Testing for No Arbitrage in Continuous Time:  
A Resolution to the Forward Premium Anomaly*†

Stefan Jacewitz, Hwagyun Kim, and Joon Y. Park‡
Texas A&M University

September 25, 2008

Abstract

In this paper, we investigate the forward premium anomaly (FPA) in a continuous time framework. Given the ultra high-frequency nature of the currency market, a continuous time approach is, in many ways, attractive. We derive a no arbitrage condition for a currency exchange market in continuous time and then test whether this condition holds. In particular, the condition implies that the volatility of spot exchange returns, as well as the risk premium requested by spot traders, are functions of differences in the market prices of risk between the two countries. Moreover, it reveals that, even in the simplest case where there is no market price of risk differential, the conventional FPA tests based on discrete time models will be invalid and subject to data misaggregation bias unless the expectation hypothesis holds continuously at all frequencies and maturities. To empirically evaluate our continuous time regression model, we employ a novel econometric methodology based on a time change from calendar to volatility time. Specifically, our method requires collection of samples at random intervals having the same level of excess exchange return volatility. This amounts to using a sampling chronometer that runs at a rate inversely proportional to the volatility. By doing so, we may effectively make the distributions of residuals into independent standard normals, nonparametrically correcting for the non-normality and time-varying stochastic volatility typically present in exchange return data. The model is estimated and tested by minimizing the distance between the sampling distributions of residuals and the standard normal distribution. We apply our methodology to exchange rates between the US dollar and six other major currencies. Our results are unambiguous and robust: In every case, the FPA disappears.

*This document is a draft and is not for distribution or citation.
†The authors would like to thank the Program for Time Series with Applications in Financial Economics (PTSAFE) at Texas A&M University for research support, and Yosoon Chang, Yongok Choi and Dahee Jeong for many helpful discussions and suggestions. Park gratefully acknowledges the financial supports from the NSF under NSF Grant No. SES-0518619.
‡Texas A&M University and Sungkyunkwan University
1 Introduction

The forward premium anomaly (FPA hereafter) is widely considered one of the most enduring and important unsolved puzzles in international finance. Moreover, testing this statement is equivalent to testing the arguably more important uncovered interest parity hypothesis, a condition which is explicitly assumed to be true in many international macroeconomic models, but often fails empirically. Indeed, the relationship between expected and actual future exchange rates is an important topic to firms, investors, as well as economists. The forward rate unbiasedness hypothesis states that expected speculative return in the forward exchange market is zero. That is, the forward exchange rate is an unbiased predictor of the future spot exchange rate. The paradox lies in the ubiquitous conclusion that, while the forward-spot spread is a predictor of the future spot return, it is a reverse predictor. That is, it mis-predicts both the magnitude and, more troublesome, the direction of movements in the return. It originated more than two decades ago ([18]). Throughout its past and including the present, this empirical puzzle has posed a dependably prolific and progressively complex topic.

Among other reasons for the popularity and the longevity of this puzzle is the nature of the currency market. The foreign exchange market is extremely active, even by financial market standards. Average daily volume in over-the-counter foreign exchange instruments was $618 billion in the year proceeding April 2007 (Foreign Exchange Committee Semi-Annual Foreign Exchange Volume Survey, April 2007). This includes $274 billion in spot transactions and $101 billion in forward transactions. For a reference point for the magnitude of this market, the average daily US GDP in 2006 was $36 billion. The average daily total foreign exchange market turnover $3.2 trillion. Including $1 trillion in spot transactions and $362 billion in forward transactions. Further, the market is highly concentrated. Nearly 86% of all transactions involved the top seven currencies (viz. the currencies which we analyze here: US dollar, euro, yen, pound sterling, Swiss franc, Australian dollar, and Canadian dollar) (Triennial Central Bank Survey 2007). The foreign exchange market is (“by far”) the largest financial market in existence ([7]).

In this paper, we study FPA in a continuous time framework. In particular, we present a continuous-time version of the FPA model, which is derived using the assumption of no arbitrage from the stochastic differential equation linking spot exchange rate changes to forward spot spread at an infinitesimal time interval. We believe that the no arbitrage condition in continuous time used to derive our model is much more realistic and relevant than the analogous condition in discrete time, given the ultra high-frequency nature of transactions in the currency market. Our model implies, among other things, that both the risk premium and the volatility of spot exchange rates are functions of differences in the market prices of risk between the two countries. If the two countries have the same market prices of risk, then our model reduces to the traditional FPA models presuming risk neutrality. However, even in this case, we make it clear that the conventional FPA tests based on discrete time models will be invalid and subject to data misaggregation bias unless the expectation hypothesis holds continuously at all frequencies and maturities.

Our continuous time model is statistically analyzed using a novel econometric methodology. The methodology relies on random sampling using a time change from calender
to volatility time. Our sampling chronometer runs at a rate inversely proportional to the volatility. More precisely, the samples are collected in our analysis at random intervals having the same level of excess exchange return volatility, rather than at some fixed frequency such as daily, weekly or monthly. Under this scheme, the samples may be regarded as being independent and identically distributed as normal. This is due to the celebrated theorem by Dambis, Dubins and Schwarz, which is well known in the theory of stochastic processes. By using a time change, we may therefore accommodate the non-normality and time-varying stochastic volatility that is typically present in exchange return data. After the time change, our model is estimated and tested by the so-called martingale method developed recently by Park (2008), which minimizes the distance between the sampling distributions of residuals and the standard normal distribution.

Our approach has several trenchant differences from some of the most recent contributions in the test of FPA. Mark and Moh [19] develop an uncovered interest parity condition (UIP hereafter) in continuous time. However, it is not derived from no arbitrage condition and assumes risk neutrality. They also do not derive the corresponding covered interest parity equation. Chaboud and Wright [8] study a discrete-time UIP condition using intraday data and report that the UIP tends to hold at an intraday level, but not over one day of the span of the data window. Again they do not consider the FPA nor adjusting for no arbitrage and risk premium. Groen and Balakrishnan [14] attempt to compensate for the risk premia in a discrete model with some success, though for most exchange rates the forward premium is not a significant predictor of the spot return. Baillie and Kilic [3] adjust the forward premia for risk simply by dividing them by the standard error. However, this ignores the stochastically volatile nature of the data.

The rest of the paper is organized as follows. In the remainder of this section, we will introduce the conventional theory and the anomaly in the empirical findings reported in previous studies of the FPA. In Section 2, we present a new continuous time version of the forward premium equation. The equation is obtained from a continuous time model of the foreign currency market derived under the no arbitrage condition. Section 3 introduces the econometric methodology used in our analysis. We provide a brief overview of the time change and martingale estimation method, which is followed by a detailed explanation about how we may implement them to estimate and test for our model. Section 4 reports all of the results for our statistical analysis. Some concluding remarks are given in Section 5. Appendix contains some additional information on our statistical analysis. Finally, a word on notation. In the paper, we deal with both continuous time processes and discrete time series. The time index will be denoted by \( t \) or \( s \) for the former, and by \( i \) and \( j \) for the latter throughout the paper.

1.1 Theory

Let \( F_{i,j} \) be the forward exchange rate, or the forward price of one unit of foreign currency, at time \( i \) with value date \( j \) periods ahead, and \( X_t \) be the current spot price in domestic currency of one unit of foreign currency, or the spot exchange rate. The standard hypothesis of unbiasedness is

\[
F_{i,j} = E_t[X_{i+j}],
\]
or, equivalently

\[ \frac{F_{i,j}}{X_i} = \frac{E_i[X_{i+j}]}{X_i} \]

for all \( i \) and \( j \), where \( E_i \) is the mathematical expectation operator conditional on the information set available at time \( i \). This is the most commonly tested implication of the hypothesis in recent literature.

