Affine Currency Pricing Model with Regime Switching

Alexander Lebedinsky

Western Kentucky University

January 4, 2007
Abstract

Recent empirical studies show that, while uncovered interest parity fails at short horizons, there is more support for UIP at longer horizons. In this paper I show how these different results for UIP can be explained with a single model. The proposed model is a discrete-time affine stochastic discount factor model that allows for switching in mean and variance of the pricing kernel. Estimates of the model show that it can reproduce complex dynamics observed in the data.
1 Introduction

A number of recent studies revisit the issue of the forward premium anomaly (FPA) and find results contradictory to what seems to have become a stylized fact of international finance literature. The anomaly is a well-documented\(^1\) empirical regularity that uncovered interest parity (UIP) does not hold and that exchange rates move in the direction opposite to what UIP predicts. However, these recent empirical studies show that the anomaly may not be as prevalent of a phenomenon as previously believed, and there is more favorable evidence for UIP when it is tested over longer horizons (Meridith and Chinn, 1998; Razzak 2002)\(^2\).

The implications of this new empirical evidence are particularly interesting because the results are based on the same currencies for which UIP was rejected in earlier studies based on higher frequency data. Because UIP is one of the crucial assumptions for such well-known and widely used models of exchange rate determination as Dornbush (1976) and Krugman (1991), it’s good news that UIP might hold over the longer horizons, particularly over a period of one year or longer. Results based on low frequency data are more relevant when we study international transmission mechanisms of the monetary policy.

While this might be good news for monetary policy modelers, these results present a new challenge for international asset pricing literature. Empirical evidence suggests that the FPA exists in the high frequency data but does not exist in the low frequency data. To explain this phenomenon with a single model, we need a currency pricing model that must incorporate the length of the forecast horizon and carry different implications for different frequencies.

Since the inception of the FPA literature, time-varying risk premium has been the most frequently proposed explanation for the anomaly. It is also possible to explain the difference in the low- and high-frequency data with time-varying risk premium: If the variance of the risk premium diminishes as we forecast exchange rates over longer horizons, that would be consistent with what we observe in the data. Why would this happen? In the short

\(^1\)See Engel (1996) for a comprehensive survey
\(^2\)Bekaert et al. (2002), on the other hand, say that empirical evidence in favor of UIP depends on currency and not on horizon.
run, market participants respond to intermittent news, so we could observe a lot of daily variation in expectations. These day-to-day changes would not have much of an effect on the long-term outlook. So, if the risk premium is linked to economic fundamentals, it would not exhibit as much variation over the long run as in the short run. Its size can remain large, but its variability should diminish as the length of the horizon increases.

The model proposed in this paper could reconcile this conflicting evidence from low- and high-frequency studies. This model generates time-varying risk premium whose variance, depending on the parameters of the model, can become smaller with longer horizons. Therefore, it is possible to explain with a single model the FPA in the high-frequency data and UIP in the low-frequency data.

The model is a discrete-time stochastic discount factor (SDF) model that incorporates a Markov-switching (MS) process. The model uses as a basis the Vasicek (1977) bond pricing model and allows for switching in mean and variance parameters of the kernel equation. In the model, regimes follow an independent Markov process with constant transition probabilities. This model is not a learning model – market participants do not learn about the current regime; they are assumed to know what it is. This model can be generalized to any number of regimes, yet it retains the simple structure of affine models.

I estimate the model using monthly data on one-month forward and spot exchange rates for three pairs of currencies: British pound (GBP), German mark (DEM) and Japanese yen (JPY) versus U.S. dollar (USD). To estimate the model, I employ the maximum likelihood approach described in Hamilton(1994). At one-month frequency, parameter estimates for the GBP/USD and JPY/USD exchange rates imply a negative relationship between interest rate differential and exchange rate depreciation. Although the sign of the implied slope coefficient matches the sample estimate, the size of the implied slope of the forward premium regression is smaller than the sample counterpart. Using the standard error for the implied slope coefficient to compute the test statistic, one could reject the hypothesis that the implied slope coefficient is equal to the sample estimate. However, the small-sample distribution of the slope coefficient simulated under the null that the model is true shows that values below −1 are quite likely.

To analyze long-run predictions of the model, I compute the term structure of implied
slope coefficients as well as the variances of the risk premium and expected exchange rate depreciation using parameter estimates of the model. The results show that predictions of the model can vary widely depending on the length of the forecast. For example, in the case of the JPY/USD exchange rate, the model implies a negative correlation between the forward premium and exchange rate depreciation for shorter forward contracts (up to four months). For longer forward contracts, the slope coefficient of the forward premium regression is zero.

The MS-SDF model is closely linked to other recent efforts in currency pricing and bond pricing literature. For example, Backus, Foresi and Telmer (2001) showed that the FPA implies an inverse relation between conditional means and conditional higher moments of the stochastic discount factor. The model that I present here can generate such a relation between conditional means and conditional second moments through switching in regimes.

There is ample empirical evidence that switching models can address the issues of nonlinearities in the data, particularly discrete jumps in level and variance. In the context of exchange rates, regime switching models were employed by Engel and Hamilton (1990) and Evans and Lewis (1995). Clarida, Sarno, Taylor and Valente (2003) incorporate a Markov-switching process in the VEC model that allows for switching in the mean and variance. They find that shifts from one regime to another are “largely due to shifts in the variance of the term structure equilibrium.” Other recent applications of the switching models are in the papers by Ang and Bekaert (2002a and 2002b) and Smith (2002), who present evidence that switching models describe well the volatility patterns in the interest rates. Allowing for switching in the inflation process, Evans (2003) shows that this type of models reflect accurately the term structure of nominal and real yields.

In the following section, I discuss in more detail the nature of the forward premium anomaly and how it can be addressed using SDF models. Following that, I describe the setup of my model in the third section, then discuss the estimation results in section four. Section five contains an economic analysis of the model.
2 Stochastic Discount Factor Models and Currency Pricing

2.1 Forward Premium Anomaly

The literature on the FPA documents the widespread failure of UIP and investigates its causes. The studies of UIP are focused around the following regression equation:

\[ s_{t+1} - s_t = \alpha_1 + \alpha_2 (f_t - s_t) + \varepsilon_t, \] (1)

where \( s_t \) and \( f_t \) are logs of the spot and the forward exchange rates defined in dollars per unit of foreign currency. An overwhelming majority of studies find \( \alpha_2 \) negative and significantly different from zero. Most find at least some evidence of the anomaly by showing that \( \alpha_2 \) is more than two standard errors away from 1, which is interpreted as a failure of the hypothesis that the forward rate is an unbiased predictor of the spot rate (the "unbiasedness hypothesis").

The sample used in this paper covers the period from January 1975 through December 2001. Least squares regression estimates of equation 1 based on this sample also show that \( \alpha_2 \) is negative and significantly different from 0 for the JPY/USD exchange rate and significantly different from 1 for the GBP/USD and DEM/USD exchange rates (shown in Table 1).

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_1 )</td>
<td>-0.0032 (0.0022)</td>
<td>0.0010 (0.0021)</td>
<td>0.0066** (0.0025)</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>-0.8750 (0.6790)</td>
<td>-0.5897(0.6763)</td>
<td>-1.4890** (0.6369)</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.005</td>
<td>0.002</td>
<td>0.016</td>
</tr>
</tbody>
</table>

The dependent variable is exchange rate depreciation \( s_{t+1} - s_t \). The explanatory variable is forward premium \( f_t - s_t \). The numbers in parentheses are OLS standard errors associated with each coefficient. Double asterisk indicates significance at 5% confidence level.

Proposing an explanation of this anomaly, Fama (1984) decomposed the forward premium into the risk premium \( p_t \) and the expected rate of depreciation \( q_t \):

\[ f_t - s_t = f_t - E_t s_{t+1} + (E_t s_{t+1} - s_t) \equiv p_t + q_t. \]
This allows us to write the OLS estimator of $\alpha_2$ as

$$\alpha_2 = \frac{\text{Cov}(q, p+q)}{\text{Var}(p+q)} = \frac{\text{Cov}(q, p) + \text{Var}(q)}{\text{Var}(p+q)}. \quad (2)$$

This decomposition is true under the assumption of rational expectations when the difference between the expectation ($E_t s_{t+1} - s_t$) and the realization of exchange rate depreciation $s_{t+1} - s_t$ is not correlated with variables known at time $t$. This expression shows that, for a model to be able to reproduce the forward premium anomaly, two conditions must hold: a) The covariance of $p_t$ and $q_t$ must be negative and b) The absolute value of the covariance of $p_t$ and $q_t$ should exceed the variance of $q_t$. Thus, variation in risk premium over time is crucial to the model’s ability to reproduce the anomaly, and its variance must be greater than the variance of the expected rate of depreciation.

### 2.2 Applying Stochastic Discount Factor Models to Currency Pricing

The intuition for using stochastic discount factor models for currency pricing presented here closely follows Backus, Foresi and Telmer (2001). First, assume there exist two economies with integrated financial and currency markets: a domestic economy whose currency is denominated in dollars and a foreign economy whose currency denominated in pounds. Then, allowing no pure arbitrage opportunities between these economies, the asset pricing relations for each of these economies can be derived as

$$b_t = E_t [c_{t+1} M_{t+1}], \quad (3)$$

where $b_t$ is the dollar value of an asset paying $c_{t+1}$ in $t + 1$, and $M_{t+1}$ is the stochastic discount factor or pricing kernel.

For the assets denominated in pounds, one can write a similar expression:

$$b_t^* = E_t [c_{t+1}^* M_{t+1}^*],$$

where asterisks denote variables pertinent to the foreign economy. If the financial markets are fully integrated, pound denominated assets can be priced using the dollar pricing
kernel. Then, the pricing relation for foreign assets using the domestic pricing kernel can be expressed in the following way:

\[ b_t^* = E_t \left[ c_{t+1}^* \left( \frac{S_{t+1}}{S_t} \right) M_{t+1} \right], \]

where \( S_t \) represents the spot exchange rate at time \( t \).

