Limits to Speculation and Nonlinearity in Deviations from Uncovered Interest Parity: Empirical Evidence and Implications for the Forward Bias Puzzle*

Hyginus Leon[a], Lucio Sarno[b,c] and Giorgio Valente[b]

a: International Monetary Fund

b: University of Warwick

c: Centre for Economic Policy Research (CEPR)

September 2003

Abstract

We examine empirically the conjecture that limits to speculation in the foreign exchange market may induce nonlinearities in the spot-forward relationship and in the process driving the deviations from the uncovered interest rate parity (UIP) condition. Our empirical results provide strong evidence of important nonlinearities which are consistent with a model of deviations from UIP with two extreme regimes: one regime with persistent but tiny deviations from UIP, and another regime where UIP holds. In a battery of Monte Carlo experiments, 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 predictability of foreign exchange excess returns documented in the literature. In turn, these findings have implications for the economic significance of the rejection of foreign exchange market efficiency.

JEL classification: F31.
Keywords: foreign exchange; uncovered interest parity; forward bias; nonlinearity.

***Acknowledgments:** This paper was partly written while Lucio Sarno was a Visiting Scholar at the International Monetary Fund and at the Central Bank of Norway. Financial support from the Economic and Social Research Council (No. L138251044) is gratefully acknowledged. The authors are indebted for useful conversations and constructive comments to John Driffill, Peter Kenen, Marcus Miller, Dagfinn Rime, Mark Taylor, Shang-Jin Wei and Mark Wohar. The authors alone are responsible for any errors that may remain and for the views expressed in the paper.

1 Introduction

The uncovered interest rate parity (UIP) condition postulates that the expected foreign exchange gain from holding one currency rather than another - the expected exchange rate change - must be just offset by the opportunity cost of holding funds in this currency rather than the other - the interest rate differential or, which is the same in the absence of arbitrage in foreign exchange markets, the forward premium. This condition represents the cornerstone parity condition for foreign exchange market efficiency and 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 when one might guess that investor would demand higher interest rates on currencies expected to fall in value. This anomaly, often termed '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, among the major floating currencies against the dollar, the spot exchange rate has usually been recorded to fall when the forward market would have predicted it to rise and viceversa (e.g., Cumby and Obstfeld, 1984; Hodrick, 1987; Bekaert and Hodrick, 1993; Lewis, 1995; Engel, 1996; Sarno and Taylor, 2002, Ch. 2, and the references therein).[1]

An alternative way of examining the properties of UIP is by examining whether future UIP deviations - or, identically, foreign exchange excess returns - are predictable using the forward premium as a predictor variable. Under the hypothesis that UIP holds (market efficiency), UIP deviations must be unpredictable. This issue was investigated, for example, by Bilson (1981), Fama (1984) and Backus, Gregory and Telmer (1993), who report evidence of strong predictability of excess returns (deviations from UIP) on the basis of the lagged forward premium when, in fact, there should not be any if UIP held.

Attempts to explain, statistically and economically, 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). It

[1] Exceptions include Bansal and Dahlquist (2000), who document that the forward bias is largely confined to developed economies and to countries for which the US 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.

is difficult to explain the rejection of UIP and the forward bias puzzle, moreover, 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, Gallant, Hussey and Tauchen, 1995; Bekaert, 1996) and from expected utility (Bekaert, Hodrick and Marshall, 1997), or else using popular models of the term structure of interest rates adapted to a multicurrency setting (Backus, Foresi and Telmer, 2001). Hence, even with the benefit of almost twenty 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 generally relies on linear frameworks in analyzing the properties of UIP deviations. However, 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; Sercu and Wu, 2000; Obstfeld and Rogoff, 2000), central bank intervention (e.g. Mark and Moh, 2002; Moh, 2002), and the existence of limits to speculation (e.g. Lyons, 2001, pp. 206-220). 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 grows it becomes large enough to generate Sharpe ratios that are large enough to attract speculative capital away from alternative trading strategies (Lyons, 2001).[2]

Although the literature has already documented that normal values of the forward premia may impact on future exchange rates differently from extreme values (e.g. Bilson, 1981; Flood and Rose, 1994; Flood and Taylor, 1996; Huisman, Koedijk, Kool and Nissen, 1998) and some authors have investigated the role of nonlinearities in the term structure of forward premia for the purpose of forecasting exchange rates (e.g. Clarida, Sarno, Taylor and Valente, 2003), to the best our knowledge, little has been done to investigate whether nonlinearities in the spot-forward relationship can help us

[2] The limits to speculation hypothesis 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 opportunties with the highest Sharpe ratio.

shed some light on the forward bias puzzle. 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 ten major US dollar exchange rates since 1985 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 important nonlinearities. While this result is not novel *per se*, our nonlinear model proves especially useful to understand the properties of deviations from UIP. In particular, consistent with the limits-to-speculation hypothesis which we use to rationalize our nonlinear spot-forward regression, we find that in the neighboroughood of UIP departures from market efficiency and hence the forward bias are statistically significant and persistent but economically too small to attract speculative capital, while for deviations from UIP which 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, estimation of the conventional linear spot-forward regressions would lead us to reject the validity of UIP and the hypothesis of no predictability of foreign exchange excess returns with parameters estimates that are very close to the ones observed using actual data and generally reported in the literature. However, the failure of UIP and the findings of a forward bias and predictability of excess returns are features that the DGP only has in the inner regime, which is the regime where deviations from UIP are tiny enough to be economically unimportant and unlikely to attract speculative capital. Our interpretation of the empirical 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 very 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 2 provides an outline of the theoretical background and introduces the limits-to-speculation hypothesis. Section 3 describes the empirical framework used to analyze the relationship between spot and forward exchange rates. In Section 4 we report and discuss our empirical results, while Section 5 provides our Monte Carlo simulation results. Section 6 concludes. Details of the estimation procedure used for linearity testing and some robustness results are provided in Appendices A and B.

2 Uncovered Interest Parity and the Forward Bias Puzzle: A Nonlinear Perspective

2.1 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. The UIP condition represents the cornerstone parity condition for foreign exchange market efficiency:

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

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.[3] Most often, however, analyses of foreign exchange market efficiency 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 (Aliber, 1973; Frenkel and Levich, 1975, 1977; Dooley and Isard, 1980; Levich, 1985; Clinton, 1988; Frankel and MacArthur, 1988; Taylor, 1987, 1989; for a survey of this evidence, see e.g. Sarno and Taylor, 2002, Ch. 2). Note that, unlike CIP, UIP is not an arbitrage condition since one of the terms in the equation, namely the exchange rate at time $t+k$, is unknown at time t and, therefore, non-zero deviations from UIP do not necessarily imply the existence of arbitrage profits due to the foreign exchange risk associated with future exchange rate movements.

