.-12m	9.18	16.13	101.80	−88.85	0.78	10.79
Gr-7d	6.67	2.26	14.00	3.00	0.56	2.48
Gr-6m	7.04	5.83	42.56	−59.33	−0.77	30.00
Gr-12m	7.23	9.74	67.60	−83.86	0.015	18.01
Jp-7d	5.40	1.78	10.38	1.34	−0.28	2.22
Jp-6m	5.72	3.84	21.47	−12.18	0.56	5.45
Jp-12m	6.01	6.91	36.91	−30.10	0.24	6.32
$g1$	2.01	83.47	447.00	−260.7	0.26	4.43
$g2$	5.83	75.65	492.10	−294.1	0.87	6.38
$x1$	1.26	3.56	11.18	−6.62	−0.56	2.54
$x2$	2.52	2.98	13.50	−2.96	1.08	4.65

U.S., Gr (Germany), and Jp (Japan) refer to the countries. $7d$, $6m$, and $12m$ refer to the weekly return on a 7-day, 6-month, and 12-month nominal bond, respectively. $g1$ and $g2$ refer to the continuous change in dollar-mark and dollar-yen exchange rates. $x1$ and $x2$ refer to the 7-day interest rate differential (U.S. interest rate rate minus foreign interest rate) between the U.S. and Germany, and the U.S. and Japan, respectively. All weekly returns and exchange rates changes are multiplied by 5200.

closing Thursday prices and include currency spot prices (nominal exchange rates) and annual Euro-currency nominal interest rates for the maturities of 7 days, 1 month, 3 months, 6 months, and 12 months. These interest rates are used to construct continuous 7-day holding-period returns. Spot currency prices and 1-month interest rates for seven additional countries are also collected to characterize the relation between expected changes in exchange rates and interest rate differentials.

Let r_t be the continuously compounded 7-day U.S. nominal risk-free interest rate and let $\tilde{r}_t$ be the foreign country interest rate between t and $t+1$ (all variables with a tilde refer to variables of the foreign country). Let $x_t = r_t - \tilde{r}_t$ be the 7-day nominal interest rate differential between the U.S. and a foreign country. Also, let e_t be the exchange rate in dollars per unit of foreign currency and $g_{t+1} = \log(e_{t+1}/e_t)$ be the continuous change in exchange rates. The continuous change for the dollar-mark and the dollar-yen exchange rate is $g1_{t+1}$ and $g2_{t+1}$, respectively.

Tables 1 and 2 provide sample information regarding bond returns, interest rate differentials, and exchange rates for all three countries. Figure 1 displays the U.S.–Germany and U.S.–Japan interest rate differentials.

1.2 Expected change in exchange rate (U.S.–Germany and U.S.–Japan)

Let f_t be the time t forward price in dollars for delivery of one unit of foreign currency at $t+1$. Covered interest rate parity implies that the

Table 2
Correlation matrix

Variable	U.S.-7d	U.S.-6m	U.S.-12m	Gr-7d	Gr-6m	Gr-12m	Jp-7d	Jp-6m	Jp-12m	$g1$	$g2$	$x1$	$x2$
U.S.-7d	1.00	0.45	0.25	0.32	0.13	0.07	0.55	0.28	0.16	−0.06	−0.10	0.79	0.86
U.S.-6m		1.00	0.89	0.19	0.41	0.42	0.30	0.31	0.33	0.22	0.12	0.28	0.29
U.S.-12m			1.00	0.11	0.38	0.41	0.20	0.23	0.27	0.25	0.17	0.13	0.11
Gr-7d				1.00	0.43	0.27	0.41	0.27	0.17	−0.04	−0.07	−0.31	0.13
Gr-6m					1.00	0.86	0.16	0.29	0.27	0.17	0.16	−0.11	0.04
Gr-12m						1.00	0.09	0.26	0.28	0.20	0.21	−0.08	0.01
Jp-7d							1.00	0.50	0.29	0.00	−0.03	0.28	0.07
Jp-6m								1.00	0.79	0.12	0.10	0.11	0.07
Jp-12m									1.00	0.17	0.15	0.04	0.02
$g1$										1.00	0.65	−0.05	−0.09
$g2$											1.00	−0.05	−0.12
$x1$												1.00	0.78
$x2$													1.00

U.S., Gr (Germany), and Jp (Japan) refer to the countries. $7d$, $6m$, and $12m$ refer to the weekly return on a 7-day, 6-month, and 12-month nominal bond, respectively. $g1$ and $g2$ refer to the continuous change in dollar-mark and dollar-yen exchange rates. $x1$ and $x2$ refer to the 7-day interest rate differential (U.S. interest rate rate minus foreign interest rate) between the U.S. and Germany, and the U.S. and Japan, respectively.

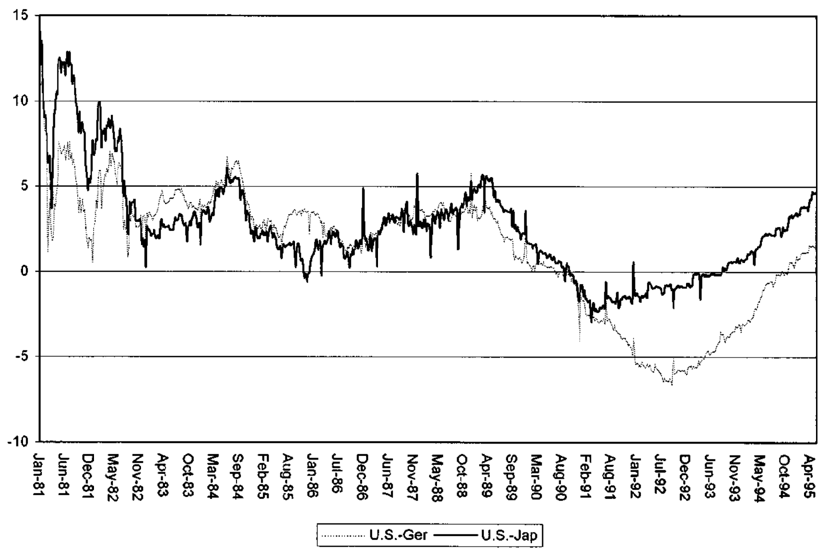

Figure 1
Annualized interest rate differential

forward premium, $\log(f_t) - \log(e_t)$, is equal to the interest rate differential at time t. Following Fama (1984), add and subtract $\log(e_{t+1})$ from the forward premium and take expectations. This implies

$$\begin{aligned} \log(f_t) - \log(e_t) &= E_t[\log(e_{t+1}) - \log(e_t)] + E_t[\log(f_t) - \log(e_{t+1})] \\ &= d_t + p_t = r_t - \tilde{r}_t. \end{aligned} \tag{1}$$

The interest rate differential is equal to the expected exchange rate depreciation, $d_t = E_t[\log(e_{t+1}) - \log(e_t)]$, plus the forward risk premium, $p_t = E_t[\log(f_t) - \log(e_{t+1})]$. The hypothesis of uncovered interest rate parity assumes that the variance of p_t is zero, which implies that d_t is equal to a constant plus the interest rate differential across countries. According to this hypothesis, currencies of high nominal interest rate countries should be expected to depreciate.

A wide array of articles explore the relationship between the exchange rate changes and the interest rate differential [see Bilson (1981), Cumby and Obstfeld (1984), Fama (1984), Hansen and Hodrick (1983), Hsieh (1984), and Hodrick and Srivastava (1986)]. The empirical regularity, well-documented in these articles, is that the slope coefficient from regressing continuous exchange rate changes (g_{t+1}) on the interest rate differential ($r_t - \tilde{r}_t$) is negative and significantly different from one. Empirical results reported in Table 3 (see row 1, Reg_1) confirm this result for the dollar-mark and dollar-yen exchange rates. This empirical finding is inconsistent with uncovered interest rate parity.

The new evidence provided in this article shows that there are important nonlinearities in the relationship between expected exchange rates and interest rate differentials. In particular, the sign and magnitude of the slope coefficient depends on whether the interest rate differential is positive or negative. This empirical finding has important economic implications which are elaborated in Section 1.4.

This new evidence is most starkly illustrated when a dummy-variable method is used to estimate the expected change in exchange rates. Let the variables x^+ and x^- be defined as follows:

$$x_t^+ = \begin{cases} x_t, & \text{if } x_t > 0 \\ 0, & \text{if } x_t < 0. \end{cases} \tag{2a}$$

$$x_t^- = \begin{cases} 0, & \text{if } x_t > 0 \\ x_t, & \text{if } x_t \leq 0. \end{cases} \tag{2b}$$

The variables x^+ and x^- separate the interest rate differential x into periods of positive and negative interest rate differentials, respectively. Table 4 gives summary statistics for x^+ and x^-.

Table 3 (see Reg_2) reports the results from the regression of g_{t+1} on a constant, x_t^+, and x_t^-. Results are similar for both exchange rates. When the interest rate differential is positive, the slope coefficients (standard error) are −5.9 (1.74) and −3.6 (1.12) for the dollar-mark and dollar-yen exchange rates, respectively. Hence the coefficient is negative and significantly different from one. When the interest rate differential is negative, however, the slope coefficient is positive [U.S.–Germany: 4.96 (2.00); U.S.–Japan: 12.32 (5.44)] and the t-ratios for the hypothesis that the slope coefficient is one is about 2.0 for both exchange rates. Hence, while uncovered interest rate parity is rejected in both positive and negative interest rate differential states, the slope coefficient is significantly negative when the interest differential is positive and is significantly positive and bigger than one when the interest differential is negative.[1] Note that the relation between expected

[1] Bilson (1981) also shows that the negative regression coefficient is due to large positive interest rate differentials. However, his analysis uses large and small absolute values of the interest rate differential as the explanatory variable. His approach cannot distinguish between positive and negative interest rate differentials, which is the focus of the above regressions.

Table 3
Currency depreciation; evidence from U.S.–Germany and U.S.–Japan

Currency	GM	JY
Reg_1		
x	−0.87	−2.39
	(0.97)	(0.89)
$100 * R^2$	0.01	0.78
Reg_2		
x^+	−5.98	−3.95
	(1.73)	(1.12)
x^-	4.96	12.32
	(2.00)	(5.44)
$100 * R^2$	1.32	1.52
$eq - test$	−10.95	−16.27
	(3.27)	(6.08)
$n_{\text{obs}}(x^+)$	526	595
$n_{\text{obs}}(x^-)$	224	155
Reg_3		
x	0.10	2.72
	(1.47)	(2.20)
x^2	−62.32	−190.97
	(29.78)	(69.38)
x^3	−188.30	1247.24
	(408.5)	(464.56)
$100 * R^2$	1.25	1.47
Wald_2	17.07	7.58
p-value	0.00	0.02

Variable x is the U.S. 7-day interest rate minus the foreign country 7-day interest rate at date t. The dependent variable in all regressions is the continuous change in the exchange rate between t and $t+1$. GM and JY refer to the dollar-mark and the dollar-yen exchange rates, respectively. Reg_1, Reg_2, and Reg_3 refer to the regression of the continuous change in exchange rate on functions of x. The standard error of each coefficient is in parentheses. All standard errors are constructed using a Newey–West lag length of 4. Reg_1 is the regression on a constant and x. Reg_2 is the regression on a constant, x^+, and x^-. The variables x^+ and x^- are defined as

$$x_t^+ = \begin{cases} x_t, & \text{if } x_t > 0 \\ 0, & \text{if } x_t < 0, \end{cases}$$

$$x_t^- = \begin{cases} 0, & \text{if } x_t > 0 \\ x_t, & \text{if } x_t \leq 0. \end{cases}$$

$eq - test$ tests the hypothesis that the slope coefficient on x^+ is equal to the coefficient on x^- (i.e., the difference in the coefficients is zero). The difference in the slope coefficients and its standard error is reported. The standard error is constructed using the delta method. Reg_3 is the regression on a constant and x, x^2, and x^3. Wald_2 refers to the test of the null hypothesis that the coefficient on x^2 and x^3 is zero. The test statistic is distributed as a χ^2 with two degrees of freedom. p-value refers to the probability value of this test. R^2 refers to the adjusted R^2.