If there are no arbitrage opportunities, it also must be true that

\[ E_t \left[ c_{t+1}^* M_{t+1}^* \right] = E_t \left[ c_{t+1}^* \left( \frac{S_{t+1}}{S_t} \right) M_{t+1} \right] \tag{4} \]

because the price of the asset should be the same regardless of which kernel is used to price it.

This relation is satisfied when \( M_{t+1} \left( \frac{S_{t+1}}{S_t} \right) = M_{t+1}^* \), and this solution is unique if asset and currency markets are complete. This result can be used to write the log of depreciation of the exchange rate in the following form:

\[ m_{t+1}^* - m_{t+1} = s_{t+1} - s_t, \tag{5} \]

where \( m \) is the log of the pricing kernel. This equation ties two markets together by describing three stochastic processes and makes one of them redundant. Specifying the stochastic behavior of \( M \) and \( M^* \) also determines how \( \frac{S_{t+1}}{S_t} \) should behave. When markets are incomplete, the choice of \( M \) and \( M^* \) satisfying equation 4 is not unique. However, one still could choose them to satisfy equation 4.

The expression for the forward premium can be derived using equation 5 and kernel equations. Consider a forward contract at period \( t \) to exchange one pound into \( F_t \) dollars at time \( t + 1 \) (so \( F_t \) is the forward rate.) This contract requires no payment at time \( t \), so \( b_t = 0 \). The net cash flow of this contract in dollars at time \( t + 1 \) is just \( F_t - S_{t+1} = c_{t+1} \). Substituting this into equation 3, we can price the forward contract as:

\[ 0 = E_t \left[ (F_t - S_{t+1}) M_{t+1} \right]. \]
Dividing by $S_t$ and using equation 4, we obtain

$$\left(\frac{E_t}{S_t}\right) E_t M_{t+1} = E_t \left( M_{t+1} \frac{S_{t+1}}{S_t} \right) = E_t M_{t+1}^*.$$

Taking the log of the last equation gives the expression for the forward premium in terms of pricing kernels:

$$f_t - s_t = \log E_t M_{t+1}^* - \log E_t M_{t+1}.$$

Until this point, no assumptions have been made about stochastic behavior of the pricing kernels. Comparing equations 6 and 5 reveals that the differences in the stochastic behavior of $E_{t+s} - s_t$ and $f_t - s_t$ will come from the difference in $\log E_t M_{t+1}$ and $E_t \log M_{t+1}$. In the next section, I set up the model of stochastic behavior for $M$ and derive the implications for the forward premium and exchange rate depreciation.

3 The MS-SDF Model

If time-varying risk premium is indeed responsible for the existence of the FPA, then only those asset pricing models that allow for variation in risk premium can be used for currency pricing. The Vasicek model in its original form would not be able to explain the anomaly because it implies that risk premium is constant. When the Vasicek model is modified to allow for switching in the parameters of the kernel equation, it can generate time-varying risk premium. Regime-switching also allows for negative correlation between the risk premium and expected exchange rate depreciation. The latter is a necessary condition to explain the FPA.

The MS-SDF model consists of the usual components of affine SDF models: a state equation and kernel equations. The model is expressed in discrete time. In the simplest, single state variable version of the model, the state equation describes the law of motion of the state variable $x_t$ as an AR(1) process:

$$x_{t+1} = \phi x_t + \varepsilon_{t+1},$$

7
where $\phi$ is a scalar and $0 < \phi < 1$. Innovation term $\varepsilon_t$ is distributed as $N(0, 1)$. The kernel equations describe how the state variable relates to the stochastic discount factor:

$$-m_{t+1} = \delta(z_t) + \beta x_t + \lambda(z_t) \varepsilon_{t+1} + \omega_{t+1}$$

(8)

$$-m^*_{t+1} = \delta^*(z_t) + \beta^* x_t + \lambda^*(z_t) \varepsilon_{t+1} + \omega^*_{t+1},$$

(9)

where $m_t = \log M_t$, $\omega_t \sim N(0, \sigma_\omega)$, $\omega^*_{t} \sim N(0, \sigma_{\omega^*})$ and asterisks denote the variables pertinent to the foreign economy. Parameters $\delta(z_t), \delta^*(z_t), \lambda^*(z_t)$ and $\lambda(z_t)$ depend on regime $z_t$. Parameters $\beta$ and $\beta^*$ are scalars that don’t change with regimes. Together these equations imply that, within a regime, the stochastic discount factor depends on the state variable $x_t$ the same way it does in the Vasicek model. However, the way this state variable affects the discount factor changes when regimes switch. In other words, the state process remains the same across regimes, but investors price risk differently in each regime.

The regimes are assumed to follow an independent first-order Markov process. The transition probabilities for regimes are $Pr(z_t = i | z_{t-1} = i) = \pi_i$, where $i = 1, 2$. Implications of the model hold for any number of regimes, but, for simplicity of exposition, assume that there are only two regimes. Also, empirical results presented below show that, with just two regimes, the model can reproduce the forward premium anomaly. The market participants are assumed to know what the current regime is to avoid complicating the model with issues of learning about the current regime.

To find the forward premium, use the model to solve for $-\log E_t M_{t+1}$ and $-\log E_t M^*_{t+1}$:

$$f_t - s_t = -\log E_t M_{t+1} + \log E_t M^*_{t+1}$$

(10)

$$= \delta(z_t) - \delta^*(z_t) + (\beta - \beta^*) x_t - \frac{1}{2} \left( \lambda(z_t)^2 + \sigma_\omega^2 - \lambda^*(z_t)^2 - \sigma_{\omega^*}^2 \right)$$

$$= \nabla \delta_t - \frac{1}{2} \left( \nabla \lambda_t^2 - \nabla \sigma_\omega^2 \right) + \nabla \beta x_t.$$
Exchange rate depreciation can be found using equations 5, 8 and 9 as

\[ s_{t+1} - s_t = -m_{t+1} + m^*_{t+1} \]

\[ = \delta (z_t) - \delta^* (z_t) + (\beta - \beta^*) x_t + [\lambda (z_t) - \lambda^* (z_t)] \xi_{t+1} + \omega_{t+1} - \omega^*_{t+1} \]

\[ = \nabla \delta_t + \nabla \beta x_t + \nabla \lambda_t \xi_t + \omega_{t+1} - \omega^*_{t+1}. \]

To simplify notation, in the last two equations \( \nabla \) denotes the difference in parameters across the two countries within a regime (e.g. \( \nabla \delta_t = \delta (z_t) - \delta^* (z_t) \)). The time subscript \( t \) on parameters implies that they are a function of the current regime. When a parameter appears with subscript \( i \), that denotes the value of that parameter in regime \( z_i \).

Conditional expectation at time \( t \) of the exchange rate depreciation can be found as

\[ E_t [s_{t+1} - s_t] = \nabla \delta_t + \nabla \beta x_t. \]

The risk premium is the difference between the forward premium and expected exchange rate depreciation:

\[ p_t = -\frac{1}{2} (\nabla \lambda_t^2 + \nabla \sigma^2_\omega). \]

Thus, the risk premium depends on the current regime. The last expression clearly demonstrates why the Vasicek model would not be able to account for the forward premium anomaly: Without switching, \( p_t \) is constant over time. Hence, the slope of the forward premium regression \( \alpha_2 \) equals 1 because \( Var (p_t) = 0 \). However, just because the model can generate a time-varying risk premium, it does not automatically imply that it can reproduce the forward premium anomaly. It is also necessary to have negative correlation between expected exchange rate depreciation and the risk premium. To see that this model is capable of reproducing the forward premium anomaly, consider the expression for the
slopes coefficient of the forward premium regression derived from this model:

$$\alpha_2 = \frac{\text{Cov} (p_t, q_t) + \text{Var} (q_t)}{\text{Var} (p_t + q_t)} = \frac{\text{Cov} (\nabla \delta_t, p_t) + \text{Var} (\nabla \delta_t) + \nabla \beta^2 \text{Var} (z_t)}{\text{Var} (\nabla \delta_t + p_t) + \nabla \beta^2 \text{Var} (z_t)}.$$ 

This expression shows that, if $\nabla \delta_t$ and $p_t$ are negatively correlated, then $\alpha_2$ could be negative when $|\text{Cov} (\nabla \delta_t, p_t)| > \text{Var} (\nabla \delta_t) + \nabla \beta^2 \text{Var} (z_t)$.

4 Estimation Results and Statistical Analysis of the Model

In this section, I present estimation results for two versions of the model: a model with two regimes and a three-regime model. I do not conduct any explicit test on which model is “better.” Instead, the aim of this paper is to show that a parsimonious regime-switching model is capable of reproducing the FPA along with other features of the data. Because the estimation results do not differ very much between the two- and three-regime models, I discuss the results from the two-regime model in detail and leave the estimates of three-regime model for the appendix.

To estimate the model, I use monthly data on three spot and one-month forward exchange rates: British pound/U.S. Dollar, German mark/U.S. dollar and Japanese yen/U.S. dollar. The data cover the period from January 1975 through December 2001. Even though the model is expressed in terms of an unobservable state variable $x_t$, it can be rewritten in terms of observable variables only and estimated using the algorithm described in Hamilton (1994). The estimation equations are

$$f_{t+1} - s_{t+1} = \nabla \delta_{t+1} + p_{t+1} + \phi [(f_t - s_t) - \nabla \delta_t - p_t] + \nabla \beta \varepsilon_{t+1},$$ (12)

$$s_{t+1} - s_t = (f_t - s_t) - p_t + \nabla \lambda_{t+1} + \nabla \omega_{t+1}.$$ (13)

---

For a reference on how to test models with different numbers of states, see Evans (2003) and Clarida, Sarno, Taylor and Valente (2003).

The dataset is from Bekaert and Hodrick (1993) updated with the data from Datastream.
Equations (12) and (13) are obtained by substituting the solution for the state variable in terms of the forward premium (10) into the state equation (7) and the exchange rate depreciation equation (11). Not all parameters of the model can be identified. Parameters $\delta_i, \delta^*_i, \beta$ and $\beta^*$ and the innovations $\omega_{t+1}$ and $\omega^*_{t+1}$ cannot be identified separately. I also impose identifying restrictions on variances, which are identified only up to a factor of proportionality. Variance of $\varepsilon_t$ is set to 1; variances of $\omega_t$ and $\omega^*_t$ are assumed to be the equal. The restrictions allow us to identify $\lambda_i$ and $\lambda^*_i$. Besides the level of these two parameters, these identifying restrictions do not have an effect on any results presented below.