Let \( x_i \) be the log of the spot exchange rate at the \( i \)-th period, and \( f_{i,j} \) be the log of the forward exchange rate at the \( i \)-th period with value date \( j \) periods ahead. Unbiasedness of the forward exchange rate would of course mean approximately that

\[ E_i[x_{i+j}] = f_{i,j}. \]

This implies that \( f_{i,j} = x_{i+j} + u_{i+j} \), where \( u_i \) is a martingale difference sequence and can be interpreted as the rational agents’ prediction error. The equality is directly implied if agents are risk neutral and rationally use information, there are no transaction costs, and the market is competitive and efficient.

The early works examining forward rate unbiasedness focused on conventional OLS regressions of the most direct implied equation in levels

\[ x_{i+j} = \alpha + \beta f_{i,j} + u_{i+j}. \tag{1} \]

Here, the null would be that \( \alpha = 0 \) and \( \beta = 1 \). The common finding was that \( \beta \) was, indeed, very close to 1 and the null of unbiasedness could not be rejected. Precisely speaking, however, this should not be regarded as an evidence for the unbiasedness, since both the spot and forward exchange rates are non-stationary. If they are integrated processes, their relationship (1) in levels defines a cointegrating regression, for which the usual OLS estimator is consistent even when the endogeneity is present and \( E_i[x_{i+j}] \neq f_{i,j} \). It is well known that the OLS procedure is super-consistent, though not fully efficient, in this case. See, e.g., Goodhart, McMahon, and Ngama [13] for a discussion on how to efficiently analyze the regression (1) in levels as a cointegrating regression.

Forward rate unbiasedness has more often been tested based on the regression

\[ x_{i+j} - x_i = \alpha + \beta (f_{i,j} - x_i) + u_{i+j}, \tag{2} \]

where \( f_{i,j} - x_i \) is the forward premium or discount depending on whether this difference is positive or negative, respectively. Throughout this paper, we will follow the literature and abuse diction by simply referring to \( f_{i,j} - x_i \) as the forward premium. It is widely accepted that both the return and the forward premium are stationary processes, so (2) may be considered as a standard stationary regression which can be analyzed by OLS. We might expect \( \alpha = 0 \) and \( \beta = 1 \), if preferences are risk neutral. Clearly, under the null, equations (1) and (2) are equivalent. Thus, the (log of) the forward premium provides an unbiased forecast of the (log of) the future spot exchange rate return.

Equation (2) is a direct implication of the Covered Interest Parity (CIP) and the Uncovered Interest Parity (UIP) conditions. The CIP, a fundamental assumption in international finance, relates the forward premium to cross-country interest rate differentials according to

\[ f_{i,j} - x_i = r_{i,j} - r^*_{i,j}, \]

where \( r_{i,j} \) and \( r^*_{i,j} \) are the domestic and foreign interest rates, respectively.
where \( r_{i,j} \) and \( r_{i,j}^* \) are the returns on zero-coupon bonds at the \( i \)-th period with maturity \( j \) periods ahead for the domestic and foreign countries respectively. This relationship is widely supported empirically. The UIP instead connects the interest rate differentials to the spot exchange rate return, and it is given by

\[
x_{i+j} - x_i = r_{i,j} - r_{i,j}^*
\]

using the same notation as above. It is assumed to be true in many macroeconomic models and yet routinely fails in tested empirically.

A stylized fact of major exchange rates is that they very closely mimic simple random walks. In fact, it is difficult to outperform the simple random walk in forecasting exchange rates. This has led to another common and well known hypothesis that exchange rates follow a random walk. If the exchange rates did follow a random walk, then it follows in particular that the best predictor of the future spot rate is the current spot rate, \( x_i = E_i[x_{i+j}] \), in the mean squared error sense. In this case, we should find in (2) that \( \alpha = 0 \) and \( \beta = 0 \). If \( \alpha = 0 \) and \( \beta \neq 0 \) in (2), then we should have \( f_{i,j} = x_i \), in which case \( \beta \) becomes unidentified.

1.2 The Anomaly

The anomalous finding in investigating the unbiasedness hypothesis is the consistent result that the forward premium is, indeed, a predictor of future spot returns. However, it is (counter-intuitively) a reverse predictor. A plethora of empirical studies have not only rejected the null of unbiasedness, but have found significantly negative estimates of the slope parameter. Thus, if these results are correct when the forward rate is below the spot rate, there is always an expected positive return to entering into a forward contract. Baillie and Bollerslev [2] believe that: “the FPA has become a well established regularity and is generally regarded as being one of the most important unresolved paradoxes in international finance, and occupies a similar role to that of the equity premium puzzle in financial economics.” Froot [12] notes that the average value of over 75 published estimates is \(-0.88\) and very few are positive. Some of these estimates can be quite low. Consider for example an estimate of \( \beta = -5.644 \) for the United States dollar/Dutch guilder exchange rate ([26]). If this value is correct, it would imply that the forward premium is, indeed, a valid predictor of the future spot return. If the current forward exchange rate were 1% higher than the current spot exchange rate, we would expect that on average there would be a \(-5%\) loss in the corresponding future spot price of the foreign currency.

These results are troubling on an intuitive and a theoretical level. If we accept this empirical regularity, it can be justified only by abandoning the rational expectations assumption or by the existence of a time-varying risk premium. Further, this time-varying risk premium must exhibit more volatility than the expected depreciation. Such a characteristic is difficult to justify economically.

Assume that \( f_{i,j} - x_i \) and \( x_{i+j} - x_i \) are jointly stationary and ergodic. If \( \hat{\beta} \) is a consistent estimator, then it follows that

\[
\hat{\beta} \to_p \beta = \frac{cov(f_{i,j} - x_i, x_{i+j} - x_i)}{var(f_{i,j} - x_i)}.
\]
However, we have $\text{cov}(f_{i,j} - x_i, x_{i+j} - x_i) = \text{cov}(f_{i,j} - x_i, \mathbb{E}_i[x_{i+j}] - x_i)$, if we maintain the assumption of rational expectations and the forecast error $x_{i+j} - \mathbb{E}_i[x_{i+j}]$ is orthogonal to $f_{i,j} - x_i$. Moreover, if we let

$$\nu_{i,j} = f_{i,j} - \mathbb{E}_i[x_{i+j}],$$

i.e., the (potentially time varying) risk premium, then

$$\text{cov}(f_{i,j} - x_i, \mathbb{E}_i[x_{i+j}] - x_i) = \text{var}(f_{i,j} - x_i) - \text{cov}(\mathbb{E}_i[x_{i+j}] - x_i, \nu_{i,j}) - \text{var}(\nu_{i,j}).$$

Therefore, we may write

$$\beta = 1 - \beta_\nu$$

with

$$\beta_\nu = \frac{\text{cov}(\mathbb{E}_i[x_{i+j}] - x_i, \nu_{i,j}) + \text{var}(\nu_{i,j})}{\text{var}(f_{i,j} - x_i)}.$$

Consequently, $\beta$ is low if $\text{cov}(\mathbb{E}_i[x_{i+j}] - x_i, \nu_{i,j}) + \text{var}(\nu_{i,j})$ is positive and large. Clearly, if market participants are risk neutral, then $\nu_{i,j}$ would be zero and $\beta = 1$.

A highly variable risk premium is difficult to justify with macroeconomic models since the consumption and other determinants of the risk premium are usually smooth. Several popular models can support a positive slope coefficient that is less than one (i.e., $[4]$). However, to the best of our knowledge, none can support a negative slope.

Further economic unpleasantness is created by the consistent finding that the implied levels of risk aversion are unreasonably high. For example, Hodrick [16] estimates the coefficient to be 60.918. Kaminsky and Peruga [17] estimate it to be 372.37. The standard value used in calibration is around 1.5.

Rejecting (or modifying) the assumption that agents have rational expectations is certainly controversial. Moreover, though several popular models can support a positive slope coefficient that is less than one. (i.e., $[4]$), none that we know of can justify a time-varying risk premium volatile enough to imply a negative slope coefficient. Thus, much work has focused on showing why the econometric techniques themselves are flawed. Much of the work has been aimed at explaining why the standard regression could give such results.