Using CIP and replacing the forward premium (or forward discount) $f^k_t - s_t$ for the interest rate

[3] In its simplest form, the efficient markets hypothesis can be reduced to a joint hypothesis that foreign exchange market participants are, in an aggregate sense, a) endowed with rational expectations and b) risk-neutral. The hypothesis can be modified to adjust for risk, so that it then becomes a joint hypothesis of a model of equilibrium returns (which may admit risk premia) and rational expectations.

differential $i_{t,k} - i^*_{t,k}$, a number of researchers have tested UIP by estimating a regression of the form:

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

where we have assumed $k = 1$ for simplicity, and υ_{t+1} is a disturbance term. Under UIP, the slope parameter β must equal unity, and the disturbance term υ_{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 and the efficient markets hypothesis (e.g. see the references in the survey of Hodrick, 1987; Lewis, 1995; Taylor, 1995; Engel, 1996). Indeed it constitutes a stylized fact that estimates of β, using exchange rates against the dollar, are often statistically insignificantly different from zero and generally closer to minus unity than 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 at a certain term k, the less the home currency is predicted to depreciate over the k periods to maturity.[4] The negative value of β is the central feature of the forward bias puzzle, and, following much previous literature, we shall refer to equation (2) as the 'Fama regression' later in the paper.[5]

It is also worth noting that the relevant literature has 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:

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

where the excess returns $ER_{t+1}^1 \equiv \Delta s_{t+1} - \left(f_t^1 - s_t \right) \equiv s_{t+1} - f_t^1$. 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

[4] 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.

[5] Attempts to locate the source of this failure of the risk-neutral efficient markets hypothesis either in the presence of stable, significant and plausible risk premia, or in some sense in the failure of rational expectations when applied to the foreign exchange market as a whole, have also met with limited success - see the surveys of Hodrick (1987), Lewis (1995) and Engel (1996), and the references therein.

negative estimate of β in equation (2), is that β^{τ} is negative and massively statistically significant. 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 5, where we will show that both these stylized facts can indeed be matched using a nonlinear model of regime-dependent UIP deviations which takes into account the existence of limits to foreign exchange speculation.

2.2 Limits to Speculation and Nonlinearity in the Fama Regression: A Brief Overview

The idea that there may be nonlinearities in deviations from the Fama regression or from UIP 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.[6] Sercu and Wu (2000) derive, in a partial equilirbium model, an expression for the spot-forward relationship where, in the presence of transactions costs, expected exchange rate changes and forward premia are imperfectly aligned even in the absence of a risk premium, inducing nonlinearity in the spot-forward relationship and implying that β may be different depending on the size of the deviation from UIP. 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 interventions are not announced and take the market by surprise.

A related, albeit different, rationalization of nonlinearity in the spot-forward relationship stems

[6] E.g. see Dumas (1992, p. 174, equation 23); see also Baldwin (1990), Hollifield and Uppal (1995) and Obstfeld and Rogoff (2000).

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. 209-220). 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 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 Sharpe ratio for a buy-and-hold equity strategy has averaged about 0.4 on an annual basis for the US[7], a currency trading strategy yielding a Sharpe ratio lower than 0.4 would not be worth taking up. 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 only when β departs from unity that the numerator of the Sharpe ratio becomes positive.[8] Indeed, 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 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 deviations from UIP are too small to attract speculative capital and, therefore, do not imply any glaring profitable 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. It is only for values of β outside the band of inaction that the forward bias will attract capital and a relationship between spot and forward rates consistent with UIP can be established. 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

[7] The 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 exact same figure of 0.4 is reported by Sharpe (1994, p. 51).

[8] The numerator is just the difference between 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 covered 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 rates strategies, also the covariances among the exchange rates considered in the currency strategy.

cost of taking up a currency strategy exceeds the marginal benefit.

The crucial implication of the above analysis shows 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 and indeed they may even move in opposite directions within a bounded interval without giving rise to any glaring profitable opportunities. 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 that 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, 1994).

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, 1994) and characterized by 'limited participation' due to the fact that information costs may limit the participation of some classes of traders in derivatives markets (e.g. 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 - investors like mutual funds, pension funds or insurance companies, who do not have a comparative advantage in implementing pure currency strategies to exploit the forward bias over, for example, proprietary bank traders - 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 size. 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. If the bounds are violated by a relatively greater amount, then progressively more agents will enter the market to effect trades. Thus, the forces pushing β within the band of inaction will increase as the deviation from the bounds increases since an increasing number of agents face profitable opportunities, implying the possibility of a smooth transition of β back towards the bounds (or unity, which is the centre of the band of inaction) such that the speed of reversion of the deviations from UIP towards zero increases with the degree of violation of the band of inaction itself (e.g. see Dumas, 1992, 1994).[9]

Overall, the arguments discussed above suggest that limits to speculation create a band where UIP does not hold and spot and forward may be unrelated or even move in different directions, but deviations from UIP can stray beyond the thresholds. Once beyond the upper or lower threshold, deviations from UIP become increasingly mean reverting with the distance from the threshold. 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 Dumas (1994), 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.

At the empirical level, although there is no study - to the best of our knowledge - which has formally investigated the role of the nonlinearities induced by the above arguments in explaining the forward bias puzzle, there are several studies which have documented that small deviations from UIP (or small forward premia) tend to behave differently from large ones. Bilson (1981) first noted that outlier observations have worse forecasting power than normal ones (defined as forward premia smaller than ten percent in absolute value). However, Flood and Rose (1994) and Flood and Taylor

[9] In other words, one may be tempted to argue that, once a glaring profit opportunity arises, each agent will invest as much as possible to exploit it. However, this is obviously not the case in real-world derivatives markets since even arbitrage may be risky for a number of reasons, including the existence of margin requirements and position limits. For example, Liu and Longstaff (2000) 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.

(1996) present evidence that abnormal observations outperform normal observations in forecasting power, and Huisman, Koedijk, Kool and Nissen (1998) report that the rejection of UIP is not as severe as is commonly found and it 'almost perfectly' holds in periods characterized by large forward premia. Clarida, Sarno, Taylor and Valente (2003) also document, in the context of a forecasting exchange rate model based on the term structure of forward premia, the existence of significant nonlinearities in the dynamic interaction between spot and forward exchange rates and deviations from UIP, which they model using a multivariate regime switching model with shifts in both the intercept and the slope parameters. Our empirical framework extends this research by providing a general econometric characterization of the Fama regression which can test the general predictions of the limits to speculation hypothesis in order to assess its potential to shed light on the forward bias puzzle and the excess returns predictability documented in the literature.