4.1 Model Estimates and Specification Tests

Table 2 presents the estimates of the model. The upper panel of the table contains the structural parameter estimates, and the bottom panel of the table shows reduced form parameters as they appear in equations 12 and 13.

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9330** (0.0136)</td>
<td>0.9631** (0.0175)</td>
<td>0.9356** (0.0185)</td>
</tr>
<tr>
<td>$\sigma_{\gamma_\omega}$</td>
<td>0.0317** (0.0009)</td>
<td>0.0331** (0.0012)</td>
<td>0.0348** (0.0011)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-8.2980** (0.3082)</td>
<td>-6.2935** (2.6586)</td>
<td>-12.8543* (7.5845)</td>
</tr>
<tr>
<td>$\lambda^*_1$</td>
<td>-8.2965** (0.3053)</td>
<td>-6.2965** (2.6586)</td>
<td>-12.8535* (7.5844)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-11.0488** (0.3355)</td>
<td>-1152.6** (0.0127)</td>
<td>8.0922* (4.5564)</td>
</tr>
<tr>
<td>$\lambda^*_2$</td>
<td>-11.0466** (0.3335)</td>
<td>-1152.6** (0.0128)</td>
<td>8.0892* (4.5564)</td>
</tr>
<tr>
<td>Pr ($z_t = 1</td>
<td>z_{t-1} = 1$)</td>
<td>0.9165** (0.0305)</td>
<td>0.9690** (0.0171)</td>
</tr>
<tr>
<td>Pr ($z_t = 2</td>
<td>z_{t-1} = 2$)</td>
<td>0.9606** (0.0156)</td>
<td>0.6807** (0.1162)</td>
</tr>
<tr>
<td>$\nabla \beta$</td>
<td>0.0010** (0.0000)</td>
<td>0.0007** (0.0000)</td>
<td>0.0009** (0.0000)</td>
</tr>
<tr>
<td>$\nabla \delta_1$</td>
<td>-0.0065* (0.0036)</td>
<td>0.0003 (0.0023)</td>
<td>-0.0039 (0.0031)</td>
</tr>
<tr>
<td>$\nabla \delta_2$</td>
<td>0.0017 (0.0021)</td>
<td>-0.0028 (0.0060)</td>
<td>0.0058** (0.0026)</td>
</tr>
<tr>
<td>$\nabla \lambda_1$</td>
<td>-0.0015 (0.0029)</td>
<td>-0.0030 (0.0021)</td>
<td>-0.0009 (0.0030)</td>
</tr>
<tr>
<td>$\nabla \lambda_2$</td>
<td>-0.0023 (0.0019)</td>
<td>0.0000 (0.0044)</td>
<td>0.0030 (0.0030)</td>
</tr>
<tr>
<td>$p_1$</td>
<td>0.0077** (0.0026)</td>
<td>0.0015 (0.0020)</td>
<td>0.0089** (0.0028)</td>
</tr>
<tr>
<td>$p_2$</td>
<td>-0.0049** (0.0019)</td>
<td>0.0012 (0.0045)</td>
<td>0.0045** (0.0022)</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>2672.02</td>
<td>2764.70</td>
<td>2658.37</td>
</tr>
</tbody>
</table>

Estimates of the model were obtained using the algorithm described in Hamilton (1994). Standard errors (in parentheses) are calculated using the outer product of the gradients. A double asterisk indicates that a parameter is significant at 5% confidence level; single indicates a 10% confidence level.
As can be seen from the table, most of the estimates are significantly different from zero. However, some of the reduced form parameters ($\nabla \delta_i, \nabla \lambda_i, p_i$) are not significantly different from zero. Switching in these parameters is crucial for the model’s ability to reproduce the forward premium anomaly. So a more relevant question is whether there is a significant difference in these parameters across regimes. Below I present results of a formal test of switching in these parameters across regimes.

To examine how well the model fits the data, we can compare the sample moments from the data and their implied estimates from the model. The model should be able to reproduce other well-known and well-documented features of the data in addition to the FPA. For example, forward premium is typically very persistent. Exchange rate depreciation shows little persistence and has a mean close to zero. Variance of exchange rate depreciation is greater than the variance of the forward premium. Table 3 shows the sample and implied moments. Expressions for these moments in terms of model parameters are derived in the appendix.

### Table 3: Sample and implied moments

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[f_t - s_t]$</td>
<td>Sample</td>
<td>Implied</td>
<td>Sample</td>
</tr>
<tr>
<td></td>
<td>-0.00199</td>
<td>-0.00175</td>
<td>0.00151</td>
</tr>
<tr>
<td></td>
<td>(0.00095)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E[s_{t+1} - s_t]$</td>
<td>-0.00151</td>
<td>-0.00101</td>
<td>0.00013</td>
</tr>
<tr>
<td></td>
<td>(0.00232)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Var(f_t - s_t)$</td>
<td>$7 \times 10^{-6}$</td>
<td>$1 \times 10^{-6}$</td>
<td>$7 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>(2 $\times 10^{-6}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Var(s_{t+1} - s_t)$</td>
<td>0.00102</td>
<td>0.00103</td>
<td>0.00109</td>
</tr>
<tr>
<td></td>
<td>(0.00006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Corr(f_t - s_t)$</td>
<td>0.84311</td>
<td>0.91288</td>
<td>0.91422</td>
</tr>
<tr>
<td>$f_{t-1} - s_{t-1}$</td>
<td>(0.01596)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Corr(s_{t+1} - s_t, s_t - s_{t-1})$</td>
<td>0.06728</td>
<td>0.01902</td>
<td>0.00771</td>
</tr>
<tr>
<td></td>
<td>(0.01092)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Implied moments are computed from the parameters of the model. Analytical solutions for implied moments are shown in the appendix. Standard errors (shown in parentheses) are computed using the delta method.

The table shows that, with few exceptions ($Var(f_t - s_t), Corr(f_t - s_t, f_{t-1} - s_{t-1})$ for GBP/USD, Corr($f_t - s_t, f_{t-1} - s_{t-1}$) for JPY/USD), implied moments are within one
standard error from the sample estimates. So, judging by how well the model matches sample moments, it works well in fitting the data.

Another specification test is presented in Table 4 below. The null hypothesis is that there is no serial correlation in the residuals within a regime. According the p-values shown in the table, the null hypothesis cannot be rejected (at a 5% confidence level) for any of the exchange rates. For GBP/USD, we can reject the null that \( \varepsilon_t \) has no autocorrelation in regime 2 at a 10% confidence level. Also for DEM/USD, the null hypothesis of no autocorrelation can be rejected for \( \nabla \omega_t \) when \( z = 2 \). Overall, this collection of tests shows no evidence of misspecification.

Table 4: p-values for the LM test of serial correlation in the residuals within a regime

<table>
<thead>
<tr>
<th>( H_0 )</th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E(\varepsilon_t, \varepsilon_{t-1}</td>
<td>z_t = z_{t-1} = 1) )</td>
<td>0.74</td>
<td>0.19</td>
</tr>
<tr>
<td>( E(\varepsilon_t, \varepsilon_{t-1}</td>
<td>z_t = z_{t-1} = 2) )</td>
<td>0.07</td>
<td>0.19</td>
</tr>
<tr>
<td>( E(\nabla \omega_t, \nabla \omega_{t-1}</td>
<td>z_t = z_{t-1} = 1) )</td>
<td>0.47</td>
<td>0.14</td>
</tr>
<tr>
<td>( E(\nabla \omega_t, \nabla \omega_{t-1}</td>
<td>z_t = z_{t-1} = 2) )</td>
<td>0.13</td>
<td>0.08</td>
</tr>
</tbody>
</table>

4.2 Regimes

Estimates of \( \Pr(z_t = i|z_{t-1} = i) \) presented in Table 2 suggest that regimes are fairly persistent. In most cases they are above 0.9. Ergodic probabilities of being in regime 1 computed from transition probabilities are 0.32, 0.91 and 0.34 for GBP/USD, DEM/USD and JPY/USD exchange rates respectively. So in the case of DEM/USD exchange rate, the estimates imply that 91% of the observations come from regime 1. This casts a bit of doubt on the results obtained from that exchange rate. Because the sample size is 323, only about 30 observations are identified with regime 2. So for DEM/USD it’s hard to place a lot of confidence in the parameter estimates for the second regime.

Table 5 contains p-values of the Wald test on the hypothesis that parameters do not change across the regimes. So, under the assumption that a model with two regimes is specified correctly, this test examines whether parameters are different across the regimes.

\(^6\)For Markov-switching models, Hamilton (1996) describes other specification tests such as ARCH and omitted variable test.
Table 5: Wald test of switching in parameters across regimes (p-values)

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nabla \delta_1 = \nabla \delta_2$</td>
<td>0.02</td>
<td>0.52</td>
<td>0.00</td>
</tr>
<tr>
<td>$p_1 = p_2$</td>
<td>0.00</td>
<td>0.96</td>
<td>0.00</td>
</tr>
<tr>
<td>$\nabla \lambda_1 = \nabla \lambda_2$</td>
<td>0.80</td>
<td>0.52</td>
<td>0.40</td>
</tr>
</tbody>
</table>

The table shows the p-values for the Wald test statistic of the null hypothesis that the parameters are the same in both regimes.

The p-values show that the null hypothesis can't be rejected for $\nabla \lambda_i$ for all three exchange rates. Also for DEM/USD exchange rate, the null hypotheses that $\nabla \delta_1 = \nabla \delta_2$ and $p_1 = p_2$ cannot be rejected. But for the other two exchange rates, the test decisively rejects the null hypothesis that risk premia do not switch between regimes.