Table 4
Sample statistics

Panel A: U.S. and Germany

Variable	Mean($x1 > 0$)	Mean($x1 \leq 0$)	Std($x1 > 0$)	Std($x1 \leq 0$)
U.S.-7d	9.413	4.593	3.213	1.612
Gr-7d	6.129	7.955	2.296	1.553
$x1$	3.283	−3.416	1.785	1.992
n_{obs}	526	224	526	224

Panel B: U.S. and Japan

Variable	Mean($x2 > 0$)	Mean($x2 \leq 0$)	Std($x2 > 0$)	Std($x2 \leq 0$)
U.S.-7d	8.800	4.719	3.493	1.682
Jp-7d	5.323	5.765	1.728	1.968
$x2$	3.477	−1.045	2.668	0.669
n_{obs}	595	155	595	155

The variable x refers to the 7-day interest rate differential between the United States and the foreign country. Mean and standard deviation of variables are for variables sorted by the state of the interest rate differential—$x > 0$ and for $x \leq 0$. n_{obs} is the number of observations in each state of the interest rate differential. The variables $x1$ and $x2$ refer to the 7-day interest rate differential between the U.S. and Germany, and U.S. and Japan, respectively. Also note that the correlation $\text{corr}_{(x1>0)}(x1_t, g1_{t+1}) = -0.151$ and $\text{corr}_{(x1\leq 0)}(x1_t, g1_{t+1}) = 0.066$. The correlation $\text{corr}_{(x2>0)}(x2_t, g2_{t+1}) = -0.140$ and $\text{corr}_{(x2\leq 0)}(x2_t, g2_{t+1}) = 0.099$.

depreciation and x^- (negative interest rate differential) is positive and, in this sense, is more consistent with the implications of uncovered interest rate parity.

A test to evaluate the equality of the slope coefficient for x^+ and x^- is also conducted. For both exchange rates, this test sharply rejects the null hypothesis that the slope coefficient is equal across x^+ and x^-. The t-ratios for this test are −3.34 (U.S.–Germany) and −2.67 (U.S.–Japan). This rejection implies that the relation between expected depreciation and the interest rate differential is significantly different across x^+ and x^-. Also note that the adjusted R^2 for the two-state regressions are 1.3% and 1.5% for the dollar-mark and the dollar-yen exchange rates, respectively, versus an adjusted R^2 of 0.01% and 0.78% found using the standard linear projection (Reg_1). Hence the incremental contribution of this simple nonlinearity is measurable.

Whereas the above approach uses a dummy-variable method to estimate the expected change in exchange rate, there are reasons to use a smooth nonparametric function. Such an approach provides an alternative test of the presence of nonlinearities and provides a check on the results from the dummy-variable-based approximation. Evidence from this nonparametric regression is also used to restrict the economic models considered in the article. Aït-Sahalia (1995), Conley et al. (1995), and Tauchen (1995) use similar methods to document interesting nonlinearities in the drift of the interest rate process.

The nonparametric function used to estimate the expected change in exchange rate is a cubic polynomial, $\sum_{l=0}^{l=3} \tau_l x_t^l$.[2] Table 3 ($\text{Reg}_3$) contains results of this estimation. Table 3 provides results from the Wald test of the hypothesis that τ_2 and τ_3 are jointly zero. For both exchange rates, this hypothesis is sharply rejected, suggesting that nonlinearities in the expected change in exchange rate are an important aspect of the data. This evidence is identical to that obtained with the dummy-variable approximation.

Note that since the forward risk premia p_t is equal to $[r_t - \tilde{r}_t] - d_t$, it immediately follows that the relation between p_t and the interest rate differential also inherits the nonlinearities documented for the expected change in exchange rate.[3]

1.3 Evidence from other currencies

To evaluate the robustness of the results documented in the key currency blocks of the United States, Germany, and Japan, empirical tests are also conducted for seven additional currencies: British pound, Belgian franc, Canadian dollar, Dutch guilder, French franc, Italian lira, and Swiss franc. The interest rate differential is constructed using 1-month interest rates because the 7-day interest rate is not available for all countries under consideration.[4] Other than this difference, the empirical exercise is identical to the one discussed above.

The empirical evidence for these currencies is provided in Table 5. The slope coefficient from the standard linear regression (see Reg_1) for all currencies other than the French franc and the Italian lira are negative and the hypothesis of uncovered interest rate parity can be rejected.

Results from the projection of the continuous change in the exchange rate on x^+ and x^- (see Reg_2) show that the slope coefficient associated with x^+ for all currencies, except for the Canadian dollar, is negative and significantly different from one.[5] For the Belgian franc,

---

[2] Adding additional terms does not alter the primary result.

[3] For example, if $d_t = \beta_0 + \beta^+ x_t^+ + \beta^- x_t^-$, then $p_t = -\beta_0 + (1 - \beta^+)x_t^+ + (1 - \beta^-)x_t^-$. Since the estimated value of β^+ is negative, it follows that $1 - \beta^+$ is positive, and since the estimated point estimate of β^- is positive and bigger than one, it follows that $1 - \beta^-$ is negative.

[4] For the United States, Germany, and Japan, both the 7-day and the 1-month interest rates are available. Using the 1-month interest rate differential yields the same empirical results as the 7-day interest rate differential. The correlation between the 7-day interest rate differential and the 1-month differential for U.S.–Germany and U.S.–Japan is 0.995 and 0.991, respectively. This evidence suggests that using the 1-month interest rate differential is not likely to make any difference to the empirical results.

[5] This regression for Italian lira was not conducted because there are only five observations when the U.S. interest rate exceeds the Italian interest rate.

Table 5
Currency depreciation; evidence from additional currencies

Currency	BP	BF	CD	DG	FF	IL	SF
Reg_1							
x	−3.19	−1.45	−2.24	−1.54	0.84	0.52	−1.68
	(1.07)	(0.91)	(0.70)	(1.01)	(0.71)	(1.32)	(0.92)
$100 * R^2$	1.13	0.19	0.90	0.27	0.17	−0.08	0.39
Reg_2							
x^+	−6.02	−11.04	−3.83	−7.48	−10.65	na	−3.85
	(1.99)	(2.77)	(3.29)	(1.51)	(2.95)	na	(1.22)
x^-	−2.14	0.02	−2.12	5.63	1.54	na	5.24
	(1.58)	(1.05)	(0.85)	(2.02)	(0.61)	na	(3.86)
$100 * R^2$	1.14	0.97	0.77	2.49	1.94	na	0.96
$eq - test$	−3.88	−11.06	−1.71	−13.10	−12.19	na	−9.10
	(3.03)	(3.25)	(3.80)	(3.01)	(3.06)	na	(4.61)
$n_{\text{obs}}(x^+)$	138	213	72	512	109	5	530
$n_{\text{obs}}(x^-)$	612	537	678	238	641	745	220
Reg_3							
x	−6.52	−3.17	−4.29	−0.99	−3.74	−7.03	0.60
	(2.08)	(1.30)	(1.58)	(1.49)	(1.21)	(4.98)	(1.41)
x^2	−15.95	−64.02	−66.50	−82.10	−45.59	−70.78	−9.10
	(20.23)	(20.07)	(107.7)	(29.66)	(12.79)	(53.77)	(39.82)
x^3	808.53	−244.05	−352.80	64.88	−49.10	−153.45	−271.56
	(580.83)	(67.07)	(1820.2)	(424.30)	(14.84)	(124.49)	(283.91)
$100 * R^2$	1.30	0.79	0.78	2.00	2.47	0.22	0.94
Wald_2	2.40	14.31	1.76	20.05	18.17	1.79	9.30
p-value	0.30	0.00	0.41	0.00	0.00	0.40	0.01
Sample statistics							
mean $-g$	−2.94	0.38	−0.88	1.84	−0.91	−4.06	2.61
std $-g$	85.57	86.01	31.50	82.74	81.84	80.46	90.16
mean $-x$	−2.55	−1.92	−1.77	1.03	−2.96	−5.86	2.56
std $-x$	3.01	3.35	1.42	3.41	5.36	3.56	3.88

The exchange rates under consideration are BP (British pound), BF (Belgian franc), CD (Canadian dollar), DG (Dutch guilder), FF (French franc), IL (Italian lira), and SF (Swiss franc). x refers to the interest rate differential on the 1-month interest rate (U.S. interest rate − foreign interest rate) at date t. g refers to the continuous change in the dollars per unit of foreign currency exchange rate between t and $t+1$. The mean and standard deviation are weekly quantities multiplied by 5200. For an explanation of all other variables and test statistics, see the notes in Table 3.

French franc, and Swiss franc, the slope coefficient on x^- (negative interest rate differential) is positive, and the hypothesis that it is equal to one cannot be rejected. For the Dutch guilder, this slope coefficient is positive and significantly greater than one.

Consider the evidence regarding nonlinearities. The equality test ($eq - test$) and the Wald_2 test show that there is considerable evidence in favor of nonlinearities in the Belgian franc, Dutch guilder, French franc, and Swiss franc. As is the case for the mark and the yen, for these currencies the relationship between expected depreciation and the interest rate differential is significantly different across positive and negative interest rate differential states. For six of the nine total currencies considered, there is a significant difference in the relationship between expected depreciation and interest rate differentials across x^+ and x^-. For the British pound and the Canadian dollar, there is little evidence in favor of nonlinearities.

While the slope coefficient on x^- for the Dutch guilder, German mark, and Japanese yen is positive and significantly greater than one, for the Belgian franc, French franc, and Swiss franc this positive slope coefficient is insignificantly different from one. This implies that for the first group of currencies uncovered interest rate parity is rejected in both x^+ and x^-, whereas it is rejected for the second group of currencies only when the interest differential is positive.

To summarize, for most currencies there are important differences in the relationship between expected depreciation and interest differential that depends on the sign of the interest differential. Typically when the differential is positive, the slope is negative and significantly different from one. However, when the interest differential is negative, the slope coefficient is positive and is either significantly greater than one or is not significantly different from one.

1.4 Economic implications

The assumption of rational expectations implies that g_{t+1} is the sum of d_t and a forecast error which is orthogonal to all information at time t. The assumption of rational expectations, along with the fact that $x_t = d_t + p_t$ [see Equation (1)] implies that $\frac{\text{cov}(g_{t+1}, x_t)}{\text{var}(x_t)}$ is equal to $\frac{\text{cov}(d_t, d_t+p_t)}{\text{var}(d_t+p_t)}$. Note that $\frac{\text{cov}(d_t, d_t+p_t)}{\text{var}(d_t+p_t)} = 1$ if and only if the variance of p_t is zero, that is if and only if uncovered interest rate parity holds.

