WP/06/136

IMF Working Paper

Nonlinearity in Deviations from Uncovered Interest Parity: An Explanation of the Forward Bias Puzzle

Lucio Sarno, Giorgio Valente, and Hyginus Leon

INTERNATIONAL MONETARY FUND

 WP/06/136

IMF Working Paper

Middle East and Central Asia Department

Nonlinearity in Deviations from Uncovered Interest Parity: An Explanation of the Forward Bias Puzzle

Prepared by Lucio Sarno, Giorgio Valente, and Hyginus Leon*

Authorized for distribution by Mohsin S. Khan

May 2006

Abstract

We provide empirical evidence that deviations from uncovered interest rate parity (UIP) display significant nonlinearities, consistent with theories based on transaction costs or limits to speculation. This evidence suggests that the forward bias documented in the literature may be less indicative of major market inefficiencies than previously thought. Monte Carlo experiments allow us to reconcile these results with the large empirical literature on the forward bias puzzle since we show that, if the true process of UIP deviations were of the nonlinear form we consider, estimation of conventional spot-forward regressions would generate the anomalies documented in previous research.

JEL Classification Numbers: C22, F31

Keywords: Forward bias puzzle; uncovered interest parity; nonlinearity

Author(s) E-Mail Address: lsarno@wbs.ac.uk; giorgio.valente@cuhk.edu.hk; hleon@imf.org

---

* Lucio Sarno is at University of Warwick and CEPR; Giorgio Valente is at Chinese University of Hong Kong; Hyginus Leon is at International Monetary Fund. This paper was written while Lucio Sarno was a visiting scholar at the IMF Institute. Financial support from the Economic and Social Research Council (Grant No. RES-000-22-0404) is gratefully acknowledged. The authors are indebted for useful conversations or constructive comments to Ulf Axelson, Magnus Dahlquist, Paul De Grauwe, Hans Dewachter, John Driffill, Bob Flood, Gordon Gemmill, Campbel Harvey, Peter Kenen, Rich Lyons, Angelo Melino, Chris Neely, Anthony Neuberger, Carol Olser, David Peel, Dagfinn Rime, Piet Sercu, Per Strömberg, Shinji Takagi, Gabriel Talmain, Mark Taylor, Timo Teräsvirta, Dan Thornton, Shang-Jin Wei, Mike Wickens, Mark Wohar, and Josef Zechner, as well as participants at various seminars and conference presentations. The authors alone are responsible for any errors that may remain.

Contents Page

I. Introduction

The uncovered interest rate parity (UIP) condition postulates that the expected foreign exchange gain from holding one currency rather than another, that is, the expected exchange rate change, must be just offset by the opportunity cost of holding funds in this currency rather than the other, namely the interest rate differential. Assuming that there are no arbitrage opportunities and, therefore, that the interest rate differential equals the forward premium, UIP then implies that the expected exchange rate change must equal the current forward premium. In this case, the risk-neutral efficient markets hypothesis holds. This condition is routinely assumed in models of international macroeconomics and finance.

In a highly influential paper, Fama (1984) noted that high interest rate currencies tend to appreciate, whereas one might suppose that investors would demand higher interest rates on currencies expected to fall in value. In turn, this result suggests that the forward premium tends to be inversely related to future exchange rate changes, in contrast to the UIP hypothesis. This anomaly, often termed the "forward bias puzzle," continues to spur a large literature. However, regardless of the increasing sophistication of the econometric techniques employed and of the increasing quality of the data sets utilized, researchers generally report results which reject UIP. In fact, for the major floating currencies against the U.S. dollar, the spot exchange rate has usually been recorded to fall when the forward market would have predicted it to rise and vice versa (e.g., Cumby and Obstfeld, 1984; Hodrick, 1987; Bekaert and Hodrick, 1993; Lewis, 1995; Engel, 1996; Sarno and Taylor, 2003; Sarno, 2005).

An alternative way of examining the properties of UIP is by investigating whether UIP deviations—or, identically, foreign exchange excess returns—are predictable using the forward premium as a predictor variable. Under the hypothesis that UIP holds (risk-neutral market efficiency), excess returns must be unpredictable. This issue was investigated, for example, by Bilson (1981), Fama (1984) and Backus, Gregory, and Telmer (1993), who report evidence of predictability of excess returns on the basis of the lagged forward premium, inconsistent with UIP.

Attempts to explain the forward bias puzzle using models of risk premia have met with limited and mixed success, especially for plausible degrees of risk aversion (e.g., Frankel and Engel, 1984; Domowitz and Hakkio, 1985; Cumby, 1988; Mark, 1988; Engel, 1996). Moreover, it is difficult to explain the rejection of UIP and the forward bias puzzle by recourse either to explanations such as learning, peso problems, and bubbles (e.g., Lewis, 1995); or by recourse to consumption-based asset pricing theories which allow for departures from time-additive preferences (Backus, Gregory, and Telmer, 1993; Bansal and others, 1995; Bekaert, 1996) and from expected utility (Bekaert, Hodrick, and Marshall, 1997); or else by using popular models of the term structure of interest rates adapted to a multi-currency setting (Backus, Foresi, and Telmer, 2001). Hence, even with the benefit of 20 years of hindsight, the forward bias puzzle has not been convincingly explained and continues to baffle the international finance profession.

In this paper, we start from noting that prior empirical research in this area has generally relied on linear frameworks in analyzing the properties of UIP deviations. This is surprising because several authors have argued that the relationship between expected exchange rates and interest rate differentials may be nonlinear for a variety of reasons, including transactions costs (see, inter alia, Baldwin, 1990; Dumas, 1992; Hollifield and Uppal, 1997; Sercu and Wu, 2000), central bank intervention (e.g., Mark and Moh, 2002; Moh, 2002), and the existence of limits to speculation (e.g., Lyons, 2001, pp. 206-20). In particular, the limits to speculation hypothesis is based on the idea that financial institutions only take up a currency trading strategy if this strategy is expected to yield an excess return per unit of risk (or a Sharpe ratio) that is higher than the one implied by alternative trading strategies, such as, for example, a simple buy-and-hold equity strategy. This argument effectively defines a band of inaction where the forward bias does not attract speculative capital and, therefore, does not imply any glaring profitable opportunity and will persist until it generates Sharpe ratios that are large enough to attract speculative capital away from alternative trading strategies (Lyons, 2001).

Although the literature has already documented that forward premia may affect future exchange rates in a nonlinear fashion (e.g., Bilson, 1981; Flood and Rose, 1996; Flood and Taylor, 1996; Bansal, 1997; Huisman and others, 1998; Bansal and Dahlquist, 2000; Clarida and others, 2003, 2006; Sercu and Vandebroek, 2005), the potential importance of nonlinearities to shed light on the forward bias puzzle remains largely under-researched. The present paper fills this gap. Our empirical framework provides a characterization of the UIP condition which allows us to test some of the general predictions of the limits to speculation hypothesis and to assess its potential to explain the forward bias puzzle and the excess returns predictability documented in the literature.

Our empirical results, obtained using five major U.S. dollar exchange rates and considering forward rates with 1- and 3-month maturity, are as follows. First, there is strong evidence that the relationship between spot and forward exchange rates is characterized by significant nonlinearities. While the detection of nonlinearities in this context is not novel per se, our empirical model proves especially useful for understanding the properties of deviations from UIP. In particular, consistent with the limits-to-speculation hypothesis which we use to motivate our nonlinear spot-forward regression, we find that, when Sharpe ratios are small, departures from market efficiency and hence the forward bias are statistically significant and persistent but economically too small to attract speculative capital, while when Sharpe ratios are large enough to attract speculative capital, the spot-forward relationship reverts rapidly towards the UIP condition.

Second, in a battery of Monte Carlo experiments we demonstrate that, if the true data generating process (DGP) governing the relationship between spot and forward exchange rates were of the nonlinear form we consider in this paper, we can replicate the empirical results generally reported in the literature. In particular, estimation of the conventional linear spot-forward regressions would lead us to reject both the validity of UIP and the hypothesis of no predictability of foreign exchange excess returns with parameter estimates that are close to the ones observed using actual data. However, the failure of UIP and the findings of a forward bias and predictability of excess returns are features that the DGP has only when expected deviations from UIP are tiny enough to be economically unimportant and unlikely to attract speculative capital.

Our interpretation of the empirical and Monte Carlo evidence in this paper is that the stylized fact that the UIP condition is statistically rejected by the data is not indicative of substantial market inefficiencies. Indeed, the inefficiencies implied by this rejection appear to be tiny, and it is not clear, on the basis of the evidence in this paper, that they are economically important.

The rest of the paper is organized as follows. Section II provides an outline of the theoretical background and introduces the limits-to-speculation hypothesis. Section III describes the empirical framework used to analyze the relationship between spot and forward exchange rates. In Section

IV we report and discuss the empirical results, while Section V provides the results of Monte Carlo simulations. Section VI concludes. Technical details on the nonlinear model employed and some robustness results are provided in Appendices I and II respectively.

II. UIP and the Forward Bias Puzzle: A Nonlinear Perspective

A. The Forward Bias Puzzle

In an efficient speculative market, prices should fully reflect information available to market participants and it should be impossible for a trader to earn excess returns to speculation. Under foreign exchange market efficiency and risk neutrality, UIP holds:

$$\Delta_k s^e_{t+k} = i_{t,k} - i^*_{t,k} \tag{1}$$

where s_t denotes the logarithm of the spot exchange rate (domestic price of foreign currency) at time t; $i_{t,k}$ and $i^*_{t,k}$ are the nominal interest rates available on similar domestic and foreign securities, respectively (with k periods to maturity); $\Delta_k s_{t+k} \equiv s_{t+k} - s_t$; and the superscript e denotes the market expectation based on information at time t. Testing UIP in its form as given by Equation (1) is tantamount to testing the joint hypothesis that foreign exchange market participants are, in an aggregate sense, i) endowed with rational expectations and ii) risk neutral.

Most often, however, empirical analyses of UIP have taken place in the context of the relationship between spot and forward exchange rates under the assumption that covered interest parity (CIP) holds: $f^k_t - s_t = i_{t,k} - i^*_{t,k}$, where f^k_t is the logarithm of the k-period forward rate (i.e., the rate agreed now for an exchange of currencies k periods ahead). Indeed, CIP is a reasonably mild assumption, given the extensive empirical evidence suggesting that CIP holds (Frenkel and Levich, 1975, 1977; Taylor, 1987; for a survey of this evidence, see, e.g., Sarno and Taylor, 2003, Ch. 2). Note that, unlike CIP, UIP is not an arbitrage condition because one of the terms in the UIP equation, namely the exchange rate at time $t+k$, is unknown at time t and, therefore, non-zero deviations from UIP ex ante do not necessarily imply the existence of arbitrage profits ex post due to the foreign exchange risk associated with future exchange rate movements.

Using CIP and replacing the interest rate differential $i_{t,k} - i^*_{t,k}$ with the forward premium (or forward discount) $f^k_t - s_t$, a number of researchers have tested UIP by estimating a regression of the form:

$$\Delta s_{t+1} = \alpha + \beta \left(f^1_t - s_t\right) + v_{t+1}, \tag{2}$$

where we have assumed $k = 1$ for simplicity, and v_{t+1} is a disturbance term. Under UIP, $\alpha = 0$, the slope parameter $\beta = 1$, and the disturbance term v_{t+1} (the rational expectations forecast error under the null hypothesis) must be uncorrelated with information available at time t (e.g. Fama, 1984).

Empirical studies based on the estimation of Equation (2) for a large variety of currencies and time periods, generally report results which reject UIP (e.g., see the references in the survey of Hodrick, 1987; Lewis, 1995; Taylor, 1995; Engel, 1996; Sarno, 2005). Indeed, it constitutes a stylized fact that estimates of β using exchange rates against the U.S. dollar are often statistically insignificantly different from zero and generally closer to minus unity than to plus unity (Froot and Thaler, 1990). The stylized fact of a negative β coefficient in this regression implies that the more the foreign currency is at a premium in the forward market, the less the home currency is predicted to depreciate.[1] The negative value of β is the central feature of the forward bias puzzle, and, following much previous literature, we refer to Equation (2) as the "Fama regression."[2] [3]

The literature has also investigated the predictability of UIP deviations (or foreign exchange excess returns) using the forward premium as a predictor variable in a linear model obtained from reparameterizing Equation (2) as follows:

$$ER_{t+1} = \alpha + \underbrace{(\beta - 1)}_{\beta^{\tau}} \left(f_t^1 - s_t\right) + v_{t+1}. \tag{3}$$

where the excess returns $ER_{t+1} \equiv \Delta s_{t+1} - \left(f_t^1 - s_t\right) \equiv s_{t+1} - f_t^1$. This regression was investigated, for example, by Bilson (1981), Fama (1984), and Backus, Gregory, and Telmer (1993) and was shown to generate strong predictability of excess returns (deviations from UIP) on the basis of the lagged forward premium. Specifically, while β^{τ} should be zero under UIP, the evidence, consistent with a negative estimate of β in Equation (2), is that β^{τ} is negative and statistically significantly different from zero. Clearly, given that Equation (3) is obtained simply from reparameterizing the Fama regression (2), the forward bias puzzle arising from Equation (2) and the predictability of excess returns documented on the basis of Equation (3) must be linked, and any explanation of the forward bias puzzle ($\beta \neq 1$) ought to be able to explain also the finding of a non-zero value of β^{τ} in Equation (3). We shall return to the link between the forward bias puzzle and the predictability of excess returns in Section V, where we will show that both these stylized facts can indeed be matched using a nonlinear model of regime-dependent UIP deviations which is rationalized by the existence of limits to foreign exchange speculation.

It is also worth noting that the framework described in this section (as well as our nonlinear framework described later) rely on the notion that spot and forward exchange rates are cointegrated and that the forward premium should, therefore, be stationary, in a similar fashion to Brenner and Kroner (1995) and Clarida and others (2003). Furthermore, another finding of the relevant literature which should be noted here and will be recalled later as appropriate is the evidence that estimates of β display significant time variation across different sample periods over the recent float.

[1] Equivalently, via the covered interest arbitrage condition, these findings indicate that the more domestic interest rates exceed foreign interest rates, the more the domestic currency tends on average to appreciate over the holding period, not to depreciate so as to offset on average the interest differential in favor of the home currency.

[2] Exceptions include Bansal and Dahlquist (2000), who document that the forward bias is largely confined to developed economies and to countries for which the U.S. interest rate exceeds foreign interest rates; Bekaert and Hodrick (2001), who, paying particular attention to small-sample distortions of tests applied to UIP and expectations hypotheses tests, provide a "partial rehabilitation" of UIP; and Flood and Rose (2002), who report that the failure of UIP is less severe during the 1990s and for countries which have faced currency crises over the sample period investigated.