3 A Nonlinear Fama Regression: The Empirical Framework

The key regression of interest in the present study is the Fama regression, given by equation (2). Indeed, this is the most commonly regression used in the literature for the purpose of testing the validity of UIP and for understanding the relationship between spot and forward exchange rates. Also, this is the regression which has generated the stylized fact of a forward bias, essentially amounting to the fact that not only estimates of β are found to be different from the theoretical value of unity but they are typically found to be either negative or statistically insignificantly different from zero.

The discussion in the previous section implies, however, that limits of speculation are likely to generate nonlinear dynamics in the relationship between spot and forward exchange rates or in deviations from UIP, which has important implications for conventional testing procedures based on regression (2). Specifically, limits to speculation may induce a band of inaction where the forward bias might be justifiable on the ground that the departures from UIP are too small to attract speculative capital and, hence, they may persist. Strictly speaking, this would imply threshold-type behavior for UIP deviations so that it is only when β is 'sufficiently' away from unity that financial institutions can exploit the forward bias and take up profitable currency trading strategies. Threshold models allow for a band of inaction where no adjustment towards UIP takes place, while outside the band the process switches abruptly to become consistent with UIP. While discrete switching of this kind is appealing, discrete adjustment to UIP would clearly be most appropriate only when financial institutions have identical features of, for example, limits to speculation, transactions costs, position limits, that is when they are essentially homogeneous. Hence, smooth rather than discrete adjustment may be more appropriate since, as suggested by Dumas (1994), Teräsvirta (1994) and Granger and Lee (1999), time aggregation and, most importantly, nonsynchronous adjustment by heterogeneous agents is likely to result in smooth aggregate regime switching.

A characterization of nonlinear adjustment in the Fama regression which allows for smooth rather than discrete adjustment is in terms of a smooth transition regression (STR) model (Granger and Teräsvirta, 1993; Teräsvirta, 1994, 1998). In the STR model, adjustment takes place in every period but the speed of adjustment varies with the extent of the deviation from UIP. An STR model may be written as follows:

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

where ε_{t+1} is a disturbance term. The transition function $\Phi \left[ER_t^1, \gamma\right]$ 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, assumed to be the excess return ER_t^1 or, equally, the deviation from UIP.[10]

A simple transition function suggested by Granger and Teräsvirta (1993) and Teräsvirta (1994, 1998), which is particularly attractive in the present context, is the exponential function:

$$\Phi \left[ER_t^1, \gamma\right] = \left\{1 - \exp\left[-\gamma \left(ER_t^1\right)^2\right]\right\}, \tag{5}$$

in which case (4) would be termed an exponential STR or ESTR model. The exponential transition function is bounded between zero and unity, $\Phi : \Re \rightarrow [0, 1]$, has the properties $\Phi[0] = 0$ and $\lim_{x \rightarrow \pm\infty} \Phi[x] = 1$, and is symmetrically inverse–bell shaped around zero. These properties of the ESTR model are 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.[11] The transition parameter γ determines the

[10] In fact, in our empirical work we employed several possible transition variables, also including the forward premium $\left(f_t^1 - s_t\right)$ and the interest rate differential $\left(i_t - i_t^*\right)$, both current and lagged. We found that the rejection of linearity of the Fama regression obtained when using the excess return, ER_t^1 was stronger and hence we only report these empirical results. However, we carried out the full empirical work reported in this paper also using as a transition variable the forward premium $\left(f_t^1 - s_t\right)$, and recorded results qualitatively very similar to the ones obtained when using the excess return, ER_t^1 as transition variable (full details available upon request).

[11] Clearly, the class of nonlinear models is infinite, and this paper focuses on the ESTR formulation primarily because

speed of transition between the two extreme regimes, with lower absolute values of γ implying slower transition.

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 inner regime corresponds to $ER_t^1 = 0$, where $\Phi(\cdot) = 0$ 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 outer regime corresponds, for a given γ, to $\lim_{ER_t^1 \to \pm\infty} \Phi\left[ER_t^1, \gamma\right]$, 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 of course 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 2. 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 excess returns are too small to attract speculative capital. However, for larger deviations from UIP at time t (of either sign), financial institutions would take up the glaring profit opportunities provided by currency trading strategies and induce reversion towards the UIP condition. 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 inner 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 5 using Monte Carlo methods.

of these attractive properties, its relative simplicity, and the fact that it seems to be the logical empirical counterpart of the theoretical considerations discussed in Section 2. A further generalization of our model, which we leave for further research, would involve having also a multiplicative function which allows for asymmetry in the dynamics of UIP deviations (Bansal, 1997).

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

$$ER_{t+1}^1 = \left[\alpha_1 + \underbrace{(\beta_1 - 1)}_{\beta^*}\left(f_t^1 - s_t\right)\right] + \left[\alpha_2 + \beta_2\left(f_t^1 - s_t\right)\right]\Phi\left[ER_t^1, \gamma\right] + \varepsilon_{t+1}. \tag{8}$$

The discussion on the effects of limits to speculation in the previous section suggests that the larger the deviation from UIP the stronger will be the tendency to move back to UIP. This implies that while $\beta^* < 0$ is admissible in equation (8), one must have $\beta_2 > 0$ and $\beta^* + \beta_2 = 0$, implying that the forward premium should have no predictive power on future excess returns in the outer regime. That is, for small UIP deviations, excess returns may be persistent and predictable ($\beta^* < 0$) using the information in the lagged forward premium, but for large UIP deviations excess returns are unpredictable ($\beta^* + \beta_2 = 0$).

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, again, 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 inner regime one may find negative and statistically significant estimates of β^τ from estimating equation (3). Again, we shall investigate this issue in Section 5 using Monte Carlo methods.

It is worth noting that 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 or LSTR model which implies asymmetric behavior of the deviations from UIP according to whether they are positive or negative. 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 shall evaluate the adequateness of the linear Fama regression performing tests of linearity against the alternative of smooth transition nonlinearity and following a decision rule due to Teräsvirta (1994, 1998) designed to select the most adequate transition function for modelling nonlinearity in the present context (see Appendix A).