Fama (1984), in the context of the linear projection, shows that the slope coefficient is also equal to $\frac{\text{cov}(d_t, d_t+p_t)}{\text{var}(d_t+p_t)}$. He also shows that the empirical finding of a negative slope coefficient (in Reg_1) has two implications: first, that the covariance between p_t and d_t is negative, and second, that

$$\text{var}(p_t) > |\text{cov}(p_t, d_t)| > \text{var}(d_t). \tag{3}$$

In particular, note the implication that the variance of the risk premium (p_t) must exceed that of expected depreciation (d_t).

Now consider the economic implications of the new empirical evidence documented above. Consider the slope coefficient $\frac{\text{cov}(d,x)}{\text{var}(x)}$ [or equivalently $\frac{\text{cov}(d,d+p)}{\text{var}(d+p)}$] conditional only on the sign of the interest rate differential. This slope coefficient is equal to the slope coefficient on x^+ when the interest differential is positive, and is equal to the slope coefficient on x^- when the interest differential is negative.[6]

[6] Averaging is done across all interest rate differential states given only the sign of the interest rate differential.

First, consider the evidence when the interest rate differential is positive. In this case the associated slope coefficient is negative and significantly different from one and the implications for the covariance and relative variance of p_t and d_t are identical to those stated in Equation (3). Next, consider the evidence when the interest differential is negative. As documented earlier, for many currencies this slope coefficient is positive and significantly greater than one (German mark, Japanese yen, and Dutch guilder). This finding has two implications: first, that the covariance between p_t and d_t should be negative, and second, that

$$\operatorname{var}(d_t) > |\operatorname{cov}(p_t, d_t)| > \operatorname{var}(p_t). \tag{4}$$

Note that unlike Equation (3), the relative variance of d_t exceeds that of p_t.[7] It then follows that when the interest rate differential is negative, the implication for the relative variances of p_t and d_t are opposite to those when the interest differential is positive (and hence also to those in the linear projection context). Finally, consider the case where the slope coefficient on x^+ is not significantly different from one. In this case $\operatorname{var}(p)$ is zero and the variance of d is equal to the variance of the interest rate differential. Hence, in this case also, the variance of expected depreciation exceeds the variance of the risk premium.

Figure 2 provides a graphical representation of the new empirical finding discussed above.[8] Note that the quantity $\frac{\operatorname{cov}(d,p+d)}{\operatorname{var}(p+d)}$ is solely determined by the correlation between p and d and the relative variances of p and d. Figure 2 also shows that the correlation between p and d has to be relatively large (in excess of 0.95) to observe the large variation in the absolute values of the slope coefficient (say from -4 to $+4$). Also note that when $\frac{\operatorname{std}(p)}{\operatorname{std}(d)} = 1$, the slope coefficient is 0.5, independent of the magnitude of the correlation. This suggests that to produce a negative $\frac{\operatorname{cov}(d,p+d)}{\operatorname{var}(p+d)}$, the magnitude by which $\frac{\operatorname{std}(p)}{\operatorname{std}(d)}$ must exceed one is relatively larger than is required for this ratio to be less than one to produce the large positive slope coefficient.

Models such as the single-factor CIR model considered by Backus et al. (1995) and Nielson and Saá-Requejo (1994) imply that d_t and

[7] Note than when the slope coefficient is greater than one $[\operatorname{cov}(d, d+p)/\operatorname{var}(p+d)] > 1$, which implies that $\operatorname{cov}(d, d+p) > \operatorname{var}(p+d)$ and $-\operatorname{cov}(p, d) > \operatorname{var}(p)$. Since the variance of p is positive, to satisfy this inequality, $\operatorname{cov}(p, d)$ must be negative. In addition, the positive slope coefficient also implies $\operatorname{cov}(d, d+p) = \operatorname{cov}(d, p) + \operatorname{var}(d) > 0$ or that $\operatorname{var}(d) > -\operatorname{cov}(d, p)$. Hence it follows that $\operatorname{var}(d) > -\operatorname{cov}(p, d) > \operatorname{var}(p)$, which is the desired result.

[8] I thank an anonymous referee for suggesting Figure 2 and its interpretation.

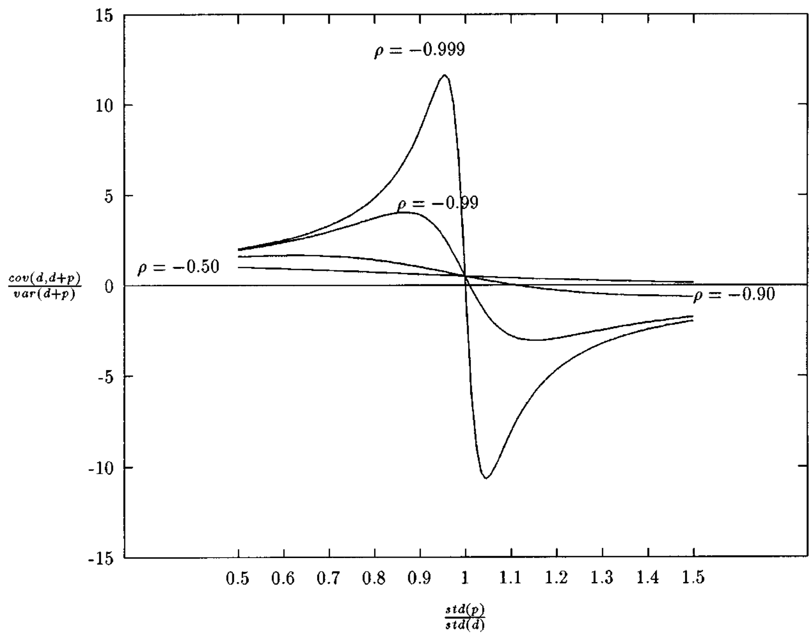

Figure 2
Theoretical slope coefficient
p refers to the forward risk premia, d refers to expected depreciation, and ρ is the correlation between d and p. The slope coefficient, $\frac{\text{cov}(d, d+p)}{\text{var}(d+p)} = \frac{\text{std}(d)/\text{std}(p)+\rho}{\text{std}(d)/\text{std}(p)+\text{std}(p)/\text{std}(d)+2\rho}$. The chosen values of ρ are -0.50, -0.90, -0.99, and -0.999.

p_t are linear in the interest rate differential. Even if such a model can account for the negative slope coefficient, it cannot account for the dependence of the slope coefficient on the sign of the interest rate differential. Explaining these nonlinearities poses a serious challenge to any economic model.

2. The Conceptual Framework

2.1 Stochastic discount factors and exchange rates

A stochastic discount factor (in equilibrium models, the marginal rate of substitution) is a random variable that "discounts" the average of the payoff at the riskless rate and adjusts the price of the payoff for its riskiness. The relationship between expected depreciation and interest rate risks can be characterized by stochastic discount factors across countries.

Consider the set of all payoffs denominated in dollars. This collection includes payoffs from investing in foreign assets and converting them back into dollars. In addition, also assume that the set of all dollar payoffs does not permit any arbitrage opportunities. It is well-recognized that the absence of arbitrage opportunities permits the creation of a variety of stochastic discount factors that price the given collection of dollar returns [see Harrison and Kreps (1979) and Ross (1978)]. For example, Hansen and Jagannathan (1991) and Hansen and Richard (1987) develop stochastic discount factors that relate to minimum variance problems, and Cox and Huang (1989) and Long (1990) rely on the solution to a growth optimal portfolio problem (a unit cost portfolio that maximizes the expected continuous return) to develop a stochastic discount factor. Let m_{t+1} be a strictly positive stochastic discount factor that prices all payoffs denominated in dollars. By definition, m satisfies $E_t[m_{t+1}R_{t+1}] = 1$ for any dollar return R_{t+1}. In addition, it also satisfies $E_t[m_{t+1}\frac{e_{t+1}}{e_t}\tilde{R}_{t+1}] = 1$, where $\tilde{R}_{t+1}$ is a return denominated in the foreign currency.

A stochastic discount factor that prices all payoffs in foreign currency units can be formed by scaling m_{t+1} by the growth in the nominal exchange rate, $\frac{e_{t+1}}{e_t}$. In particular, the stochastic discount factor $\tilde{m}_{t+1} = m_{t+1}\frac{e_{t+1}}{e_t}$ prices all payoffs denominated in foreign currency units—since $E[m_{t+1}\frac{e_{t+1}}{e_t}\tilde{R}_{t+1}] = 1$ it also follows that $E[\tilde{m}_{t+1}\tilde{R}_{t+1}] = 1$. The relationship between m_{t+1}, $\tilde{m}_{t+1}$, and $\frac{e_{t+1}}{e_t}$ can be stated as

$$\frac{e_{t+1}}{e_t} = \frac{\tilde{m}_{t+1}}{m_{t+1}}. \tag{5}$$

Backus et al. (1995) use a similar argument to derive Equation (5). Note that knowing any two of the three variables m_{t+1}, $\tilde{m}_{t+1}$, and e_{t+1}/e_t is sufficient to restrict the third via Equation (5).

Economic models, such as the two-country term structure models and general equilibrium models, imply particular m and $\tilde{m}$ and then use Equation (5) to derive restrictions on the expected depreciation and the forward risk premium. For example, Backus et al. (1995) and Nielson and Saá-Requejo (1994) use the single-factor CIR model to restrict m and $\tilde{m}$ and derive implications for d and p. A similar approach is pursued in this article—implications of term structure models are used to restrict m and $\tilde{m}$, and Equation (5) is used to derive the implications for the expected depreciation and the forward risk premium.

2.2 Exchange rates and term structure models

There are several reasons to explore the ability of term structure models to provide an explanation of the forward premium puzzle. First, as stated earlier, a variety of economic models have not been successful in explaining this puzzle. Second, empirical evidence suggests that interest rate movements (risks) are important in understanding the observed behavior of forward risk premia. Since term structure models are designed to capture interest rate risks, it stands to reason that these models may also explain the forward premium puzzle.

A property of many term structure models, such as Vasicek (1977), CIR, Duffie and Kan (1995), and Longstaff and Schwartz (1992), is that the discount factor m and $\tilde{m}$ in these models is characterized solely by risks contained in domestic nominal default-free interest rate movements. Consequently they provide a convenient framework to relate d and p to interest rate risks.

To characterize the discount factor associated with many term structure models, let H_{t+1} be an arbitrage-free vector of unit cost returns on domestic nominal default-free bonds. A stochastic discount factor that prices all payoffs in H_{t+1} can be constructed by solving the growth optimal problem, as in Long (1990). The motivation for choosing the growth optimal discount factor is that with additional assumptions, the growth optimal discount factor coincides with the discount factor of many term structure models.

The growth optimal portfolio (ω_t') is the solution to the following problem:

$$\max_{\omega_t'} \; E_t[\log(\omega_t' H_{t+1})], \quad s.t \quad \omega_t' \iota = 1. \tag{6}$$

The unit cost return on the growth optimal portfolio is R^b_{t+1} and equals $\omega_t' H_{t+1}$. Moreover, the growth optimal return is unique and strictly positive [see Bansal and Lehmann (1996), Long (1990), and Merton (1993)]. The first-order conditions associated with the growth optimal problem satisfy $E_t[(1+h_{t+1})/R^b_{t+1}] = 1$ for all $1+h_{t+1} \in H_{t+1}$. It then follows that $m_{t+1} = 1/R^b_{t+1}$ is a stochastic discount factor that prices all portfolios in H_{t+1}. Adler and Dumas (1983), Bansal and Lehmann (1996), Cox and Huang (1989), Hakansson (1977), Long (1990), and Merton (1993) provide additional details regarding the properties of the growth optimal portfolio.