4.3 Does the Model Reproduce the FPA at Short Horizons?

Table 6 compares the sample estimates of the slope coefficient of the forward premium at a one-month horizon and the slope coefficient implied by the model. The lower panel of the table shows the p-values of the Wald test on the null hypotheses that $\alpha_2 = 0$, $\alpha_2 = -0.5$ and $\alpha_2 = 1$. In two out of three cases, parameter estimates of the model imply that the slope coefficient is negative (however, not significantly different from zero). In the case of DEM/USD exchange rate, the model suggests that the anomaly is not present, even though the OLS estimate of the slope coefficient is negative. This is consistent with the results from the previous section that showed no evidence of changes in parameters across regimes for DEM/USD.

For JPY/USD and GBP/USD, one could not reject the null hypothesis $\alpha_2 = -0.5$, and for JPY/USD one could not reject the null hypothesis that $\alpha_2 = -1$ at a 95% confidence level. But what’s most apparent is that the implied estimates of $\alpha_2$ are quite far from their OLS counterparts in all three cases.

Instead of relying solely on asymptotic theory, I further explore statistical properties of the implied slope coefficient of the forward premium regression under the null of MS-SDF. To examine the distribution of $\alpha_2$ in a large sample under the null of MS-SDF, I simulated distributions of implied $\alpha_2$ for each exchange rate. For each set of parameter estimates, I generated 100,000 draws from the distribution $N(\theta, \Sigma)$, where $\theta$ is the parameter vector
Table 6: Comparison of the OLS estimate and implied value of the Fama Coefficient

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2$ OLS estimate</td>
<td>-0.875 (0.679)</td>
<td>-0.590 (0.676)</td>
<td>-1.489 (0.637)</td>
</tr>
<tr>
<td>$\alpha_2$ Implied</td>
<td>-0.038 (0.357)</td>
<td>0.992 (0.192)</td>
<td>-0.211 (0.386)</td>
</tr>
</tbody>
</table>

$H_0$

<table>
<thead>
<tr>
<th>$\alpha_2$</th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2 = 0$</td>
<td>0.91</td>
<td>0.00</td>
<td>0.60</td>
</tr>
<tr>
<td>$\alpha_2 = -0.5$</td>
<td>0.19</td>
<td>0.00</td>
<td>0.48</td>
</tr>
<tr>
<td>$\alpha_2 = -1$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
</tr>
</tbody>
</table>

The top panel in the table compares OLS estimates of the slope coefficient in the forward premium regression and its value implied by parameter estimates of the two-regime MS-SDF model. The bottom panel of the table contains the p-values of the Wald test statistics for the null hypotheses that $\alpha_2 = 0, -0.5$ and -1.

Table 7: Large sample empirical distribution of $\alpha_2$ under the null of SDF-MS model

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2 &lt; 0$</td>
<td>0.46</td>
<td>0.00</td>
<td>0.60</td>
</tr>
<tr>
<td>$\alpha_2 &lt; -0.5$</td>
<td>0.09</td>
<td>0.00</td>
<td>0.18</td>
</tr>
<tr>
<td>$\alpha_2 &lt; -1$</td>
<td>0.01</td>
<td>0.00</td>
<td>0.02</td>
</tr>
</tbody>
</table>

100,000 draws were generated from the distribution $N(\theta, \Sigma)$, where $\theta$ is the vector of parameter estimates of the two-regime MS-SDF model $\theta = [\phi, \nabla \beta, \nabla \lambda_1, \nabla \lambda_2, p_1, p_2, \sigma_{\nabla \omega}, \pi_1, \pi_2, \nabla \delta_1, \nabla \delta_2]$ and $\Sigma$ is the covariance matrix of the parameters. Then, for each draw of the parameter vector, $\alpha_2$ is computed using ergodic probabilities of each regime. The table represents the proportion of draws that resulted in $\alpha_2$ below a certain value.

$\theta = [\phi, \nabla \beta, \nabla \lambda_1, \nabla \lambda_2, p_1, p_2, \sigma_{\nabla \omega}, \pi_1, \pi_2, \nabla \delta_1, \nabla \delta_2]$ and $\Sigma$ is the covariance matrix of the parameters. Then, for each draw of the parameter vector, I computed analytically the implied slope coefficient. Because an analytical solution for the slope coefficient depends on ergodic probabilities of each regime, these results can be interpreted as the large-sample distribution of the slope coefficient under the null of MS-SDF.

Table 7 shows probabilities of observing different values of $\alpha_2$ in a large sample under the null of MS-SDF. These results appear to be in line with the results of the Wald test. For DEM/USD, the probability of observing $\alpha_2 < 0$ is (almost) zero. For GBP/USD and JPY/USD exchange rates, the probability of $\alpha_2$ being as small as the OLS estimate is still very low.

The discrepancy between the implied and the actual value of $\alpha_2$ could be attributed to small-sample properties of the OLS estimate. To explore this possibility, I simulate distributions of the small-sample estimates of $\alpha_2$ under the null hypothesis that the model
correctly characterizes the data. Taking the model parameters as given, I generate 100,000 datasets with 323 observations each to match the length of the actual sample. Then for each of those datasets, I compute $\alpha_2$ using the OLS. Table 8 shows the probability of observing various $\alpha_2$ in a sample with 323 observations. Under the null of MS-SDF and a set of parameter estimates, these results suggest that the large negative values of the OLS estimate of $\alpha_2$ could be due to the small size of the sample. For GBP/USD and JPY/USD, the probabilities of observing $\alpha_2 < 1$ are 14% and 18% respectively. For DEM/USD, it is also plausible to observe a negative slope coefficient in a small sample.

Table 8: Empirical distribution of small sample estimate of $\alpha_2$ under the null of SDF-MS model

<table>
<thead>
<tr>
<th>$H_0$</th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2 &lt; -0$</td>
<td>0.72</td>
<td>0.07</td>
<td>0.67</td>
</tr>
<tr>
<td>$\alpha_2 &lt; -0.5$</td>
<td>0.45</td>
<td>0.01</td>
<td>0.40</td>
</tr>
<tr>
<td>$\alpha_2 &lt; -1$</td>
<td>0.14</td>
<td>0.00</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Using the parameter estimates of the two-regime MS-SDF model, 100,000 datasets were generated. Each dataset contains 323 observations. For each dataset, OLS estimate of the slope coefficient $\alpha_2$ is computed. The table shows proportion of datasets that produced $\alpha_2$ below a certain value.

5 Term Structure of the Risk Premium and Slope of Forward Premium Regression

The MS-SDF model specifies the dynamics of forward premia ($f_{n,t} - s_t$) for any horizon of the forward contract and exchange depreciation ($s_{t+n} - s_t$) over any $n$ periods. Thus, it allows us to investigate the implied term structure of the risk premium and expected exchange rate depreciation. With this model we can construct implied forward premia for horizons $n > 1$ using information on one-month forward premium and the estimate of the current regime. Given a set of parameter estimates, we also can construct implied unconditional variances of the forward premia, expected exchange rate depreciation and risk premia for any horizon. Hence, we can derive the term structure of the implied slope coefficient of the forward premium regression as well. Therefore, using the estimates of the model based on one-month forward premia and exchange rate depreciation, we can examine predictions of the model for the UIP over longer horizons. In this section, I compare the
two-regime and three-regime versions of the model.

The expressions for the $n$-period forward premium and exchange rate implied by the MS-SDF model can be derived as:

$$f_{n,t} - s_t = A_{n,t} + B_n x_t,$$

(14)

$$s_{t+n} - s_t = D_{n,t} + B_n x_t + \xi_{t+n},$$

(15)

where $n$ denotes the length of the horizon. Coefficients $A_{n,t}$ and $D_{n,t}$ are functions of model parameters and depend on the current regime and expectations of future regimes. Coefficient $B_n$ does not change across regimes and depends only on the length of the horizon. Derivation of these coefficients is shown in the appendix.

The $n$-period equivalent of the forward premium regression equation 1 can be written as

$$\frac{(s_{t+n} - s_t)}{n} = \alpha_{1,n} + \alpha_{2,n} \frac{(f_{n,t} - s_t)}{n} + \epsilon_{t+n},$$

where $\alpha_{1,n}$ and $\alpha_{2,n}$ are intercept and slope coefficients for $n$-period horizon.

Using expressions 14 and 15, we can derive the implied slope of the forward premium regression for any horizon:

$$\alpha_{2,n} = \frac{Cov(A_{n,t}, D_{n,t}) + (B_n)^2 Var(x)}{Var(A_{n,t}) + (B_n)^2 Var(x)}.$$ 

(16)

The expressions for the per-period risk premium $p_{n,t}$ for $n$-period horizon can be found as

$$np_{n,t} = f_{n,t} - s_t - E_t (s_{t+n} - s_t)$$

$$= A_{n,t} - D_{n,t},$$

(17)
and expected exchange rate depreciation is $q_{n,t}$

\[
n q_{n,t} = E_t (s_{t+n} - s_t) = D_{n,t} + B_n x_t.
\]

Therefore, at long horizons $\alpha_{2,n}$ is negative when $\text{Cov}(A_{n,t}, D_{n,t}) < 0$ and $|\text{Cov}(A_{n,t}, D_{n,t})| > (B_n)^2 \text{Var}(x)$. It is easy to show that the intuition for explaining the FPA for the one-period horizon translates into long horizons: The variance of the risk premium has to be larger than the variance of expected exchange rate depreciation when $\alpha_{2,n} < 0$.

Two-Regime MS-SDF: Figures 1, 2 and 3 show the term structure of the slope coefficient for up to three years ($n = 36$). To explain the changes in the implied slope coefficient over different horizons, I also show the graphs of the term structure of the covariance between $p_{n,t}$ and $q_{n,t}$, the ratio of their variances and the variance of the risk premium alone. When covariance between $p_{n,t}$ and $q_{n,t}$ goes to zero, $\alpha_{2,n}$ should be positive. For the slope coefficient $\alpha_{2,n}$ to be negative, the variance of the risk premium must be greater than the variance of the expected rate of depreciation, so the ratio $\text{Var}(p)/\text{Var}(q)$ must be less than 1.