4 Empirical results

4.1 Data, Summary Statistics and the Fama Regression

Our data set comprises weekly observations of spot and 4- and 13-week (or 1- and 3-month) forward US dollar exchange rates among a broad set of countries (Canadian dollar, Japanese yen, UK sterling, German mark, the euro, Swiss franc, Singaporean dollar, Swedish krona, Norwegian krona, and Danish krona). The sample period spans from January 4 1985 to December 31 2002 for all exchange rates except for the German mark (7 January 1986 to 31 December 1998) and the euro (5 January 1999 to 31 December 2002). Following much previous literature (e.g. Hansen and Hodrick, 1980, p. 852), data are Tuesdays of every week, taken from *Datastream*. 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 1- and 3-month forward exchange rates, f_t^1 and f_t^3 respectively, both at weekly and monthly frequency. The core of the empirical work is based on s_t and f_t^1 at weekly frequency, while we shall 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 Table 1, we report sample moments for several combinations of weekly spot and 1-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 the forward premium, the depreciation rate and the excess return have a mean close to zero, both economically and statistically, with a large standard deviation. However, while the first-order autocorrelation coefficient of the depreciation rate is very small in size (never higher than 9 percent) 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 0.855 for Canada) and massively statistically significant, and the corresponding first-order autocorrelation coefficient of the excess return is small (in the range between 0.69 for Singapore and 0.131 for the euro) but sometimes statistically significant. These results are consistent with the stylized facts that the forward premium is a highly persistent process, the depreciation rate is near white noise (or the exchange rate is a near random walk process), and the excess return is mildly serially correlated (e.g. Backus, Gregory and Telmer, 1993).[12]

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 in nine of the ten regressions estimated and it is often statistically insignificantly different from zero. The exception is the German mark, where the estimate of β is positive (about 0.32) and statistically significant, but this estimate does not comprise the theoretical value of unity 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 Telmar, 1993), we find that, for each exchange rate, β^{τ} is massively statistically significant, 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.

4.2 Linearity Tests

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 nonlinearity, using the excess return, ER_t^1 as the transition variable. We followed the Teräsvirta (1994, 1998) decision rule to select the most adequate transition function for modelling nonlinearity in the present context (see Appendix A). 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 ESTR and LSTR formulations led us to conclude that an ESTR 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 2, and with the evidence documented in the literature that exchange rate movements may behave differently depending on the absolute size of

[12] We also tested for unit root behavior of the spot rate and the forward rate time series examined by calculating standard augmented Dickey-Fuller (ADF) test statistics. In each case, the number of lags was chosen such that no residual autocorrelation was evident in the auxiliary regressions. In keeping with the very large number of studies of unit root behavior for these time series, we were in each case unable to reject the unit root null hypothesis at conventional nominal levels of significance. On the other hand, differencing the 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.

their deviation from zero (e.g. Bilson, 1981; Flood and Rose, 1994; Flood and Taylor, 1996; Huisman, Koedijk, Kool and Nissen, 1998). Thus, the strength of the reversion to UIP depends on the extent of the disequilibrium, i.e. the size of the excess return.

4.3 The Nonlinear Fama Regression: Estimation Results

Given the results from the linearity tests, we proceed to 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 shall then test formally below). In estimation, we followed the recommendation of Granger and Teräsvirta (1993) and Teräsvirta (1994, 1998) of standardizing the transition parameter γ (and indeed the transition variable ER_t^1) by dividing it by the sample variance of the transition variable, $\widehat{\sigma}_{ER}^2$, and using a starting value of unity for the estimation algorithm. Also, since this standardization applies to the transition variable, which becomes $ER_t^1/\widehat{\sigma}_{ER}^2$, the transition variable has a natural interpretation in terms of Sharpe ratio, which tightens the link between the empirical framework and the limits-to-speculation hypothesis.[13]

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 strongly 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 strong rejection 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).[14] 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 $\Phi(\cdot) = 1$. 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 positive value of β was found. In turn, these values imply that, for small excess returns (the transition variable), UIP does not hold and we observe a forward bias, while for increasingly larger values of excess returns (when deviations from UIP are expected to attract speculative capital),

[13] We repeated the estimation procedure several times using different sets of starting values for the parameters in order to ensure that the results were robust to the specification search rule and that a global optimum was achieved.

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

reversion to UIP can occur very rapidly. These findings also imply that, since reversion to UIP occurs rapidly for large excess returns, the bulk of the observations of the deviations from UIP is in the inner 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 5.

The estimated transition parameters also imply well-defined and very fast transition functions. These are shown in Figure 1, which displays the plots of the estimated transition functions, $\Phi(\cdot)$ against the transition variable ER_t for each exchange rate. The limiting case of $\Phi(\cdot) = 1$ is attained in each case except one (the euro), which is very 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 very close to zero, i.e. in the inner regime.

A battery of diagnostic tests is reported in Panel B of Table 4. Using a likelihood ratio (LR) test for the null hypothesis that $\alpha_2 = -\alpha_1$ and $\beta_2 = 1 - \beta_1$, reported in the first column, in no case could we reject at the five percent significance level the validity of these restrictions. Indeed, these restrictions imply an equilibrium log-level in the model which is exactly consistent with deviations from UIP being equal to zero (market efficiency), in the neighborhood of which we observe a value of the slope parameter consistent with the forward bias (always negative except for Germany), while the slope parameter tends increasingly fast to the theoretical value of unity with the absolute size of the deviations from UIP. The second, third and fourth columns of the Panel B of Table 4 report tests for residual serial correlation, constructed as in Eithreim and Teräsvirta's (1996). In no case, these tests were found to be statistically significant at conventional nominal levels of significance. For each of the estimated nonlinear Fama regressions, we then tested the null hypothesis of no remaining nonlinearity (F_{NRN}), again constructed as in Eithreim and Teräsvirta's (1996) and reported in the last 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.

Overall, the nonlinear estimation results uncover strong evidence in favor of the presence of nonlinearities in the relationship between spot and forward exchange rates, with UIP deviations adjusting towards their zero equilibrium level at a speed which depends upon the absolute size of the deviation itself. The estimated models are in every case statistically well determined and consistent with the priors established by the limits to speculation arguments made in Section 2.2, and pass a battery of diagnostic tests. The bottom line is that our model is consistent with the forward bias characterizing only small departures from UIP, while for large excess returns UIP holds. 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 inner regime, UIP does *not* hold most of the time. However, our model implies that UIP does not hold when 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 massive 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. Before turning to a finer interpretation of our results, we first discuss some robustness exercises.

4.4 Robustness Results

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 B. We re-estimated the nonlinear Fama regression (4) for each exchange rate examined using a 3-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 1-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 B1 show that the estimates obtained are, in each case, qualitatively very similar to the results reported using a 1-month forward contract at the weekly frequency given in Table 4. 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 is strong.

We also addressed thoroughly the question of the robustness of our linearity tests results. The main concern involves the possibility of a spurious rejection of the linearity hypothesis under the test statistics F_L, F_3, F_2, and F_1 applied to the Fama regression (2) in finite sample. We addressed this issue by executing a battery of Monte Carlo experiments constructed using $5,000$ replications in each 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 it is possible that one would reject the linear Fama regression when in fact that were 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 a sample size 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 very 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, all of the test statistics display no 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 an exponential (ESTR) and a 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).

The main result arising from these simulations for our purposes is that it is unlikely, in light of the documented size and power properties, that we are detecting spurious nonlinearities in this paper in that 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.