The forward rate is consistently found to be highly persistent, but still stationary. Baillie and Bollerslev [1] suggest that the temporal dependencies exhibited by the forward premium are well described by a fractionally integrated process. Cheung uses a Kalman filter, treating the risk premium $\nu_{i,j}$ as an unobserved variable. The risk premium is assumed to follow an ARMA process. Bekaert, Hodrick, and Marshall [6] attempt to base hypothesis testing on the small-sample distributions rather than the test statistics’ asymptotic distributions. Their results strengthen the evidence against the expectations hypothesis of the term structure of interest rates. Baillie and Bollerslev [2] assert that the FPA is merely a statistical artifact created by the persistence in the forward premium and the small sample sizes. Thus, the finding is due to the relatively slow convergence rate of $\hat{\beta}$.

It appears that there are two basic reactions to the FPA: economic and econometric. Those who subscribe to the former assert that the econometric tools and therefore the results are correct, thus the underlying economic theory must be incorrect. Advocates of the latter, however, claim that the standard economic theory is correct and the econometric
methods are then invalid. Thus, economists assume that econometricians are correct and seek to adjust their fundamental models (e.g., agents are not rational), while the econometricians assume that the economists are correct and adjust their techniques. This paper suggests that both the economic and econometric theory require adjustment. Here, we suggest an appropriate model (which has been done before in various ways), we also suggest an econometric technique suitable to our economic framework (which has also previously been done to varying degrees of success), then we tackle this issue via re-estimating the appropriate model using an appropriate technique (which has not).

2 A Continuous Time Model of the Foreign Currency Market and Forward Exchange Rates

As Hansen and Hodrick (1980) put, “[a]ny discussion of the efficiency of a market requires a specification of the preferences (for risk) and information sets of economic agents,... and the costs inherent in transactions.” Given the highly active and competitive nature of foreign exchange markets and the availability of much improved and inexpensive transactions technology, we model our valuation framework in continuous time, with negligible transaction costs, where information flows via a standard Brownian motion. In absence of consensus on risk preference modeling in general equilibrium asset pricing models, we adopt no arbitrage condition following much of the finance literature (see Duffie (2002) for details). This method has an advantage of imposing a competitive equilibrium restriction and allowing one to model a risk adjustment for asset returns.

Specifically, consider a continuous-time economy that consists of two countries referred to as the domestic country and the foreign country. Assume that there is no arbitrage in international markets and that the currency market clears continuously. Let $W$ be a standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$, and fix the standard filtration $(\mathcal{F}_t)$ generated by $W$. We define the spot currency exchange rate $X_t$ by the domestic value per unit of the foreign currency at time $t$. Further, let this price process, $X_t$, follow an Ito process driven by the Brownian motion $W$.

Assume that there exist locally riskless money-market accounts $B$ and $B^*$ exclusively available in the domestic and foreign currency, respectively, with the following laws of motion

$$ dB_t = r_t B_t dt $$
$$ dB^*_t = r^*_t B^*_t dt, $$

where we set $B_0 = B^*_0 = 1$ and $r_t$ and $r^*_t$ are the instantaneous short-term interest rates for each country at time $t$.

If we further assume that markets are complete, we are guaranteed the existence of the unique and equivalent martingale measures $\hat{P}$ and $\hat{P}^*$ for the domestic and foreign countries, respectively. In what follows, we will denote respectively by $\hat{E}_t$ and $\hat{E}_t^*$ the $\mathcal{F}_t$-conditional expectations with respect to $\hat{P}$ and $\hat{P}^*$. We can now derive a risk-adjusted uncovered interest parity equation.
Lemma 1 If there exist no arbitrage across countries and markets are complete, spot exchange return and foreign and domestic spot interest rates will be related according to

$$
\frac{X_t}{X_0} = \exp \left( - \int_0^t r_s ds \right) \frac{D^*_t}{D_t},
$$

(3)

where $D_t$ and $D^*_t$ are the Radon-Nikodym derivatives of $\tilde{P}^*$ and $\tilde{P}$, respectively, with respect to $\mathbb{P}$ on $\mathcal{F}_t$, i.e., $D_t = \mathbb{E}_t[\mathcal{d}\tilde{P}/\mathcal{d}\mathbb{P}]$ and $D^*_t = \mathbb{E}_t[\mathcal{d}\tilde{P}^*/\mathcal{d}\mathbb{P}]$.

Proof of Lemma 1 Suppose that there exists a domestic asset whose current price is $S_0$ at time 0 and pays $S_t$ at $t$. Assuming that there is no arbitrage, there exists an equivalent martingale measure $\tilde{P}$ such that the deflated process is a martingale. That is,

$$
S_0 = \tilde{E}_0 \left[ \frac{S_t}{B_t} \right].
$$

(4)

For the same asset, in the foreign country, we have

$$
\frac{S_0}{X_0} = \tilde{E}^*_0 \left[ \frac{S_t}{X_t} \right],
$$

i.e.,

$$
S_0 = \tilde{E}^*_0 \left[ \frac{S_t (X_0/X_t)}{B_t^*} \right]
$$

(5)

under no arbitrage condition.

It follows from (4) and (5) that

$$
\int_A S_0 d\tilde{P} = \int_A \frac{S_t}{B_t} d\tilde{P}
$$

and

$$
\int_A S_0 d\tilde{P}^* = \int_A \frac{S_t (X_0/X_t)}{B_t^*} d\tilde{P}^*
$$

for any $A \in \mathcal{F}_0$. Moreover, we have

$$
\int_A S_0 d\tilde{P}^* = \int_A S_0 \frac{D^*_0}{D_0} d\tilde{P} = \int_A S_0 d\tilde{P}
$$

and

$$
\int_A \frac{S_t (X_0/X_t)}{B_t^*} d\tilde{P}^* = \int_A \frac{S_t (X_0/X_t)}{B_t^*} \frac{D^*_t}{D_t} d\tilde{P}.
$$

Therefore, we have

$$
\int_A \frac{S_t}{B_t} d\tilde{P} = \int_A \frac{S_t (X_0/X_t) D^*_t}{B_t^*} d\tilde{P}
$$

(6)

for all $A \in \mathcal{F}_0$. 

---

7
To finish the proof, note that (6) holds for all \( \mathcal{F}_t \)-measurable \( S_t \), since the market is assumed to be complete. Consequently, we may readily deduce that
\[
\frac{1}{B_t} = \frac{X_0/X_t}{D_t^*} D_t^*.
\]
and the stated result follows immediately. \( \square \)

One can easily see that the equation (3) becomes the usual uncovered interest parity theorem if we assume risk neutrality. Given the completeness of the market and absence of arbitrage, the market prices of risk \( \lambda_t \) and \( \lambda_t^* \) exist uniquely and we may write the Doléans exponentials
\[
D_t = \exp \left( -\int_0^t \lambda_s dW_s - \frac{1}{2} \int_0^t \lambda_s^2 ds \right), \quad (7)
\]
\[
D_t^* = \exp \left( -\int_0^t \lambda_s^* dW_s - \frac{1}{2} \int_0^t \lambda_s^{*2} ds \right), \quad (8)
\]
due to Girsanov’s theorem. It is well known that the Novikov condition suffices for these exponentials to be martingale and have finite variance. Plugging (7) and (8) into (3) and then applying Ito’s lemma gives
\[
d \ln X_t = \left( r_t - r_t^* - \frac{1}{2} (\lambda_t - \lambda_t^*) (\lambda_t + \lambda_t^*) \right) dt + (\lambda_t - \lambda_t^*) dW_t. \quad (9)
\]
It is clear from (9) that, as opposed to the usual UIP equation, the stochastic differential equation of \( X_t \) will also include an additional term to \( r_t - r_t^* \), which is the risk premium related to investing in foreign currency.\(^1\) Note that the instantaneous volatility term is solely determined by the difference of market prices of risks for the two countries’ currencies and the risk premium together with logarithmic adjustment term is also expressed entirely in terms of the risk premium of each country. Since the instantaneous volatility term is the difference of the two market prices of risk, modeling risk premium for holding foreign currency reduces to specifying a market price of risk for either the domestic or foreign country. One caveat of this setup is that our model abstracts away from many issues such as capital control, market integration, and other politico-economic risk factors which can affect exchange rate dynamics in a non-trivial way.