[3] Note that the vast majority of studies in this context estimate the Fama regression using ordinary least squares (OLS). This can be problematic in the presence of an omitted risk premium in the regression, in which case OLS would yield biased and inconsistent estimates of β due to a simultaneity problem (Fama, 1984; Liu and Maddala, 1992). Recently, Barnhart, McKnown, and Wallace (1999, 2002) have shown that two conditions are needed for the simultaneity problem to arise: (i) the forward rate must be a function of an unobservable omitted variable, such as predictable excess returns; (ii) the term containing the forward rate in the estimated regression is stationary or, if nonstationary, can be normalized to a stationary variable. Under these conditions, Barnhart, McKnown, and Wallace (1999, 2002) document the severity of the simultaneity problem in a variety of spot-forward regressions.

The range of estimates of β over time is large and includes both negative and positive values (e.g., Baillie and Bollerslev, 2000).

B. The Rationale for Nonlinearity in the Fama Regression

The idea that there may be nonlinearities in the spot-forward relationship is not novel. For example, the work of Dumas (1992) on general equilibrium models of exchange rate determination in a spatially separated world with international trade costs generated a variety of exchange rate equations where nominal exchange rates are shown to depend nonlinearly on their fundamentals in a way that reversion towards international parity conditions is a function of the size of the deviation from the parity conditions themselves—e.g., see Dumas (1992, p. 174, Equation 23); see also Baldwin (1990), Hollifield and Uppal (1997) and Obstfeld and Rogoff (2000). Sercu and Wu (2000) derive, in a partial equilibrium model, an expression for the spot-forward relationship where, in the presence of transaction costs, expected exchange rate changes and forward premia are imperfectly aligned, inducing nonlinearity in the spot-forward relationship. Mark and Moh (2002) and Moh (2002) study continuous-time models where UIP is a stochastic differential equation which has a solution where the exchange rate is a nonlinear function of the interest differential, modelled according to a jump-diffusion process regulated by occasional central bank intervention. This model records some success in matching some of the moments in the data and is capable of shedding some light on the forward bias puzzle when central bank intervention is not announced and takes the market by surprise.

A related, albeit different, rationalization of nonlinearity in the spot-forward relationship stems from the limits-to-speculation hypothesis. A rich account of the implications of limits to speculation for market efficiency tests and the nonlinear behavior of deviations from UIP is provided by Lyons (2001, Ch. 7, pp. 209-20). The line of reasoning is that financial institutions will only take up a currency trading strategy if the strategy yields a Sharpe ratio at least equal to an alternative investment strategy, say a buy-and-hold equity strategy. As it is well known, the Sharpe ratio is commonly defined as $(E[R_s] - R_f)/\sigma_s$, where $E[R_s]$ is the expected return on the strategy, R_f is the risk-free interest rate, and σ_s is the standard deviation of the returns to the strategy. In essence, the Sharpe ratio may be seen as the expected excess return from speculation per unit of risk. Given that the realized Sharpe ratio for a buy-and-hold equity strategy has averaged about 0.4 on an annual basis for the United States over the last 50 years or so,[4] a buy-and-hold currency trading strategy yielding a Sharpe ratio lower than 0.4 would not be worth taking up. Although the specific example in Lyons (2001) is based on comparisons of unconditional buy-and-hold Sharpe ratios (static strategies) with currency and equity strategies, the same logic is applicable to comparisons between two (or more) dynamic, conditional strategies, where one would examine conditional Sharpe ratios at a point in time and engage in market-timing activities.

Noting that under the null hypothesis that UIP holds (i.e., foreign exchange market efficiency), $\alpha = 0$ and $\beta = 1$ in Equation (2) and the Sharpe ratio of currency strategies is zero, then it is

[4] The average excess return (the numerator of the Sharpe ratio) is about 0.7 and the annualized standard deviation of returns (the denominator) equals about 0.17 (see Lyons, 2001, p. 210). The figure of 0.4 is also reported by Sharpe (1994, p. 51).

only when β departs from unity that the numerator of the Sharpe ratio takes nonzero values.[5] Indeed, according to Lyons's calculations, it is only when $\beta \leq -1$ or $\beta \geq 3$ that the Sharpe ratio for currency strategies is about the same as the average from a buy-and-hold equity strategy, i.e., 0.4 (see Lyons, 2001, p. 210). This argument effectively defines a band of inaction for the forward bias such that if $-1 < \beta < 3$ financial institutions would have no incentive to take up the currency strategy since a buy-and-hold equity strategy would have a higher return per unit of risk; within this band of inaction the forward bias and expected deviations from UIP are too small to attract speculative capital and, therefore, do not imply any glaring profit opportunity.

In its essence, the limits-to-speculation argument implies that, within a certain band of β (and, consequently of the Sharpe ratio), the forward bias does not attract capital and hence may potentially persist for a long time. In some sense, this argument suggests that limits to speculation and the existence of an opportunity cost of speculative capital create a band for the deviations from UIP where the marginal cost of taking up a currency strategy exceeds the marginal benefit.[6] The crucial implication of the above analysis is that when limits to speculation of the kind described by Lyons (2001) are taken into account, the spot exchange rate and forward exchange rate need not move together at all times and, indeed, they may even move in opposite directions within a bounded interval without giving rise to any glaring profit opportunities. This argument is also consistent with the evidence that β (and hence the forward bias) is in fact highly time-varying, inducing parameter instability in standard linear regression models which assume a constant value of β over time (Baillie and Bollerslev, 2000). Arguments of this sort may be used to motivate the adoption of threshold-type models of the type originally proposed by Tong (1990) to empirically characterize the spot-forward relationship or the behavior of deviations from UIP: these threshold models would allow for a band within which β may differ from unity and may be positive, zero, or even negative, while outside the band the process switches abruptly to become exactly consistent with UIP and $\beta = 1$. Strictly speaking, assuming instantaneous allocation of speculative capital to currencies at the edges of the band of inaction then implies that the thresholds become reflecting barriers.

Nevertheless, while threshold-type models are appealing in this context, various arguments can be made to rationalize multiple-threshold or smooth, rather than single-threshold or discrete, nonlinear adjustment in deviations from UIP. First, the thresholds may be interpreted more broadly to reflect the opportunity cost of speculative capital, proportional transactions costs, and the tendency of traders or financial institutions to wait for sufficiently large Sharpe ratios before entering the market and trading (see, for example, Sofianos, 1993, Neal, 1996, and Dumas, 1992).

Second, one may argue that the assumption of instantaneous trade at the edges of the band of inaction should be replaced with the presumption that it takes some time to observe a profitable trading opportunity and execute transactions and that trade is infrequent (Dumas, 1992) and characterized by "limited participation" due to the fact that information costs may limit the participation of some

[5] The numerator is just the expected foreign exchange excess return, $E[\Delta s_{t+1} - (f_t - s_t)]$, where $(f_t - s_t)$ represents a position in foreign exchange fully hedged in the forward market, essentially taking up the equivalent role of the risk-free rate in the context of Sharpe ratios for equity strategies. The denominator is determined by the exchange rate variances and, in the case of multiple exchange rate strategies, also the covariances among the exchange rates considered in the currency strategy.

[6] The limits to speculation hypothesis proposed by Lyons (2001) is also inspired by the limits to arbitrage theory of Shleifer and Vishny (1997). Shleifer and Vishny's model allows for agency frictions in professional money management to lead to less aggressive trading than in a frictionless world, so that only limited speculative capital is allocated to the trading opportunities with the highest Sharpe ratio.

classes of traders in derivatives markets (Grossman and Weiss, 1983; Hirshleifer, 1988). Essentially, limited participation models assume that agents adjust their portfolios infrequently, with a different subset of agents adjusting in each period. Limited participation in the foreign exchange market by nonfinancial corporations and unleveraged investors[7] implies that their portfolio shifts will be gradual, rather than abrupt (Lyons, 2001, p. 218).

Third, in a market with heterogeneous agents who face different levels of position limits, agents essentially face bands of different sizes. For relatively small deviations of β from the edges of the band of inaction, only some traders or institutions will be able or willing to effect trades. As deviations from the edges of the band get larger, then progressively more agents will enter the market to effect trades. Thus, the forces pushing β toward the band of inaction will increase as the deviations from the edges of the band increase since an increasing number of agents face profitable opportunities. Consequently, we should observe a smooth transition toward UIP, with the speed of reversion of the deviations from UIP toward zero increasing with the degree of violation of the band of inaction itself (see Dumas, 1992).[8]

Overall, the arguments discussed above suggest that limits to speculation create a band within which UIP does not hold and where spot and forward rates may be unrelated or even move in opposite directions; further, deviations from UIP can stray beyond the values implied by the edges of the band. Once Sharpe ratios are large enough to attract speculative capital, deviations from UIP become increasingly mean reverting with the size of the Sharpe ratios. Under certain restrictive conditions (including, inter alia, identical limits to speculation and position limits, and homogeneity of agents), the reversion to UIP ($\beta = 1$) may be discrete, but in general it is smooth, and Teräsvirta (1994) and Granger and Lee (1999) suggest that even in the former case, time aggregation will tend to smooth the transition between regimes. Hence, smooth rather than discrete adjustment may be appropriate in the present context, and time aggregation and nonsynchronous adjustment by heterogeneous agents are likely to result in smooth aggregate regime switching. This is indeed the kind of behavior we shall try to capture in our empirical framework, as discussed in the next section. Also, note that this behavior is consistent with the evidence that spot and forward exchange rates are cointegrated (Brenner and Kroner, 1995; Clarida and others, 2003) since the wedge between exchange rate movements and the forward premium cannot grow to infinity without agents regarding departures from UIP as economically important and exploiting them. It is this speculative process that binds together spot and forward rates in the long run, in the same spirit as the general arguments in Brenner and Kroner (1995). This modeling strategy is also consistent with the evidence of parameter instability reported in Baillie and Bollerslev (2000), providing a rationale for the result of instability of the parameter β over time: under the working hypothesis of nonlinearity induced by limits to speculation, β is a function of time-varying Sharpe ratios and, therefore, is expected to vary over time.

[7]For example, investors such as mutual funds, pension funds, or insurance companies, who do not have a comparative advantage over proprietary bank traders in implementing pure currency strategies to exploit the forward bias.

[8]In other words, one may be tempted to argue that, once a glaring profit opportunity is expected, the investor will invest as much as possible to exploit it. However, this is obviously not the case in real-world financial markets due to, among other things, the existence of margin requirements and position limits. For example, Liu and Longstaff (2004) demonstrate that, as an effect of the existence of margin requirements, it is not optimal to take unlimited positions in arbitrage and it is often optimal to take smaller positions in arbitrage than margin constraints would allow.

III. A Nonlinear Fama Regression

A. The Empirical Framework

One characterization of nonlinear adjustment in the Fama regression which allows for smooth adjustment is in terms of a smooth transition regression (STR) model (Granger and Teräsvirta, 1993; Teräsvirta, 1994, 1998). An STR model of spot and forward rates may be written as follows:

$$\Delta s_{t+1} = \left[\alpha_1 + \beta_1 \left(f_t^1 - s_t\right)\right] + \left[\alpha_2 + \beta_2 \left(f_t^1 - s_t\right)\right] G\left[ER_t^e, \gamma\right] + \varepsilon_{t+1}, \tag{4}$$

where ε_{t+1} is a disturbance term. The transition function $G\left[ER_t^e, \gamma\right]$ determines the degree of reversion to zero of the deviations from UIP and is itself governed by the parameter γ, which effectively determines the speed of reversion to UIP, and the transition variable. The transition variable is assumed to be the expected excess return ER_t^e; specifically, ER_t^e is the deviation from UIP at time $t+1$ that is expected by market participants conditional on information available at time t. Therefore, all terms on the right-hand side of Equation (4), namely $f_t^1 - s_t$ and ER_t^e, are known at time t.

A simple transition function suggested by Granger and Teräsvirta (1993) and Teräsvirta (1994, 1998) is the exponential function:

$$\begin{aligned} G\left[ER_t^e, \gamma\right] &= \left\{1 - H\left[ER_t^e, \gamma\right]\right\} \\ &= \left\{1 - \exp\left[-\gamma\left(ER_t^e\right)^2\right]\right\}, \end{aligned} \tag{5}$$

in which case (4) would be termed an exponential STR or ESTR model. The exponential transition function G is bounded between zero and unity, i.e., $G : \Re \rightarrow [0,1]$, has the properties $G[0] = 0$ and $\lim_{x \to \pm\infty} G[x] = 1$, and is symmetrically inverse-bell shaped around zero. These properties of the ESTR model are particularly attractive in the present context because they allow a smooth transition between regimes and symmetric adjustment of the deviations from UIP above and below the equilibrium level, consistent with the limits to speculation hypothesis. Note also that, given the definition in Equation (5), which relates G to H linearly, $H : \Re \rightarrow [0,1]$, has the properties $H[0] = 1$ and $\lim_{x \to \pm\infty} H[x] = 0$, and is symmetrically bell shaped around zero. Obviously, $G = 0$ when $H = 1$, and $G = 1$ when $H = 0$.

The transition parameter γ determines the speed of transition between the two extreme regimes, with lower absolute values of γ implying slower transitions. The "lower" and "upper" regimes are defined as the regimes corresponding to the two extreme values of the transition function, where $G(\cdot) = 0$ and $G(\cdot) = 1$, respectively. Investigation of the properties of the model in these two extreme regimes sheds light on the stability and dynamic properties of the STR model. The arguments in the spirit of limits to speculation suggest the restrictions $\alpha_2 = -\alpha_1$ and $\beta_2 = 1 - \beta_1$. Under these restrictions (which we test formally in our empirical work), the lower regime corresponds to $ER_t^e = 0$, where $G(\cdot) = 0$ and Equation (4) becomes a standard linear Fama regression of the form:

$$\Delta s_{t+1} = \left[\alpha_1 + \beta_1 \left(f_t^1 - s_t\right)\right] + \varepsilon_{t+1}. \tag{6}$$

The upper regime corresponds, for a given γ, to $\lim_{ER_t^e \to \pm\infty} G\left[ER_t^e, \gamma\right]$, where (4) becomes a different linear Fama regression with parameters exactly consistent with UIP:

$$\begin{aligned}\Delta s_{t+1} &= \alpha_1 + \alpha_2 + (\beta_1 + \beta_2)\left(f_t^1 - s_t\right) + \varepsilon_{t+1} \\ &= \left(f_t^1 - s_t\right) + \varepsilon_{t+1}. \end{aligned} \tag{7}$$

This formulation of the nonlinear Fama regression has several virtues. First, the model nests the standard linear Fama regression, to which it would collapse in the absence of nonlinearity. Second, under the restrictions $\alpha_2 = -\alpha_1$ and $\beta_2 = 1 - \beta_1$, which are formally testable using standard statistical inference, this specification captures the behavior of the deviations from UIP which is implied by the theoretical considerations discussed in Section II. Deviations from UIP may be persistent and consistent with the well-known forward bias when they are in the neighborhood of UIP, that is, when expected excess returns are too small to attract speculative capital. In this lower regime, UIP does not hold but departures from UIP are economically small. However, for larger expected deviations from UIP (of either sign) that move exchange rates toward the upper regime, financial institutions would take up the glaring profit opportunities provided by currency trading strategies and induce reversion towards the UIP condition.