4.5 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 know forward bias within a certain range of excess returns (e.g. Lyons, 2001). According to the limits-to-speculation hypothesis, for small excess returns 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 excess returns 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 the spot and forward exchange rates of ten dollar exchange rates have behaved in this fashion over the 1985-2002 sample period. This evidence was found to be robust both to the frequency of the data and to the maturity of the forward contract used in our sensitivity analysis.

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. It is rather remarkable how the average Sharpe ratio across the ten dollar exchange rates examined is very close to 0.4, the value Lyons (2001, p. 210) suggests as a useful benchmark in thinking of a minimum threshold level of Sharpe ratios above which speculative capital might begin to be allocated to currency trading strategies rather than, say, a buy-and-hold equity strategy.[15] The asymptotic standard errors of these average Sharpe ratios, reported in parentheses in Table 5, also make apparent the huge variability of Sharpe ratios in the foreign exchange market. The last column in Table 5 gives the number of observations for which the Sharpe ratio is below 0.4 (roughly its mean), indicating that most of the observations are indeed below the average Sharpe ratio, for each exchange rate - in the range between 54 percent for the Swiss franc and 71 percent for the Singaporean dollar. If any Sharpe ratio below 0.4 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.

Although a Sharpe ratio of 0.4 may be a threshold level where some financial institutions may begin to consider taking up speculative positions in the foreign exchange market, institutions' heterogeneity and their tendency to wait for sufficiently large mispricing before entering the market imply that

[15] This is because a US buy-and-hold equity strategy yields on average a Sharpe ratio of about 0.4 (see Sharpe, 1994; Lyons, 2001).

switching to trading in the foreign exchange market is not discrete but smooth, which is the kind of behavior captured by our proposed nonlinear model, where reversion to UIP (or dissipation of the forward bias) occurs at an increasingly higher speed the larger the foreign exchange excess return. 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 the upper part of Figure 2, we plot the estimated transition function of the dollar-yen and dollar-sterling, as representative exchange rates, over the sample. The plots make clear how the model implies that the spot-forward regressions (or the deviations from UIP) are in the inner regime (say when $\Phi(\cdot) \leq 0.5$) most of the time. The inner regime is the one characterized by a very persistent forward bias, which, however, is associated with low and economically unimportant Sharpe ratios, plotted in the bottom part of Figure 2. 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 (the foreign exchange excess return) is relatively low. On the other hand, although fewer observations are in the outer regime (say when $\Phi(\cdot) > 0.5$), in this regime the deviations from UIP are characterized by very little persistence, suggesting that speculative forces induce very 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 inner regime), but our analysis reveals that these observations are characterized by small and economically unimportant departures from UIP. Put another way, the massive 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.

In the graphs of the Sharpe ratios given in the bottom part of Figure 2, we also draw two straight lines, one corresponding to the value of 0.4 suggested by Lyons and one corresponding to the minimum level of the Sharpe ratio which will lead to a shift from the inner regime to the outer regime - defined as the value of the transition function $\Phi(\cdot) > 0.5$. The full calculations are given, for each exchange rate, in the middle column of Table 5. Clearly, while 0.4 (the value that Lyons suggested conservatively as a minimum Sharpe ratio necessary to attract speculative capital) is not sufficient to induce the shift to the outer regime, the range of the minimum level required goes from 0.6 for Switzerland to 1.28 for the euro. Indeed, this evidence seems remarkably consistent with the argument made by Lyons (2001, p. 215), on the basis of interviews with several priopprietary traders and desk managers, that restoration of the UIP equilibrium condition 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 about one and certainly not any value below 0.5.

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

Given our discussion in Section 3 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 resulted in recording a forward bias ($\beta \neq 1$) when in fact the forward bias may characterize only statistically and 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).

5.1 Matching the Forward Bias

We executed a number of Monte Carlo experiments based on an artificial DGP identical to the estimated nonlinear Fama regression (2), calibrated on the estimates reported in Table 4, with independent and identically distributed Gaussian innovations.[16] 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 observations available to us. 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, matching the number of observations available for these exchange rates. For each generated sample

[16] The assumption of Gaussianity is fairly mild given that statistical tests revealed no evidence of non-normality or heteroskedasticity in the estimated residual series from our nonlinear Fama regressions.

of observations we then estimated the standard linear Fama regression (2). In Panel A of Table 6, we recall 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 percentile 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 percentile 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.

Overall, our simulations show that, if the data were generated from a nonlinear spot-forward model characterized by an inner regime with tiny but economically small UIP deviations and an outer regime where UIP holds exactly, estimation of the conventional linear Fama regression would yield a massive rejection of UIP, the finding of forward bias, and estimates which are very close to the ones observed in actual data.

5.2 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 2, since 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, for each of the $5,000$ generated samples of observations we estimated a regression of form (3). In Panel B of Table 6, 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 $\bar{t}(\delta)^{MC}$ together with the 5th and 95th percentile 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-statistic for the significance of estimate of β^{τ} recorded on average would be very close to the one estimated on actual exchange rate data, it would be strongly statistically significant and it would lie in the interval between the 5th and 95th percentile of the empirical distribution of $t(\delta)^{MC}$, for each exchange rate examined. On average, the t-statistics recorded are indeed massively significant. Finally, in the last column of Panel B, we report the p-value from a formal test statistic for the null hypothesis that $t(\delta) = \bar{t}(\delta)^{MC}$, termed t_1. The p-value is generally very 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 ($\bar{t}(\delta)^{MC}$) one would obtain from estimating the predictability regression (3) using the artificial data we generated.

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 the validity of UIP, to record a forward bias (β different from unity and possibly negative), and to find evidence of predictability of excess returns using the information in the lagged forward premium. However, these three features - violation of UIP, forward bias, and predictability of excess returns - are features which the DGP we study only has in the inner regime, which is a regime where deviations from UIP are tiny enough to be economically unimportant and likely to be unable to attract speculative capital.

6 Conclusions

Our empirical results provide strong confirmation that ten major real bilateral dollar exchange rates are linked to forward premia in a nonlinear fashion in the context of a model for UIP deviations which allows for time-variation in the forward bias and nonlinear reversion towards UIP. The nonlinearities we uncover are consistent with a model of deviations from UIP with two extreme regimes: an inner regime with persistent but tiny deviations from UIP, and an outer 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 they are not large enough to attract speculative capital.

The estimated models imply an equilibrium level of the spot-forward regression in the neighborhood of which there is statistically significant forward bias, but such forward bias dissipates with the absolute size of the deviation from the UIP condition. This 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 relevant literature.

Our results therefore allow us to end this study by making, with some degree of caution, three statements. First, the statistical rejection of UIP recorded by the literature may be less indicative of major inefficiencies in the foreign exchange market than it 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 will be more persistent for small deviations from UIP, that is the closer 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 and represent an interesting avenue for further research.