We can now derive a continuous-time analogue of the uncovered interest parity equation via forward spot spread. We define \( Y_t = \ln X_t \) as the log exchange rate, and let as before \( F_{t,s} \) be the forward price of \( X_t \) for delivery at \( s > t \).

**Theorem 2** If there exist no arbitrage across countries, the domestic and foreign instantaneous interest rate differential and the forward premium will be related according to
\[
r_t - r_t^* = \pi_t, \quad (10)
\]
\(^1\)Backus, Foresi, and Telmer (2001) in their discrete-time setup with an affine term structure derived a similar expression. But, we do not rely on any particular term structure model nor specific Ito processes to derive our expression.
where $\pi_t$ is the instantaneous forward premium at time $t$ defined as

$$
\pi_t = \lim_{s \to t^+} \frac{\ln F_{t,s} - \ln X_t}{s - t}.
$$

Furthermore, we have

$$
dY_t = \left[ \pi_t + \left( \lambda_t - \lambda_t^* \right)^2 / 2 - \lambda_t \left( \lambda_t - \lambda_t^* \right) \right] dt + \left( \lambda_t - \lambda_t^* \right) dW_t 
$$

(11)

in terms of the instantaneous forward premium and risk premium.

**Proof of Theorem 2** If we denote $Z(t,s)$ and $Z^*(t,s)$ as the zero coupon bond prices for the domestic and foreign countries with maturity $s > t$, a simple arbitrage argument implies that

$$
\frac{F_{t,s}}{X_t} = \frac{Z^*(t,s)}{Z(t,s)},
$$

(12)
i.e.,

$$
\ln F_{t,s} - \ln X_t = \ln Z^*(t,s) - \ln Z(t,s). 
$$

(13)

Recall that

$$
r_t = - \lim_{s \to t^+} \frac{\ln Z(t,s)}{s - t}.
$$

Therefore, it follows by definition that

$$
r_t - r_t^* = \lim_{s \to t^+} \frac{\ln Z^*(t,s) - \ln Z(t,s)}{s - t} 
$$

(14)

$$
= \lim_{s \to t^+} \frac{\ln F_{t,s} - \ln X_t}{s - t}.
$$

If we combine (9) and (10), we have our equation of exchange rate determination (11). □

Note that $\pi_t$, the instantaneous forward premium at time $t$ represents the premium paid to an agent who agrees to a contract to exchange currencies at a certain price an infinitesimally small time in the future. Therefore, (14) is simply an instantaneous version of the well known CIP equation. That is, for a small, discrete interval $[t, t + \delta]$, (14) is

$$
(r_t - r_t^*) \delta \simeq \ln F_{t,t+\delta} - \ln X_t.
$$

In this sense, (11) can be understood as a continuous-time version of forward premium equation with time-varying risk premium and volatility. As mentioned earlier, our continuous-time framework presumes that currency trading occurs at a very fine time scale and markets clear almost continuously. These presumptions are made wholly credible by the nature of the currency market. With hundreds of dealers all over the world in the USD/EUR currency pair alone, the foreign exchange market is clearly a twenty-four hour market. Moreover, quotes are available on a per second basis.

A discrete-time relationship between, say, a forward premium of maturity one month and currency exchange rate changes over a month, and can be derived by integrating (11).
However, we encounter several issues in empirically evaluating the forward premium equation. First, it is subject to several econometric biases which we will explain in next section. Further, there is another important, economic issue. To illustrate this, suppose that (11) holds at every $t$ and we want to analyze a discrete-time relation for one-month forward premium using the data collected at a fixed time of every month. If we denote the time by $t_i$ for month $i$ and let $x_i = \ln X_{t_i}$, then we have

$$
x_{i+1} - x_i = \int_{t_i}^{t_{i+1}} \left[ \pi_s + \frac{(\lambda_s - \lambda_s^*)^2}{2} - \lambda_s (\lambda_s - \lambda_s^*) \right] ds + \int_{t_i}^{t_{i+1}} (\lambda_s - \lambda_s^*) dW_s \tag{15}
$$

A usual forward premium equation used for this estimation is

$$
x_{i+1} - x_i = \alpha + \beta (f_{i,i+1} - x_i) + u_{i+1}, \tag{16}
$$

where $f_{i,i+1}$ is the log one month forward rate at month $i$. Even aside from the presence of risk premium and various econometric issues arising from time varying stochastic volatilities in the error, (16) in general is not compatible with (15), unless some form of the expectations hypothesis holds. Expectations hypothesis implies that $f_{i,i+1}$ should be a summation of the expected instantaneous forward rates over a month interval plus some constant. Thus, testing (16) amounts to a joint test of the forward premium anomaly and the expectations hypothesis, whereas (15) is designed solely for the forward premium anomaly. This is a subtle, but important difference: In order to tackle the forward premium equation in pure form, our example suggests that we have to use a theoretical relationship that is compatible with market clearing intervals. Of course, if market prices of risk are constant for both countries, and there is no additional sources of shocks affecting currency markets, then the expectations hypothesis will hold in each country and therefore this is no concern. But even in this case there still does exist an econometric problem, i.e. need for bias correction due to temporal dependence. However, given ample evidence against the expectations hypothesis (See Campbell and Shiller (1991) for example), and time-varying nature of volatilities, it is more appropriate to use (15) when testing the forward premium anomaly.\(^2\)

## 3 Econometric Methodology

### 3.1 Econometric Model

The econometric model allows for the possibility of econometric error in the forward premium equation in the form of $\omega_t dV_t$ where $V_t$ is a standard Brownian motion independent of $W_t$. Consequently, we have

$$
dY_t = \left[ \pi_t + (\lambda_t - \lambda_t^*)^2 / 2 - \lambda_t (\lambda_t - \lambda_t^*) \right] dt + (\lambda_t - \lambda_t^*) dW_t + \omega_t dV_t \tag{17}
$$

$$
= \left[ \pi_t + (\lambda_t - \lambda_t^*)^2 / 2 - \lambda_t (\lambda_t - \lambda_t^*) \right] dt + \sigma_t dU_t,
$$

\(^2\)We do not argue that joint testing is unimportant. On the other hand, this is a very fundamental problem that requires a careful consideration. To this end, however, a full specification of the term structure models for both countries is necessary.
where
\[
\sigma_t^2 = (\lambda_t - \lambda_t^*)^2 + \omega_t^2
\]
and \(U_t\) is a standard Brownian motion.

The expression (17) has several interesting features distinguishing it from its conventional discrete time counterpart. First, the mean of spot returns is a function of each country’s market prices of risk. Further, the instantaneous volatility term is solely determined by the difference of market prices of risks for the two countries’ currencies. Thus, if agents are risk neutral (i.e., \(\lambda_t = \lambda_t^* = 0\)) or agents are risk averse such that the market prices of risk are always equal across countries (i.e., \(\lambda_t = \lambda_t^* \neq 0\)), then (17) reduces to an expression similar to the standard forward premium equation. The appropriate continuous-time analogue to the traditional model considers the case where \(\lambda_t = \lambda_t^*\). In this case, we are left with the equation
\[
dY_t = \pi_t dt + \omega_t dV_t,
\]
which may be estimated if properly time-aggregated. However, it seems reasonable that risk attitudes, and therefore the price of risk, might vary across countries. In such a case, the standard FPA specification will yield spurious estimations.

Our model (17) can be estimated if we specify the market prices of risk for two countries. Of course, correctly specifying and accurately estimating the market price of risk is no trivial task. There is a large and active literature on the estimation of the market price of risk \(\lambda_t\). Innovation in the estimation of this latent variable is beyond the scope of this paper. Instead, to maintain focus on the FPA, we use several simple or widely used models of the market price of risk. We will use the following models of the market price of risk.