The discount factor $m_{t+1} = 1/R^b_{t+1}$, with additional assumptions, also coincides with the stochastic discount factor of many term structure models. An implication of many term structure models is that nominal default-free bond returns satisfy a linear factor structure where the number of factors is very small relative to the number of available bonds [see Brennan and Schwartz (1979), CIR, Vasicek (1977), Duffie and Kan (1995), and Longstaff and Schwartz (1992)]. This "dimensionality reduction" implies that R^b_{t+1} is a portfolio of a few basis bond returns. *Assumption 1* is that domestic bond returns satisfy a single-factor structure.[9] Assumption 1 implies that R^b_{t+1} is composed of a portfolio of the risk-free rate and the return on a risky default-free nominal bond. Consequently, $R^b_{t+1} = (1 + r_t) + \omega_t[h_{t+1} - r_t]$, where $h_{t+1} - r_t$ is a risky excess return between t and $t + 1$.[10] *Assumption 2* is that bond excess returns are conditionally normal,

$$h_{t+1} - r_t = c_t + \sigma_t u_{t+1}, \tag{7}$$

where c_t and σ_t^2 are the conditional mean and variance of the excess return. The innovation u_{t+1} is $\mathcal{N}(0, 1)$ and represents the risks associated with interest rate movements.[11] *Assumption 3* is that the conditional mean and variance are time invariant functions of the short interest rate, consequently $c_t = c(r_t)$ and $\sigma_t = \sigma(r_t)$.

Note that the growth optimal discount factor $1/R^b_{t+1}$ can also be stated as $\exp\{-\log(R^b_{t+1})\}$. The assumption of normality ensures that the conditional mean of $\log(R^b_{t+1})$ is $r_t + \omega_t c_t - \omega_t^2 \sigma_t/2$, and the one-step innovation in $\log(R^b_{t+1})$ is $\omega_t \sigma_t u_{t+1}$. It then follows that

$$\log(R^b_{t+1}) = r_t + \omega_t c_t - \omega_t^2 \sigma_t/2 + \omega_t \sigma_t u_{t+1}. \tag{8}$$

Equation (8) and the above stated implications for the conditional mean and variance of $\log(R^b_{t+1})$ are approximate in discrete time and are exact in continuous time. Ingersoll (1987), Long (1990), and Merton (1993) derive similar results and provide a discussion of the discrete and continuous-time arguments relating to this issue. Since $E_t[\log(R^b_{t+1})] = r_t + \omega_t c_t - \omega_t^2 \sigma_t/2$, it follows that the portfolio ω_t that maximizes $E_t[\log(R^b_{t+1})]$ is equal to c_t/σ_t^2. Substituting this optimal portfolio weight into Equation (8) implies

$$\exp[-\log(R^b_{t+1})] \equiv m_{t+1} = \exp(-r_t)\exp(-\lambda_t^2/2 - \lambda_t u_{t+1}), \tag{9}$$

where

$$\lambda_t = \frac{c_t}{\sigma_t} \tag{10}$$

[9] Using more factors does not alter the argument, although it considerably increases the number of parameters to be estimated. Other articles that consider a single-factor model include Backus et al. (1995), Gibbons and Ramaswamy (1993), Nielson and Saá-Requejo (1994), and Pearson and Sun (1994). Also see Ferson and Harvey (1993) and Knez et al. (1994).

[10] The risk-free rate herein refers to the weekly arithmetic risk-free rate and should be distinguished from the weekly continuously compounded rate defined earlier. It is assumed that the weekly continuous risk-free rate is approximately equal to the weekly arithmetic risk-free rate, that is, if r is the weekly arithmetic risk-free rate, then $\log(1 + r) \approx r$.

[11] In deriving the results in this section, a local approximation (exact in continuous time) is used. Under this approximation the gross return also satisfies $\log(1 + h_{t+1}) = r_t + c_t - \sigma_t^2/2 + \sigma_t u_{t+1}$—Equation (7) could also be stated in terms of this expression.

is the market price of risk in the domestic bond returns.[12] Assumption 3, that $c_t = c(r_t)$ and $\sigma_t = \sigma(r_t)$, implies that the market price of risk is also a function of the short rate, that is $\lambda_t = \lambda(r_t) = c(r_t)/\sigma(r_t)$. The market price of risk determines the slope of the mean standard deviation frontier on bond returns. Since the above derivation applies to any country under consideration, it follows that $\tilde{m} = \exp(-\tilde{r}_t)\exp(-\tilde{\lambda}_t^2/2 - \tilde{\lambda}_t\tilde{u}_{t+1})$, where $\tilde{\lambda}_t = c(\tilde{r}_t)/\sigma(\tilde{r}_t)$ can be derived using analogous arguments.

Different single-factor term structure models impose different restrictions on λ_t. For example, λ_t is a constant in the Vasicek (1977) model and is proportional to $\sqrt{r_t}$ in the CIR model—additional models are discussed below. In the context of continuous-time term structure models, Huang (1991), Malliaris and Brock (1991), and Vasicek (1977) derive an expression that is equivalent to Equation (9).

To derive implications for the expected depreciation, note that Equation (5) implies that $d_t = E_t[\log(\tilde{m}) - \log(m)]$. Consequently Equation (9) and its analog for the foreign country imply that

$$d_t = [r_t - \tilde{r}_t] - ([\tilde{\lambda}_t^2 - \lambda_t^2]/2), \tag{11}$$

and it then follows that

$$p_t = [\tilde{\lambda}_t^2 - \lambda_t^2]/2. \tag{12}$$

Recall, via Assumption 3, that the market price of risk in each country is a function of its short interest rate, that is, $\lambda_t = \lambda(r_t)$ and $\tilde{\lambda}_t = \tilde{\lambda}(\tilde{r}_t)$. Economically Equations (11) and (12) capture the intuition that the expected depreciation and the forward risk premium are primarily determined by interest rate risks across countries. Also note that Equations (9) and (11) impose joint restrictions on the expected depreciation and the domestic term-structure of nominal interest rates.

The above derivation of Equation (9) implicitly assumes that the stochastic discount factor that prices a subset of assets, namely bond returns, is also the discount factor that prices all dollar denominated assets. An alternative approach that does not make that assumption, but which yields identical implications for d, p, and the domestic term structure can also be developed. In this approach the discount factor that prices all dollar payoffs is proportionally decomposed into a discount factor that captures domestic interest rate risks and a component that represents risks other than interest rate risks. An additional assumption that the volatility of risks other than interest rate risks is

[12] Note that c_t and σ_t depend on the maturity of the nominal bond return. However, $\lambda_t = c_t/\sigma_t$ does not depend on the maturity of the bond.

constant is sufficient to deliver Equations (11) and (12).[13] The last assumption implies that conditional movements in the forward risk premium are entirely due to interest rate risks, although unconditionally the forward risk premium is determined by all risks in the economy.

3. Specification

The key variable to estimate for each country under consideration is λ_t. Once λ_t is known, m_{t+1} is also known for each country under consideration [Equation (9)]. There are two restrictions that follow from knowing λ_t (and hence, m_{t+1}) for each country. First, the expected depreciation process must satisfy Equation (11). Second, each country's m_{t+1} should also price a rich collection of domestic bond returns. In estimation, the primary focus will be on the first restriction and, to a limited extent, on the second restriction.

3.1 Nested models

The $\lambda(r)$ for each country is estimated by modeling the conditional mean and the conditional variance for a risky bond return for each country. The return process, as stated in Equation (7), is $h_{t+1} - r_t = c_t + \sigma_t u_{t+1}$. For each country under consideration, it is additionally assumed that

$$c_t = \mu + \delta r_t \tag{13}$$

$$\sigma_t = \kappa r_t^{\gamma}, \tag{14}$$

with $\kappa > 0$. The last two equations also imply that

$$\lambda_t = \frac{\mu + \delta r_t}{\kappa r_t^{\gamma}} \tag{15}$$

It is assumed that the stationary process for r satisfies $r > 0$, so that the conditional variance is always strictly positive.[14] Given this specification for λ, the implications for expected depreciation follow from Equation (11). In addition, m_{t+1} is fully characterized as it depends on the innovation, u_{t+1}, and λ_t [see Equation (9)].

[13] Specifically, let m be the benchmark discount factor that prices all dollar payoffs and $1/R^b_{t+1}$ be a benchmark growth optimal discount factor that only prices a subset—domestic bond returns. Then a proportional decomposition immediately follows from the pricing restriction; $m_{t+1}R^b_{t+1} = 1 + v_{t+1}$, which implies that $m_{t+1} = [1/R^b_{t+1}](1 + v_{t+1})$. Note that both $1 + v$ and $1/R^b$ are strictly positive, and that v_{t+1} is a martingale difference sequence. v/R^b is orthogonal to domestic bond returns and represents risks outside domestic bond markets. Expected depreciation, using $d = E[\log(\tilde{m}) - \log(m)]$, can be primarily restricted by $E_t[\log(R^b)]$ and its counterpart for the foreign country if additional assumptions regarding the stochastic properties of $\log(1 + v)$ are made.

[14] For a discussion of the restrictions that ensure a positive interest rate process see Conley et al. (1995) and Cox, Ingersoll, and Ross (1985).

There are several reasons for considering this specification for the conditional mean and variance. First, as detailed below, a variety of parametric single-factor models are nested within this specification. Second, in the context of models for the level of the interest rate, Aït-Sahalia (1995), Chan et al. (1992), Conley et al. (1995), and Tauchen (1995) consider a similar specification; their results provide some guidance regarding the reasonableness of the estimated parameters and this specification. Finally, as elaborated below, this simple specification can accommodate a variety of relationships between $\lambda(r)$ and r—some of which may be important in accounting for the forward premium puzzle.

Consider the various term structure models that are nested by this specification. The specification $\delta = 0$ and $\gamma = 0$ corresponds to the specification of Vasicek (1977) [also see Campbell, Lo, and MacKinlay (1995)] and implies that $\lambda = \frac{\mu}{\kappa}$. Note that the mean excess return and the conditional variance are constants. The CIR specification corresponds to $\mu = 0$ and $\gamma = 0.5$, which implies that $\lambda_t = \frac{\delta}{\kappa}\sqrt{r_t}$ [see Campbell, Lo, and MacKinlay (1995)]. Additionally, $\gamma = 1$ corresponds to the specification considered in Brennan and Schwartz (1979), and $\gamma > 1$ is considered in Chan et al. (1992), Conley et al. (1995), and Tauchen (1995).

3.2 Market price of risk and the forward premium puzzle

Many of the specifications discussed above are incapable of explaining the forward premium puzzle. Any specification with constant λ [such as Vasicek (1977) and the specification $\mu = 0$ and $\gamma = 1$] implies that the variance of p_t is zero [see Equation (12)] and that uncovered interest parity holds. The single-factor CIR model also cannot explain the forward premium puzzle. In this model, $p = [(\frac{\tilde{\delta}}{\tilde{\kappa}})^2\tilde{r} - (\frac{\delta}{\kappa})^2 r]/2$, and substituting this into Equation (11) shows that an increase in the U.S. rate will tend to increase d (depreciate the dollar). Further, assuming symmetry (i.e., δ and κ are identical across countries), as in Backus, Foresi, and Telmer (1995), implies that $d = (1 + \{(\frac{\delta}{\kappa})^2/2\})[r - \tilde{r}]$—the slope coefficient is strictly positive in the single-factor CIR model.