Although our results have been shown to be robust to a number of relevant tests, some 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 precise test of this specific hypothesis, but rather a test of its general predictions. 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 forward premia which is able capture 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. 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 inner regime rather than being, for example, in the middle of the inaction range around unity. Explaining this finding requires 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.[17] Finally, we have judged the economic significance of deviations from UIP as being small in this paper on the basis of their size and persistence properties and, hence, the implied size and persistence of Sharpe ratios from currency strategies. This choice was made both because of the simplicity and intuitive sense of Sharpe ratio calculations and because this is the reasoning used in the limits to speculation hypothesis which motivates our empirical work. However, our conclusions with respect to the economic importance of UIP deviations would benefit from further analysis designed to measure specifically the economic value of deviations from UIP using measures other than Sharpe ratios in order to assess the robustness of our findings.

[17] Lyons (2001, p. 216-8) provides a first step along those lines, proposing a meta-model that determines the exchange rate when pure currency speculation does not occur.

Table 1. Summary Statistics

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

	mean		standard deviation		AR(1)	
Canada	-0.0009	(0.0002)	0.0192	(0.0008)	0.855	(0.042)
Japan	0.0024	(0.0003)	0.0036	(0.0011)	0.614	(0.094)
UK	-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)
Singapore	0.0013	(0.0002)	0.0026	(0.0010)	0.421	(0.093)
Sweden	-0.0023	(0.0004)	0.0041	(0.0022)	0.831	(0.036)
Norway	-0.0026	(0.0004)	0.0043	(0.0020)	0.775	(0.044)
Denmark	-0.0013	(0.0004)	0.0037	(0.0021)	0.797	(0.053)

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

	mean		standard deviation		AR(1)	
Canada	-0.0001	(0.0001)	0.0069	(0.0033)	-0.060	(0.040)
Japan	0.0008	(0.0006)	0.0159	(0.0049)	0.077	(0.030)
UK	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)
Singapore	0.0002	(0.0002)	0.0071	(0.0034)	0.003	(0.044)
Sweden	0.0001	(0.0005)	0.0141	(0.0051)	0.028	(0.034)
Norway	0.0003	(0.0004)	0.0138	(0.0043)	0.031	(0.029)
Denmark	0.0004	(0.0005)	0.0146	(0.0040)	0.061	(0.027)

(continued ...)

(... Table 1 continued)

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

	mean		standard deviation		AR(1)	
Canada	-0.0007	(0.0003)	0.0073	(0.0021)	0.090	(0.045)
Japan	0.0002	(0.0007)	0.0159	(0.0049)	0.117	(0.028)
UK	-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)
Singapore	0.0010	(0.0003)	0.0077	(0.0029)	0.069	(0.054)
Sweden	-0.0023	(0.0007)	0.0147	(0.0050)	0.088	(0.033)
Norway	-0.0029	(0.0007)	0.0145	(0.0044)	0.098	(0.029)
Denmark	-0.0018	(0.0007)	0.0152	(0.0042)	0.117	(0.027)

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 1 January 1985 to 31 December 2002 for all exchange rates except for the German mark (7 January 1986 to 31 December 1998) and the euro (5 January 1999 to 31 December 2002). Figures in parentheses are standard errors calculated by using an autocorrelation and heteroskedasticity consistent matrix of residuals (Newey and West, 1987).

Table 2. Forward Premium (Fama) Regressions

	α	SE(α)	β	SE(β)	s.e.	T	$t(\delta)$
Canada	-0.0004	(0.0002)	-0.2497	(0.0866)	0.006	939	-9.264
Japan	0.0015	(0.0005)	-0.2865	(0.1586)	0.015	939	-6.742
UK	-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
Singapore	0.0006	(0.0003)	-0.3101	(0.1288)	0.007	939	-13.249
Sweden	0.0001	(0.0005)	0.0254	(0.2348)	0.014	939	-7.384
Norway	0.0001	(0.0005)	-0.0725	(0.1378)	0.013	939	-8.155
Denmark	0.0003	(0.0005)	-0.1187	(0.1065)	0.014	939	-8.176

Notes: The table shows the results from estimating, by ordinary least squares, the conventional forward premium (Fama) regression (2): $\Delta s_{t+1} = \alpha + \beta \left(f_t^1 - s_t \right) + e_{t+1}$. 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 e_{t+1}; and T is the number of usable observations. The last column reports the t-statistic (namely $t(\delta)$) for the parameter β^τ in regression (3): $ER_{t+1}^1 = \alpha + \beta^\tau \left(f_t^1 - s_t \right) + e_{t+1}$, where $ER_{t+1}^1 \equiv \Delta s_{t+1} - \left(f_t^1 - s_t \right) \equiv s_{t+1} - f_t^1$, and $\beta^\tau = \beta - 1$.

Table 3. Linearity Tests on the Fama Regression

	F_L	F_3	F_2	F_1
Canada	0.035	0.078	0.016	0.798
Japan	0.020	0.077	0.006	0.142
UK	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
Singapore	1.21×10^{-28}	0.524	0.027	0.029
Sweden	9.87×10^{-11}	0.317	0.004	0.098
Norway	2.91×10^{-4}	0.304	0.002	0.977
Denmark	0.012	0.948	0.040	0.710

Notes: The table reports the p-values from applying the linearity testing procedure suggested by Teräsvirta (1994, 1998). F_L, F_3, F_2 and F_1 statistics are linearity tests constructed as described in Appendix A. The transition variable is the time-t excess return, $ER_t^1 \equiv \left(s_t - f_{t-1}^1 \right)$. The p-values were calculated using the appropriate F distribution.

Table 4. Nonlinear Fama Regressions: ESTR Estimation Results

Panel A. Parameter estimates

	$\alpha_1 = -\alpha_2$	SE	$\beta_1 = 1 - \beta_2$	SE	γ	SE(γ)	
Canada	-0.0005	(0.0003)	-0.5213	(0.1677)	0.3871	(0.0331)	[0]
Japan	0.0018	(0.0008)	-0.5719	(0.2278)	0.4014	(0.0501)	[0.0010]
UK	-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]
Singapore	0.0010	(0.0005)	-0.5981	(0.2035)	0.1081	(0.0050)	[0.0340]
Sweden	-0.0004	(0.0006)	-0.3305	(0.1369)	0.2893	(0.0893)	[0.0002]
Norway	-0.0004	(0.0006)	-0.4582	(0.1805)	0.2514	(0.0200)	[0]
Denmark	0.0002	(0.0006)	-0.3870	(0.1408)	0.2358	(0.0119)	[0.0016]