In short, the model allows for departures from UIP at all points where the transition function $G(\cdot) = 1 - H(\cdot) \neq 1$. There is only one value of the transition function and one set of parameter restrictions that allow us to conclude that UIP holds (rational expectations and risk neutrality are both valid). In essence, the model may be seen as a characterization of the relation between spot and forward rates that allows for departures from UIP, while yielding UIP under specific restrictions. The conditions $\alpha_2 = -\alpha_1$ and $\beta_2 = 1 - \beta_1$ guarantee global stability, as exchange rate changes are moving in response to the lagged forward premium in a way to restore UIP under those conditions. Put another way, the model is consistent with the evidence of cointegration between spot and forward exchange rates, because the long-run tendency of the model is the UIP condition, where $\Delta s_{t+1} - \left(f_t^1 - s_t\right) = 0$. Because Δs_{t+1} is stationary, and each of f_t^1 and s_t is nonstationary, the only way in which $\Delta s_{t+1} - \left(f_t^1 - s_t\right)$ can be stationary is if f_t^1 and s_t co-move in the long run and cointegrate with a long-run parameter of unity (e.g., Brenner and Kroner, 1995).

Note that the transition function, $G(\cdot)$ is defined as a function of the conditional expected excess return, ER_t^e. Strictly speaking, the arguments presented in Section II.B are in terms of (conditional or unconditional) Sharpe ratios, which would require us to use the expected excess return divided by the standard deviation of the excess return. We use survey data on exchange rate expectations to construct a proxy for the conditional expected excess return. As described in the data section below, survey data to proxy a conditional standard deviation are not available. Consequently, in the empirical work, we standardize the excess return by its unconditional standard deviation (see Section IV.B). This proxy is, however, imperfect in that the Sharpe ratio is calculated using the conditional expected excess return in the numerator and the unconditional standard deviation in the denominator. We shall test for robustness of our core results using an alternative proxy for the risk factor in the denominator of the Sharpe ratio, namely, the implied volatility from currency options.

We note that, if the true DGP of the spot-forward relationship is indeed nonlinear of the form (4), then β, as given in Equation (2), will lie in the interval between β_1 and $(\beta_1 + \beta_2) = 1$. It seems plausible that if the distribution of UIP deviations is consistent with the majority of observations being in the lower regime (where β may be negative and the forward bias is expected to be persistent), one may well find negative values of β from estimating the linear Fama regression (2). We shall investigate exactly this issue in Section V using Monte Carlo methods.

It is also instructive to reparameterize the nonlinear Fama regression (4) in terms of deviations from UIP by subtracting the forward premium, $(f_t^1 - s_t)$ from both sides of Equation (4) as follows:

$$ER_{t+1} = \left[\alpha_1 + \underbrace{(\beta_1 - 1)}_{\beta^*}(f_t^1 - s_t)\right] + \left[\alpha_2 + \beta_2 (f_t^1 - s_t)\right] G\left[ER_t^{'}, \gamma\right] + \varepsilon_{t+1}. \tag{8}$$

The discussion on the effects of limits to speculation in the previous section suggests that the larger the expected deviation from UIP, the larger the effect of speculative forces in generating reversion toward UIP. This implies that while $\beta^* < 0$ is admissible in Equation (8), one must have $\beta_2 > 0$ and $\beta^* + \beta_2 = 0$ for the forward premium to have no predictive power on future excess returns in the upper regime.

Note that Equation (8) may be seen as the nonlinear analogue of (and indeed nests) the predictability regression (3), exactly like Equation (4) is the nonlinear analogue of the Fama regression (2). Hence, Equation (8) also has implications for conventional tests of predictability of excess returns using the forward premium as a predictor variable based on a linear model obtained from reparameterizing the Fama regression (2). Clearly, if the true DGP of UIP deviations is indeed nonlinear of the form (4) (or (8)), then β^τ as given in Equation (3) will lie in the interval between β^* and $(\beta^* + \beta_2) = 0$. Whether β^τ is closer to β^* or to $(\beta^* + \beta_2)$ will depend on the distribution of UIP deviations, but it seems at least possible that if the distribution of UIP deviations is consistent with the majority of observations being in the lower regime one may find negative and statistically significant estimates of β^τ from estimating Equation (3). Again, we shall investigate this issue in Section V using Monte Carlo methods.

Granger and Teräsvirta (1993) and Teräsvirta (1994) also suggest the logistic function as a plausible transition function for some applications, resulting in a logistic STR (LSTR) model, which implies asymmetric behavior of the deviations from UIP according to whether they are positive or negative, which could potentially arise in the context of this paper due to, for example, short-sale restrictions. Hence, we do test for nonlinearities arising from the LSTR formulation as a test of specification of the estimated models in the section discussing the empirical analysis. Also, as a preliminary to our estimation of a nonlinear Fama regression, we evaluate the adequateness of the linear Fama regression by performing tests of linearity against nonlinearity of smooth-transition type (for both ESTR and LSTR formulations) and by testing the hypothesis of symmetry directly.

However, we wish to emphasize that, while our empirical analysis is inspired by the limits to speculation hypothesis, we do not claim to provide a direct test of this specific hypothesis, but rather a test of its general predictions for the relationship between spot and forward exchange rates. Our approach is best interpreted as an empirical characterization of the spot-forward relationship motivated by the limits to speculation hypothesis or simply as an empirical investigation of parsimonious models of foreign exchange excess returns. In particular, although we have focused on a specific nonlinear formulation of the relationship between exchange rates and forward premia capable of capturing some of the key predictions of the limits to speculation hypothesis, we cannot discard the possibility that the source of nonlinearities we document below is a factor other than limits to speculation.

B. The Solution of the Model When the Market Expectation Equals the Mathematical Expectation

In this subsection, we further clarify the properties of the model and the treatment of market (subjective) expectations and mathematical expectations. Taking expectations of Equation (8) and defining the mathematical expectation conditional on information at time t as $E_t(y_{t+1}) = y_t^m$ yields

$$ER_t^m = \left[\alpha_1 + (\beta_1 - 1)\left(f_t^1 - s_t\right)\right] + \left[\alpha_2 + \beta_2\left(f_t^1 - s_t\right)\right] G\left[ER_t^e, \gamma\right]. \quad (9)$$

Then, noting that the market expectation needs not be equal to the mathematical expectation, we can define

$$ER_t^e = ER_t^m + \eta_t \quad (10)$$

with $\eta_t = 0$ if rational expectations hold. Substituting (10) into (9) gives

$$ER_t^m = \left[\alpha_1 + (\beta_1 - 1)\left(f_t^1 - s_t\right)\right] + \left[\alpha_2 + \beta_2\left(f_t^1 - s_t\right)\right] G\left[ER_t^m + \eta_t, \gamma\right]. \quad (11)$$

We can now formally derive the conditions under which there is a unique solution for ER_t^e in this model. To this end, let us impose the conditions under which the theoretical implications of the limits to speculation hypothesis hold, i.e., $\alpha_2 = -\alpha_1$ and $\beta_2 = 1 - \beta_1$, and with expectations formed rationally, i.e., $ER_t^e = ER_t^m$ and $\eta_t = 0$. Equation (12) then becomes:

$$ER_t^m = \left[\alpha_1 + (\beta_1 - 1)\left(f_t^1 - s_t\right)\right] H\left[ER_t^m, \gamma\right] \quad (12)$$

where $H\left[ER_t^m, \gamma\right] = 1 - G\left[ER_t^m, \gamma\right] = \exp\left[-\gamma\left(ER_t^m\right)^2\right]$.

Building on the results of Peel and Venetis (2005), the solution with respect to ER_t^m is as follows:

$$ER_t^m = \exp\left\{-0.5W\left[2\gamma\left[\alpha_1 + (\beta_1 - 1)\left(f_t^1 - s_t\right)\right]^2\right]\right\}\left[\alpha_1 + (\beta_1 - 1)\left(f_t^1 - s_t\right)\right] \quad (13)$$

where $W(x)$ is the Lambert W function (also called Omega function), defined implicitly as the solution of $W(x)e^{W(x)} = x$; see Appendix I for a derivation of this equation. The solution captures the expectations-consistent process and embodies the notion that expectations are a monotonic function of the current forward premium, $\left(f_t^1 - s_t\right)$. The expectations-consistent process implies that adjustment to UIP is faster the greater the expected deviation from UIP; essentially, on the basis of information on the forward premium, $\left(f_t^1 - s_t\right)$ as well as the parameters α_1 and $(\beta_1 - 1)$, agents form expectations of UIP deviations next period, which may be seen as underlying a carry-trade type strategy of the kind commonly followed by foreign exchange speculators. Speculative forces would therefore be increasing in the current forward bias, related to the expected deviation from UIP according to Equation (13), which is the underlying rationale of the arguments described in Section II.B—see also the discussion in Peel and Venetis (2005) on arbitrage and expectations consistency in the context of arbitrage in international goods markets and purchasing power parity.

The Lambert W function does not admit a closed-form expression, and so it would have to be approximated in empirical modeling if one wants to take to the data Equation (13) (see Corless and others, 1996, and the references therein). However, this is not a problem in the context of this paper because in our empirical model we employ survey data on exchange rate expectations to proxy the transition variable, ER_t^e. As discussed later, these expectations data are well known to be correlated with the forward premium, a fact that is consistent with the above theoretical result that expectations of UIP deviations are a monotonic function of current departures from UIP.

IV. Empirical Results

A. Data, Summary Statistics, and the Fama Regression

Our data set comprises weekly observations of spot and 4- and 13-week (or one- and three-month) forward U.S. dollar exchange rates against the Japanese yen, the U.K. sterling, the German mark, the euro, and the Swiss franc. Due to data availability considerations and the advent of the euro on January 1, 1999, the sample period spans from January 4, 1985 to December 31, 2002, for all exchange rates except for the German mark (January 7, 1986, to December 31, 1998) and the euro (January 5, 1999, to December 31, 2002). Following previous literature (e.g., Hansen and Hodrick, 1980, p. 852), data are taken on Tuesdays of every week, from *Datastream.* To keep the notation simple, a four-week change in a variable is stated as a change from t to $t+1$, and f_t^j is the forward rate for a contract with j months (or j months to maturity). From this data set, we constructed the time series of interest, namely the logarithm of the spot exchange rate, s_t and the logarithm of the one- and three-month forward exchange rates, f_t^1 and f_t^3, respectively, both at the weekly and monthly frequency. The core of the empirical work is based on s_t and f_t^1 at the weekly frequency, whereas we use the weekly f_t^3 as well as monthly data for s_t, f_t^1, and f_t^3 in our robustness analysis.

In order to construct a proxy for the expected excess return, ER_t^e, we use survey data on exchange rate expectations from Money Market Services (MMS). MMS reports the median forecasts of survey respondents. The survey provides data on a weekly basis for both one-month and three-month expectations of the dollar exchange rates examined. Participants to the survey include exchange rate dealers, banking and corporate economists, as well as market economists. Because MMS reports the median forecasts, this measure masks individual heterogeneity. However, to analyze issues related to market efficiency of the kind studied in this paper, one has to resort to an aggregate expectation of the exchange rate and, therefore, the MMS measure seems appropriate. Unfortunately, MMS does not report a measure of dispersion of the MMS forecasts nor does it report data on forecasts of the exchange rate volatility. This prevents us from constructing the ideal Sharpe ratio, namely the conditional expected excess return divided by the conditional expected standard deviation. Essentially, MMS only provides us with the numerator of the Sharpe ratio, which we shall standardize by dividing it by the unconditional standard deviation in the core results and by the conditional implied volatility in our robustness checks.

The advantage of using the MMS forecasts is that the properties of these exchange rate expectations are well-documented in the literature (see the references in Takagi, 1991; Osterberg, 2000; Sarno and Taylor, 2001). Specifically, expected changes in exchange rates have been found to have a tendency to underpredict actual exchange rate movements, implying that part of the actual exchange rate changes are unexpected. The evidence also suggests that "unbiasedness"—the notion that survey measures are unbiased forecasts of actual future outcomes—is rejected in time series evidence but is not rejected in cross-sectional evidence (Chinn and Frankel, 2002). Another feature of some importance in relating survey measures to the rational expectations hypothesis is that MMS forecasts are not consistent with the "orthogonality" hypothesis, implying that survey data do not fully incorporate all available information, in contrast with the rational expectations paradigm (see Takagi, 1991; Osterberg, 2000). Finally, exchange rate expectations are correlated with the forward premium, suggesting that at least some of the variation in the forward premium is due to expected depreciation—or, in other words, the variation in the forward premium cannot be solely due to risk premia (Chinn and Frankel, 2002).[9]

In Table 1, we report sample moments for several combinations of weekly spot and one-month forward exchange rates, including the forward premium $f_t^1 - s_t$ (Panel A), the depreciation rate $s_{t+1} - s_t$ (Panel B), and the excess return $s_{t+1} - f_t^1$ (Panel C). The summary statistics confirm the stylized facts that each of $f_t^1 - s_t$, $s_{t+1} - s_t$ and $s_{t+1} - f_t^1$ have a mean close to zero with a large standard deviation. However, while the first-order autocorrelation coefficient of the depreciation rate is very small in size (never higher than 0.09) and generally statistically insignificantly different from zero, the first-order autocorrelation coefficient of the forward premium is generally large (in the range between 0.439 for Germany and 0.761 for the United Kingdom) and statistically significantly different from zero, and the corresponding first-order autocorrelation coefficient of the excess return is small (in the range between 0.053 for Germany and 0.131 for the euro) but often statistically significant. These results are consistent with the stylized facts that the forward premium is a highly persistent process, and the depreciation rate shows weak serial correlation (or the exchange rate is a near random walk process).[10]

As a preliminary exercise, we estimated the conventional Fama regression (2) for each exchange rate examined. The results, reported in Table 2, are consistent with the existence of forward bias in that, while the constant term α is very close to zero and often statistically insignificant, β is estimated to be negative for all but one of the exchange rate regressions estimated, and it is often statistically insignificantly different from zero. The only exception is Germany, where the estimate of β is positive (about 0.32) and statistically significant, but this estimate does not comprise the theoretical value of unity that is implied by UIP when examining the standard errors. In the last column of Table 2 we also report the t-statistics for the significance of the parameter associated with the forward premium—namely β^{τ}—in a predictability regression of the form (3). Consistent with a large literature (e.g., Fama, 1984; Backus, Gregory, and Telmer, 1993), we find that, for each exchange rate, β^{τ} is statistically significantly different from zero, indicating a departure from market efficiency (under which $\beta^{\tau} = 0$) and that the forward premium, which is an element of the market participants' information set, can be used to predict foreign exchange excess returns.