<table>
<thead>
<tr>
<th>Model</th>
<th>(\lambda_t)</th>
<th>(\lambda_t - \lambda_t^*)</th>
<th>(\omega_t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>constant</td>
<td>constant</td>
<td>variable</td>
</tr>
<tr>
<td>Model 2</td>
<td>variable</td>
<td>constant</td>
<td>variable</td>
</tr>
<tr>
<td>Model 3</td>
<td>variable</td>
<td>variable</td>
<td>variable</td>
</tr>
</tbody>
</table>

The constant specifications of \(\lambda_t\) and \(\lambda_t^*\) simply treat them as unknown parameters. For the variable specifications of \(\lambda_t\) and \(\lambda_t^*\), we define them as
\[
\lambda_t = \frac{\mu_t - r_t^f}{\nu_t} \quad \text{and} \quad \lambda_t^* = \frac{\mu_t^* - r_t^f^*}{\nu_t^*},
\]
where \(\mu_t\) and \(\mu_t^*\) are returns from the risky assets and \(r_t^f\) and \(r_t^f^*\) are the risk-free rates in two countries. Therefore, \(\lambda_t\) and \(\lambda_t^*\) are the excess returns per unit of risk in two countries. For the estimation of our model, we use the S&P 500 for \(\mu_t\) and the three-month treasury bill rate as the risk-free rate \(r_t^f\).

Note that our model (17) allows for the presence of time-varying stochastic volatility in the errors. The instantaneous volatility \(\sigma_t\) is in general stochastic and varies over time. This is very consistent with common characteristics of financial data. This aspect of financial data is both widely accepted in the financial literature and largely ignored by the FPA literature. There are at least two important characteristics that we need to pay particular
attentions in dealing with the financial data. First, the distribution of the return process is far from normal. The peakedness and fat-tails found the return process are ubiquitous throughout financial data. Further, there is time-heterogeneity in the volatility of the return process, another common feature of financial data.

3.2 Time Change

To effectively deal with the time-varying stochastic volatility in the error process of (17), we use a time change. The idea is based on the widely known theorem by Dambis, Dubins and Schwarz (DDS hereafter). To state and interpret the DDS theorem more precisely, we need to introduce some additional concepts and notations. For a continuous martingale $M_t$, we define a time change $T_t$ by

$$T_t = \inf \{ s > 0 \mid \langle M \rangle_s > t \},$$

where $\langle M \rangle_t$ is the quadratic variation process of $M_t$. If $\langle M \rangle_t$ is continuous and strictly increasing a.s. as in many of the models commonly used for practical applications, the time change $T_t$ is nothing but the time inverse of the quadratic variation process $\langle M \rangle_t$. The DDS theorem states that

$$M_{T_t} = Z_t \quad \text{and} \quad M_t = Z_{\langle M \rangle_t},$$

where $Z_t$ is the standard Brownian motion, which is often referred to as the DDS Brownian motion.

Loosely put, the DDS theorem implies that all continuous martingales are essentially Brownian motions if we use chronometers given by their quadratic variation processes. The standard Brownian motion has a quadratic variation process, which is deterministic and given exactly by the actual time. Other more general continuous martingales have quadratic variation processes that are stochastic and varying across their different realizations. Indeed, it is well known that the standard Brownian motion is the only continuous martingale whose quadratic variation process is given by the actual time. It follows from the DDS theorem that if we use the chronometer inversely proportional to its quadratic variation process, any continuous martingale reduces to the standard Brownian motion, or equivalently, that any continuous martingale can be thought of the standard Brownian motion whose sample paths are read using the chronometer given by its quadratic variation process.

In the paper, we apply the DDS theorem to the continuous martingale process

$$M_t = \int_0^t \sigma_s dU_s,$$

i.e., the error process in our model (17), whose quadratic variation process is given by

$$\langle M \rangle_t = \int_0^t \sigma_s^2 ds.$$

Note that we do not impose any restriction on the instantaneous volatility $\sigma_t$. In particular, we allow $\sigma_t$ to be stochastic and varying over time in any arbitrary fashion. Furthermore, it
is easy to see that \( \langle M \rangle_t \) is continuous, and strictly increasing as long as \( \sigma_t \) is non-vanishing except for a set of Lebesgue measure zero.

Now we consider our model (17) under the time change, which is given by

\[
dY_t = \left[ \pi_t + (\lambda - \lambda^*)^2 / 2 - \lambda (\lambda - \lambda^*) T_t \right] dT_t + \sigma T_t dU_t
\]

and

\[
\sigma T_t dU_t = dM_t = dZ_t,
\]

where \( M_t \) and \( Z_t \) are defined earlier. By employing a time change from calendar time \( t \) to volatility time \( T_t \), the error process in our model has become the standard Brownian motion. Therefore, the non-normality and time-varying stochastic volatility in the errors of our original model disappear. Once the clock has changed, the error process has the well behaved features of a constant volatility, independence in increments and Gaussianity. In the next subsection, we will explore this to test for the parameters in the model more effectively.

In general, the error process \( M_t \) is not observed. Clearly, it is unobserved in our model (17) unless \( \lambda_t \) and \( \lambda^*_t \) are fully specified and observable. The quadratic variation \( \langle M \rangle_t \) of the error process, however, is observed without specifying \( \lambda_t \) and \( \lambda^*_t \). In fact, we have

\[
\langle Y \rangle_t = \langle M \rangle_t,
\]

since the term including \( \pi_t, \lambda_t \) and \( \lambda^*_t \) is of bounded variation and its quadratic variation vanishes at all \( t > 0 \). Therefore, we may obtain the quadratic variation of \( M_t \) directly from the exchange rate process \( Y_t \).

In the paper, we set \( M_t = Y_t - \int_0^t \pi_s ds \), i.e., \( \lambda_t = \lambda^*_t = 0 \) for all \( t > 0 \), to compute \( \langle M \rangle_t \). For any \( t > 0 \), \( \langle M \rangle_t \) is estimated by the realized variance

\[
\sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})^2,
\]

where \( 0 = t_0 < \cdots < t_n = t \), and the time change \( T_t \) is obtained from the estimated \( \langle M \rangle_t \). For each \( t > 0 \), the realized variance converges in probability to the quadratic variation as \( \max_i |t_i - t_{i-1}| \to 0 \). Therefore, the estimated time change is also expected to converge in probability to \( T_t \). The reader is referred to Park (2008) for the technical details. See Figure 1 for the estimated quadratic variation process and time change. In the subsequent explanation of our methodology, we simply assume to ease the exposition that \( \langle M \rangle_t \) and \( T_t \) are directly observed.

### 3.3 Martingale Estimation

For a fixed \( \Delta > 0 \), it follows from (18) and (19) that

\[
Y_{T_{i} \Delta} - Y_{T_{i-1} \Delta} = \int_{T_{i-1} \Delta}^{T_{i} \Delta} \left[ \pi_t + (\lambda_t - \lambda^*_t)^2 / 2 - \lambda_t (\lambda_t - \lambda^*_t) \right] dt + \varepsilon_i
\]

(20)
where
\[ \varepsilon_i = (Z_{T_i\Delta} - Z_{T_{(i-1)\Delta}}) \]
that are independent and identically distributed as \( N(0, \Delta) \) for \( i = 1, \ldots, N \).

For Model 1 introduced earlier, we consider the regression model given by
\[ Y_{T_i\Delta} - Y_{T_{(i-1)\Delta}} = \alpha_0 + \beta_0 \int_{T_{(i-1)\Delta}}^{T_i\Delta} \pi_t dt + \gamma_0 (T_i\Delta - T_{(i-1)\Delta}) + \varepsilon_i, \tag{21} \]
where \( \alpha_0 = 0, \beta_0 = 1 \) and
\[ \gamma_0 = (\lambda_t - \lambda_t^* )^2 / 2 - \lambda_t (\lambda_t - \lambda_t^* ) , \]
which is assumed to be a constant for all \( t \). This follows directly from (20). Note that the true parameter values \( \alpha_0, \beta_0 \) and \( \gamma_0 \) in (21) are identified by the conditions that \( (\varepsilon_i / \sqrt{\Delta}) \) are independent standard normals. It is clear that the normalized regression error \( (\varepsilon_i / \sqrt{\Delta}) \) is distributed as independent standard normals for no other values of these parameters.