It is possible to characterize the properties of $\lambda(r)$ and $\tilde{\lambda}(\tilde{r})$ to explain the negative slope coefficient and the documented nonlinearities. Consider the symmetric case where the functional form of $\lambda(r)$ is assumed to be identical across countries. Also, let $\lambda_r^2(r)$ be the derivative of the squared market price of risk with respect to r [i.e., $\frac{d}{dr}(\lambda^2(r)) \equiv \lambda_r^2(r)$]. A mean value approximation around $\tilde{r} = r = \bar{r}$ implies that $p = (\lambda_r^2(\bar{r})/2)[\tilde{r} - r]$ and $d = (1 + (\lambda_r^2(\bar{r})/2))\ [r - \tilde{r}]$.[15]

[15] Consider the expansion $\lambda^2(r) = \lambda^2(\bar{r}) + \lambda_r^2(\bar{r})(r - \bar{r})$. The analogous approximation for the foreign country is $\tilde{\lambda}^2(r) = \tilde{\lambda}^2(\bar{r}) + \tilde{\lambda}_r^2(\bar{r})(\tilde{r} - \bar{r})$. As the functional forms are assumed to be identical across

The slope coefficient $\frac{\text{cov}(d,x)}{\text{var}(x)}$ is equal to $1 + (\lambda_r^2(\bar{r})/2)$. For the slope coefficient to be negative, $1 + (\lambda_r^2(\bar{r})/2)$ must be less than zero, that is, $\lambda_r^2(\bar{r})/2$ must be negative and greater than one in absolute value. Hence the squared market price of risk must be decreasing in the level of the interest rate.

Since $\lambda^2(r) = (c(r)/\sigma(r))^2$, its interest rate derivative, λ_r^2, satisfies

$$\lambda_r^2 = 2\,\lambda^2 \left[\frac{c_r}{c} - \frac{\sigma_r}{\sigma}\right]. \tag{16}$$

The sign of λ_r^2 is determined by the difference of the interest elasticity of the conditional mean $\frac{c_r}{c}$ and the interest rate elasticity of the conditional standard deviation $\frac{\sigma_r}{\sigma}$.[16] For $\lambda^2(r)$ to be decreasing (i.e., for $\lambda_r^2(r) < 0$), the elasticity of the standard deviation should exceed that of the conditional mean. The specification in Equations (13) and (14) ensures that the elasticity of the conditional mean is $\frac{\delta}{\mu+\delta r}$ and the elasticity of the standard deviation is $\frac{\gamma}{r}$. This implies that

$$\left[\frac{c_r}{c} - \frac{\sigma_r}{\sigma}\right] = \left[\frac{\delta}{\mu + \delta r} - \frac{\gamma}{r}\right]. \tag{17}$$

From Equations (16) and (17) it then follows that λ_r^2 is negative if $\frac{\delta r}{\mu+\delta r} < \gamma$, and vice versa. Note that that if $\mu = 0$, then $\lambda^2(r)$ is decreasing in r if $\gamma > 1$ and increasing in r if $\gamma < 1$, independent of all other parameters.

Based on the above discussion, one should suspect that when there is symmetry across countries, the specification $\mu = 0$ and $\gamma > 1$ may account for the negative slope coefficient. While this condition may explain the negative slope coefficient it is unlikely to account for the documented nonlinearities. For example, with $\mu = 0$ and $\gamma = 1.5$, the market price of risk $\lambda^2(r) = (\delta/\kappa)^2(1/r)$ is clearly decreasing in r. From Equation (12) it follows that

$$p = \left(\frac{\delta}{\kappa}\right)^2 \frac{1}{(2r\tilde{r})}[r - \tilde{r}].$$

In the data, the sign of the coefficient that determines the relationship between the forward risk premium and the interest rate differential depends on the sign of the interest rate differential. The coefficient, $(\frac{\delta}{\kappa})^2\frac{1}{(2r\tilde{r})}$, which determines the relation between p and the interest

countries, it follows that $\lambda^2(\bar{r}) = \tilde{\lambda}^2(\bar{r})$ and $\bar{r}\lambda^2(\bar{r}) = \bar{r}\tilde{\lambda}^2(\bar{r})$. Using Equation (11) implies that $d = [r - \tilde{r}] + (\lambda_r^2(\bar{r})/2)[r - \tilde{r}]$.

[16] To derive Equation (16) note that the derivative of λ^2 w.r.t to r is $2\lambda\lambda_r$. Also note that $\lambda_r = \frac{c_r c}{c\sigma} - \frac{\sigma_r c}{(\sigma)(\sigma)}$. Substituting this expression into $2\lambda\lambda_r$ leads to Equation (16).

differential, is always strictly positive, implying that this specification may not quantitatively account for the documented nonlinearities.

One possible explanation of the documented nonlinearities is a particular type of asymmetry across countries. Suppose there is asymmetry across countries in the form of the squared market price of risk across countries. In particular, assume that $\lambda_r^2(r) < 0$ and $\tilde{\lambda}_{\tilde{r}}^2(\tilde{r}) > 0$, implying that $\lambda_r^2(r) - \tilde{\lambda}_{\tilde{r}}^2(\tilde{r}) < 0$. In addition, also assume that $1 + (\lambda_r^2(r)/2) < 0$. A mean value expansion of d around $\tilde{r} = r = \bar{r}$ implies[17]

$$d = -\phi_0 + \{1 + (\lambda_r^2(\bar{r})/2)\}[r - \tilde{r}] + ([\lambda_r^2(\bar{r}) - \tilde{\lambda}_{\tilde{r}}^2(\bar{r})]/2)\tilde{r},$$

where $\phi_0 = \tilde{\lambda}^2(\bar{r})/2 - \lambda^2(\bar{r})/2 - ([\tilde{\lambda}_{\tilde{r}}^2(\bar{r}) - \lambda_r^2(\bar{r})]/2)(\bar{r})$, and $[\lambda_r^2(\bar{r}) - \tilde{\lambda}_{\tilde{r}}^2(\bar{r})] < 0$. Note that if symmetry is assumed across countries, then $\phi_0 = 0$ and the second term in the expression for d is zero since $\lambda_r^2(\bar{r}) = \tilde{\lambda}_{\tilde{r}}^2(\bar{r})$.

To see how asymmetry affects the relationship between d and the interest differential x, consider the expression for the slope coefficient,

$$\frac{\mathrm{cov}(d, x)}{\mathrm{var}(x)} = [1 + (\lambda_r^2(\bar{r})/2)] + ([\lambda_r^2(\bar{r}) - \tilde{\lambda}_{\tilde{r}}^2(\bar{r})]/2)\psi, \quad \psi = \frac{\mathrm{cov}(\tilde{r}, x)}{\mathrm{var}(x)}.$$

Unlike the case of symmetry, the implied slope coefficient is additionally affected by $([\lambda_r^2(\bar{r}) - \tilde{\lambda}_{\tilde{r}}^2(\bar{r})]/2)\psi$. The sign of the slope coefficient also depends on the sign of ψ. In particular if $\psi \geq 0$ when the interest differential is positive, then the slope coefficient will be large and negative. On the other hand, if ψ is large and negative when the interest differential is negative, then the slope coefficient can be positive. The unconditional slope can be negative either because the unconditional $\psi \geq 0$, or $\psi < 0$ and the effect of the second term—$([\lambda_r^2(\bar{r}) - \tilde{\lambda}_{\tilde{r}}^2(\bar{r})]/2)\psi$—is smaller than the first. Note that these are sufficient conditions to account for the observed state dependence in the slope coefficient—they are not necessary.

The plausibility of the above argument depends on the behavior of ψ. In the data, the unconditional estimates of ψ are -0.193 and 0.038 for U.S.–Germany and U.S.–Japan, respectively. When the interest differential is positive, ψ for U.S.–Germany and U.S.–Japan are

[17] The mean value expansion and Equation (12) implies that $p = \phi_0 + (\tilde{\lambda}_{\tilde{r}}^2(\bar{r})/2)\tilde{r} - (\lambda_r^2(\bar{r})/2)r$, where ϕ_0 is a constant that does not affect any of the subsequent arguments. The expression for ϕ_0 is $\phi_0 = \tilde{\lambda}^2(\bar{r})/2 - \lambda^2(\bar{r})/2 + ([\tilde{\lambda}_{\tilde{r}}^2(\bar{r}) - \lambda_r^2(\bar{r})]/2)(-\bar{r})$. The expression for p, after adding and subtracting $(\lambda_r^2(\bar{r})/2)\tilde{r}$, can be restated as $p = \phi_0 + (\tilde{\lambda}_{\tilde{r}}^2(\bar{r})/2)\tilde{r} - (\lambda_r^2(\bar{r})/2)\tilde{r} + (\lambda_r^2(\bar{r})/2)\tilde{r} - (\lambda_r^2(\bar{r})/2)r = \phi_0 + ([\tilde{\lambda}_{\tilde{r}}^2(\bar{r}) - \lambda_r^2(\bar{r})]/2)\tilde{r} + (\lambda_r^2(\bar{r})/2)[\tilde{r} - r]$. Equation (11) and the expression for p implies that $d = -\phi_0 + \{1 + (\lambda_r^2(\bar{r})/2)\}[r - \tilde{r}] + ([\lambda_r^2(\bar{r}) - \tilde{\lambda}_{\tilde{r}}^2(\bar{r})]/2)\tilde{r}$.

0.292 and 0.146, respectively, and when the differential is negative, ψ is -0.476 and -1.666.[18] Hence the sign and magnitude of ψ in the data are consistent with the example of asymmetry given above.

One convenient way to accommodate asymmetries across countries is to include a country-specific μ in the conditional mean of the excess return. Consider the case where $\mu > 0$, $\delta > 0$, and $\gamma > 1$. With this specification $\frac{\delta r}{\mu+\delta r} < 1 < \gamma$, and for all r, $\lambda_r^2 < 0$ [see Equation (17)]. Now consider the case where $\mu < 0$, $\delta > 0$, and $\gamma > 1$. In this case, for $\frac{-\mu}{\delta} < r < \frac{(-\mu)\gamma}{\delta(\gamma-1)}$, $\lambda_r^2 > 0$, and for r outside this range, $\lambda_r^2 \leq 0$.[19] Hence the sign and size of μ affects the slope $\lambda^2(r)$. Thus a specification that allows μ to be different across countries can accommodate considerable differences in the relationship between $\lambda^2(r)$ and r across countries, and may account for nonlinearities in the forward currency risk premium.

4. Estimation Method and Results

4.1 Estimation method and diagnostics

Estimation is conducted under the hypothesis that bond excess returns are conditionally normal, a single factor is adequate to characterize the cross section of domestic bond excess returns for each country, risks outside of bond markets do not affect expected depreciation (at least conditionally), and the assumed specification in Equations (13)–(15). Under this null hypothesis, the primary restriction of interest is Equation (11), and an additional restriction is that m for each country should price domestic bond returns. Variables related to the United States, Germany, and Japan are subscripted by 1, 2, and 3, respectively.