B. The Nonlinear Fama Regression

In order to evaluate the validity of the assumption of linearity in the conventional Fama regressions reported in Table 2, we performed tests of linearity against the alternative of smooth transition

[9] However, all of these properties are based on empirical evidence that originates from the assumption of linearity. If nonlinearity and regime dependence characterize the unknown true DGP linking exchange rate expectations to risk premia and forward premia, then this affects the reliability of standard measures of correlations as well as unbiasedness and orthogonality tests based on linear regressions (see Granger and Teräsvirta, 1993).

[10] Asymptotic standard errors were calculated using an autocorrelation and heteroskedasticity-consistent matrix of residuals throughout the paper (Newey and West, 1987). We also tested for unit root behavior of the spot rate and the forward rate time series examined by calculating several unit root test statistics. We were in each case unable to reject the unit root null hypothesis for spot and forward exchange rates, whereas the forward premium was found to be stationary at conventional nominal levels of significance. On the other hand, differencing the spot and forward exchange rate time series did appear to induce stationarity in each case. Hence, the unit root tests (not reported but available from the authors upon request) clearly indicate that each of the spot and forward rates time series examined is a realization from a stochastic process integrated of order one, whereas the forward premium is stationary, consistent with the evidence that spot and forward exchange rates cointegrate (Brenner and Kroner, 1995).

nonlinearity, using the expected annualized excess return, ER_t^e as the transition variable. We followed the Teräsvirta (1994, 1998) decision rule to select the most adequate transition function for modeling nonlinearity in the present context. This is a testing procedure designed to test the hypothesis of linearity and to select the most adequate nonlinear function between a logistic and an exponential function. As shown by the results in Table 3, the general linearity test F_L strongly rejects the null hypothesis of linearity. Employing the Teräsvirta rule to discriminate between exponential and logistic formulations led us to conclude that an exponential function (ESTR model) is the most adequate parametric formulation (given that F_2 yields the lowest p-value). This finding is consistent with our priors and the limits-to-speculation hypothesis, discussed in Section II.

Given the results from the linearity tests, we estimate the nonlinear Fama regression (4) by nonlinear least squares under the restrictions $\alpha_2 = -\alpha_1$ and $\beta_2 = 1 - \beta_1$ (which we test formally below). In estimation, we followed the recommendation of Granger and Teräsvirta (1993) and Teräsvirta (1994, 1998) of standardizing the transition variable ER_t^e by dividing it by the sample standard deviation of the transition variable, $\widehat{\sigma}_{ER^e}$, and using a starting value of unity for the estimation algorithm. Also, since this standardization applies to the transition variable, which now becomes $\left(ER_t^e\right)/\widehat{\sigma}_{ER^e}$, the transition variable has a natural interpretation in terms of annualized Sharpe ratio, which tightens the link between the empirical framework and the limits-to-speculation hypothesis.

The results, reported in Panel A of Table 4, indicate that the Fama regression is indeed highly nonlinear. The estimated transition parameter appears to be significantly different from zero, in each equation, both on the basis of the individual asymptotic standard errors as well as on the basis of the Skalin's (1998) parametric bootstrap likelihood ratio test (see the p-values in square brackets in the last column of Panel A of Table 4).[11] The estimates of the slope parameters β_1 and β_2 are correctly signed according to our priors based on the limits to speculation hypothesis, namely, we find a negative estimated value of β_1 and a large positive value of β_2 such that UIP holds exactly when the transition function $G(\cdot) = 1$. The only exception is Germany, where we record a positively signed estimate of β_1 equal to about 0.15, which seems reasonable given that in estimation of the linear Fama regression we found that Germany was the only country for which a statistically significant positive value of β was found.[12] In turn, these values imply that, for small Sharpe ratios (the transition variable), UIP does not hold and we observe a forward bias, while for increasingly large Sharpe ratios (which are likely to attract speculative capital), reversion to UIP can occur rapidly. These findings also imply that, since reversion to UIP occurs rapidly for large Sharpe ratios, the bulk of the observations of the deviations from UIP is in the lower regime, potentially generating substantial persistence in the forward bias, as predicted by the limits to speculation hypothesis. We shall return to the analysis of the distribution of deviations from UIP in Section V.

The estimated transition parameters also imply well-defined transition functions, as shown in Figure 1, which displays the plots of the estimated transition functions, $G(\cdot)$, against the annualized Sharpe ratio ER_t^e/σ_{ER} for each exchange rate. The speed of the transition functions is made clear by the

[11] Because the Skalin test of the null hypothesis that $\gamma = 0$ in the transition function may also be construed as a test of nonlinearity, these results confirm the presence of nonlinearity in the Fama regression for each exchange rate examined.

[12] The standard errors and test statistics in Table 4 are calculated using an autocorrelation and heteroskedasticity consistent matrix of residuals (Newey and West, 1987), since the overlap of the weekly sampling induces moving average terms in the residuals of order 3 for regressions based on a one-month forward rate—and 12 for regressions based on a three-month (13-week) forward rate, reported in the robustness section below.

evidence that the limiting case of $G(\cdot) = 1$ is attained for each exchange rate except the euro, which is impressive given that we are dealing with weekly data. The transition functions also confirm how most of the observations of the deviations from UIP are close to zero, i.e., in the lower regime.[13]

A battery of diagnostic tests is reported in Panel B of Table 4. We report a likelihood ratio test ($LR1$) for the null hypothesis of no asymmetric response of exchange rate changes to positive and negative values of the forward premium. This test, which was calculated using an asymmetric generalization of the nonlinear Fama regression reported in Panel A of Table 4, may be interpreted as a further test of the adequateness of the chosen nonlinear model and of its ability to account for the nonlinearity in the data. For each exchange rate, we are unable to reject the hypothesis of no asymmetry in the STR model, therefore justifying the estimation of an STR model that imposes symmetric adjustment, as given by Equations (4)-(5). Using a likelihood ratio ($LR2$) test for the null hypothesis that $\alpha_2 = -\alpha_1$ and $\beta_2 = 1 - \beta_1$, reported in the second column, we could not reject the validity of these restrictions at the 5 percent significance level. As discussed in Section II.B, these restrictions imply an equilibrium log-level in the model which is exactly consistent with deviations from UIP being equal to zero. For each of the estimated nonlinear Fama regressions, we then tested the null hypothesis of no remaining nonlinearity (F_{NRN}), constructed as in Eitrheim and Teräsvirta's (1996) and reported in the third column of Panel B. The null hypothesis of no remaining nonlinearity could not be rejected for any of the estimated models, indicating that our parsimonious generalization of the Fama regression appears to capture satisfactorily the nonlinearity in the spot-forward relationship.[14] We also tested for the stability of the model by constructing the appropriate test proposed by Eitrheim and Teräsvirta's (1996) for each nonlinear model. This is a test for the hypothesis of no structural break in the parameters (F_{NSB}) specifically designed for smooth transition models. The results, reported in the last column of Panel B, suggest no structural break in the parameters of the model, with p-values reasonably larger than the conventional 5 percent.

Overall, the nonlinear estimation results uncover strong evidence of nonlinearities in the relationship between spot and forward exchange rates, with UIP deviations adjusting toward their zero equilibrium level at a speed which depends upon the size of the Sharpe ratio. The estimated models are in each case consistent with the priors established by the limits to speculation arguments made in Section II.B. The bottom line is that our model is consistent with the forward bias characterizing only small departures from UIP. It is worth emphasizing, however, that this model does not imply that UIP holds all the time. On the contrary, given the persistence of the forward bias in the lower regime, UIP does *not* hold most of the time. The model implies that UIP does not hold when expected departures from UIP are economically small enough to be ignored by investors who are not willing or able to trade for such excess returns. If this is the case, then one may argue that the rejections of UIP routinely recorded in the literature are indeed primarily statistical, rather than economic, rejections of the theoretical link between exchange rates and interest rates (or forward rates). Before turning to a finer interpretation of our results, we discuss some robustness exercises.

[13] Note that the Sharpe ratios on the horizontal axis are positive on both sides of the midpoint, zero, to reflect the possibility of taking long or short positions in the foreign exchange market. Specifically, if the agent believes that the forward rate undervalues (overvalues) the dollar, that agent will take a long (short) position forward, resulting in positive excess returns.

[14] To shed further light on the properties of the model, one can employ the numerical 'skeleton' procedure typically suggested in this context (e.g. Teräsvirta and Anderson, 1992; Teräsvirta, 1998; Franses and van Dijk, 2000; see also De Grauwe and Grimaldi, 2005). Results on the implementation of this procedure are available upon request.

C. Robustness Checks

In this section we report several robustness checks carried out in order to evaluate the sensitivity of the empirical results reported in the previous sections. In particular, we assessed the robustness of our results to the choice of the maturity of the forward contract and to the choice of the frequency of the data. The results are reported in Appendix II. We re-estimated the nonlinear Fama regression (4) for each exchange rate examined using a three-month forward contract at the weekly frequency (see Panel A of Table B1) to assess the robustness to the choice of the forward contract maturity and then using a one-month forward contract at the monthly frequency (see Panel B of Table B1) to assess the robustness to the frequency of the data. The results reported in Table A1 show that, in each case, the estimates obtained are qualitatively similar to the results reported in Table 4 for a one-month forward contract at the weekly frequency. For the monthly data (Panel B of Table B1), the estimates of the slope parameters and the transition parameters are larger than for weekly data, as one would expect. However, the sign of the parameters is consistent with the results in Table 4, and their statistical significance and the evidence of nonlinearity are strong.

Another check of the robustness of our core results relates to the definition of the transition variable in the model. As noted earlier, the transition function depends on the conditional expected excess return, ER_t^e standardized by the unconditional standard deviation, σ_{ER}. One would ideally wish to have a conditional Sharpe ratio (conditional "expected" excess return divided by the conditional standard deviation of the excess return). Our proxy is, however, imperfect in that the Sharpe ratio is calculated with a conditional term on the numerator and an unconditional term on the denominator. This is because survey data to proxy a conditional standard deviation are not available. We test for robustness of our core results using an alternative proxy for the risk factor in the denominator of the conditional Sharpe ratio, namely, the implied volatility from currency options.[15] The resulting proxy for annualized Sharpe ratios is now conditional (both in terms of the numerator, the expected excess return, and the denominator, the square root of implied volatility). However, it is still imperfect in some sense because the numerator is an expectation coming from survey data, whereas the denominator is implied by market prices under the specific assumption that the Black-Scholes model applies to currency options. Nevertheless, this setting helps us to provide further evidence on the robustness of our core results.

Using this alternative proxy for the transition variable, we re-estimated the STR model for all five exchange rates examined and found that the restrictions implied by the limits to speculation hypothesis hold in each case. We also obtained convergence for each exchange rate model and parameter estimates that are similar to the ones reported in Table 4. In general, therefore, the use of this alternative, fully conditional proxy for the Sharpe ratio in the transition function of the STR model delivers results that are consistent with the core results in Table 4.[16]

We also addressed thoroughly the robustness of our linearity tests results. The main concern involves the possibility of a spurious rejection of the linearity hypothesis when the test statistics F_L, F_3, F_2, and F_1 are applied to the Fama regression (2) in finite samples. We addressed this issue by executing a battery of Monte Carlo experiments, constructed using $5,000$ replications in each

[15] Weekly data on implied volatility were obtained for the sample period from January 3, 1986, to December 31, 2002, from the Philadelphia Exchange (PHLX), the main currency options exchange in the US according to the 2004 Bank for International Settlements (BIS) *Survey on Foreign Exchange and Derivatives Market Activity*.

[16] These results are not reported to conserve space since they resemble closely the results in Table 4.

experiment and with identical random numbers across experiments. The aim of the experiments is to evaluate the empirical size and power properties of these tests and to gauge the extent to which one would reject the linear Fama regression when in fact that was the true DGP (empirical size) and the extent to which the tests would detect nonlinearity when in fact the true DGP is a nonlinear Fama regression (empirical power). In setting up the DGP for each of the linear and nonlinear Fama regressions, we calibrated the DGP on our results for the dollar-yen—exactly as reported in Table 2 for the linear Fama regression and in Table 4 for the nonlinear ESTR Fama regression. Given that our actual sample comprises 940 data points in total, we carried out the simulations for sample sizes of 470 (half of the actual sample) and 940 (the actual sample size) artificial data points. Our simulations results—reported in Table B2 for each of the 10, 5, and 1 percent significance levels—indicate satisfactory empirical size and power properties for each of the test statistics F_L, F_3, F_2, and F_1. In terms of empirical size, none of the test statistics displays evidence of substantial size distortion at any of the three significance levels considered. In terms of empirical power, the general linearity test F^L rejects about 73 (66) percent of the times with 940 (470) observations at the 5 percent significance level when the true DGP is nonlinear. This is not the theoretical level of 95 percent but it is high enough to judge the test as satisfactory. The test statistics F_3, F_2, and F_1 are less powerful than F_L but they appear to be satisfactory in discriminating between exponential (ESTR) and logistic (LSTR) specifications, as evidenced by the much higher power of F_2 (linearity versus ESTR) relative to F_3 and F_1 (linearity versus LSTR).[17]

The main result arising from these simulations for our purposes is that it seems unlikely, in light of the documented size and power properties, that we are detecting spurious nonlinearities in this paper since we find no tendency of the linearity tests employed to over-reject the null hypothesis of linearity when the true DGP is a linear Fama regression.

D. Interpreting the Empirical Results

Our empirical results provide clear evidence that the relationship between spot and forward exchange rates is characterized by important nonlinearities. While this result is not novel per se, we considered a nonlinear model which may be viewed as a generalization of the conventional Fama spot-forward regression and which therefore may be used to understand the properties of deviations from UIP. Our nonlinear spot-forward regression was rationalized on the basis of the argument that the existence of limits to foreign exchange speculation can allow deviations from UIP to be both persistent and consistent with the well known forward bias within a certain range (e.g. Lyons, 2001). According to the limits-to-speculation hypothesis, for small Sharpe ratios the forward bias does not attract speculative capital, which can be more profitably invested in alternative investment opportunities for the same level of risk. However, as Sharpe ratios become larger, agents take up positions in currency trading strategies, which induce the spot-forward relationship to revert exactly to UIP. Our nonlinear model parsimoniously captures this behavior and our estimation results uncover robust evidence that five major spot and forward dollar exchange rates have behaved in this fashion over the 1985–2002 sample period.

[17] Another check we carried out relates to the robustness of our choice of the exponential function over the sample. Specifically, we carried out the linearity tests in two different subsamples for the three exchange rates for which we have a longer full sample—Japanese yen, U.K. sterling, and Swiss franc. Splitting the sample in two subsamples at the end of 1993, we found that for each subsample the linearity tests confirm the choice of an exponential smooth transition model, corroborating our full sample results. These test results are available upon request.