The graphs of the implied slope coefficient in the long-horizon forward premium regression show that $\alpha_{2,n}$ behaves very differently across currencies and along horizons. GBP/USD and JPY/USD exchange rates exhibit very similar patterns. For both of these exchange rates, as $n$ increases, the implied slope coefficient increases to about 1.4 at $n = 2$ and then declines with higher maturity. At $n > 20$, $\alpha_{2,n}$ is close to zero. So, at shorter horizons (two to four months), UIP should hold, and at longer horizons UIP does not hold. This happens because, as $n$ increases, $\text{Var}(q)$ shrinks relatively to $\text{Var}(p)$, so the slope coefficient $\alpha_{2,n}$ goes to 0. Even though one of the conditions for the FPA is satisfied – variability of the risk premium is far greater than variability of exchange rate depreciation – covariance between these two variables becomes smaller, and the slope coefficient is close to zero as $n$ rises.

In the case of DEM/USD exchange rate, $\alpha_{2,1}$ is close to unity. For $n > 1$, $\alpha_{2,n}$ is close to zero. The graph shows that covariance between the risk premium and exchange rate
Figure 1: Two-regime model: Term Structure of implied slope coefficient $\alpha_{2,n}$, $\text{Cov}(q_{n,t}, p_{n,t})$, $\frac{\text{Var}(p_{n,t})}{\text{Var}(q_{n,t})}$ and $\text{Var}(p_{n,t})$ for GBP/USD
Figure 2: Two-regime model: Term Structure of implied slope coefficient $\alpha_{2,n}$, $\text{Cov}(q_{n,t}, p_{n,t})$, $\frac{\text{Var}(p_{n,t})}{\text{Var}(q_{n,t})}$ and $\text{Var}(p_{n,t})$ for DEM/USD.
Figure 3: Two-regime model: Term Structure of implied slope coefficient $\alpha_{2,n}$, $Cov\left(q_{n,t}, p_{n,t}\right)$, $\frac{Var\left(p_{n,t}\right)}{Var\left(q_{n,t}\right)}$ and $Var\left(p_{n,t}\right)$ for JPY/USD
depreciation is positive for all \( n \). Variance of expected exchange rate depreciation declines relatively to the variance of the risk premium and the slope coefficient becomes close to zero.

Three-Regime MS-SDF: Figures 4, 5 and 6 show the term structure of \( \alpha_{2,n} \) for the three regime model.

For GBP/USD exchange rate, \( \alpha_{2,n} \) is slightly above zero for all \( n \). Variance of the risk premium is much larger than the variance of expected exchange rate depreciation. Thus, according to the model, most of the variation in the forward premium is explained by the risk premium. And because after \( n > 2 \), covariance between \( p \) and \( q \) is positive the slope coefficient is also positive. So the model suggests that we should not observe the FPA at any horizon.

In the case of DEM/USD exchange rate, \( \alpha_{2,1} \) is 1.08. As \( n \) grows, \( \alpha_{2,n} \) converges to a value around 0.91. Even though \( \text{Cov}(p,q) \) is negative, its absolute value is smaller than
Figure 5: Three-regime model: Term Structure of implied slope coefficient $\alpha_{2,n}$, $\text{Cov}(q_{n,t}, p_{n,t})$, $\frac{\text{Var}(p_{n,t})}{\text{Var}(q_{n,t})}$ and $\text{Var}(p_{n,t})$ for DEM/USD.
Figure 6: Three-regime model: Term Structure of implied slope coefficient $\alpha_{2,n}$, $\text{Cov}(q_{n,t},p_{n,t})$, $\frac{\text{Var}(p_{n,t})}{\text{Var}(q_{n,t})}$ and $\text{Var}(p_{n,t})$ for JPY/USD

![Graph showing $\alpha_{2,n}$, $\text{Cov}(c_{n,t},p_{n,t})$, $\frac{\text{Var}(p_{n,t})}{\text{Var}(q_{n,t})}$ and $\text{Var}(p_{n,t})$ over time for JPY/USD.](image-url)
the \( \text{Var}(q) \), so the slope coefficient remains positive. Variance of expected exchange rate depreciation is also much greater than the variance of risk premium. So, most of variation in the forward premium is due to changes in expected exchange rate depreciation. Therefore, the slope coefficient in the forward premium regression should be close to 1.

The graph of the term structure of \( \alpha_{2,n} \) for JPY/USD exchange rate shows that the slope coefficient is negative but converges to zero with longer maturities. Even though the variance of the risk premium is always bigger than the variance of \( q \) (which is a necessary condition for the FPA), the covariance between \( p \) and \( q \) diminishes and the slope coefficient becomes smaller. So the model implies the FPA at shorter horizons but not at longer horizons.

Predictions of the three-regime model appear more in line with some empirical studies on the long-horizon UIP than the predictions of two-regime MS-SDF. Using monthly data on the forward and spot rates, Razzak (2002) finds that, at a one-month horizon, \( \alpha_{2,1} \) is -1.94, 1.21 and -1.38 for GBP/USD, DEM/USD and JPY/USD respectively. At a one-year horizon he finds \( \alpha_{2,12} \) to be 1.00, 1.67 and -0.18 for GBP/USD, DEM/USD and JPY/USD respectively. The slope coefficients implied by MS-SDF for the same horizons are 1.08 and 0.74 for DEM/USD exchange rate and -0.09 and -0.03 for JPY/USD.

6 Conclusion

Although the forward premium anomaly has been documented in many studies, recent evidence shows that uncovered interest parity might hold over horizons of one year and longer. This paper proposes a currency pricing model that provides an explanation for different evidence on UIP for different horizons. The model is an affine stochastic discount factor model that allows for switching in the mean and variance parameters of kernel equation. Introduction of switching into the model generates time-varying risk premium and enables the model to reproduce the FPA at shorter horizons. For certain parameter values of the model, variability of the risk premium diminishes at longer horizons and the model predicts that UIP holds.

Two sets of model estimates were presented in this paper: for two-regime and three-
regime versions of the model. The estimates are obtained using the data on the one-month forward premium and exchange rate depreciation for three exchange rates. Estimation results can be summarized as follows:

- Comparing the slope coefficient of the one-month forward premium regression with the OLS estimate from the sample shows that, in two out of three cases, the implied coefficient is negative, while the OLS estimate is negative in all three cases. The absolute value of the implied slope coefficient is less than the sample estimate. This, however, could be due to insufficient length of the sample. A Monte Carlo experiment shows that, if the model is true, in small samples it is possible to observe OLS estimates of the slope coefficient much larger than the true value.

- From the model, we can derive an implied slope coefficient of the forward premium regression for different horizons of the forward contracts. When I use parameter estimates of the model to compute the term structure of the slope coefficient, the results show that the slope coefficient changes with the length of forward contracts. Based on the estimates of the three-regime model, for DEM/USD exchange rate UIP should hold at a short as well as a long horizon. Results for JPY/USD show that the FPA exists at shorter horizons, but at longer horizons the implied slope coefficient is close to zero. Both of these implications of the model are supported by empirical findings in the forward premium literature. For GBP/USD exchange rate, the model fails to reproduce the term structure of the slope coefficient in the forward premium regression.

- Using the model, I compute implied moments for forward premia and exchange rate depreciation. Comparing results of the model and sample moments shows that the model matches the means, variances and autocorrelations of those variables. Most of the implied moments fall within one standard error of the sample estimates. The LM test of residuals shows no evidence of serial correlation within a regime in all but one case.

The model proposed in this paper can be extended in a number of ways. A natural extension would be to incorporate greater number of regimes. A more interesting extension
would be to use time-varying instead of constant transition probabilities. Bekaert et al. (2001), for example, studied term structure of U.S. interest rates using a model where observable variables are correlated with probabilities of being in a particular regime.

Nevertheless, results presented here demonstrate that a switching model with just a few regimes can explain very complex dynamics of the data, while retaining a fairly simple structure.

References


A Derivation of the Affine Coefficients

In this section of appendix I show how to derive $A_{n,t} = A_n(z_t)$, $B_{n,t} = B_n(z_t)$ and $D_{n,t} = D_n(z_t)$. In appendix I use this notation to distinguish clearly between the variables (such as $x_t$) and coefficients that are a function of regime $z_t$ (such as $A_n(z_t)$).

\[ f_{n,t} - s_t = A_n(z_t) + B_n(z_t) x_t, \]

\[ E_t [s_{t+n} - s_t] = D_n(z_t) + B_n(z_t) x_t, \]

where $f_{n,t}$ is logarithm of the $n$-period forward rate, $s_t$ is the log of the spot rate, $z_t$ is the regime and $x_t$ is the state variable. The approach shown here closely follows Evans (2003).

To derive $A_n(z_t), B_n(z_t)$ and $D_n(z_t)$, first we need to derive the solution for an $n$-period discount bond:

\[ -b_{n,t} = r_{n,t} = K_n(z_t) + L_n(z_t) x_t, \]  \hspace{1cm} (A1)

where $r_{n,t}$ is the yield on a $n$-period discount bond. Once we know $K_n(z_t), L_n(z_t)$, we can find $A_n(z_t) = K_n(z_t) - K_n^*(z_t)$ and $B_n = L_n(z_t) - L_n^*(z_t)$.