The parameters of interest are $\theta = \{\mu_k, \delta_k, \kappa_k, \gamma\}$, $k = 1, 2, 3$. In estimation, the additional restriction that γ is the same across all countries is also imposed. This assumption implies that interest rate elasticity of the variance is identical across countries, although the level of volatility can be different across countries at the same level of

[18] Note that ψ is the regression coefficient from regressing $\tilde{r}$ on $x = r - \tilde{r}$ for the appropriate state of the interest rate differential.

[19] This can be proved as follows. From Equation (17), $\frac{\delta r}{\mu+\delta r} - \gamma$ determines the sign of the slope of λ^2. Throughout, it is assumed that $\delta > 0$, $\gamma > 1$. First, consider the case where $\mu > 0$, then $\frac{\delta r}{\mu+\delta r} - \gamma < 0$, for all r, since $\frac{\delta r}{\mu+\delta r}$ is always less than one. It then follows that $\lambda_r^2(r) < 0$. Now consider the case where $\mu < 0$. Note that $\lambda(r) > 0$ when $r > \frac{-\mu}{\delta}$. For $r > \frac{-\mu}{\delta}$, the quantity $\frac{\delta r}{\mu+\delta r} - \gamma > 0$ when $\delta r > (\mu + \delta r)\gamma$, that is, when $r < \frac{(-\mu)\gamma}{\delta(\gamma-1)}$. Hence $\lambda_r^2 > 0$ when $\frac{-\mu}{\delta} < r < \frac{(-\mu)\gamma}{\delta(\gamma-1)}$. Using the same argument, for r greater than $\frac{(-\mu)\gamma}{\delta(\gamma-1)}$, $\lambda_r^2 < 0$. When $r < \frac{-\mu}{\delta}$, note that $\lambda(r)$ is negative, and $\lambda_r^2 < 0$ since $\frac{\delta r}{\mu+\delta r}$ is always negative [see Equation (17)]. When $r = \frac{-\mu}{\delta}$, then $\lambda(r) = 0$ and $\lambda_r^2 = 0$ [see Equation (16)].

the interest rate (as κ may be different across countries). This restriction also imposes parsimony in terms of estimated parameters. With this assumption, θ has a maximum of 10 parameters. The parameter vector θ is estimated using generalized method of moments (GMM) as developed in Hansen (1982) and Hansen and Singleton (1982).

Let $c_{k,t} = \mu_k + \delta_k r_{k,t}$, $\sigma_{k,t} = \kappa_k r_{k,t}^{\gamma}$, and $\lambda_{k,t} = \frac{c_{k,t}}{\sigma_{k,t}}$. Under the null hypothesis, consider the following four errors, which are martingale difference sequences:

$$\sigma_{k,t}\ u_{k,t+1} = h_{k,t+1} - r_{k,t} - c_{k,t}, \quad k = 1, 2, 3 \tag{18}$$

$$\eta_{k,t+1} = \sigma_{k,t}^2 u_{k,t+1}^2 - \sigma_{k,t}^2, \quad k = 1, 2, 3 \tag{19}$$

$$\zeta_{1,t+1} = g1_{t+1} - [(r_{1,t} - r_{2,t}) + (\lambda_{1,t}^2 - \lambda_{2,t}^2)/2] \tag{20}$$

$$\zeta_{2,t+1} = g2_{t+1} - [(r_{1,t} - r_{3,t}) + (\lambda_{1,t}^2 - \lambda_{3,t}^2)/2] \tag{21}$$

where $g1_{t+1}$ ($g2_{t+1}$) is the continuous change in the dollar-mark (dollar-yen) exchange rate. The first two errors are related to the specification of the conditional mean and conditional variance of the bond return and follow from Equation (8). Errors $\zeta_{1,t+1}$ and $\zeta_{2,t+1}$ are related to the restrictions on the conditional mean of the continuous change in the exchange rate and follow from Equation (12). The bond excess return used to estimate the parameter vector and hence the market price of risk, λ_t, is the bond return on the 12-month (the longest available maturity) bond for each country under consideration.[20]

To estimate the parameter vector θ and test the model, 23 orthogonality conditions are constructed by exploiting the martingale difference property of these errors. The orthogonality conditions used are the error $\sigma_{k,t}u_{k,t+1}$ is orthogonal to $\{1, r_{k,t}\}$ and the error $\eta_{k,t+1}$ is orthogonal to $\{1, r_{k,t}, \sigma_{k,t-1}^2 u_{k,t}^2\}$ for $k = 1, 2, 3$. These 15 orthogonality conditions evaluate the ability of the assumed specification to explain the bond return process for each country. An additional eight orthogonality conditions are used to evaluate the ability of the model to explain the conditional drift in the exchange rate process. Hence $\zeta_{1,t+1}$, the error associated with the continuous change in the dollar-mark exchange rate, is set orthogonal to $\{1, r_{1,t} - r_{2,t}, (r_{1,t} - r_{2,t})^2, (r_{1,t} - r_{2,t})^3\}$ and $\zeta_{2,t}$ is set orthogonal to the analogous functions of the interest rate differential across the United States and Japan. Recall that the same

[20] In a single-factor model the return on any nominal default-free risky bond can be used to estimate λ_t. By choosing the 12-month return, the pricing restrictions of m_{t+1} are made consistent with the longest available maturity. Moreover, m_{t+1} is, by construction, also consistent with the shortest available maturity—the 7-day risk-free rate [see Equation (9), and recognize that $\exp(-\lambda_t^2/2 - \lambda_t u_{t+1})$ integrates to one]. Consequently this procedure allows the model the best chance to price all the intermediate maturities.

set of functions were used to characterize the nonlinear relationship between expected depreciation and the interest rate differential. In estimation, the GMM optimal weight matrix was constructed with the Newey and West (1987) procedure using four lags.[21]

Given the estimated parameter vector $\hat{\theta}$, a variety of diagnostics are constructed to evaluate the ability of the model to explain the forward premium puzzle. For a given model, the average of any given orthogonality condition should not be statistically different from zero. The t-ratio of the average of the orthogonality condition provides a test to evaluate this implication. This test provides additional information regarding the dimensions in which a given model fails.

For a given $\hat{\theta}$, the process for the market price of risk for each country follows from Equation (15) since it is a known function of the interest rate. Additionally, the process for $d_t(\hat{\theta})$ and $p_t(\hat{\theta})$ can be readily computed using Equations (12) and (11). The second diagnostic evaluates whether the unconditional variance and covariance of these processes satisfy the requirements of Fama (1984) stated in Equation (3). The third diagnostic provides direct information regarding the size and sign of the slope coefficient. Expected depreciation, as implied by the model, $d_t(\hat{\theta})$, is regressed on a constant and the interest rate differential to evaluate whether the slope coefficient implied by the model is negative. In addition, $d_t(\hat{\theta})$ is also regressed on a constant, x^+, and x^- to evaluate whether the estimated model is capable of producing the documented nonlinearities in expected depreciation.

As stated earlier, knowledge of $\lambda_{k,t}$ is sufficient to characterize $m_{k,t+1}$ [see Equation (9)]. The fourth diagnostic evaluates whether $m_{k,t+1}$ is capable of pricing a richer collection of bond returns. This diagnostic provides evidence regarding the reasonableness of the estimated market price of risk from the perspective of the term structure of each country under consideration. This diagnostic is implemented by evaluating whether pricing restrictions such as

$$E_t[m_{k,t+1}(1 + h^n_{k,t+1})] = 1 \quad k = 1, 2, 3$$

hold for bond payoffs not included in estimation of θ. $1 + h^n_{k,t+1}$ is the unit cost return on an n-maturity nominal bond for country k. For each country under consideration, four payoffs are included in this test: the return on the 6-month bond, the return on the 6-month bond multiplied by its lagged return, the 12-month bond return multiplied by its lag, and the pricing error on the 6-month bond multiplied by its lag. With three countries there are 12 pricing restrictions (orthogonal-

[21] Sensitivity analysis was conducted, and the inclusion of additional lags did not affect the results.

ity conditions) to consider. Since there are no parameters to estimate, the usual GMM criterion is used to test if these 12 orthogonality conditions are satisfied in the data. The test statistic is distributed as a χ^2 with 12 degrees of freedom [for a discussion of this test, see Ogaki (1990)].

4.2 Estimation results

Tables 1 and 2 provide information regarding the various variables used in estimation. Note that the mean and variances of the short rate and bond returns are quite different across the United States, Germany, and Japan. The mean and variance of the short rate and bond returns are highest for the United States and smallest for Japan. The correlation between the United States short rate and the interest rate differential is 0.79 in the case of the United States and Germany, and 0.86 for the United States and Japan. These rather high correlations suggest that much of the variation in the differential is due to the relatively high variation in the U.S. interest rate vis-à-vis Germany and Japan.

4.2.1 Restricted model, $\mu = 0$. Table 6 (see column labeled Col 1) presents estimation results for the case where μ is assumed to be zero for all countries. In addition, since γ is constrained to be the same across countries, considerable symmetry is imposed across countries. As mentioned earlier, this specification can potentially account for the negative slope coefficient when $\gamma > 1$.

The model has a p-value of 0.002 (see χ^2 and p-value). Note that for all countries, δ is positive and reasonably well estimated. A positive value of δ implies that the mean excess return (interest rate risk premium) is increasing in the short rate. This suggests that when short rates are high, the longer maturity interest rate (12-month interest rate) is expected to fall. Intuitively, a positive δ reflects the effects of mean reversion in interest rates. Note that κ is also well estimated for all countries, and its value is the highest for the United States, which implies that for any given level of interest rate, the conditional volatility of bond returns is highest in the United States vis-à-vis Germany and Japan. The point estimate of γ is 1.199 and is precisely estimated. The two standard error range for γ (1.04 to 1.35) is greater than one and precludes the CIR specification of $\gamma = 0.5$ and the Brennan and Schwartz (1979) specification of $\gamma = 1$.

Figures 3–5 show the plot of λ^2 as a function of the short rate for the United States, Germany, and Japan, respectively. These graphs (see the curves labeled spec 1) show that the squared market price of risk is decreasing in the interest rate for each country under consideration. The downward sloping squared market prices of risk indicate that

Table 6
Model estimation results

Parameter	Col 1	Col 2	Col 3
μ_1		0.003	0.004
		(0.002)	(0.002)
μ_2		−0.022	−0.009
		(0.003)	(0.005)
μ_3		−0.006	−0.003
		(0.004)	(0.001)
δ_1	0.177	0.119	0.064
	(0.043)	(0.048)	(0.059)
δ_2	0.126	0.413	0.245
	(0.030)	(0.061)	(0.053)
δ_3	0.108	0.225	0.144
	(0.026)	(0.070)	(0.028)
κ_1	2.614	1.928	1.967
	(0.489)	(0.449)	(0.448)
κ_2	1.977	1.461	1.490
	(0.408)	(0.360)	(0.365)
κ_3	1.775	1.272	1.319
	(0.411)	(0.350)	(0.354)
γ	1.199	1.080	1.075
	(0.080)	(0.098)	(0.095)
χ^2	36.69	21.43	27.37
p-value	0.002	0.065	0.038
Wald$_3$ p-value		0.000	0.000

The estimated model is $E_t[y_{t+1}] = \mu + \delta r_t$ and $\sigma_t^2(y_{t+1}) = \kappa^2 r_t^{2\gamma}$. y_{t+1} is the 7-day excess return on a 12-month nominal discount bill, and r_t is the 7-day nominal risk-free rate. In estimation, y and r are multiplied by 52. The parameters μ_1, δ_1, and κ_1 refer to the U.S., μ_2, δ_2, and κ_2 refer to Germany, and μ_3, δ_3, and κ_3 refer to Japan. Standard errors are reported in parentheses below the point estimate. χ^2 refers to the GMM criterion value which is distributed as a χ^2 with 16, 13, and 16 degrees of freedom for Col 1, Col 2, Col 3, respectively. The p-value refers to probability in the right tail. Wald$_3$ p-value refers to the p-value for the Wald test of the hypothesis that $\mu_1 = \mu_2 = \mu_3 = 0$. This statistic is distributed as a χ^2 with 3 degrees of freedom.

the interest elasticity of the standard deviation exceeds the interest elasticity of the conditional mean [see Equation (17)].