One aspect of the rationale behind this model is, therefore, that financial institutions decide to allocate capital on the basis of Sharpe ratios and that this process induces the nonlinear dynamics we observe in the data. In Table 5, we report the average annualized Sharpe ratios (first column), for each exchange rate examined, calculated as the average realized standardized excess returns over the sample period. The Sharpe ratios range from 0.14 for the Japanese yen to 0.88 for U.K. sterling. It is interesting how the average Sharpe ratio across the five dollar exchange rates examined is 0.48, the value reported by Lyons (2001, p. 214) as the Sharpe ratio obtaining from employing a currency strategy (with equal weights) on the six most liquid currencies.

Given that the transition function we estimate is bounded between zero and unity and may be viewed as the probability of being in one of the two extreme regimes (one regime with persistent but tiny deviations from UIP, and another regime where UIP holds), it is instructive to graph the estimated transition functions over time. In Figure 2, we plot the estimated transition function for each exchange rate, over the sample. The plots make clear how the model implies that the spot-forward regressions (or the deviations from UIP) are in the lower regime (defined, for simplicity, as the case where $G(\cdot) \leq 0.5$) most of the time. The lower regime is the one characterized by a very persistent forward bias, which, however, is associated with low and economically unimportant Sharpe ratios. In some sense, therefore, these findings suggest that the forward bias does characterize the majority of the observations in the data, but only those observations where financial institutions' speculative capital is unlikely to be attracted by currency trading strategies because the size of the inefficiency is relatively low. On the other hand, although fewer observations are in the upper regime (say when $G(\cdot) > 0.5$), in this regime deviations from UIP are characterized by low persistence, suggesting that speculative forces induce fast reversion to UIP. Interestingly, therefore, rejections of UIP in a linear framework can be explained by the dominance of the observations for which UIP does not hold in the data (the lower regime), but our analysis reveals that these observations are characterized by small and economically unimportant departures from UIP. Put another way, the statistical rejections of UIP typically recorded in the relevant literature may indicate that exchange rates have on average been relatively close to UIP, rather than implying that UIP and foreign exchange market efficiency are strongly violated.[18], [19]

It is instructive to calculate the size of the annualized Sharpe ratio such that transition function $G(\cdot) = 0.5$, which we term the minimum Sharpe ratio (*min* SR) such that one may be in the upper regime.[20] The calculations are given, for each exchange rate, in the middle column of Table 5. Clearly, while 0.4 (the value that Lyons suggests conservatively as a minimum Sharpe ratio necessary to attract speculative capital) is not sufficient to induce the shift to the upper regime, the range of the minimum level required goes from about 1.14 for Germany to about 2.46 for the euro. Indeed, this evidence seems consistent with the argument made by Lyons (2001, p. 215), on the basis of interviews with several proprietary traders and desk managers, that restoration

[18] Note that the simplifying assumption that $G(\cdot) > 0.5$ is indicative of being in the upper regime is quite mild. In fact, as noted by a referee, given our estimates of γ in Table 4, deviations from UIP can still be quite substantial when $G(\cdot) = 0.5$. This means that assuming $G(\cdot) > 0.5$ (rather than requiring a value larger than 0.5) is sufficient to move to the upper regime will lead to a conservative estimate of the importance of limits to speculation. In other words, assuming that UIP deviations only become economically small for higher values of $G(\cdot)$ would further reinforce our point that the majority of the observations in the data are in the lower regime.

[19] Also, note that these transition functions imply that the parameter linking future exchange rate changes to the current forward premium, which depends on the transition function, displays a lot of variation over time, as one would expect under the limits-to-speculation hypothesis and as documented in other research focusing on the instability of the parameter β in the Fama regression (e.g., Baillie and Bollerslev, 2000).

[20] Given $\widehat{\gamma}$ reported in Table 4, this is simply the value of SR that solves $\left\{1 - \exp\left[-\gamma SR^2\right]\right\} = 0.5$.

of the UIP equilibrium condition through allocation of speculative capital is likely to require an extremely large amount of order flow and that these large amounts generally occur when traders or desk managers are facing Sharpe ratios of at least unity. These Sharpe ratios can hardly be achieved through bias trading in a one-exchange-rate setting of the type considered here, as illustrated in the last column of Table 5, which reports that some 68 to 97 percent of the observations are below the minimum Sharpe ratios we calculate. While it may be possible to achieve Sharpe ratios of this magnitude, this happens rarely in a one-exchange-rate setting. These Sharpe ratios are likely to require much more sophisticated multi-currency portfolio strategies based on currency overlay programs. If any Sharpe ratio below the minimum Sharpe ratio we calculate is consistent with the forward bias being too small to attract speculative capital, then one would expect that most of the time deviations from UIP are indeed characterized by forward bias.

An important caveat is, however, that, while our empirical analysis is inspired by the limits to speculation hypothesis, we do not claim to be able to provide a direct test of this specific hypothesis, but rather a test of its general predictions for the nonlinear linkages between spot and forward exchange rates. Our approach enables us to capture some of the key predictions of the limits to speculation hypothesis, but the nonlinearity recorded in our research may be due to factors other than or in addition to limits to speculation. It is possible to rationalize nonlinearities, for example, on the basis of arguments based on discrete intervention of central banks (e.g., Mark and Moh, 2002), general equilibrium models of exchange rate determination in segmented international capital markets and international trade costs (Dumas, 1992; Hollifield and Uppal, 1997; Obstfeld and Rogoff, 2000), and other standard transactions-costs arguments (e.g., Sercu and Wu, 2000). We leave to future research the design of a framework where researchers can formally discriminate among different theories or frictions that predict the existence of nonlinearities on spot-forward regressions.

E. Forecasting

The primary purpose of this paper is to improve our understanding of the behavior of UIP deviations and the relationship between spot and forward exchange rates and to test theoretical predictions that such a relationship should be nonlinear. However, the nonlinear empirical models examined can be used as predictive models for the returns from currency speculation to shed further light on their ability to explain exchange rate movements over time and their performance relative to the standard linear Fama regression. Given that the selected nonlinear models fit better than their linear counterparts, it may be possible to build a better trading strategy than is available using the standard Fama regression model.

We evaluate predictive performance in two ways. First, we report evidence using the "projection" statistic recently suggested by Evans and Lyons (2005). This test statistic has a known asymptotic distribution under the null hypothesis that the spot exchange rate follows a random walk. The results are obtained using recursive estimates of the linear and nonlinear models using a growing number of observations. The statistic is calculated by estimating the parameter δ in the regression: $\widehat{\Delta s_{t+1|t}} = \delta_0 + \delta \Delta s_{t+1} + \xi_{t+1}$, where the hat denotes the one-step-ahead (one-week-ahead) prediction from the (linear or nonlinear) Fama regression—see Evans and Lyons (2005) for further details on the properties of this test. We compute standard errors for δ using the Newey-West autocorrelation and heteroskedasticity consistent correction. Under the null hypothesis that s_t follows a random walk $\Delta s_{t+1|t} = 0$ and $\delta = 0$. The results, reported in Table 6, indicate that using the predictions from the linear Fama regression yields a correctly signed estimate of δ (positive). However, the estimates of δ are statistically insignificantly different from zero (except for the German mark). Using the predictions from the nonlinear Fama regression model gives larger estimates of δ, which are also statistically significantly different from zero for three out of five exchange rates. Bearing in mind that these are weekly predictions of exchange rate changes, which are well known to contain a large unforecastable component, the estimates of δ obtained using predictions from the nonlinear models compare well to the estimates reported in the work of Evans and Lyons (2005), even though the information used here only includes publicly available information.

Second, we calculate unconditional annualized Sharpe ratios corresponding to each $\overline{R}^2$ for both the linear and nonlinear Fama regressions. The annualized Sharpe ratios are calculated as: $SR = \sqrt{0.4^2 + \overline{R}^2}/\sqrt{1 - \overline{R}^2}$, following Cochrane (1999, pp. 65-66, 75-76) and Gallant, Hansen, and Tauchen (1990). The value chosen to proxy for the unconditional annualized Sharpe ratio of a U.S. buy-and-hold strategy is 0.4 (the first term in the square root in the numerator), following Sharpe (1994) and Lyons (2001). The derivation of this formula for the Sharpe ratio stems from the implementation of a simple market timing strategy based on a general predictability model (Cochrane, 1999, p. 65-66).[21] The results, reported in Table 7 show a clear difference between the linear and nonlinear Fama regression. While the linear regression delivers an $\overline{R}^2$ in the range from 0.2% to 7%, the nonlinear regression provides us with a $\overline{R}^2$ ranging from 3.7% to about 17%. These values of the $\overline{R}^2$ imply Sharpe ratios for the nonlinear model ranging from 0.45 to about 0.63. These Sharpe ratios appear respectable if one considers that the underlying strategy is a pure currency strategy only involving two currencies (one exchange rate). In practice, however, financial institutions typically engage in multicurrency forward-bias strategies involving more than two currencies. Hence, assuming that banks use more sophisticated strategies than the one discussed here, and hence might achieve higher Sharpe ratios than the ones we record on one-exchange-rate strategies, would further support our interpretation of the results in this section.

As noted by an anonymous referee, Deutsche Bank frequently reports Sharpe ratios close to 0.9 for a simple multicurrency strategy. In practice, speculative capital is often attracted by (relatively simple) forward bias trading strategies. Notably, Galati and Melvin (2004) show that simple carry trades aimed at exploiting the forward bias are one source of the surge in trading in the foreign exchange market observed in recent years. In light of our conversations with foreign exchange chief dealers, these facts may be explained by the fact that the decision to implement active currency strategies is, in practice, not predicated exclusively on the assessment of the Sharpe ratio but also on the basis of the correlation of returns to the active hedge with returns to the underlying portfolio, as well as the historical track record of active currency managers in generating positive total returns. Therefore, the Sharpe ratio of, say, a forward bias strategy may not be appealing over the long term, but if the return streams from the two strategies are uncorrelated, then the performance of the portfolio is improved as a result and capital may be attracted by Sharpe ratios lower than the minimum Sharpe ratios reported in Table 5. The relative performance of the two parts of the portfolio strategy may be important in the determination of the relative risk budget that investors will allocate to each part (e.g., how much active currency management does a pension fund want on top of its equity portfolio). This suggests further ways in which the approach taken

[21] Cochrane (1999) shows how the unconditional Sharpe ratio is directly related to the adjusted R^2 in the regression model used.

in this paper could be improved in future research by explicitly modelling returns in equity, bond and foreign exchange markets simultaneously while allowing a role for limits to speculations and for cross correlation across returns.

V. Can We Match the Stylized Facts in Spot-Forward Regressions? Some Monte Carlo Evidence

Given our discussion in Section III of the possibility of explaining the observed anomalies in spot-forward regressions if in fact the true DGP driving deviations from UIP is nonlinear, it seems worthwhile investigating whether we can match the stylized facts in spot-forward regressions using a DGP calibrated according to our estimated nonlinear Fama models. This may help us understand why much previous research estimating the linear Fama regression (2) has recorded a forward bias ($\beta \neq 1$) when in fact the forward bias may characterize only economically small departures from UIP. This exercise may also shed some light on the finding of excess returns predictability on the basis of the lagged forward premium, given that the regression typically used by researchers (Equation (3)) is a reparameterization of the Fama regression (2).

A. Matching the Forward Bias

We executed a battery of Monte Carlo experiments based on an artificial DGP identical to the estimated nonlinear Fama regression (4), calibrated on the estimates reported in Table 4, with independent and identically distributed Gaussian innovations. Initializing the artificial series at zero, we generated $5,000$ samples of $1,040$ observations and discarded the first 100, leaving $5,000$ samples of 940 observations, matching exactly the total number of actual observations used in this study. For the German mark and the euro we carried out the simulations by generating $5,000$ samples of 778 and 309 observations and discarded the first 100, leaving $5,000$ samples of 678 and 209 observations respectively, again matching the number of observations available for these exchange rates. The discarding of the first 100 artificial data is carried out to reduce the impact of the DGP initialization. For each generated sample of observations we then estimated the standard linear Fama regression (2). In Panel A of Table 8, we repeat in the first two columns the estimates of α and β obtained from the actual data (taken from Table 2), while in the third and fourth columns we report the average of the $5,000$ estimates obtained from the estimation of the Fama regression on the artificial data, say $\overline{\alpha}^{MC}$ and $\overline{\beta}^{MC}$, together with their 5th and 95th percentiles from the empirical distributions (reported in parentheses).

The results of these Monte Carlo investigations reveal that, if the true DGP were indeed of the nonlinear form (4) and researchers estimated a linear Fama regression, the estimates of α and β recorded on average would be very close to the ones estimated on actual data. In fact, the estimates of α and β recorded on actual data are, for each exchange rate examined, in the interval between the 5th and 95th percentiles of the empirical distribution of α^{MC} and β^{MC} obtained from estimating the Fama regression on the simulated data. In the last two columns, we report the p-values from a formal test statistic of the null hypotheses that $\overline{\alpha}^{MC} = \alpha$ and $\overline{\beta}^{MC} = \beta$ respectively. The p-values are generally very high, indicating that the estimates of α and β obtained from the actual data are indeed statistically insignificantly different from the average estimates $\overline{\alpha}^{MC}$ and $\overline{\beta}^{MC}$ one would obtain from estimating the linear Fama regression using the artificial data we generated.

B. Matching the Predictive Power of the Forward Premium on Future Excess Returns

We also investigated the ability of our nonlinear Fama regression to explain the puzzling finding that estimation of regressions of the form (3) typically yield the result that the forward premium can predict future excess returns. As argued in Section II, because regression (3) is obtained from reparameterizing the Fama regression (2), it is plausible that if nonlinearity in the true DGP of the spot-forward relationship can shed light on the forward bias puzzle arising from Equation (2) it should also shed some light on the predictability arising from Equation (3). Hence, using the same artificial data described in the previous sub-section, we estimated a regression of the form (3) for each of the 5,000 generated samples. In Panel B of Table 8, we report in the first column the t-statistic for the significance of the parameter associated with the forward premium in regression (3), namely β^{τ} (as given in Table 2), while in the second column we report the corresponding average of the 5,000 t-statistics obtained from the estimation of regression (3) on the artificial data, say $\overline{t}(\beta^{\tau})^{MC}$, together with the 5th and 95th percentiles from its empirical distribution (reported in parentheses).