A.1 Solution for One-Period Discount Bond

The equilibrium pricing relation is

\[ 1 = E [\exp (m_{t+1} + b_{n-1,t+1} - b_{n,t}) | I_t]. \]  \hspace{1cm} (A2)

Because $-m_{t+1}$ depends only on the regime at time $t$ and is conditionally normal, it is easy to find the solution for a one-period discount bond:

\[ 1 = E [\exp (m_{t+1} - b_{1,t}) | I_t]. \]

Therefore, using equations 7 and 8, we can derive the solution for the price of one-period
bond as

\[ 1 = \exp \left\{ E (m_{t+1} - b_{t,1} | \mathcal{I}_t) + \frac{1}{2} Var (m_{t+1} | \mathcal{I}_t) \right\} \]

\[ = \exp \left\{ E (m_{t+1} - b_{t,1} | \mathcal{I}_t) + \frac{1}{2} \lambda (z_t)^2 + \frac{1}{2} \sigma^2 \right\} \]

\[ = \exp \left\{ -\delta (z_t) - \beta x_t - b_{t,1} + \frac{1}{2} \lambda (z_t)^2 + \frac{1}{2} \sigma^2 \right\}. \]

Taking the logs of the last expression, we can find the yield on a one-period discount bond as

\[ y_t = -b_{t,1} = \delta (z_t) + \beta x_t - \frac{1}{2} \lambda (z_t)^2 - \frac{1}{2} \sigma^2. \]  

(A3)

A.2 Solution for n-Period Discount Bond

Using the law of iterated expectations, we can write the equilibrium pricing equation A2 for a n-period discount bond as

\[ 1 = \mathbb{E} \left[ \mathbb{E} [ \exp (m_{t+1} + b_{n-1,t+1} - b_{n,t}) | \mathcal{I}_t, z_{t+1}] | I_t \right]. \]  

(A4)

The inner expectation conditional on \( \mathcal{I}_t \) and \( z_t \) can be written as

\[ \exp \left\{ \mathbb{E} [(m_{t+1} + b_{n-1,t+1} - b_{n,t}) | \mathcal{I}_t, z_{t+1}] + \frac{1}{2} Var (m_{t+1} + b_{n-1,t+1} | \mathcal{I}_t, z_{t+1}) \right\}. \]

Using equations 8 and A1, the variance term in this expression can be written as

\[ Var \left[ (\lambda (z_t) \epsilon_{t+1} + L_{n-1} (z_{t+1})) \epsilon_{t+1} + \omega_{t+1} | I_t, z_t \right] = (\lambda (z_t) + L_{n-1} (z_{t+1}))^2 + \sigma^2 \]

\[ = \lambda (z_t)^2 + 2 \lambda (z_t) L_{n-1} (z_{t+1}) + L_{n-1} (z_{t+1})^2 + \sigma^2. \]
Using the solution for the yield on a one-period bond, we can rewrite equation A4 as

\[ 1 = E \left[ \exp \left\{ E \left[ (b_{n-1,t+1} - b_{n,t} - y_t) | I_t, z_{t+1} \right] + \lambda (z_t) L_{n-1} (z_{t+1}) + \frac{1}{2} (L_{n-1} (z_{t+1}))^2 \right\} | I_t \right] \]

or,

\[ 1 = E \left[ \exp \left\{ E \left[ \delta_{t+1,n} | I_t, z_{t+1} \right] + \Gamma_{n-1} (z_{t+1}, z_t) \right\} | I_t \right], \tag{A5} \]

where

\[ \delta_{t+1,n} \equiv b_{n-1,t+1} - b_{n,t} - y_t \]

is the excess return and \( \Gamma_{n-1} (z_{t+1}, z_t) = Cov (b_{n-1,t+1}, m_{t+1}) + \frac{1}{2} Var (b_{n-1,t+1}) : \)

\[ \Gamma_{n-1} (z_{t+1}, z_t) \equiv \lambda (z_t) L_{n-1} (z_{t+1}) + \frac{1}{2} (L_{n-1} (z_{t+1}))^2. \tag{A6} \]

To keep it simple, consider a case when there are only two regimes. Thus, there are two possibilities for \( E \left[ \delta_{t+1,n} | I_t, z_{t+1} \right] : E \left[ \delta_{t+1,n} | I_t, z_{t+1} = 1 \right] \) and \( E \left[ \delta_{t+1,n} | I_t, z_{t+1} = 2 \right] \). Let vector \( \Phi_{n,t} \)

\[ \Phi_{n,t} = \begin{bmatrix} E \left[ \delta_{t+1,n} | I_t, z_{t+1} = 1 \right] \\ E \left[ \delta_{t+1,n} | I_t, z_{t+1} = 2 \right] \end{bmatrix}. \]

To use this vector with equation A5 define a \( 1 \times 2 \) vector \( \ell_z \), which is equal to [1, 0] when \( z = 1 \) and [0, 1] when \( z = 2 \). With this addition to the notation, equation A5 becomes

\[ 1 = E \left[ \exp \{ \ell_{z_{t+1}} \Phi_{n,t} + \Gamma_{n-1} (z_{t+1}, z_t) \} | I_t \right]. \tag{A7} \]

For each possible realization of regime \( z_{t+1} \), we can write the expectation conditional...
on $\mathcal{I}_t$ as

$$1 = \sum_{\tilde{z}=1}^{2} \Pi_{\tilde{z},z} \exp(\ell_{\tilde{z}} \Phi_{n,t} + \Gamma_{n-1}(\tilde{z}, z)),$$

where

$$\Pi_{\tilde{z},z} \equiv \Pr(z_{t+1} = \tilde{z}, | z_t = z),$$

So we can stack the equations for each possible regime today ($z$):

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \Pi_{11} e^{\Gamma_{n-1}(1,1)} & \Pi_{12} e^{\Gamma_{n-1}(1,2)} \\ \Pi_{21} e^{\Gamma_{n-1}(2,1)} & \Pi_{22} e^{\Gamma_{n-1}(2,2)} \end{bmatrix} \begin{bmatrix} e^{\ell} \Phi_{n,t} \\ e^{\ell_2} \Phi_{n,t} \end{bmatrix}$$

or

$$1 = P_{n-1} \exp_v \Phi_{n,t},$$

where $\exp_v$ means exponential operator applied to the vector. Provided that $P_{n-1}$ is invertible, we can write

$$\Phi_{n,t} = \ln_v \left( (P_{n-1})^{-1} 1 \right). \quad (A8)$$

### A.3 Term Premium

The right hand side of the expression depends only on the current regime, transition probabilities and $\Gamma$: It does not depend on the state variable $x_t$. This means that the vector of expectations $\Phi_{n,t}$ does not depend on $x_t$, even though $x_t$ is in the information set $\mathcal{I}_t$.

The term premium $\theta_{n,t} \equiv E[ \delta_{t+1,n} | \mathcal{I}_t]$ can be found as

$$\theta_{n,t} = \ell_{z_t} \Pi \Phi_{n,t}, \quad (A9)$$

32
where Π is the matrix of transition probabilities with \( \Pi_{i,j} = \Pr(z_{t+1} = i | z_t = j) \). Therefore,

\[
\theta_{n,t} = \ell_{z_t} \Pi \ln\left((\mathcal{P}_{n-1})^{-1} \mathbf{1}\right).
\]

This implies that the term premium \( \theta_{n,t} \equiv E[\delta_{n,t+1} | \mathcal{I}_t] \) also depends only on the current regime. To solve for \( K_{n,t} \) and \( L_{n,t} \), express \( \theta_{n,t} \) using equation A1

\[
\begin{align*}
\theta_{n,t} &= K_n(z_t) + L_n(z_t)x_t - E[K_{n-1}(z_{t+1}) + L_{n-1}(z_{t+1})x_{t+1} | \mathcal{I}_t] - K_1(z_t) - L_1(z_t)x_t \\
&= K_n(z_t) - E[K_{n-1}(z_{t+1}) | \mathcal{I}_t] - \delta(z_t) + \frac{1}{2} \lambda(z_t)^2 + \frac{1}{2} \sigma_{\omega}^2 - \beta x_t + \\
&\quad + L_n(z_t)x_t - \phi x_t E[L_{n-1}(z_{t+1})] \\
&= K_n(z_t) - E[K_{n-1}(\tilde{z}) | \mathcal{I}_t] - \delta(z_t) + \frac{1}{2} \lambda(z_t)^2 + \frac{1}{2} \sigma_{\omega}^2 + (L_n(z_t) - \phi E[L_{n-1}(\tilde{z})] - \beta) x_t.
\end{align*}
\]

A.4 Solution for \( L_n \)

Because \( \theta_{n,t} \) does not depend on \( x_t \), the following must be true:

\[
L_n(z_t) - \phi E[L_{n-1}(\tilde{z})] - \beta = 0.
\]

Therefore

\[
L_n(z_t) = \phi E[L_{n-1}(\tilde{z})] + \beta.
\]

Notice, that when \( n = 1 \), \( L_1(z_t) = \beta \), so it does not depend on \( z_t \). Thus,

\[
L_2(z_t) = \phi E[L_1(\tilde{z})] + \beta = (\phi + 1) \beta.
\]

Similarly, for any maturity \( n \) it can be shown that \( L_n(z_t) = L_n = \frac{1 - \phi^n}{1 - \phi} \). This also means that \( \Gamma_{n-1}(z_{t+1}, z_t) = \Gamma_{n-1}(z_t) \): It does not depend on the regime at time \( t + 1 \).
A.5 Solution for $K_n(z_t)$

Because $L_n$ does not depend on the current regime, expected excess return conditional on information at time $t$ and $z_t$ (equation A5) becomes

$$1 = \exp \Gamma_n - 1 (z_t) E \{ E [\delta_{t+1,n} | I_t, z_{t+1}] | I_t \}.$$

Rearranging and expanding $\delta_{t+1,n}$ gives us

$$\exp \{ -\Gamma_{n-1}(z_t) \} = E \{ \exp \{ (K_n(z_t) + L_n x_t - K_{n-1}(z_{t+1}) - L_{n-1} x_t - K_1(z_t) - L_1 x_t) | I_t, z_{t+1} \} \}.$$

Once again, rearranging and taking logs,

$$-\Gamma_{n-1}(z_t) = K_n(z_t) + L_n x_t - L_{n-1} x_t - K_1(z_t) - L_1 x_t + \ln E \{ \exp \{ -K_{n-1}(z_{t+1}) \} | I_t \}.$$

So the solution for $K_n(z_t)$ is

$$K_n(z_t) = K_1(z_t) - \Gamma_{n-1}(z_t) - \ln E \{ \exp \{ -K_{n-1}(z_{t+1}) \} | I_t \}$$

(A11)

When $n = 1$, we know from equation A3 that $K_1(z_t) = \delta(z_t) - \frac{1}{2} \lambda(z_t)^2 - \frac{1}{2} \sigma_w^2$. So for each regime, the solution for $K_n(z)$ is

$$K_n(1) = K_1(1) - \Gamma_{n-1}(1) - \ln [p \exp \{ -K_{n-1}(1) \} + (1 - p) \exp \{ -K_{n-1}(2) \}],$$

$$K_n(2) = K_1(2) - \Gamma_{n-1}(2) - \ln [(1 - q) \exp \{ -K_{n-1}(1) \} + q \exp \{ -K_{n-1}(2) \}].$$

Notice that to compute $A_n(z_t) = K_n(z_t) - K_n^*(z_t)$, we need to know $\delta(z_t)$ and $\delta^*(z_t)$. These parameters are not identified individually. Because I use the forward premium to estimate the model, only the difference $\delta(z_t) - \delta^*(z_t)$ can be identified. To identify these
parameters individually, I use the data on for one-month Eurodollar rates for the same
sample period. The mean and the variance of the interest rate allow to identify $\delta(z_t)$ and
$\delta^*(z_t)$ individually.