Panel B. Diagnostic tests

	LR	$LM(1)$	$LM(4)$	$LM(8)$	F_{NRN}
Canada	0.940	0.149	0.060	0.306	0.944
Japan	0.901	0.190	0.173	0.342	0.903
UK	0.914	0.971	0.957	0.951	0.969
Germany	0.937	0.414	0.636	0.855	0.940
Euro	0.897	0.314	0.179	0.161	0.998
Switzerland	0.449	0.133	0.591	0.774	0.994
Singapore	0.918	0.709	0.066	0.241	0.999
Sweden	0.907	0.695	0.791	0.718	0.897
Norway	0.297	0.345	0.726	0.505	0.547
Denmark	0.911	0.100	0.455	0.831	0.955

Notes: *Panel A.* The table shows the results from the nonlinear forward premium regression $\Delta s_{t+1} = \left[\alpha_1 + \beta_1\left(f_t^1 - s_t\right)\right] + \left[\alpha_2 + \beta_2\left(f_t^1 - s_t\right)\right]\Phi\left[ER_t^1, \gamma\right] + \varepsilon_{t+1}$, where $\alpha_2 = -\alpha_1$, $\beta_2 = 1 - \beta_1$ and $\Phi\left[\left(f_{t-1}^1 - s_t\right), \gamma\right] = \left\{1 - \exp\left[-\gamma\left(ER_t^1\right)^2\right]\right\}$. Values in parentheses (SE) are asymptotic standard errors 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.* LR is the likelihood ratio test for the null hypothesis that $\alpha_2 = -\alpha_1$, $\beta_2 = 1 - \beta_1$. $LM(q)$ is the LM-type test statistic for the null hypothesis of no residual autocorrelation up to order q, constructed as in Eitrheim and Teräsvirta (1996). F_{NRN} is the test for the null hypothesis of no remaining nonlinearity, constructed as in Eitrheim and Teräsvirta (1996). For all test statistics, we report p-values.

Table 5. Sharpe Ratios

	average Sharpe ratio		*min* Sharpe	%Obs Sharpe ≤ 0.40
Canada	0.4053	(0.3299)	0.6952	59.0
Japan	0.3985	(0.3381)	0.6792	58.2
UK	0.3928	(0.3523)	1.2248	61.7
Germany	0.4018	(0.3296)	0.5877	59.7
Euro	0.4157	(0.3133)	1.2769	54.5
Switzerland	0.4074	(0.3237)	0.6042	56.7
Singapore	0.3413	(0.3987)	1.3143	71.2
Sweden	0.4032	(0.3382)	0.8034	58.8
Norway	0.4069	(0.3408)	0.8629	59.6
Denmark	0.4080	(0.3281)	0.8911	57.6

Notes: The table reports the in the first column the mean of the annualized Sharpe ratios calculated as realized standardized excess returns over the sample period. Values in parentheses are asymptotic standard errors. *min* Sharpe is the minimum value of the Sharpe ratio which leads to a shift from the inner regime to the outer regime defined as the value of the transition function $\Phi(\cdot) > 0.5$. The last column reports the percentage of observations where the annualized Sharpe ratio is lower than 0.40.

Table 6. Monte Carlo Results: Matching the Stylized Facts

Panel A. Matching the forward bias puzzle

	α	β	$\overline{\alpha}^{MC}$		$\overline{\beta}^{MC}$		$t(\alpha)$	$t(\beta)$
Canada	-0.0004	-0.2497	-0.0004	(-0.0008,0.0003)	-0.1100	(-0.3284,0.1102)	0.968	0.293
Japan	0.0015	-0.2865	0.0013	(0.0002,0.0025)	-0.1588	(-0.4737,0.1492)	0.848	0.501
UK	-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
Singapore	0.0006	-0.3101	0.0009	(0.0004,0.0013)	-0.4279	(-0.5918,-0.2653)	0.302	0.238
Sweden	0.0001	0.0254	-0.0003	(-0.0012,0.0005)	-0.0222	(-0.2414,0.1986)	0.412	0.720
Norway	0.0001	-0.0725	-0.0003	(-0.0013,0.0004)	-0.1408	(-0.3583,0.0754)	0.444	0.607
Denmark	0.0003	-0.1187	0.0002	(-0.0006,0.0010)	-0.1193	(-0.3419,0.1048)	0.777	0.996

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

	$t(\delta)$	$\overline{t}(\delta)^{MC}$		t_1
Canada	-9.264	-8.423	(-6.381,-10.539)	0.509
Japan	-6.742	-6.147	(-4.382,-8.044)	0.592
UK	-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
Singapore	-13.249	-14.354	(-12.265,-16.574)	0.405
Sweden	-7.384	-7.743	(-5.789,-9.780)	0.768
Norway	-8.155	-8.645	(-6.705,-10.698)	0.689
Denmark	-8.176	-8.310	(-6.350,-10.355)	0.912

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 of the coefficients calculated α, β obtained from estimating when 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 fractiles 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.* $t(\delta)$ 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). $\overline{t}(\delta)^{MC}$ is the average of the empirical distribution of the t-statistic for the significance of the parameter δ on forward premium in a predictive regression of the form in $ER_t = v + \delta\left(f_t^1 - s_t\right) + error$, calculated under the same DGP as above using 5,000. Values in parentheses correspond to the 5th and 95th fractiles of the empirical distribution of the test statistics $t(\delta)$. t_1 is the p-value of the test statistic for the null hypothesis that $\overline{t}(\delta)^{MC} = t(\delta)$.

A Appendix: Linearity Testing

This appendix describes the procedure employed to test for linearity briefly discussed in Section 3 and employed in Section 4.2.

Assuming that a plausible transition variable is ER_t^1, the appropriate auxiliary regression for the linearity tests against a STR alternative, which is an important preliminary to the specification and estimation of nonlinear Fama regression (4), is the following:

$$\widehat{e}_{t+1} = \boldsymbol{\vartheta}'_0 \mathbf{A}_{t+1} + \boldsymbol{\vartheta}'_1 \mathbf{A}_{t+1} ER_t^1 + \boldsymbol{\vartheta}'_2 \mathbf{A}_{t+1} ER_t^1 + \boldsymbol{\vartheta}'_3 \mathbf{A}_{t+1} ER_t^1 + innovations, \tag{A1}$$

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 the 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, ER_t^1 (see Granger and Teräsvirta, 1993; Teräsvirta, 1994, 1998). The general test for linearity against STR is then the ordinary F-test of the null hypothesis:

$$H_{0L} : \boldsymbol{\vartheta}'_1 = \boldsymbol{\vartheta}'_2 = \boldsymbol{\vartheta}'_3 = \mathbf{0} \tag{A2}$$