The simulation results suggest that, if the true DGP were of the nonlinear form (4) (or its reparameterized form (8)) and one estimated a predictability regression of the form (3), the t-statistics for the significance of β^{τ} would be very close on average to the ones estimated on actual data, they would be statistically significant and would lie in the interval between the 5th and 95th percentiles of the empirical distribution of $t(\beta^{\tau})^{MC}$, for each exchange rate examined. On average, the t-statistics recorded are indeed statistically significantly different from zero. Finally, in the last column of Panel B, we report the p-value from a formal test statistic for the null hypothesis that $t(\beta^{\tau}) = \overline{t}(\beta^{\tau})^{MC}$, termed t_1. The p-value is generally high, indicating that the t-statistic obtained from estimating Equation (3) on actual exchange rate data is statistically insignificantly different from the average t-statistic ($\overline{t}(\beta^{\tau})^{MC}$) one would obtain from estimating the predictability regression (3) using the artificial data we generated.

C. Summing Up the Monte Carlo Results

Overall, therefore, our Monte Carlo experiments suggest that if the true DGP governing the relationship between the spot and forward exchange rates were of the nonlinear form we consider in this paper, estimation of the Fama regression (2) and the predictability regression (3) would lead us to reject UIP, to record a forward bias (β different from unity and generally negative), and to find evidence of predictability of excess returns using the information in the lagged forward premium, with estimates which are very close to the ones observed in actual data. However, these three features—violation of UIP, forward bias, and predictability of excess returns—are features which the DGP we study has only in the lower regime, which is a regime where expected excess returns (deviations from UIP) are tiny enough to be economically unimportant and unlikely to attract speculative capital.

VI. Conclusions

Our empirical results provide confirmation that major bilateral dollar exchange rates are linked nonlinearly to forward premia in the context of a model for UIP deviations which allows for time-variation in the forward bias and nonlinear mean reversion towards UIP. The nonlinearities we uncover are consistent with a model of deviations from UIP with two extreme regimes: a lower regime with persistent but tiny deviations from UIP, and an upper regime where UIP holds. In some sense, this characterization of UIP deviations suggests that, while UIP does not hold most of the time, deviations from UIP are generally economically small but they may be persistent as long as expected foreign exchange excess returns are not large enough to attract speculative capital. This evidence is consistent with recent theoretical contributions on the nature of exchange rate dynamics in the presence of limits to speculation in the foreign exchange market.

In a number of Monte Carlo experiments calibrated on the estimated nonlinear models, we show that if the true data generating process of UIP deviations were of the nonlinear form we consider, estimation of conventional spot-forward regressions would generate the well-known forward bias puzzle and the kind of predictability of foreign exchange excess returns documented in the literature.

Our results therefore allow us to end this study by making three statements, albeit with some degree of caution. First, the statistical rejection of UIP recorded by the literature may be less indicative of major inefficiencies in the foreign exchange market than has often been thought. Second, the forward bias puzzle may be explained by the assumption of linearity which is standard in the relevant literature. In our fitted nonlinear models, the forward bias is more persistent the closer expected exchange rates are to the UIP equilibrium. Somewhat paradoxically, therefore, rejections of UIP in a linear context may indicate that exchange rates have on average been relatively close to the UIP equilibrium, rather than implying that UIP does not hold at all. Third, the limits-to-speculation hypothesis and the implied nonlinearities in the relationship between spot and forward exchange rates appear to be of some importance in understanding the properties of departures from the foreign exchange market efficiency condition.

Although the results have been shown to be robust to a number of relevant tests, several caveats are in order. While our empirical analysis is inspired by the limits to speculation hypothesis, we do not claim that this paper provides a direct test of this specific hypothesis, but rather a test of its general predictions for the relationship between spot and forward exchange rates. Our approach is best interpreted as an empirical characterization of the spot-forward relationship motivated by the limits to speculation hypothesis or simply as an empirical investigation of parsimonious models of foreign exchange excess returns. In particular, although we have focused on a specific nonlinear formulation of the relationship between exchange rates and forward premia capable of capturing some of the key predictions of the limits to speculation hypothesis, experimentation with alternative nonlinear characterizations of the relationship is on the agenda for future research both to assess the robustness of our results and to further tighten the link between theory and empirical testing in a way that can allow us to discriminate among different arguments capable of rationalizing the existence of nonlinearities. Experimentation with multivariate versions of the nonlinear models used here that allow estimation of a general nonlinear error-correction model linking the dynamics of spot exchange rates and the full term structure of forward exchange rates, in the spirit of Clarida and others (2003), is also a promising area of research. This development is likely to enhance the predictive power of our nonlinear model, which, given its simplicity, might ignore important information. However, vector smooth-transition error correction models of this kind are well known to present great difficulties in terms of estimation and often trade off the increased explanatory power with a decrease in the interpretability of estimation outcomes (e.g., Granger and Teräsvirta, 1993). Further, while our results shed some light on why researchers have typically recorded rejections of UIP and why the forward bias may persist, our framework does not explain why β is negative in the lower regime rather than being, for example, in the middle of the inaction range. Explaining this finding may require further theoretical models where trading activities that move exchange rates are not driven just by pure currency strategies, as it is implicitly assumed under UIP (Lyons, 2001, p. 216-18). Finally, our analysis has been confined to a single-currency setting, given our intention to shed light on the forward bias and foreign exchange excess returns predictability anomalies, which have generally—indeed virtually exclusively—been studied in a single-exchange-rate setting. However, future research may extend our framework to more complex multiple-exchange-rate strategies.

Derivation of Equation (13)

Having rewritten our modelling framework in the form (12), we are able to provide the solution in Equation (13) by following the same steps used by Peel and Venetis (2005) in the context of an autoregressive exponential smooth transition process. Specifically, recall that the Lambert W function has the property $W(x)e^{W(x)} = x$. This implies that $xe^{x} = a$ has a solution $x = W(a)$. Let $ER_t^m = \zeta$ and $\varphi = \left[\alpha_1 + (\beta_1 - 1)\left(f_t^1 - s_t\right)\right]$. Then, $\zeta = e^{-\gamma\zeta^2}\varphi \iff \zeta e^{\gamma\zeta^2} = \varphi$. Squaring both sides and multiplying by 2γ yields $2\gamma\zeta^2 e^{2\gamma\zeta^2} = 2\gamma\varphi^2$. Setting $2\gamma\zeta^2 = \vartheta$, $\vartheta e^{\vartheta} = 2\gamma\varphi^2$ and therefore $\vartheta = W\left(2\gamma\varphi^2\right)$ or $\zeta = \sqrt{\frac{W(2\gamma\varphi^2)}{2\gamma}}$. Since $e^{-W(x)} = W(x)/x$, $e^{-0.5W(x)} = \sqrt{\frac{W(x)}{x}}$ we get $\zeta = \sqrt{\frac{W(2\gamma\varphi^2)}{2\gamma}} = \exp\left[-0.5W\left(2\gamma\varphi^2\right)\right]\varphi$, which is Equation (13): $ER_t^m = \exp\left\{-0.5W\left[2\gamma\left[\alpha_1 + (\beta_1 - 1)\left(f_t^1 - s_t\right)\right]^2\right]\right\}\left[\alpha_1 + (\beta_1 - 1)\left(f_t^1 - s_t\right)\right]$.

Robustness Results

Table B1. Nonlinear Forward Premium Regressions: ESTR estimation results

Panel A. Robustness to the forward contract maturity: 3-month forward rates (weekly data)

	$\alpha_1 = -\alpha_2$	$SE_{\alpha_1=-\alpha_2}$	$\beta_1 = 1 - \beta_2$	$SE_{\beta_1=1-\beta_2}$	γ	SE_γ	
Japan	0.0090	(0.0028)	-1.0557	(0.3390)	0.4668	(0.0593)	[0.0032]
UK	-0.0017	(0.0017)	-0.8365	(0.3537)	0.2930	(0.0196)	[0.0410]
Germany	0.0043	(0.0023)	0.5062	(0.2407)	0.6483	(0.1364)	[0.0038]
Euro	-0.0008	(0.0002)	-1.3739	(0.6188)	0.0884	(0.0152)	[0.1106]
Switzerland	0.0060	(0.0027)	-0.9255	(0.3698)	0.4907	(0.0579)	[0.0001]

Panel B. Robustness to the frequency of the data: monthly data (1-month forward rates)

	$\alpha_1 = -\alpha_2$	$SE_{\alpha_1=-\alpha_2}$	$\beta_1 = 1 - \beta_2$	$SE_{\beta_1=1-\beta_2}$	γ	SE_γ	
Japan	0.0012	(0.0008)	-2.4755	(0.9766)	0.7465	(0.2841)	[0.0021]
UK	-0.0026	(0.0028)	-2.8954	(1.1746)	0.3391	(0.0756)	[0]
Germany	0.0018	(0.0043)	0.0330	(0.7126)	1.7824	(0.7556)	[0.0002]
Euro	0.0005	(0.0041)	-3.8552	(1.5574)	0.0889	(0.0120)	[0.0306]
Switzerland	0.0069	(0.0034)	-2.1883	(1.0222)	0.2529	(0.1131)	[0.0464]

Notes: *Panel A.* The table reports the results from estimating the nonlinear forward premium regression $\Delta_3 s_{t+3} = \left[\alpha_1 + \beta_1\left(f_t^3 - s_t\right)\right] + \left[\alpha_2 + \beta_2\left(f_t^3 - s_t\right)\right] G\left[ER_t^e, \gamma\right] + \varepsilon_{t+1}$, where $\alpha_2 = -\alpha_1$, $\beta_2 = 1 - \beta_1$ and $G\left[ER_t^e, \gamma\right] = \left\{1 - \exp\left[-\gamma\left(ER_t^e\right)^2\right]\right\}$, using weekly data. *Panel B.* The table reports the results from estimating the nonlinear forward premium regression $\Delta s_{t+1} = \left[\alpha_1 + \beta_1\left(f_t^1 - s_t\right)\right] + \left[\alpha_2 + \beta_2\left(f_t^1 - s_t\right)\right] G\left[ER_t^e, \gamma\right] + \varepsilon_{t+1}$, where $\alpha_2 = -\alpha_1$, $\beta_2 = 1 - \beta_1$ and $G\left[ER_t^e, \gamma\right] = \left\{1 - \exp\left[-\gamma\left(ER_t^e\right)^2\right]\right\}$ using monthly data. Note that ER_t^e is defined as the expectation, formed at time t, of UIP deviations at time $t+3$. For both Panels A and B, values in parentheses (SE_x) are asymptotic standard errors for the parameter x, calculated using an autocorrelation and heteroskedasticity consistent matrix of residuals (Newey and West, 1987); values in brackets are p-values for the null hypothesis that $\gamma = 0$ calculated by parametric bootstrap as in Skalin (1998), using 5,000 replications. 0 denotes p-values lower than 10^{-5}.

Table B2. Empirical Size and Power Properties of the Linearity Tests

Panel A. Empirical size

	10%	5%	1%
	T=470		
F_L	0.0942	0.0434	0.0106
F_3	0.0908	0.0462	0.0098
F_2	0.0996	0.0480	0.0094
F_1	0.0976	0.0480	0.0108
	T=940		
F_L	0.0926	0.0438	0.0080
F_3	0.0940	0.0464	0.0070
F_2	0.0942	0.0460	0.0092
F_1	0.0916	0.0464	0.0092

Panel B. Empirical power

	10%	5%	1%
	T=470		
F_L	0.7492	0.6656	0.4944
F_3	0.3136	0.2294	0.1194
F_2	0.5824	0.4728	0.2906
F_1	0.2384	0.1562	0.0694
	T=940		
F_L	0.8000	0.7286	0.5810
F_3	0.2818	0.1998	0.0914
F_2	0.5986	0.4892	0.3102
F_1	0.2238	0.1482	0.0586

Note: *Panel A.* The table reports the results of a Monte Carlo experiment where the null hypothesis of linearity is true (i.e., the true DGP is the standard forward premium regression (2)) and it has been calibrated on the parameters estimated for the Japanese yen reported in Table 2. Figures are probabilities of rejection for different significance levels (i.e., 10%, 5%, and 1%) and different sample sizes $T = 470, 940$. *Panel B.* The table reports the results of a Monte Carlo experiment where the null hypothesis of linearity is false and the true DGP is the nonlinear forward premium regression (4) calibrated on the parameters estimated for the Japanese yen reported in Table 4. Figures are probabilities of rejection for different significance levels (i.e., 10%, 5%, and 1%) and different sample sizes $T = 470, 940$. For both Panels A and B, probabilities are constructed using $5,000$ replications in each experiment with identical random numbers across experiments.

Table 1. Summary Statistics

Panel A. Forward premium, $f_t^1 - s_t$

	mean		standard deviation		AR(1)	
Japan	0.0024	(0.0003)	0.0036	(0.0011)	0.614	(0.094)
United Kingdom	-0.0022	(0.0003)	0.0031	(0.0013)	0.761	(0.057)
Germany	0.0004	(0.0005)	0.0040	(0.0014)	0.439	(0.059)
Euro	0.0004	(0.0005)	0.0021	(0.0008)	0.602	(0.079)
Switzerland	0.0014	(0.0003)	0.0031	(0.0009)	0.724	(0.055)

Panel B. Depreciation rate, $s_{t+1} - s_t$

	mean		standard deviation		AR(1)	
Japan	0.0008	(0.0006)	0.0159	(0.0049)	0.077	(0.030)
United Kingdom	0.0003	(0.0004)	0.0145	(0.0054)	0.025	(0.033)
Germany	0.0008	(0.0006)	0.0154	(0.0045)	0.053	(0.031)
Euro	-0.0005	(0.0010)	0.0135	(0.0047)	0.090	(0.054)
Switzerland	0.0007	(0.0005)	0.0164	(0.0045)	0.048	(0.028)

Panel C. Return from currency speculation (excess return), $s_{t+1} - f_t^1$

	mean		standard deviation		AR(1)	
Japan	0.0002	(0.0007)	0.0159	(0.0049)	0.117	(0.028)
United Kingdom	-0.0025	(0.0061)	0.0151	(0.0051)	0.057	(0.031)
Germany	-0.0001	(0.0007)	0.0152	(0.0045)	0.053	(0.033)
Euro	0.0010	(0.0014)	0.0139	(0.0048)	0.131	(0.053)
Switzerland	0.0007	(0.0007)	0.0168	(0.0045)	0.085	(0.026)

Notes: One-month log-forward and log-spot exchange rates, f_t^1 and s_t, are expressed as dollars per unit of foreign currency. Data are Tuesdays of every week, taken from *Datastream.* The sample period spans from January 4, 1985, to December 31, 2002, for all exchange rates except for the German mark (January 7, 1986, to December 31, 1998) and the euro (January 5, 1999, to December 31, 2002). Figures in parentheses are standard errors calculated by using an autocorrelation and heteroskedasticity consistent matrix of residuals, with three lags (Newey and West, 1987).