A.6 Solution for $D_{n,t}$

To derive coefficients in the expression for expected exchange rate depreciation over longer
horizon

$$E_t [s_{t+n} - s_t] = D_t (z_t) + B_n x_t,$$

first note that we can write exchange rate depreciation over $n$ periods as

$$s_{t+n} - s_t = m_{t+1} + m_{t+2} + \ldots + m_{t+n+1}.$$

Thus, expected exchange rate depreciation is

$$E_t [s_{t+n} - s_t] = E_t [m_{t+1} + m_{t+2} + \ldots + m_{t+n+1}] .$$

Therefore, using kernel equations and equation 15 we can express the expectation in terms
of model parameters:

$$E_t [s_{t+n} - s_t] = E_t \left( \sum_{j=1}^{n} \Delta \Pi^{n-j} \nu_t + B_n x_t \right) .$$

So the expression for $D_t (z_t)$ is

$$D_{n,t} = \sum_{j=1}^{n} \Delta \Pi^{n-j} \nu_t$$

where $\Delta' = \begin{bmatrix} \nabla \delta_1 \\ \nabla \delta_2 \end{bmatrix}$. 

35
The per-period risk premium for $n$-period horizon and its variance can be found as

\[
n p_{n,t} = f_{n,t} - s_t - E_t (s_{t+n} - s_t) \\
       = A_n (z_t) - D_t (z_t),
\]

\[
Var (p_{n,t}) = \frac{1}{n^2} Var (A_n (z_t) - D_t (z_t)).
\]

Similarly, per-period expected exchange rate depreciation over $n$-period horizon $q_{n,t}$ can be found as

\[
n q_{n,t} = E_t (s_{t+n} - s_t) = \\
       = D_t (z_t) + B_n x_t,
\]

\[
Var (q_{n,t}) = \frac{1}{n^2} Var (D_t (z_t)) + \frac{(B_n)^2}{n^2} Var (x_t).
\]

Covariance between the risk premium and expected exchange rate depreciation is

\[
Cov (p_{n,t}, q_{n,t}) = \frac{1}{n^2} Cov (A_n (z_t), D_t (z_t)) - \frac{1}{n^2} Var (D_t (z_t)).
\]

**B Derivation of Moments From Model Parameters**

To derive an analytical solution for $\alpha_2$, notice that in a large sample $p_t$ and $\nabla \delta_t$ can be thought of as discrete random variables. The variance of a discrete random variable $x \in \{x_1, x_2\}$ with $Pr [x = x_1] = \pi$ is $Var (x) = \pi (1 - \pi) (x_1 - x_2)^2$. Thus,

\[
\alpha_2 = \frac{\rho (1 - \rho) (\nabla \delta_1 - \nabla \delta_2) (p_1 - p_2) + \rho (1 - \rho) (\nabla \delta_1 - \nabla \delta_2)^2 + \frac{\nabla \beta^2}{1 - \phi^2}}{\rho (1 - \rho) (\nabla \delta_1 - \nabla \delta_2 + p_1 - p_2)^2 + \frac{\nabla \beta^2}{1 - \phi^2}},
\]

where $\rho$ is ergodic probability that $z_t = 1$. So, if risk premium $p_t$ has a greater variance than the expected exchange rate depreciation $\nabla \delta_t$, the implied slope coefficient of
the forward premium regression is negative.

Because the mean of the state variable $x$ is normalized at zero, unconditional means of the forward premium are easy to find using ergodic probabilities of regimes:

\[
E[s_{t+1} - s_t] = E[\nabla \delta_t] + E[\nabla \beta x_t] + E[\nabla \lambda_t \varepsilon_{t+1} + \nabla \omega_{t+1}]
\]

\[
= \rho \nabla \delta_1 + (1 - \rho) \nabla \delta_2,
\]

\[
E[f_t - s_t] = \rho [\nabla \delta_1 + p_1] + (1 - \rho) [\nabla \delta_1 + p_2].
\]

Regimes are assumed to be independent, so variance of the forward premium is just the sum of the variances of the state variable $x$ and the binary random variable $d_t + p_t$:

\[
Var(f_t - s_t) = Var(\nabla \delta_t + p_t) + Var(\nabla \beta x_t)
\]

\[
= \rho (1 - \rho) (\nabla \delta_1 + p_1 - \nabla \delta_2 - p_2)^2 + \nabla \beta^2 \frac{1}{1 - \phi^2}.
\]

To find autocorrelation of the forward premium, first consider the autocovariance of a discrete random variable $d_t$ that depends on the current regime.

\[
Cov(\nabla \delta_t, \nabla \delta_{t-1}) = E[\nabla \delta_t \nabla \delta_{t-1}] - E[\nabla \delta_t]^2
\]

\[
= \rho \pi_1 \nabla \delta_1^2 + \rho (1 - \pi_1) \nabla \delta_1 \nabla \delta_2 + (1 - \rho) (1 - \pi_2) \nabla \delta_1 \nabla \delta_2 +
\]

\[
+ (1 - \rho) \pi_2 \nabla \delta_2^2 - E[\nabla \delta_t]^2
\]

\[
= -\rho (1 - \pi_1) (\nabla \delta_1 - \nabla \delta_2)^2 + E[\nabla \delta_1^2] - E[\nabla \delta_t]^2
\]

\[
= -\rho (1 - \pi_1) (\nabla \delta_1 - \nabla \delta_2)^2 + Var(\nabla \delta_t)
\]

\[
= \rho (\pi_1 - \rho) (\nabla \delta_1 - \nabla \delta_2)^2
\]

\[
= (1 - \rho) (1 - \pi_2) (\nabla \delta_1 - \nabla \delta_2)^2.
\]

The last equality is true because we should be able to express these moments with either of the two transition probabilities. Using the above expression, autocovariance of the forward
premium can be found as

\[
Cov(f_t - s_t, f_{t-1} - s_{t-1}) = Cov(\nabla \delta_t + p_t, \nabla \delta_{t-1} + p_{t-1}) + \nabla \beta^2 Cov(x_t, x_{t-1})
\]

\[
= \rho (\pi_1 - \rho) (\nabla \delta_1 + p_1 - \nabla \delta_2 - p_2)^2 + \nabla \beta^2 \frac{\phi}{1 - \phi^2}.
\]

And the autocorrelation of \( f_t - s_t \) is

\[
Corr(f_t - s_t, f_{t-1} - s_{t-1}) = \frac{\rho (\pi_1 - \rho) (\nabla \delta_1 + p_1 - \nabla \delta_2 - p_2)^2 + \nabla \beta^2 \frac{\phi}{1 - \phi^2}}{\rho (1 - \rho) (\nabla \delta_1 + p_1 - \nabla \delta_2 - p_2)^2 + \nabla \beta^2 \frac{1}{1 - \phi^2}}.
\]

Variance of the \( s_{t+1} - s_t \) also involves the variance of the innovation terms \( \epsilon_{t+1} \) and \( \omega_{t+1} - \omega^*_t \):

\[
Var(s_{t+1} - s_t) = Var(\nabla \delta_t) + Var(\nabla \beta x_t) + Var(\nabla \lambda_t \epsilon_{t+1}) + Var(\nabla \omega_{t+1})
\]

\[
= \rho (1 - \rho) (\nabla \delta_1 - \nabla \delta_2)^2 + \nabla \beta^2 \frac{1}{1 - \phi^2} + \\
+ \rho \nabla \lambda_1^2 + (1 - \rho) \nabla \lambda_2^2 + Var(\nabla \omega_t).
\]

Autocovariance of \( s_{t+1} - s_t \) can be found as

\[
Cov(s_{t+1} - s_t, s_{t} - s_{t-1}) = Cov(\nabla \delta_t, \nabla \delta_t) + \nabla \beta^2 Cov(x_t, x_{t-1})
\]

\[
= \rho (\pi_1 - \rho) (\nabla \delta_1 - \nabla \delta_2)^2 + \nabla \beta^2 \frac{\phi}{1 - \phi^2}
\]

using the assumption that innovations \( \epsilon_{t+1} \) and \( \omega_{t+1} - \omega^*_t \) are independent.
### C Estimation Results of the Three-Regime Model