Diagnostics reported in Table 7 (see column labeled Col 1) show that this configuration of the model satisfies the condition $\text{var}(p) > |\text{cov}(p, d)| > \text{var}(d)$ for both the currencies (see rows 1–6). The annualized standard deviation of d (see rows 1–4) for Germany and Japan is 0.628% and 0.467%, respectively, and for p is 1.067% and 0.735%, respectively. The correlation between d and p for Germany and Japan is −0.96 and −0.85, respectively. Given the decreasing λ^2, not surprisingly, the slope coefficient is negative (see rows 7 and 8) for both countries. Hence this model specification is capable of accounting for the negative slope coefficient.

This model, however, cannot explain the documented nonlinearity in expected depreciation. It implies a negative regression slope coef-

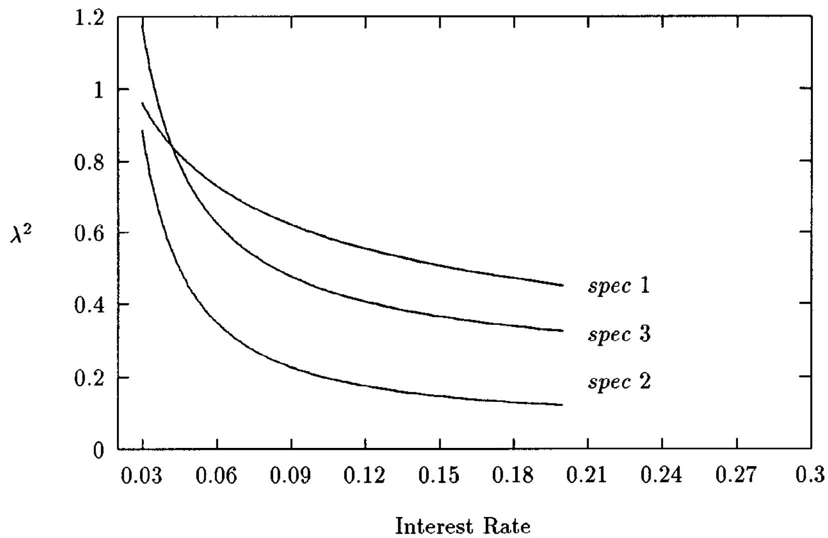

Figure 3
U.S. squared market price of risk
The x axis is the weekly interest rate multiplied by 52, and λ^2 is the squared market price of risk multiplied by 52. The graph for spec 1, spec 2, and spec 3 correspond to the parameter configuration of Col 1, Col 2, and Col 3, respectively, reported in Table 6.

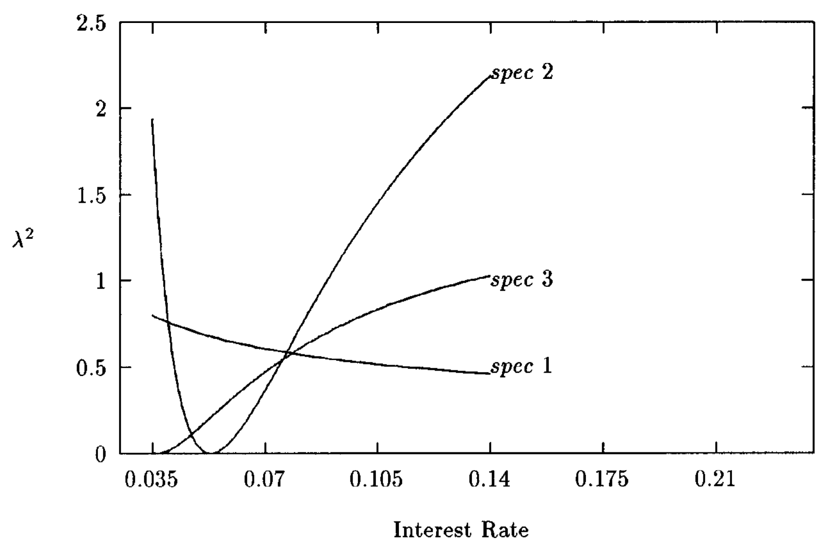

Figure 4
German squared market price of risk
The x axis is the weekly interest rate multiplied by 52, and λ^2 is the squared market price of risk multiplied by 52. The graph for spec 1, spec 2, and spec 3 correspond to the parameter configuration of Col 1, Col 2, and Col 3, respectively, reported in Table 6.

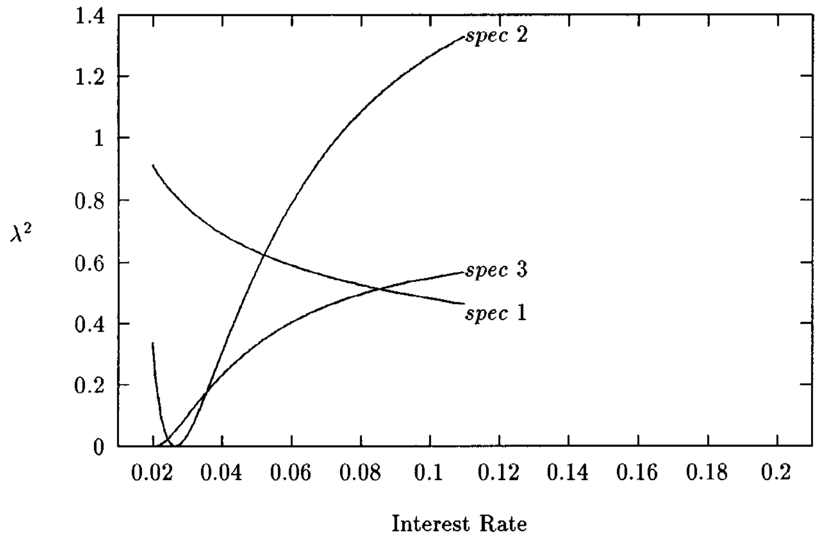

Figure 5
Japanese squared market price of risk
The x axis is the weekly interest rate multiplied by 52, and λ^2 is the squared market price of risk multiplied by 52. The graph for spec 1, spec 2, and spec 3 correspond to the parameter configuration of Col 1, Col 2, and Col 3, respectively, reported in Table 6.

ficient when the interest rate differential is negative (see rows 9 and 12), whereas this coefficient is positive in the data. Table 8 reports the t-ratio of the average orthogonality conditions associated with the continuous change in exchange rate. It is evident that the orthogonality conditions associated with the nonlinear functions of the interest rate differential are not satisfied by this model (t-ratio of two or greater) for both exchange rates. Consistent with the intuition stated earlier in this article, the inability of this model to capture the nonlinearity might be attributed to the similarities in the squared market price of risk across countries.

Table 7 (see column labeled Col 1) also reports (see row 13) the test associated with pricing a richer collection of bond returns. The p-value of this test is 0.16, indicating that the estimated market price of risk for each country is not inconsistent with pricing the additional domestic bond returns.

4.2.2 Models with country-specific μ-estimates. As discussed earlier, one possible explanation of the documented nonlinearities in expected depreciation is that there are particular types of asymmetries across countries. To better accommodate this possibility, μ is included in the specification.

Table 7
Model diagnostics

Variable	Col 1	Col 2	Col 3
1 : std(d_1)	0.0453	0.3671	0.1627
2 : std(d_2)	0.0337	0.2713	0.1672
3 : std(p_1)	0.0770	0.3737	0.1703
4 : std(p_2)	0.0530	0.2808	0.1782
5 : corr(d_1, p_1)	−0.9603	−0.9954	−0.9779
6 : corr(d_2, p_2)	−0.8499	−0.9946	−0.9868
7 : β_1	−1.0118	−1.3787	−0.4783
8 : β_2	−0.4216	−2.3798	−1.5834
9 : β_1^+	−0.1906	−4.7585	−1.5902
10 : β_1^-	−1.9517	2.4904	0.7944
11 : β_2^+	−0.1334	−3.5564	−2.1590
12 : β_2^-	−3.1527	8.7714	3.8712
13 : $\chi^2(12)$	16.62	21.71	19.84
p-value	0.1644	0.0409	0.0702

The diagnostics in this table are constructed using the parameter point estimates reported in Table 6. The standard deviations are weekly standard deviations multiplied by 52. The annualized standard deviation is the reported standard deviation divided by $\sqrt{52}$. d_1 (d_2) refers to expected depreciation of the dollar-mark (dollar-yen) exchange rate. Analogously, p_1 and p_2 refer to the forward risk premia w.r.t the U.S. and Germany, and the U.S. and Japan, respectively. β_1 (β_2) is the slope coefficient from regressing d_1 (d_2) on the interest rate differential between the U.S. and Germany (U.S. and Japan). β_1^+ and β_1^- (β_2^+ and β_2^-) refer to the slope coefficients from the projection of d_1 (d_2) on a constant, $x1^+$, and $x1^-$ (constant, $x2^+$, and $x2^-$). $\chi^2(12)$ refers to the GMM criterion value for additional pricing restrictions that are evaluated at the estimated parameters. p-value is probability in the right tail.

Table 6 (see column labeled Col 2) reports the estimated parameter values. This model specification has a p-value of 0.065 (see χ^2 and p-value). The hypothesis that $\mu_1 = \mu_2 = \mu_3 = 0$ is strongly rejected with a p-value of 0.000. Hence there is considerable evidence in favor of including μ. As above, γ is estimated quite precisely at 1.080 (standard error, 0.098). For all countries, δ is significantly positive and κ is the highest for the United States. Note that at low interest rates, the risk premium can become negative for Germany and Japan.

Diagnostics in Table 7 (see column labeled Col 2) reveal that the p_t and d_t series implied by the model satisfy Equation (3) [$\text{var}(p) > |\text{cov}(p, d)| > \text{var}(d)$]. The annualized standard deviation of d (see rows 1–4) for Germany and Japan is 5.09% and 3.76%, respectively, and for p is 5.18% and 3.89%, respectively. The correlation between d and p is −0.995 and −0.994 for Germany and Japan, respectively. This model also implies that the linear projection slope coefficient is negative (rows 7 and 8).