To estimate the parameter \( \theta = (\alpha, \beta, \gamma)' \in \Theta \), we employ the martingale estimation method proposed recently by [23]. To introduce the method, we define
\[ z_i(\theta) = \frac{1}{\sqrt{\Delta}} \left[ Y_{T_i\Delta} - Y_{T_{(i-1)\Delta}} - \alpha - \beta \int_{T_{(i-1)\Delta}}^{T_i\Delta} \pi_t dt - \gamma (T_i\Delta - T_{(i-1)\Delta}) \right] , \]
and let
\[ z^d_i(\theta) = (z_i(\theta), z_{i-1}(\theta), \ldots, z_{i-d+1}(\theta))' , \]
the vector consisting of \( d \)-number of consecutive values of \( z_i(\theta) \) for each \( i \). Furthermore, we signify by \( \Pi_N(\cdot, \theta) \) the empirical distribution function of \( (z^d_i(\theta)) \) for each \( \theta \in \Theta \) and by \( \Pi_0(\cdot) \) the distribution function of \( (z^d_i(\theta_0)) \).

The martingale estimator (MGE) \( \hat{\theta}_N \) of the parameter \( \theta \) is defined as
\[ \hat{\theta}_N = \arg\min_{\hat{\theta} \in \Theta} \int_{-\infty}^{\infty} \left[ \Pi_N (z, \theta) - \Pi_0 (z) \right]^2 \varpi (dz) , \]
where \( \varpi \) is some weight measure, and \( N \) is the number of observations selected after the time change. It is shown in [23] that we have as \( N \to \infty \)
\[ \sqrt{N} (\hat{\theta}_N - \theta_0) \to_d N(0, \Omega) \]
under suitable regularity conditions, where \( \Omega = B^{-1} A B^{-1} \) with
\[ A = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{\Pi}_0 (x) \Sigma (x, y) \hat{\Pi}_0 (y)' \varpi (dx) \varpi (dy) \]
\[ B = \int_{-\infty}^{\infty} \hat{\Pi}_0 (z) \hat{\Pi}_0 (z)' \varpi (dz) , \]
\( \hat{\Pi}_0 \) is the derivative of \( \Pi_0 \), and \( \Sigma \) is the covariance kernel of the limit Gaussian process given in [23].
The motivation for the MGE is simple and straightforward. We just find the value of \( \theta \in \Theta \), for which the empirical distribution of \((z_i^d(\theta))\) is closest to the distribution of \((z_i^d(\theta_0))\). Recall that the distribution of \((z_i(\theta))\) becomes independent standard normal only when \( \theta = \theta_0 \). Therefore, \( \theta_0 \in \Theta \) is uniquely given by the distribution of \((z_i^d(\theta_0))\). Of course, the MGE may be viewed as a usual minimum distance estimator, where the distance is given by the Cramér-von Mises (CvM) distance between the empirical distribution of the sample \((z_i(\theta))\) with the unknown parameter value \( \theta \in \Theta \) and the distribution under the true parameter value \( \theta_0 \in \Theta \). In the paper, we use the MGE’s with \( d = 1 \) and \( d = 2 \).

For the weight measure \( \varpi \), we use the measure given by \( \Pi_0 \), i.e., \( \varpi(dz) = d\Pi_0(z) \), as suggested by [23]. In this case, there are simple ways to compute the CvM distance. To introduce them, we let \( \Phi \) be the standard normal distribution function, and define

\[
t_i(\theta) = \Phi(z_i(\theta)).
\]

For \( d = 1 \), the MGE can be obtained by numerically solving

\[
\hat{\theta}_N = \arg\min_{\theta \in \Theta} \sum_{i=1}^{N} \left( t_i(\theta) - \frac{2i - 1}{2N} \right)^2.
\]

Similarly, for \( d = 2 \), we may actually compute the MGE in

\[
\hat{\theta}_N = \arg\min_{\theta \in \Theta} \frac{1}{N} \sum_{i=2}^{N} \sum_{j=2}^{N} \left[ 1 - \max(t_i(\theta), t_j(\theta)) \right] \left[ 1 - \max(t_{i-1}(\theta), t_{j-1}(\theta)) \right] - \frac{1}{2} \sum_{i=2}^{N} (1 - t_i^2(\theta)) \left( 1 - t_{i-1}^2(\theta) \right).
\]

For these numerical optimization problem, we set the initial values of parameters \( \alpha, \beta \) and \( \gamma \) to be given by \( \alpha = 0, \beta = 1 \) and \( \gamma = 0 \).

For Model 2, we consider the regression model

\[
Y_{T_{i-1}} - Y_{T_{(i-1)\Delta}} = \alpha_0 + \beta_0 \int_{T_{(i-1)\Delta}}^{T_{i\Delta}} \pi_t dt + (\gamma_0^2/2)(T_{i\Delta} - T_{(i-1)\Delta}) + \gamma_0 \int_{T_{(i-1)\Delta}}^{T_{i\Delta}} \lambda_t dt + \varepsilon_i,
\]

where \( \alpha_0 = 0, \beta_0 = 1 \) and \( \gamma_0 = -(\lambda_t - \lambda_t^*) \), which is assumed to be constant. Finally, for Model 3,

\[
Y_{T_{i\Delta}} - Y_{T_{(i-1)\Delta}} = \frac{1}{2} \int_{T_{(i-1)\Delta}}^{T_{i\Delta}} (\lambda_t - \lambda_t^*)^2 dt + \int_{T_{(i-1)\Delta}}^{T_{i\Delta}} \lambda_t (\lambda_t - \lambda_t^*) dt = \alpha_0 + \beta_0 \int_{T_{(i-1)\Delta}}^{T_{i\Delta}} \pi_t dt + \varepsilon_i,
\]

where \( \alpha_0 = 0 \) and \( \beta_0 = 1 \). For Models 2 and 3, the parameters can be estimated from regressions (22) and (23) by the MGE in the same manner as for Model 1.
4 Empirical Results

Let us concretize the various institutional concepts related to the present question. A forward is a contract between parties to exchange currency at a future date (three days or more in the future) at an exchange rate agreed upon today. This three day minimum comes from the practical fact that spot purchases of currency are for delivery in two days. We will refer to the day on which the currency is to be delivered as the value date. A one month forward contract corresponds to 30 days from the current value date, if that day is a business day. If not, it corresponds to the nearest business day that is more than 30 days in the future ([5]). Forward transactions are often non-standard, however there do exist standard contracts for one month, two months, three months, six months, and one year which are widely available.

4.1 Data

Data was gathered from Barclays Bank PLC and retrieved via Datastream. We use daily observations for the US dollar, euro, yen, pound sterling, Swiss franc, Australian dollar, and Canadian dollar. Nearly 86% of all currency exchange transactions involve these 7 currencies. The periods are from January 2, 1984 to December 31, 2007, or 6,261 observations for the yen, pound, and franc. The Australian and Canadian dollar data spans January 1, 1985 to December 31, 2007, or 6,000 observations. For modeling purposes, we let the US be the domestic country in all cases.

We provide two estimates for the Euro. One is the usual Euro data. The other uses the German mark as a proxy for the Euro before its introduction, as is common in the literature. It is well know that when estimating the drift term of a diffusion process, the time span must be large. However, the Euro was introduced on January 1, 1999, giving less than half of the observations before December 31, 2007 than are available for the other currencies, or 2,345 observations. To extend the time span, we use the German mark as a proxy. Still, our first observation of German mark forward rates is January 1, 1997, or 2,868 observations prior to December 31, 2007. The standard errors for estimations based on either data for the Euro are larger than for other currencies. This may be a product of the smaller time spans, the mark as an imperfect proxy of the Euro, or both.