Table 2. Forward Premium (Fama) Regressions

	α	SE(α)	β	SE(β)	s.e.	T	$t(\beta^{\tau})$
Japan	0.0015	(0.0005)	-0.2865	(0.1586)	0.015	939	-6.742
United Kingdom	-0.0003	(0.0004)	-0.3098	(0.2588)	0.014	939	-6.075
Germany	0.0004	(0.0006)	0.3212	(0.1495)	0.015	677	-4.712
Euro	-0.0001	(0.0008)	-0.8883	(0.4422)	0.013	208	-3.963
Switzerland	0.0012	(0.0006)	-0.3786	(0.1645)	0.016	939	-7.036

Notes: The table shows the results from estimating, by ordinary least squares, the conventional forward premium (Fama) regression (2): $\Delta s_{t+1} = \alpha + \beta\left(f_t^1 - s_t\right) + \upsilon_{t+1}$. Values in parentheses (SE(α) and SE(β)) are asymptotic standard errors calculated using an autocorrelation and heteroskedasticity consistent matrix of residuals (Newey and West, 1987). s.e. is the standard deviation of the residual υ_{t+1}; and T is the number of usable observations. The last column reports the t-statistic (namely $t(\beta^{\tau})$) for the parameter β^{τ} in regression (3): $ER_{t+1} = \alpha + \beta^{\tau}\left(f_t^1 - s_t\right) + \upsilon_{t+1}$, where $ER_{t+1} \equiv \Delta s_{t+1} - \left(f_t^1 - s_t\right) \equiv s_{t+1} - f_t^1$, and $\beta^{\tau} = \beta - 1$.

Table 3. Linearity Tests on the Fama Regression

	F_L	F_3	F_2	F_1
Japan	0.020	0.077	0.006	0.142
United Kingdom	5.29×10^{-16}	0.219	0.038	0.365
Germany	3.10×10^{-17}	0.087	0.044	0.061
Euro	5.09×10^{-5}	0.268	0.027	0.224
Switzerland	0.011	0.983	0.001	0.546

Notes: The table reports the p-values from applying the linearity testing procedure suggested by Teräsvirta (1994, 1998). Assuming that a plausible, generic transition variable is q_t, the appropriate auxiliary regression for the linearity tests against a STR alternative is the following: $\widehat{e}_{t+1} = \boldsymbol{\vartheta}_0'\mathbf{A}_{t+1} + \boldsymbol{\vartheta}_1'\mathbf{A}_{t+1}(q_t) + \boldsymbol{\vartheta}_2'\mathbf{A}_{t+1}(q_t)^2 + \boldsymbol{\vartheta}_3'\mathbf{A}_{t+1}(q_t)^3 + innovations$, where $\widehat{e}_{t+1}$ is the estimated disturbance retrieved from the linear model being tested for linearity (in the present context it is the residual from each of the Fama regression models reported in Table 2), and $\mathbf{A}_t$ denotes the vector of explanatory variables in the model being tested, which in our case simply amounts to the lagged forward premium. The transition variable, q_t used in our tests is the expected excess return for one month ahead based on information at time t, ER_t^e. The general test for linearity against STR is then the ordinary F-test of the null hypothesis: $H_{0L}: \boldsymbol{\vartheta}_1' = \boldsymbol{\vartheta}_2' = \boldsymbol{\vartheta}_3' = \mathbf{0}$. The choice between a LSTR and an ESTR model is based on a sequence of nested tests within H_{0L}. First, the null hypothesis H_{0L} must be rejected using an ordinary F-test (F_L). Then the following three hypotheses are tested sequentially: $H_{03}: \boldsymbol{\vartheta}_3' = \mathbf{0}$; $H_{02}: \boldsymbol{\vartheta}_2' \mid \boldsymbol{\vartheta}_3' = \mathbf{0}$; $H_{01}: \boldsymbol{\vartheta}_1' \mid \boldsymbol{\vartheta}_2' = \boldsymbol{\vartheta}_3' = \mathbf{0}$. Again, an F-test is used, with the corresponding test statistics denoted F_3, F_2, and F_1, respectively. The decision rule is as follows: if the test of H_{02} has the smallest p-value, an ESTR is chosen, otherwise an LSTR is selected. The p-values were calculated using the appropriate F distribution.

Table 4. Nonlinear Fama Regressions: ESTR Estimation Results

Panel A. Parameter estimates

	$\alpha_1 = -\alpha_2$	$\text{SE}_{\alpha_1=-\alpha_2}$	$\beta_1 = 1 - \beta_2$	$\text{SE}_{\beta_1=1-\beta_2}$	γ	SE_γ	
Japan	0.0018	(0.0008)	-0.5719	(0.2278)	0.4014	(0.0501)	[0.0010]
United Kingdom	-0.0004	(0.0005)	-0.4708	(0.1557)	0.1209	(0.0061)	[0.0480]
Germany	0.0008	(0.0008)	0.1540	(0.0633)	0.5348	(0.1522)	[0.0048]
Euro	-0.0003	(0.0010)	-1.0608	(0.5844)	0.1148	(0.0188)	[0]
Switzerland	0.0018	(0.0009)	-1.0072	(0.3254)	0.5130	(0.0583)	[0]

Panel B. Diagnostic tests

	$LR1$	$LR2$	F_{NRN}	F_{NSB}
Japan	0.115	0.901	0.903	0.547
United Kingdom	0.341	0.914	0.969	0.495
Germany	0.713	0.937	0.940	0.466
Euro	0.134	0.897	0.998	0.502
Switzerland	0.304	0.449	0.994	0.593

Notes: *Panel A.* The table shows the results from the nonlinear forward premium regression $\Delta s_{t+1} = \left[\alpha_1 + \beta_1 \left(f_t^1 - s_t\right)\right] + \left[\alpha_2 + \beta_2 \left(f_t^1 - s_t\right)\right] G\left[ER_t^e, \gamma\right] + \varepsilon_{t+1}$, where $\alpha_2 = -\alpha_1$, $\beta_2 = 1 - \beta_1$ and $G\left[ER_t^e, \gamma\right] = \left\{1 - \exp\left[-\gamma \left(ER_t^e\right)^2 / \widehat{\sigma}_{ER^e}^2\right]\right\}$. Values in parentheses (SE_x) are asymptotic standard errors for the parameter x, calculated using an autocorrelation and heteroskedasticity consistent matrix of residuals (Newey and West, 1987). Values in brackets are p-values for the null hypothesis that $\gamma = 0$, calculated by the parametric bootstrap procedure suggested by Skalin (1998) using $5,000$ replications. 0 denotes p-values lower than 10^{-5}. *Panel B.* $LR1$ is the likelihood ratio test for the joint null hypothesis that $\beta_1^+ = \beta_1^-$ and $\beta_2^+ = \beta_2^-$ from the unrestricted nonlinear model: $\Delta s_{t+1} = \left[\alpha_1 + \beta_1^+ \left(f_t^1 - s_t\right)^+ + \beta_1^- \left(f_t^1 - s_t\right)^-\right] + \left[\alpha_2 + \beta_2^+ \left(f_t^1 - s_t\right)^+ + \beta_2^- \left(f_t^1 - s_t\right)^-\right] G\left[ER_t^e, \gamma\right] + \varepsilon_{t+1}$, where $\left(f_t^1 - s_t\right)^+$ and $\left(f_t^1 - s_t\right)^-$ are forward premia ($f_t^1 - s_t > 0$) and forward discounts ($f_t^1 - s_t < 0$) respectively. $LR2$ is the likelihood ratio test for the null hypothesis that $\alpha_2 = -\alpha_1$, $\beta_2 = 1 - \beta_1$. F_{NRN} is the test for the null hypothesis of no remaining nonlinearity, constructed as in Eitrheim and Teräsvirta (1996). F_{NSB} is a test for the null hypothesis of no structural break in the model's parameters, constructed as in Eitrheim and Teräsvirta's (1996). For all test statistics, we report p-values, calculated using an autocorrelation and heteroskedasticity consistent matrix of residuals.

Table 5. Sharpe Ratios

	SR	*min* SR	%Obs SR $\leq$ *min* SR
Japan	0.14	1.32	82
United Kingdom	0.88	2.40	97
Germany	0.41	1.14	68
Euro	0.72	2.46	97
Switzerland	0.25	1.17	83
Average	*0.48*	*1.69*	*83*

Notes: The first column of the table reports the mean of the annualized Sharpe ratio (SR) implied by our weekly data. SR is calculated as the realized average excess returns ($E\left[\Delta s_{t+1} - (f_t - s_t)\right] = E\left[ER\right]$) divided by the standard deviations of excess returns (σ_{ER}) over the sample period on an annual basis: $\frac{E[ER]}{\sigma_{ER}} \times \sqrt{52}$. *min* SR (second column) is the minimum value of the annualized Sharpe ratio which leads to a shift from the lower regime to the upper regime, defined here as the value of the transition function $\Phi\left(\cdot\right) \equiv \left\{1 - \exp\left[-\gamma SR^2\right]\right\} = 0.5$. The last column reports the percentage of observations where the annualized Sharpe ratio is lower than or equal to the minimum Sharpe ratio, as given by *min* SR. The last row reports averages across countries for each column.

Table 6. Forecast Comparisons

	Linear Fama		Nonlinear Fama	
Japan	0.002	(0.001)	0.009*	(0.004)
United Kingdom	0.002	(0.002)	0.012	(0.011)
Germany	0.007*	(0.003)	0.009*	(0.004)
Euro	0.016	(0.010)	0.027	(0.015)
Switzerland	0.003	(0.002)	0.025*	(0.010)

Notes: Figures reported denote estimates of the parameter δ in the equation $\widehat{\Delta s_{t+1|t}} = \delta_0 + \delta \Delta s_{t+1} + \xi_{t+1}$, where the hat denotes the one-step-ahead (one-week-ahead) prediction from the (linear or nonlinear) Fama regression; see Evans and Lyons (2005). Figures in parentheses are standard errors calculated using the Newey-West (1987) autocorrelation and heteroskedasticity consistent covariance correction with three lags. The asterisk denotes estimates of δ that are statistically significantly different from zero at the five percent significance level.

Table 7. Unconditional Sharpe Ratios

	Linear Fama		Nonlinear Fama	
	$\overline{R}^2$	SR	$\overline{R}^2$	SR
Japan	0.040	0.465	0.108	0.548
United Kingdom	0.002	0.403	0.078	0.508
Germany	0.031	0.445	0.037	0.453
Euro	0.070	0.498	0.108	0.548
Switzerland	0.050	0.470	0.169	0.629

Notes: The table reports unconditional annualized Sharpe ratios corresponding to each $\overline{R}^2$ for both the linear and nonlinear Fama regression. The formula applied to compute the annualized Sharpe ratios is: $SR = \sqrt{0.4^2 + \overline{R}^2} / \sqrt{1 - \overline{R}^2}$, as in Cochrane (1999, pp. 65-66, 75-76). The value chosen to proxy for the unconditional annualized Sharpe ratio of a US buy-and-hold strategy is 0.4 (first term in the square root on the numerator), as reported in the main text and consistent with Sharpe (1994) and Lyons (2001).

Table 8. Monte Carlo Results: Matching the Stylized Facts

Panel A. Matching the forward bias puzzle

	α	β	$\overline{\alpha}^{MC}$		$\overline{\beta}^{MC}$		$t(\alpha)$	$t(\beta)$
Japan	0.0015	-0.2865	0.0013	(0.0002,0.0025)	-0.1588	(-0.4737,0.1492)	0.848	0.501
United Kingdom	-0.0003	-0.3098	-0.0003	(-0.0014,0.0007)	-0.3094	(-0.6607,0.0429)	0.979	0.999
Germany	0.0004	0.3212	0.0005	(-0.0003,0.0015)	0.4141	(0.1718,0.6508)	0.783	0.522
Euro	-0.0001	-0.8883	-0.0002	(-0.0018,0.0012)	-0.8502	(-1.638,-0.0807)	0.881	0.936
Switzerland	0.0012	-0.3786	0.0012	(0.0002,0.0022)	-0.3674	(-0.6955,-0.042)	0.966	0.955

Panel B. Matching the predictive power of the forward premium on future excess returns

	$t(\beta^{\tau})$	$\overline{t}(\beta^{\tau})^{MC}$		t_1
Japan	-6.742	-6.147	(-4.382,-8.044)	0.592
United Kingdom	-6.075	-6.166	(-4.373,-8.040)	0.935
Germany	-4.712	-4.088	(-2.396,-5.857)	0.558
Euro	-3.963	-4.046	(-2.206,-6.022)	0.943
Switzerland	-7.036	-6.947	(-5.096,-8,896)	0.939

Notes: *Panel A.* α, β are the estimates of the standard Fama regression (2), taken from Table 2. $\overline{\alpha}^{MC}, \overline{\beta}^{MC}$ denote the average of the empirical distribution (based on $5,000$ replications) of the coefficients α, β obtained from estimating the standard forward premium regression (2) using artificial data under a true DGP which is a nonlinear forward premium regression of the form (4), using $5,000$ replications. Values in parentheses are the 5th and 95th percentiles of the empirical distribution of the parameters α^{MC}, β^{MC} respectively. $t(\alpha)$ and $t(\beta)$ are the p-values of the test statistic for the null hypothesis that $\overline{\alpha}^{MC} = \alpha$ and $\overline{\beta}^{MC} = \beta$, respectively. *Panel B.* $t(\beta^{\tau})$ is the estimated t-statistic for the significance of β^{τ} in the regression of excess returns on the lagged forward premium, defined in Equation (3). $\overline{t}(\beta^{\tau})^{MC}$ is the average of the empirical distribution of the t-statistic for the significance of the parameter β^{τ} on forward premium in a predictive regression of the form $ER_{t+1} = \alpha + \beta^{\tau}\left(f_t^1 - s_t\right) + error$, calculated under the same DGP as above using 5,000 replications. Values in parentheses correspond to the 5th and 95th percentiles of the empirical distribution of the test statistics $t(\beta^{\tau})$. t_1 is the p-value of the test statistic for the null hypothesis that $\overline{t}(\beta^{\tau})^{MC} = t(\beta^{\tau})$.

References

Backus, D.K., A.W. Gregory, and C.I. Telmer, 1993, "Accounting for Forward Rates in Markets for Foreign Currency," *Journal of Finance* Vol. 48, pp. 1887-1908.

Backus, D.K., S. Foresi, and C.I. Telmer, 2001, "Affine Term Structure Models and the Forward Premium Anomaly," *Journal of Finance Vol.* 56, pp. 279-304.

Baillie, R.T., and T. Bollerslev, 2000, "The Forward Premium Anomaly is Not as Bad as You Think," *Journal of International Money and Finance* Vol. 19, pp. 471-488.

Baldwin, R., 1990, "Re-Interpreting the Failure of Foreign Exchange Market Efficiency Tests: Small Transaction Costs, Big Hysteresis Bands," NBER Working Paper No. 3319 (Cambridge, Massachusetts: National Bureau of Economic Research).