In addition, this specification is also capable of producing the observed nonlinearity in the expected exchange rate. The slope coefficient is negative when the interest rate differential is positive and is

Table 8
Error diagnostics

Variable	Col 1	Col 2	Col 3
$\zeta_{1,t+1}$	−0.571	0.294	−0.213
$\zeta_{1,t+1}x1_t$	−0.016	0.513	−0.424
$\zeta_{1,t+1}x1_t^2$	−2.291	−0.314	−1.320
$\zeta_{1,t+1}x1_t^3$	−1.083	−0.352	0.776
$\zeta_{2,t+1}$	0.135	0.554	0.333
$\zeta_{2,t+1}x2_t$	−1.388	−0.059	−0.443
$\zeta_{2,t+1}x2_t^2$	−2.127	−0.175	−1.012
$\zeta_{2,t+1}x2_t^3$	−2.008	1.448	−1.017

Reported is the t-ratio on the average of the orthogonality conditions associated with the difference (forecast error) between the realized continuous change in exchange rate and the expected continuous change in exchange rate as implied by a given model. $\zeta_{1,t+1}$ and $\zeta_{2,t+1}$ represent the forecast errors for the dollar-mark and the dollar-yen nominal exchange rates, respectively. Under the null hypothesis of the model, these forecast errors should be orthogonal to various functions of the known interest rate differential, x_t. Consequently the average of the various orthogonality conditions should not be statistically different from zero.

positive when the interest rate differential is negative (Table 7, rows 9–12). The finding that the slope coefficient is positive when the interest differential is negative implies that the relative variances of p, d, and their covariance also satisfies Equation (4) [$\text{var}(d) > |\text{cov}(p, d)| > \text{var}(p)$] when the interest rate differential is negative. Table 8 shows that the orthogonality conditions associated with the nonlinear functions of the interest rate differential are satisfied with t-ratios well below two.

The estimates imply that the squared market price of risk for the United States falls rapidly with an increase in the level of the interest rate, while for Germany and Japan it increases with the interest rate at high interest rates (see Figures 3–5, curve labeled spec 2). This model has exactly the type of asymmetries across countries discussed in Section 3.2 — arguments presented therein explain why the asymmetry in the estimated model is able to account for the documented nonlinearities.

To understand the differences in the behavior of the squared market price of risk, recall that when $\mu>0$ and $\gamma>1$ (with $\delta > 0$), the squared market price of risk is monotone decreasing (see the last paragraph of Section 3.2); for the United States, μ is positive (along with $\gamma>1$), which explains the decreasing λ^2. On the other hand, μ is negative for Germany and Japan, which implies that the squared market price of risk is increasing for $\frac{-\mu}{\delta}<r<\frac{-\mu\gamma}{\delta(\gamma-1)}$ and decreasing outside this range (see the last paragraph of Section 3.2). This explains the observed shape of the squared market price of risk for Germany and Japan.

To evaluate whether the addition of pricing restrictions associated with a richer collection of bond returns alters any of the above results, three additional pricing restrictions are added and the model is reestimated. These pricing restrictions are that $m_{k,t+1}$, which is completely characterized by θ, and the estimated bond excess return process should also price the domestic holding period return on a 6-month bond for each country. Consequently there are 26 orthogonality conditions included in the estimation. Results reported in Table 6 (see column labeled Col 3) and diagnostics in Table 4 (see column labeled Col 3) show that this change does not alter the results discussed above. The nature of the asymmetries across countries are similar to the previous case.

To get a sense of the reasonableness of the estimated parameters, note that in the context of the United States, support for values of $\gamma \geq 1$ is provided by results contained in Chan et al. (1992), Conley et al. (1995), and Tauchen (1995).[22] For Germany, Dahlquist and Gray (1995) document a γ very close to one. Hence the parameter values required to explain the negative slope coefficient and the documented nonlinearities are consistent with values documented in other studies.

5. Conclusion

A long-standing puzzle is the empirical finding that the slope coefficient of the regression of change in exchange rate on interest differential across countries is negative, and uncovered interest parity is violated. One substantive economic implication of this finding, as shown by Fama (1984), is that the variance of the forward risk premium must exceed that of expected depreciation. New empirical evidence documented in this article suggests that for many currencies the opposite is true when the interest differential is negative. This implication is based on the empirical finding that the slope coefficient for many currencies systematically depends on the sign of the interest differential. It is negative when the differential is positive and positive when the differential is negative.

This article relies on the implications of single-factor term structure models to account for the negative slope coefficient and the documented nonlinearities. It is shown that to account for the negative slope coefficient, the market price of interest rate risk should be de-

[22] These articles estimate the volatility of the short rate, and in this article, the volatility of the 12-month bond return is estimated. There is very little difference between the two approaches. Over small intervals of time, the random part of the bond return is the duration-adjusted change in its interest rate. Hence the conditional volatility of the bond return is essentially determined by changes in its interest rate.

creasing in the interest rate. Asymmetry across the United States and foreign countries, where the market price of risk is decreasing in the interest rate for the United States and increasing in its interest rate for the foreign country, provides one set of sufficient conditions to account for the documented nonlinearities. Empirical results confirm this intuition. Many single-factor parametric term structure models cannot account for the negative slope coefficient, nor the documented nonlinearities.

An important extension of this article would be to develop a framework with quantitative implications that provide a detailed link between the market price of risk and aggregate economic variables. Such as extension would require a monetary general equilibrium model where the behavior and differences of monetary policy and/or the real risks across countries determine the differences in the behavior of market prices of risk across countries and consequently determine expected depreciation. One plausible explanation for asymmetry across the United States and the foreign countries (Germany and Japan) is that there are differences in the conduct of monetary policy across these economies. A model that incorporates such differences in monetary policy is the focus of current research.

References

Adler, M., and B. Dumas, 1983, "International Portfolio Choice and Corporation Finance: A Synthesis," *Journal of Finance*, 38, 925–984.

Aït-Sahalia, Y., 1995, "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385–426.

Bilson, J. F. O., 1981, "The 'Speculative Efficiency' Hypothesis," *Journal of Business*, 54, 435–452.

Backus, D., S. Foresi, and C. Telmer, 1995, "The Forward Premium Anomaly: Three Examples in Search of a Solution," working paper, Carnegie Mellon University.

Backus, D., A. Gregory, and C. Telmer, 1993, "Accounting for Forward Rates in Markets for Foreign Currency," *Journal of Finance*, 5, 1887–1908.

Bansal, R., R. Gallant, R. Hussey, and G. Tauchen, 1995, "Nonparametric Structural Estimation of Models for High Frequency Currency Market Data," *Journal of Econometrics*, 66, 251–287.

Bansal, R., and B. Lehmann, 1996, "Growth Optimal Portfolio Restrictions on Asset Pricing Models," working paper, Duke University.

Bekaert, G., 1996, "The Time Variation of Risk and Return in Foreign Exchange Markets: A General Equilibrium Perspective," *Review of Financial Studies*, 9, 427–470.

Brennan, M., and E. Schwartz, 1979, "A Continuous Time Approach to the Pricing of Bonds," *Journal of Banking and Finance*, 3, 133–155.

Campbell, J., A. Lo, and C. MacKinlay, 1995, *The Econometrics of Financial Markets*, unpublished manuscript.

Chan, K. C., G. A. Karolyi, F. Longstaff, and A. Sanders, 1992, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47, 1209–1227.

Conley, T., L. P. Hansen, E. Luttmer, and J. Scheinkman, 1995, "Short-Term Interest Rates as Subordinated Diffusions," working paper, University of Chicago.

Cox, J. C., and C. Huang, 1989, "Optimal Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process," *Journal of Economic Theory*, 49, 33–83.

Cox, J., J. Ingersoll, and S. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.

Cumby, R. E., and M. Obstfeld, 1984, "International Interest-Rate and Price-Level Linkages Under Flexible Exchange Rates: A Review of Recent Evidence," in J. F. O. Bilson and R. Marston (ed.), *Exchange Rates: Theory and Practice*, University of Chicago Press, Chicago.

Dahlquist, M., and S. Gray, 1995, "Regime-Switching and Interest Rates in the European Monetary System," working paper, Duke University.

Duffie, D., and R. Kan, 1993, "Yield Factor Model of Interest Rates," working paper, Stanford University.

Fama, E., 1984, "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 14, 319–338.

Ferson, W. E., and C. R. Harvey, 1993, "The Risk and Predictability of International Equity Returns," *Review of Financial Studies*, 6, 527–566.

Hakansson, N. H., 1971, "Multiperiod Mean-Variance Analysis: Toward a General Theory of Portfolio Choice," *Journal of Finance*, 26, 857–884.

Hansen, L. P., 1982, "Large Sample Properties of Generalized Method of Moment Estimators," *Econometrica*, 50, 1029–1054.

Hansen, L. P., and R. J. Hodrick, 1983, "Risk-Averse Speculation in the Forward Foreign Exchange Market: An Econometric Analysis of Linear Models," in J. A. Frenkel (ed.), *Exchange Rates and International Macroeconomics*, University of Chicago Press, Chicago.

Hansen, L. P., and R. Jagannathan, 1991, "Implications of Security Market Data for Models of Dynamic Economics," *Journal of Political Economy*, 99, 225–262.

Hansen, L. P., and K. J. Singleton, 1982, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1285.

Harrison, J. M., and D. Kreps, 1979, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20, 381–408.

Hsieh, D., 1984, "Tests of Rational Expectations and No Risk Premium in Forward Exchange Markets," *Journal of International Economics*, 17, 173–184.

Hodrick, R. J., 1987, *The Empirical Evidence of the Efficiency of the Forward and Futures Foreign Exchange Markets*, Hardwood Academic Publishers, Chur, Switzerland.

Hodrick, R. J., 1989, "Risk, Uncertainty and Exchange Rates," *Journal of Monetary Economics*, 23, 433–459.

Hodrick, R. J., and S. Srivastava, 1986, "The Covariation of Risk Premiums and Expected Future Spot Exchange Rates," *Journal of International Money and Finance*, 5, 5–22.

Huang, C., 1991, *An Overview of Modern Financial Economics*, unpublished manuscript.

Ingersoll, E. J., Jr., 1987, *Theory of Financial Decision Making*, Rowman & Littlefield, New Jersey.

Knez, P., B. Litterman, and J. Scheinkman, 1994, "Explorations into Factors Explaining Money Market Returns," *Journal of Finance*, 49, 1861–1882.

Long, J. B., Jr., 1990, "The Numeraire Portfolio," *Journal of Financial Economics*, 26, 29–69.

Longstaff, F., and E. Schwartz, 1992, "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *Journal of Finance*, 47, 1259–1282.

Lucas, R. E., Jr., 1982, "Interest Rates and Currency Prices in a Two-Country World," *Journal of Monetary Economics*, 10, 335–360.

Mark, N., 1988, "Time Varying Betas and Risk Premia in the Pricing of Forward Exchange Contracts," *Journal of Financial Economics*, 22, 335–354.

Malliaris, A. G., and W. A. Brock, 1991, *Stochastic Methods in Economics and Finance*, Elsevier Science Publishers, Amsterdam, Holland.

Merton, R., 1993, *Continuous-Time Finance*, Basil Blackwell, Oxford, U.K.

Newey, W., and K. West, 1987, "A Simple, Positive Semi-Definite, Hetroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 729–751.

Nielson, L., and J. Saá-Requejo 1994, "Exchange Rates and the Term Structure Dynamics and the Pricing of Derivative Securities," working paper, INSEAD.

Ogaki, M., 1992, "Generalized Method of Moments: Econometric Applications," working paper, University of Rochester.

Pearson, N. D., and T. S. Sun 1994, "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model," *Journal of Finance*, 49, 1279–1304.

Ross, S. A., 1978, "A Simplified Approach to the Valuation of Risk Streams," *Journal of Business*, 51, 453–475.

Tauchen, G., 1995, "New Minimum Chi-Square Methods in Empirical Finance," working paper, Duke University.

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.