4.2 Estimation

The theoretical time-change results hold for any constant value $\Delta$. Of course, in practice some $\Delta$’s are better than others. For instance, a time-change based on a $\Delta$ smaller than any of the realized volatilities between observations will yield the original data, which has been shown to be non-normal and to potentially exhibit stochastic volatility. Conversely, if we choose $\Delta$ to be the total realized volatility, we are left with one observation. Currently, there is no literature on the optimal value of $\Delta$. We approach this selection in two simple ways.

First, we select a $\Delta$ based on minimizing the CvM statistic of the returns subject to maintaining at least 30 post-time-change observations and at least an average of 30
observations per random-time period. The motivation for this being that the time-change of the return process theoretically and ideally will yield a normal distribution. We thus select $\Delta$ which gives returns that are closest to that ideal. Second, we must also maintain some number of observations for our estimates to be reliable. Hence, our imposed post-time-change observation quantity requirements. The second selection criteria involves minimizing the estimated standard errors. This is to increase the precision of the estimates. The same quantity requirements are imposed.

In the limit, the quadratic variation achieved that surpasses $\Delta$ will be infinitesimally close to $\Delta$. In practice, we suspect that assigning the random time based on this rule will lead to an upward bias since the actual realized volatility will always be larger than $\Delta$. To address this issue, we select the time change based on minimizing the distance between the realized volatility and $\Delta$.

Further, since whatever selection process we use, the realized volatility will never be exactly $\Delta$, we can not expect $\varepsilon_i \sim N(0, \Delta)$ to hold exactly within each random-time period. To accommodate for this reality, we instead use as the variance of $\varepsilon_i$ the actual increment of the quadratic variation for the time interval between $T_{(i-1)\Delta}$ and $T_{i\Delta}$. So, the distribution will be normalized based on the standard deviation of the estimation error rather than the theoretical limit. In the following examples, $\Delta$ was selected by minimizing in the range of $N = [50, n/50]$, thus ensuring at least 50 post-time change observations as well as 50 observations within each random time interval.

<table>
<thead>
<tr>
<th>Model</th>
<th>$1/(12n) + \sum (F - G)^2$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euro</td>
<td>-0.0042 (0.0089)</td>
<td>1.4524 (1.2181)</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Yen</td>
<td>0.0005 (0.0029)</td>
<td>0.1908 (0.1156)</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>Pound</td>
<td>0.0027 (0.0012)</td>
<td>0.2498 (0.1406)</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td>Swiss Franc</td>
<td>0.0195 (0.0109)</td>
<td>-1.4618 (1.3236)</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td>Australian Dollar</td>
<td>0.0005 (0.0040)</td>
<td>-0.0216 (0.2383)</td>
<td>101</td>
<td></td>
</tr>
<tr>
<td>Canadian Dollar</td>
<td>0.0038 (0.0028)</td>
<td>0.2666 (0.2822)</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Model 1 minimize cvm</td>
<td>$1/(12n) + \sum (F - G)^2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>$t$</td>
</tr>
<tr>
<td>Euro</td>
<td>-0.0136</td>
<td>1.4917</td>
<td>-0.0002</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>(0.0136)</td>
<td>(0.4176)</td>
<td>(0.0003)</td>
<td></td>
</tr>
<tr>
<td>Yen</td>
<td>0.0039</td>
<td>0.5026</td>
<td>-0.0002</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>(0.0031)</td>
<td>(0.3127)</td>
<td>(0.0004)</td>
<td></td>
</tr>
<tr>
<td>Pound</td>
<td>-0.0100</td>
<td>2.0565</td>
<td>0.0006</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>(0.0047)</td>
<td>(0.6931)</td>
<td>(0.0002)</td>
<td></td>
</tr>
<tr>
<td>Swiss Franc</td>
<td>-0.0088</td>
<td>2.0550</td>
<td>-0.0009</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.2861)</td>
<td>(0.0003)</td>
<td></td>
</tr>
<tr>
<td>Australian Dollar</td>
<td>-0.0086</td>
<td>1.0844</td>
<td>0.0005</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>(0.0094)</td>
<td>(0.3631)</td>
<td>(0.0003)</td>
<td></td>
</tr>
<tr>
<td>Canadian Dollar</td>
<td>-0.0086</td>
<td>0.7423</td>
<td>-0.0002</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>(0.0023)</td>
<td>(0.1526)</td>
<td>(0.0001)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 1 minimize se</th>
<th>$1/(12n) + \sum (F - G)^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Euro</td>
<td>-0.0076</td>
</tr>
<tr>
<td></td>
<td>(0.0044)</td>
</tr>
<tr>
<td>Yen</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>(0.0072)</td>
</tr>
<tr>
<td>Pound</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.0117)</td>
</tr>
<tr>
<td>Swiss Franc</td>
<td>0.0020</td>
</tr>
<tr>
<td></td>
<td>(0.0105)</td>
</tr>
<tr>
<td>Australian Dollar</td>
<td>0.0095</td>
</tr>
<tr>
<td></td>
<td>(0.0068)</td>
</tr>
<tr>
<td>Canadian Dollar</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>(0.0021)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 2</th>
<th>$1/(12n) + \sum (F - G)^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Euro</td>
<td>-0.0217</td>
</tr>
<tr>
<td></td>
<td>(0.0077)</td>
</tr>
<tr>
<td>Yen</td>
<td>0.0038</td>
</tr>
<tr>
<td></td>
<td>(0.0034)</td>
</tr>
<tr>
<td>Pound</td>
<td>-0.0058</td>
</tr>
<tr>
<td></td>
<td>(0.0047)</td>
</tr>
<tr>
<td>Swiss Franc</td>
<td>0.0055</td>
</tr>
<tr>
<td></td>
<td>(0.0085)</td>
</tr>
<tr>
<td>Australian Dollar</td>
<td>-0.0055</td>
</tr>
<tr>
<td></td>
<td>(0.0057)</td>
</tr>
<tr>
<td>Canadian Dollar</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.0042)</td>
</tr>
</tbody>
</table>
### 4.3 Estimates

Tables 1-2 give our estimations of the simplest model, Model 1. Estimates of the bias coefficient, as well as the constant coefficient, and the variance of the error term are provided. Further, we present the selected value of $\Delta$ and the subsequent number of post-time-change observations.

Table 1 presents the results of our estimation given by selecting the value of $\Delta$ which minimizes the CvM statistic while keeping 150 post-time-change observations. The null of a slope coefficient of one can not be rejected for all currencies. No currencies have a ’biasedness’ coefficient that is significantly different from 1. Further, the euro, pound, franc, and Australian dollar are all significantly positive. The constant parameter is only significantly positive for the pound, but again quite small. The variance of the error term is significant for the yen, franc, and Australian dollar.

Similarly, our Table 2 gives the results of our estimation given by selecting the value of $\Delta$ selected by minimizing the bivariate CvM statistic that maintains 150 post-time-change observations. The estimate is then made using the two-dimensional martingale estimator. Again, the null can not be rejected for any currencies. None of the constant coefficients are significant. The yen, the franc, the Australian dollar, and the Canadian dollar all have significant error variances.