Bansal, R., 1997, "An Exploration of the Forward Premium Puzzle in Currency Markets," *Review of Financial Studies* Vol. 10, pp. 369-403.

———, and M. Dahlquist, 2000, "The Forward Premium Puzzle: Different Tales from Developed and Emerging Economies," *Journal of International Economics* Vol. 51, pp. 115-144.

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

Barnhart, S.W., R. McKnown, and M.S. Wallace, 1999, "Non-Informative Tests of the Unbiased Forward Exchange Rate," *Journal of Financial and Quantitative Analysis* Vol. 34, pp. 265-291.

———, 2002, "Some Answers to Puzzles in Testing Unbiasedness in the Foreign Exchange Market," *Applied Financial Economics* Vol. 12, pp. 687-696.

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

———, and R. J. Hodrick, 1993, "On Biases in the Measurement of Foreign Exchange Risk Premiums," *Journal of International Money and Finance* Vol. 12, pp. 115-138.

———, 2001, "Expectations Hypotheses Tests," *Journal of Finance* Vol. 56, pp. 1357-1394.

———, and D. Marshall, 1997, "The Implications of First-Order Risk Aversion for Asset Market Risk Premiums," *Journal of Monetary Economics* Vol. 40, pp. 3-39.

Bilson, J.F.O., 1981, "The 'Speculative Efficiency' Hypothesis," *Journal of Business* Vol. 54, pp. 435-451.

Brenner, R.J., and K.F. Kroner, 1995, "Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets," *Journal of Financial and Quantitative Analysis* Vol. 30, pp. 23-42.

Chinn, M.D., and J.A. Frankel, 2002, "Survey Data on Exchange Rate Expectations: More Currencies, More Horizons, More Tests," in *Monetary Policy, Capital Flows and Financial Market*

Developments in the Era of Financial Globalisation: Essays in Honour of Max Fry, ed. by D. Dickinson and W. Allen (London: Routledge), pp. 145-67.

Clarida, R.H., L. Sarno, M.P. Taylor, and G. Valente, 2003, "The Out-of-Sample Success of Term Structure Models as Exchange Rate Predictors: A Step Beyond," *Journal of International Economics* Vol. 60, pp. 61-83.

———, 2006, "The Role of Asymmetries and Regime Shifts in the Term Structure of Interest Rates," *Journal of Business*, forthcoming.

Cochrane, J.H., 1999, "Portfolio Advice for a Multifactor World," *Federal Reserve Bank of Chicago Economic Perspectives* Vol. 23, pp. 59-77.

Corless, R.M., G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth, 1996, "On the Lambert W Function," *Advances in Computational Mathematics* Vol. 5, pp. 329-359.

Cumby, R.E., 1988, "Is It Risk? Explaining Deviations from Uncovered Interest Parity," *Journal of Monetary Economics* Vol. 22, pp. 279-299.

Cumby, R.E., and M.E. Obstfeld, 1984, "International Interest Rate and Price Level Linkages under Flexible Exchange Rates: A Review of Recent Evidence," in *Exchange Rate Theory and Practice*, ed. by J.F.O. Bilson and R.C. Marston (Chicago: University of Chicago Press), pp. 121-151.

De Grauwe, P., and M. Grimaldi, 2005, *The Exchange Rate in a Behavioural Finance Framework* (Princeton: Princeton University Press), forthcoming.

Domowitz, I., and C. Hakkio, 1985, "Conditional Variance and the Risk Premium in the Foreign Exchange Market," *Journal of International Economics* Vol. 19, pp. 47-66.

Dumas, B., 1992, "Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World," *Review of Financial Studies* Vol. 5, pp. 153-180.

Eitrheim, O., and T. Teräsvirta, 1996, "Testing the Adequacy of Smooth Transition Autoregressive Models," *Journal of Econometrics* Vol. 74, pp. 59-75.

Engel, C., 1996, "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence," *Journal of Empirical Finance* Vol. 3, pp. 123-192.

Evans, M.D.D., and R.K. Lyons, 2005, "Meese-Rogoff Redux: Micro-Based Exchange Rate Forecasting," *American Economic Review Papers and Proceedings* Vol. 95, pp. 405-414.

Fama, E.F., 1984, "Forward and Spot Exchange Rates," *Journal of Monetary Economics* Vol. 14, pp. 319-338.

Flood, R.B., and A.K. Rose, 1996, "Fixes: Of the Forward Discount Puzzle," *Review of Economics and Statistics* Vol. 78, pp. 748-752.

———, 2002, "Uncovered Interest Parity in Crisis," *International Monetary Fund Staff Papers* Vol. 49, pp. 252-266.

Flood, R.B., and M.P. Taylor, 1996, "Exchange Rate Economics: What's Wrong with the Conventional Macro Approach?," in *The Microstructure of Foreign Exchange Markets*, ed. by J.A. Frankel, G. Galli, and A. Giovannini (Chicago: University of Chicago Press), pp. 261-294.

Frankel, J.A., 1980, "Tests of Rational Expectations in the Forward Exchange Market," *Southern Economic Journal* Vol. 46, pp. 1083-1101.

———, and C. Engel, 1984, "Do Asset Demand Functions Optimize Over the Mean and Variance of Real Returns? A Six Currency Test," *Journal of International Economics* Vol. 17, pp. 309-323.

Franses, P.H., and D. van Dijk. 2000, *Non-Linear Time Series Models in Empirical Finance* (Cambridge: Cambridge University Press).

Frenkel, J.A., and R.M. Levich, 1975, "Covered Interest Arbitrage: Unexploited Profits?," *Journal of Political Economy* Vol. 83, pp. 325-338.

———, 1977, "Transaction Costs and Interest Arbitrage: Tranquil versus Turbulent Periods," *Journal of Political Economy* Vol. 85, pp. 1209-1226.

Froot, K.A., and R.H. Thaler, 1990, "Foreign Exchange," *Journal of Economic Perspectives* Vol. 4, pp. 179-192.

Galati, G., and M. Melvin, 2004, "Why Has FX Trading Surged? Explaining the 2004 Triennial Survey," *BIS Quarterly Review* (December), pp. 67-74.

Gallant, A.R., L.P. Hansen, and G. Tauchen, 1990, "Using Conditional Moments of Asset Payoffs to Infer the Volatility of Intertemporal Marginal Rates of Substitution," *Journal of Econometrics* Vol. 45, pp. 141-179.

Granger, C.W.J., and T-H. Lee, 1999, "The Effect of Aggregation on Nonlinearity," *Econometric Reviews* Vol. 18, pp. 259-269.

Granger, C.W.J., and T. Teräsvirta, 1993, *Modelling Nonlinear Economic Relationships* (Oxford: Oxford University Press).

Grossman, S., and L. Weiss, 1983, "A Transactions-Based Model of the Monetary Transition Mechanism," *American Economic Review* Vol. 73, pp. 871-880.

Hansen, L.P., and R.J. Hodrick, 1980, "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy* Vol. 88, pp. 829-853.

Hirshleifer, D., 1988, "Residual Risk, Trading Costs, and Commodity Futures Risk Premia," *Review of Financial Studies* Vol. 1, pp. 173-193.

Hollifield, B., and R. Uppal, 1997, "An Examination of Uncovered Interest Parity in Segmented International Commodity Markets," *Journal of Finance* Vol. 52, pp. 2145-2170.

Hodrick, R.J., 1987, *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*, (London: Harwood).

Huisman, R., K. Koedijk, C. Kool, and F. Nissen, 1998, "Extreme Support for Uncovered Interest Parity," *Journal of International Money and Finance* Vol. 17, pp. 211-228.

Lewis, K.K., 1995, "Puzzles in International Financial Markets," in *The Handbook of International Economics* Vol. 3., ed. by G.M. Grossman and K. Rogoff (Amsterdam: North-Holland), pp. 1913-1971.

Liu, J., and F. Longstaff, 2004, "Losing Money on Arbitrages: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities," *Review of Financial Studies* Vol. 17, pp. 611-641.

Liu, P.C., and G.S. Maddala, 1992, "Rationality of Survey Data and Tests for Market Efficiency in the Foreign Exchange Markets," *Journal of International Money and Finance* Vol. 11, pp. 366-381.

Lyons, R.K., 2001, *The Microstructure Approach to Exchange Rates* (Cambridge: MIT Press).

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

———, and Y-K. Moh, 2002, "Continuous-Time Market Dynamics, ARCH Effects, and the Forward Premium Anomaly," unpublished working paper, University of Notre Dame and Tulane University.

Moh, Y-K., 2002, "Exchange Rate Dynamics in a Continuous-Time Model of Uncovered Interest Parity and the Regulated Jump-Diffusion Interest Differential," unpublished working paper, Tulane University.

Neal, R., 1996, "Direct Tests of Index Arbitrage Models," *Journal of Financial and Quantitative Analysis* Vol. 31, pp. 541-562.

Newey, W.K., and K.D. West, 1987, "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* Vol. 55, pp. 703-708.

Obstfeld, M., and K. Rogoff, 2000, "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?," in *NBER Macroeconomics Annual,* ed. by B. Bernanke and K. Rogoff (Cambridge: MIT Press), pp. 339-390.

Osterberg, W., 2000, "New Results on the Rationality of Survey Measures of Exchange-rate Expectations," *Federal Reserve of Cleveland Economic Review* (Q1), pp. 14-21.

Peel, D.A., and I.A. Venetis, 2005, "Smooth Transition Models and Arbitrage Consistency," *Economica* Vol. 72, pp. 413-430.

Sarno, L., 2005, "Towards a Solution to the Puzzles in Exchange Rate Economics: Where Do We Stand?," *Canadian Journal of Economics* Vol. 38, pp. 673-708.

———, and M.P. Taylor, 2001, *The Microstructure of the Foreign-Exchange Market: A Selective Survey of the Literature.* (Princeton: Princeton University Press).

———, 2003, *The Economics of Exchange Rates* (Cambridge: Cambridge University Press).

Sercu, P., and M. Vandebroek, 2005, "What UIP Tests on Extreme Samples Reveal About the Missing Variable," *Journal of International Money and Finance* Vol. 24, pp. 1237-1260.

Sercu, P., and X. Wu, 2000, "Uncovered Interest Arbitrage and Transaction Costs: Errors-in-Variables versus Hysteresis Effects," unpublished working paper, University of Leuven and City University of Hong Kong.

Sharpe, W.F., 1994, "The Sharpe Ratio," *Journal of Portfolio Management* Vol. 21, pp. 49-58.

Shleifer, A., and R. Vishny, 1997, "The Limits to Arbitrage," *Journal of Finance* Vol. 52, pp. 35-56.

Skalin, J., 1998, "Testing Linearity Against Smooth Transition Autoregression Using Parametric Bootstrap," unpublished working paper, Stockholm School of Economics.

Sofianos, G., 1993, "Index Arbitrage Profitability," *Journal of Derivatives* Vol. 1, pp. 6-20.

Tagaki, S., 1991, "Exchange Rate Expectations: A Survey of Survey Studies," *International Monetary Fund Staff Papers* Vol. 38, pp. 156-183.

Taylor, M.P., 1987, "Covered Interest Parity: A High-Frequency, High-Quality Data Study," *Economica* Vol. 54, pp. 429-438.

———, 1995, "The Economics of Exchange Rates," *Journal of Economic Literature* Vol. 33, pp. 13-47.

Teräsvirta, T., 1994, "Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association* Vol. 89, pp. 208-218.

———, 1998, "Modelling Economic Relationship with Smooth Transition Regressions," in *Handbook of Applied Economic Statistics*, ed. by D.E.A. Giles and A. Ullah (New York: Marcel Dekker), pp. 507-552.

———, and H.M. Anderson, 1992, "Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models," *Journal of Applied Econometrics* Vol. 7, pp.119-36.

Tong, H., 1990, *Nonlinear Time Series: A Dynamical System Approach* (Oxford: Clarendon Press).

Figure 1. Estimated Transition Functions vs. Sharpe Ratios

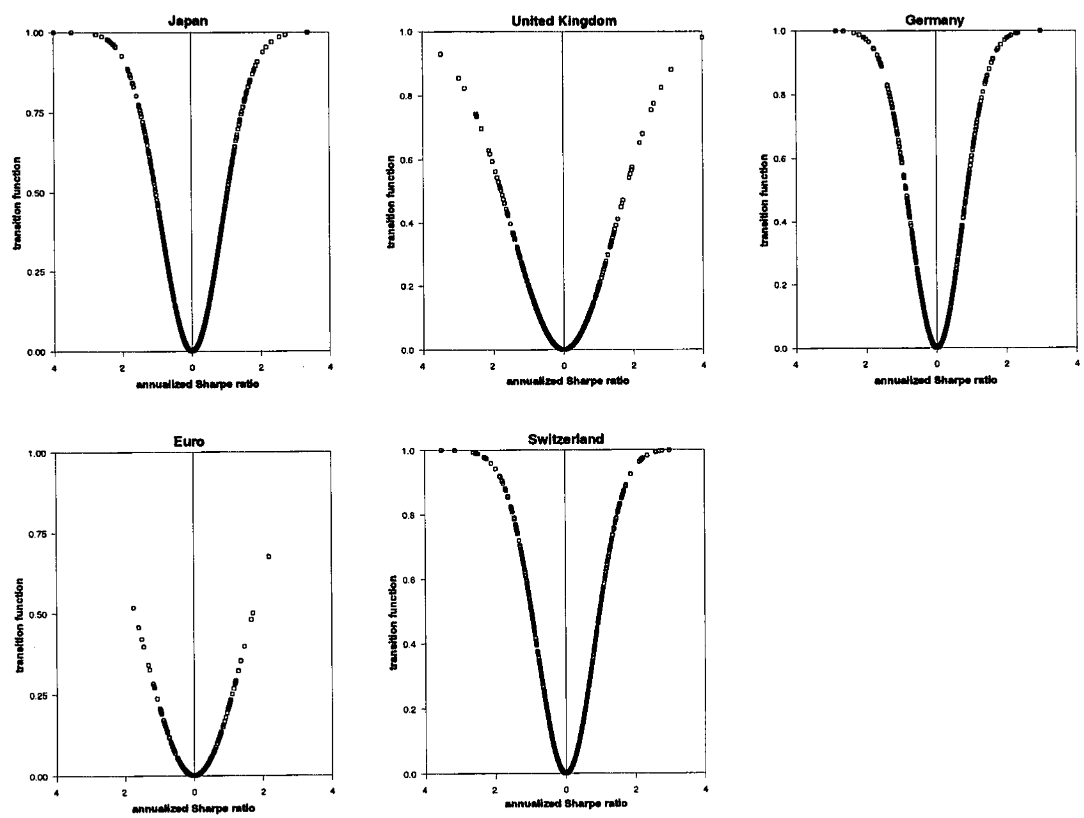

Figure 2. Estimated Transition Functions and Sharpe Ratios