Table 9: Estimation results of the three-regime switching model

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9443** (0.0148)</td>
<td>0.9659** (0.0170)</td>
<td>0.9500** (0.0195)</td>
</tr>
<tr>
<td>$\sigma_{\omega-\omega}$</td>
<td>0.0311** (0.0010)</td>
<td>0.0331** (0.0012)</td>
<td>0.0345** (0.0012)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>47.4296** (0.0002)</td>
<td>0.6053 (1.3134)</td>
<td>9.9666** (0.0093)</td>
</tr>
<tr>
<td>$\lambda_2^*$</td>
<td>47.4297** (0.0002)</td>
<td>0.6071 (1.3122)</td>
<td>9.9661** (0.0093)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-3.8833** (0.4244)</td>
<td>0.5930 (5.8554)</td>
<td>-4.0748** (0.2482)</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>-3.8823** (0.4247)</td>
<td>0.6007 (5.8853)</td>
<td>-4.0784** (0.2476)</td>
</tr>
<tr>
<td>$\lambda_3^*$</td>
<td>0.0033 (0.3196)</td>
<td>-4.3258** (0.1848)</td>
<td>-0.1056 (5.7066)</td>
</tr>
<tr>
<td>$\lambda_3^*$</td>
<td>0.0102 (0.3194)</td>
<td>-4.3274** (0.1848)</td>
<td>-0.1030 (5.7059)</td>
</tr>
<tr>
<td>$\Pr(z_t=1</td>
<td>z_{t-1}=1)$</td>
<td>0.8566** (0.0570)</td>
<td>0.9695** (0.0135)</td>
</tr>
<tr>
<td>$\Pr(z_t=2</td>
<td>z_{t-1}=1)$</td>
<td>0.0738 (0.0520)</td>
<td>0.0100 (0.0075)</td>
</tr>
<tr>
<td>$\Pr(z_t=1</td>
<td>z_{t-1}=2)$</td>
<td>0.0793 (0.0697)</td>
<td>0.1682 (0.4131)</td>
</tr>
<tr>
<td>$\Pr(z_t=2</td>
<td>z_{t-1}=2)$</td>
<td>0.5134** (0.1336)</td>
<td>0.5023 (0.5112)</td>
</tr>
<tr>
<td>$\Pr(z_t=3</td>
<td>z_{t-1}=1)$</td>
<td>0.0274 (0.0191)</td>
<td>0.5770* (0.3347)</td>
</tr>
<tr>
<td>$\Pr(z_t=3</td>
<td>z_{t-1}=2)$</td>
<td>0.0304* (0.0188)</td>
<td>0.0000 (0.3269)</td>
</tr>
<tr>
<td>$\nabla \beta$</td>
<td>0.0008** (0.0000)</td>
<td>0.0006 (0.0000)</td>
<td>0.0007** (0.0000)</td>
</tr>
<tr>
<td>$\nabla \delta_1$</td>
<td>-0.0046 (0.0041)</td>
<td>0.0004 (0.0022)</td>
<td>0.0067** (0.0027)</td>
</tr>
<tr>
<td>$\nabla \delta_2$</td>
<td>0.0186** (0.0061)</td>
<td>0.0004 (0.0635)</td>
<td>-0.0101* (0.0057)</td>
</tr>
<tr>
<td>$\nabla \delta_3$</td>
<td>-0.0021 (0.0022)</td>
<td>-0.0084 (0.0110)</td>
<td>-0.0006 (0.0151)</td>
</tr>
<tr>
<td>$\nabla \lambda_1$</td>
<td>-0.0001 (0.0040)</td>
<td>-0.0018 (0.0020)</td>
<td>0.0004 (0.0029)</td>
</tr>
<tr>
<td>$\nabla \lambda_2$</td>
<td>-0.0060 (0.0093)</td>
<td>-0.0077 (0.0323)</td>
<td>0.0036 (0.0049)</td>
</tr>
<tr>
<td>$\nabla \lambda_3$</td>
<td>-0.0070** (0.0025)</td>
<td>0.0016 (0.0063)</td>
<td>-0.0026 (0.0118)</td>
</tr>
<tr>
<td>$p_1$</td>
<td>0.0057** (0.0009)</td>
<td>0.0011 (0.0012)</td>
<td>0.0044** (0.0008)</td>
</tr>
<tr>
<td>$p_2$</td>
<td>-0.0232** (0.0009)</td>
<td>0.0045** (0.0012)</td>
<td>0.0146** (0.0009)</td>
</tr>
<tr>
<td>$p_3$</td>
<td>-0.0001 (0.0009)</td>
<td>-0.0079** (0.0013)</td>
<td>-0.0003 (0.0008)</td>
</tr>
</tbody>
</table>

Log-likelihood 2701.12 2799.40 2686.71

Estimates of the model were obtained using the algorithm described in Hamilton (1995). Standard errors (in parentheses) are calculated using the outer product of the gradients. A double asterisk indicates that a parameter is significant at 5% level. A single asterisk indicates that significance at 10% level.
Table 10: Sample and implied moments

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample</td>
<td>Implied</td>
<td>Sample</td>
</tr>
<tr>
<td>( E[f_t - s_t] )</td>
<td>-0.0020</td>
<td>-0.0017</td>
<td>0.0015</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0012)</td>
<td></td>
</tr>
<tr>
<td>( E[s_{t+1} - s_t] )</td>
<td>-0.0015</td>
<td>-0.0011</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.0021)</td>
<td>(0.0024)</td>
<td></td>
</tr>
<tr>
<td>( Var(f_t - s_t) )</td>
<td>6.49E-06</td>
<td>8.02E-06</td>
<td>6.82E-06</td>
</tr>
<tr>
<td></td>
<td>(1.52E-06)</td>
<td>(3.01E-06)</td>
<td></td>
</tr>
<tr>
<td>( Var(s_{t+1} - s_t) )</td>
<td>0.0010</td>
<td>0.0010</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td></td>
</tr>
<tr>
<td>( Corr(f_t - s_t, f_{t-1} - s_{t-1}) )</td>
<td>0.8431</td>
<td>0.8772</td>
<td>0.9142</td>
</tr>
<tr>
<td></td>
<td>(0.0385)</td>
<td>(0.0764)</td>
<td></td>
</tr>
<tr>
<td>( Corr(s_{t+1} - s_t, s_t - s_{t-1}) )</td>
<td>0.0673</td>
<td>0.0198</td>
<td>0.0077</td>
</tr>
<tr>
<td></td>
<td>(0.0133)</td>
<td>(0.0037)</td>
<td></td>
</tr>
</tbody>
</table>

Implied moments are computed from the parameters of the model. Analytical solutions for implied moments are shown above. Standard errors (shown in parentheses) are computed using the delta method.

Table 11: p-values for the LM test of serial correlation in the residuals within a regime

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E(\varepsilon_t, \varepsilon_{t-1}</td>
<td>z_t = z_{t-1} = 1) )</td>
<td>0.11</td>
<td>0.02</td>
</tr>
<tr>
<td>( E(\varepsilon_t, \varepsilon_{t-1}</td>
<td>z_t = z_{t-1} = 2) )</td>
<td>0.16</td>
<td>0.01</td>
</tr>
<tr>
<td>( E(\varepsilon_t, \varepsilon_{t-1}</td>
<td>z_t = z_{t-1} = 3) )</td>
<td>0.12</td>
<td>0.02</td>
</tr>
<tr>
<td>( E(\nabla \omega_t, \nabla \omega_{t-1}</td>
<td>z_t = z_{t-1} = 1) )</td>
<td>0.15</td>
<td>0.01</td>
</tr>
<tr>
<td>( E(\nabla \omega_t, \nabla \omega_{t-1}</td>
<td>z_t = z_{t-1} = 2) )</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td>( E(\nabla \omega_t, \nabla \omega_{t-1}</td>
<td>z_t = z_{t-1} = 3) )</td>
<td>0.09</td>
<td>-</td>
</tr>
</tbody>
</table>

40
Table 12: Three-regime MS-SDF: Wald test of switching in parameters across regimes (p-values)

<table>
<thead>
<tr>
<th>$H_0$</th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nabla \delta_1 = \nabla \delta_2 = \nabla \delta_3$</td>
<td>0.13</td>
<td>0.90</td>
<td>0.32</td>
</tr>
<tr>
<td>$p_1 = p_2 = p_3$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\nabla \lambda_1 = \nabla \lambda_2 = \nabla \lambda_3$</td>
<td>0.27</td>
<td>0.83</td>
<td>0.80</td>
</tr>
</tbody>
</table>

The table shows the p-values for the Wald test statistic of the null hypothesis that the parameters are the same in all three regimes.

Table 13: Comparison of the OLS estimate and implied value of the Fama Coefficient

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS estimate</td>
<td>-0.875 (0.679)</td>
<td>-0.590 (0.676)</td>
<td>-1.489 (0.637)</td>
</tr>
<tr>
<td>Implied from the model</td>
<td>0.002 (0.547)</td>
<td>1.077 (0.241)</td>
<td>-0.081 (0.463)</td>
</tr>
</tbody>
</table>

The top panel in the table compares OLS estimates of the slope coefficient in the forward premium regression and it’s value implied by parameter estimates of the three-regime MS-SDF model. The bottom panel of the table contains the p-values of the Wald test statistic for the null hypotheses that $\alpha_2 = 0$, $-0.5$ and $-1$.

Table 14: Large-sample empirical distribution of $\alpha_2$ under the null of SDF-MS model

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2 &lt; 0$</td>
<td>0.45</td>
<td>0.11</td>
<td>0.44</td>
</tr>
<tr>
<td>$\alpha_2 &lt; -0.5$</td>
<td>0.19</td>
<td>0.02</td>
<td>0.21</td>
</tr>
<tr>
<td>$\alpha_2 &lt; -1$</td>
<td>0.07</td>
<td>0.00</td>
<td>0.08</td>
</tr>
</tbody>
</table>

100,000 draws were generated from the distribution $N(\theta, \Sigma)$, where $\theta$ is the vector of parameter estimates of the three-regime MS-SDF model $\theta = [\phi, \nabla \beta, \nabla \lambda_1, \nabla \lambda_2, p_1, p_2, \sigma \nabla \omega, \pi_1, \pi_2, \nabla \delta_1, \nabla \delta_2]$ and $\Sigma$ is the covariance matrix of the parameters. Then, for each draw of the parameter vector, $\alpha_2$ is computed using ergodic probabilities of each regime. The table represents proportion of draws that resulted in $\alpha_2$ below a certain value.
Table 15: Empirical distribution of small sample estimate of $\alpha_2$ under the null of SDF-MS model

<table>
<thead>
<tr>
<th></th>
<th>GBP/USD</th>
<th>DEM/USD</th>
<th>JPY/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2 &lt; -0$</td>
<td>0.53</td>
<td>0.08</td>
<td>0.60</td>
</tr>
<tr>
<td>$\alpha_2 &lt; -0.5$</td>
<td>0.27</td>
<td>0.03</td>
<td>0.39</td>
</tr>
<tr>
<td>$\alpha_2 &lt; -1$</td>
<td>0.10</td>
<td>0.01</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Using the parameter estimates of the three-regime MS-SDF model, 100,000 datasets were generated. Each dataset contains 323 observations. For each dataset, OLS estimate of the slope coefficient $\alpha_2$ is computed. The table shows proportion of datasets that produced $\alpha_2$ below a certain value.