We next follow a similar approach to the slightly more realistic Model 2. We now
Table 1: Model 3 Univariate Martingale Estimates

<table>
<thead>
<tr>
<th></th>
<th>beta</th>
<th>alpha</th>
<th>sigma</th>
<th>gamma</th>
<th>K</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euro</td>
<td>1.0333</td>
<td>0.00057</td>
<td>0.00077</td>
<td>-0.00137</td>
<td>0.00106</td>
<td>154</td>
</tr>
<tr>
<td></td>
<td>(0.57054)</td>
<td>(0.00462)</td>
<td>(0.00054)</td>
<td>(0.00111)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yen</td>
<td>0.40215</td>
<td>0.00006</td>
<td>0.00118</td>
<td>0.00028</td>
<td>0.00319</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>(0.18983)</td>
<td>(0.00922)</td>
<td>(0.00065)</td>
<td>(0.00035)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pound</td>
<td>0.85674</td>
<td>0.0008</td>
<td>-0.00013</td>
<td>-0.00058</td>
<td>0.00283</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>(0.10665)</td>
<td>(0.00607)</td>
<td>(0.00003)</td>
<td>(0.00033)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swiss franc</td>
<td>0.94965</td>
<td>-0.00583</td>
<td>0.00194</td>
<td>-0.00011</td>
<td>0.00392</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>(0.21987)</td>
<td>(0.00842)</td>
<td>(0.00048)</td>
<td>(0.00038)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australian dollar</td>
<td>0.90088</td>
<td>-0.00048</td>
<td>-0.00086</td>
<td>0.00008</td>
<td>0.00308</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>(0.23097)</td>
<td>(0.00392)</td>
<td>(0.00048)</td>
<td>(0.00042)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canadian dollar</td>
<td>0.78906</td>
<td>0.00178</td>
<td>-0.00022</td>
<td>0.00001</td>
<td>0.00094</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>(0.46278)</td>
<td>(0.00232)</td>
<td>(0.00029)</td>
<td>(0.0004)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Model 3 Bivariate Martingale Estimates

<table>
<thead>
<tr>
<th></th>
<th>beta</th>
<th>alpha</th>
<th>sigma</th>
<th>gamma</th>
<th>K</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euro</td>
<td>0.74434</td>
<td>0.00106</td>
<td>-0.00021</td>
<td>0.00139</td>
<td>0.00095</td>
<td>171</td>
</tr>
<tr>
<td></td>
<td>(0.85603)</td>
<td>(0.00741)</td>
<td>(0.00091)</td>
<td>(0.00077)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yen</td>
<td>0.86998</td>
<td>-0.00036</td>
<td>0.00215</td>
<td>-0.00087</td>
<td>0.00339</td>
<td>156</td>
</tr>
<tr>
<td></td>
<td>(0.11227)</td>
<td>(0.00754)</td>
<td>(0.00048)</td>
<td>(0.00035)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pound</td>
<td>0.96112</td>
<td>0.00057</td>
<td>-0.00063</td>
<td>-0.00016</td>
<td>0.00246</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>(0.16848)</td>
<td>(0.0054)</td>
<td>(0.00023)</td>
<td>(0.00019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swiss franc</td>
<td>0.8946</td>
<td>0.0134</td>
<td>0.00097</td>
<td>-0.00043</td>
<td>0.00386</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td>(0.21582)</td>
<td>(0.00568)</td>
<td>(0.00049)</td>
<td>(0.00028)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australian dollar</td>
<td>0.89134</td>
<td>0.00259</td>
<td>-0.00151</td>
<td>0.00059</td>
<td>0.00324</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>(0.24115)</td>
<td>(0.00519)</td>
<td>(0.00021)</td>
<td>(0.00034)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canadian dollar</td>
<td>0.8232</td>
<td>0.00202</td>
<td>-0.00052</td>
<td>0.00006</td>
<td>0.00105</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>(0.21272)</td>
<td>(0.0027)</td>
<td>(0.00021)</td>
<td>(0.00054)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

estimate an additional parameter γ, which corresponds to the domestic constant price of risk. The ordering of the tables is identical to that of Model 1. In general, the estimates become closer to 1 and the accuracy (in terms of standard errors) increases as the models become more sophisticated.

When we minimize the CvM statistic keeping at least 150 observations (Table 3), cannot reject unbiasedness for any currencies. All currencies are significantly positive. No currencies have a significant constant term, variance of the error term, or market price of risk.

When we minimize the two dimensional CvM statistic keeping at least 150 observations (Table 4), cannot reject unbiasedness for any currencies. All currencies, except the euro, are significantly positive. The Swiss franc has a significant constant term. No currency has a significant variance of the error term. The franc has a significant and positive market
price of risk.

Finally, we consider Model 3, with a roughly estimated market price of risk in Table 5. These results are perhaps the most realistic, and consequently, the most precise. We minimize the two dimensional CvM statistic keeping at least 150 observations. The estimated slope coefficient for the exchange rate of each currency is not significantly different from 1. Additionally, except for the Euro, all are significantly greater than 0. The franc again is found to have a significant constant parameter. The yen, the pound, the franc, the Australian and Canadian dollars all have significant variances in their respective error terms. Finally, the yen alone has a significant market price of risk.

5 Conclusion

The main goal of this paper is to make use of both economic and econometric theories in a coherent way to sharpen our understanding about the dynamic behavior of foreign currency markets. To this end, we develop a continuous-time model of forward and spot exchange rates and apply a new econometric approach to identify it.

Specifically we pay attention to the point that if the foreign exchange rate market clears continuously, the traditional forward premium anomaly regression may be misspecified. This would render traditional estimates invalid even if econometric methodology were correct. To address this, we derive a continuous time forward premium equation, linking the spot returns to both the instantaneous forward premium and the market prices of risk. Alas foreign exchange rates data at a high frequency level exhibit all of the complications commonly found in financial data. Among others, most significant would be the time-varying volatility of foreign currency. To correct for the stochastic volatility found in the exchange rates, we change from a standard clock to a volatility chronometer. Once this change is made, the error process should be identically and independently normally distributed. Using this property, we apply a minimum distance method and verify if the no arbitrage conditions hold in the foreign exchange market.

Our results are unambiguous and robust. When the appropriate economic model is accompanied by the appropriate econometric technique, the forward premium anomaly disappears. The instantaneous forward exchange market is thus speculatively efficient. That is, the continuously compounded instantaneous forward premium is an unbiased predictor of the risk adjusted spot return. The results suggest that even a simple model of market price of risk is useful to resolve the forward premium puzzle, but more sophisticated models appear to perform better. In this sense, our paper is broadly consistent with the current trend in both economics and finance literatures which focus on modeling dynamic behaviors of risk premia.

The next direction for this research is extending the analysis beyond the instantaneous forward exchange rate. This may be somewhat more challenging since we can no longer use (14), and instead must derive a general continuous-time forward premium equation involving the term structure of interest rates for international bond markets.
6 Appendix

6.1 The Nelson-Siegel Procedure

The need to estimate $f_{t,t}$ remains. Indeed, there is no need to restrict ourselves to an instantaneous forward exchange rate. We can further expand our analysis to any maturity. This elucidates the need for a continuous forward exchange rate equation.

Forward exchange rate premiums as a function of maturity have shapes which are invariably monotonic, humped, or ‘S’ shaped. If we assume spot exchange rates are generated by a differential equation, then forward exchange rates, being forecasts, will be the solution to the equations. Thus, the forward exchange rate $f_t$, a function of the time until maturity $\tau$, would be of the form:

$$f_t(\tau) = \beta_1t + \beta_2e^{-\lambda_t \tau} + \beta_3 \lambda_t e^{-\lambda_t \tau}$$

where $\beta_1$ are parameters to be estimated and the term $\lambda_t$ represents the decay rate. Such a model readily produces the shapes required of a forward exchange rate curves. This is simply the functional form of a forward interest rate applied to forward exchange rates. The Nelson-Siegel procedure is a widely used parsimonious method for interpolating forward interest rates. The Nelson-Siegel approximation is especially popular among central banks for estimating forward interest rates.

If we follow Diebold and Li, we choose approximately 30 months as a medium term maturity, which they give as $\lambda_t = 0.0609$. This allows the model to be estimated using ordinary least squares. This simplifies and adds numerical robustness since many potentially challenging numerical optimizations are replaced with trivial least squares. Using Nelson-Siegel approximation, we can estimate a $\beta$ for every possible maturity for each day $t$. Forward prices are very close to the actual forward prices.

References