The choice between a LSTR and an ESTR model is based on a sequence of nested tests within (A2). First, the null hypothesis H_{0L} in (A2) must be rejected using an ordinary F-test (F_L). Then the following hypotheses are tested:

$$H_{03} \quad : \quad \boldsymbol{\vartheta}'_3 = \mathbf{0} \tag{A3}$$

$$H_{02} \quad : \quad \boldsymbol{\vartheta}'_2 \mid \boldsymbol{\vartheta}'_3 = \mathbf{0} \tag{A4}$$

$$H_{01} \quad : \quad \boldsymbol{\vartheta}'_1 \mid \boldsymbol{\vartheta}'_2 = \boldsymbol{\vartheta}'_3 = \mathbf{0}. \tag{A5}$$

Again, an F-test is used, with the corresponding test statistics denoted F_3, F_2, and F_1, respectively, and the decision rule is as follows: after rejecting H_{0L} in (A2), the three hypotheses (A3)-(A5) are tested using F-tests; if the test of (A4) has the smallest p-value, an ESTR is chosen, otherwise an LSTR is selected (see Granger and Teräsvirta, 1993; Teräsvirta, 1994, 1998).

B Appendix: 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	$\beta_1 = 1 - \beta_2$	SE	γ	SE(γ)	
Canada	-0.0016	(0.0008)	-0.6354	(0.1511)	0.3622	(0.0246)	[0]
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]
Singapore	0.0034	(0.0014)	-0.7480	(0.2336)	0.6381	(0.0719)	[0.0040]
Sweden	-0.0012	(0.0020)	-0.7698	(0.3770)	0.6815	(0.1762)	[0]
Norway	-0.0012	(0.0020)	-0.6928	(0.2725)	0.3721	(0.0332)	[0.0002]
Denmark	0.0009	(0.0018)	-0.6494	(0.2503)	0.3414	(0.0288)	[0.0008]

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

	$\alpha_1 = -\alpha_2$	SE	$\beta_1 = 1 - \beta_2$	SE	γ	SE(γ)	
Canada	-0.0029	(0.0016)	-1.7866	(0.8232)	0.8874	(0.3563)	[0.0004]
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]
Singapore	0.0032	(0.0012)	-1.4998	(0.3920)	0.2735	(0.0338)	[0.0004]
Sweden	0.0008	(0.0035)	-2.2279	(1.0775)	2.7569	(0.8617)	[0.0002]
Norway	-0.0029	(0.0034)	-2.5385	(1.0363)	1.2298	(0.4919)	[0]
Denmark	0.0009	(0.0028)	-1.8015	(0.7895)	0.6128	(0.2791)	[0.0012]

Notes: *Panel A.* The table reports the results from estimating the nonlinear forward premium regression $\Delta s_{t+1} = \left[\alpha_1 + \beta_1 \left(f_t^3 - s_t\right)\right] + \left[\alpha_2 + \beta_2 \left(f_t^3 - s_t\right)\right] \Phi \left[\left(f_{t-3}^3 - s_t\right), \gamma\right] + \varepsilon_{t+1}$, where $\alpha_2 = -\alpha_1$, $\beta_2 = 1 - \beta_1$ and $\Phi\left[\left(f_{t-3}^3 - s_t\right), \gamma\right] = \left\{1 - \exp\left[-\gamma \left(f_{t-3}^3 - s_t\right)^2\right]\right\}$, using weekly data. *Panel B.* The table reports the results from estimating the nonlinear forward premium regression $\Delta s_{t+1} = \left[\alpha_1 + \beta_1 \left(f_t^1 - s_t\right)\right] + \left[\alpha_2 + \beta_2 \left(f_t^1 - s_t\right)\right] \Phi \left[\left(f_{t-1}^1 - s_t\right), \gamma\right] + \varepsilon_{t+1}$, where $\alpha_2 = -\alpha_1$, $\beta_2 = 1 - \beta_1$ and $\Phi\left[\left(f_{t-1}^1 - s_t\right), \gamma\right] = \left\{1 - \exp\left[-\gamma \left(f_{t-1}^1 - s_t\right)^2\right]\right\}$ using monthly data. For both Panels A and B, values in parentheses (SE) are asymptotic standard errors 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 5000 replications. 0 denotes p-values lower than 10^{-5}.

Table B2. Empirical Size and Power Properties of the Linearity Tests on Forward Premium Regression: Monte Carlo results

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 trie 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.

References

[1] Aliber, R.Z. (1973), "The Interest Rate Parity Theorem: A Reinterpretation," *Journal of Political Economy*, 81, 1451-59.

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

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

[4] Baldwin, R. (1990), "Re-Interpreting the Failure of Foreign Exchange Market Efficiency Tests: Small Transaction Costs, Big Hysteresis Bands," *National Bureau of Economic Research Working Paper N. 3319*.

[5] Bansal, R. (1997), "An Exploration of the Forward Premium Puzzle in Currency Markets," *Review of Financial Studies*, 10, 369-403.

[6] Bansal, R. and Dahlquist, M. (2000), "The Forward Premium Puzzle: Different Tales from Developed and Emerging Economies," *Journal of International Economics*, 51, 115-44.

[7] Bansal, R., Gallant, A.R., Hussey, R. and Tauchen, G. (1995), "Nonparametric Estimation of Structural Models for High Frequency Currency Market Data," *Journal of Econometrics*, 66, 251-87.

[8] Bekaert, G. (1996), "The Time-Variation of Risk and Return in Foreign Exchange Markets: A General Equilibrium Perspective," *Review of Financial Studies*, 9, 427-70.

[9] Bekaert, G. and Hodrick, R.J. (1993), "On Biases in the Measurement of Foreign Exchange Risk Premiums," *Journal of International Money and Finance*, 12, 115-38.

[10] Bekaert, G. and Hodrick, R.J. (2001), "Expectations Hypotheses Tests," *Journal of Finance*, 56, 1357-94.

[11] Bekaert, G., Hodrick, R.J. and Marshall, D. (1997), "The Implications of First-Order Risk Aversion for Asset Market Risk Premiums," *Journal of Monetary Economics*, 40, 3-39.

[12] Bilson, J.F.O. (1981), "The "Speculative Efficiency" Hypothesis," *Journal of Business*, 54, 435-51.

[13] Fama, E.F. (1984), "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 14, 319-38.

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

[15] Clinton, K. (1988), "Transactions Costs and Covered Interest Arbitrage: Theory and Evidence," *Journal of Political Economy*, 96, 358-70.

[16] Cumby, R.E. (1988), "Is It Risk? Explaining Deviations from Uncovered Interest Parity," *Journal of Monetary Economics*, 22, 279-99.

[17] Cumby, R.E. and Obstfeld, M.E. (1984), "International Interest Rate and Price Level Linkages under Flexible Exchange Rates: A Review of Recent Evidence," in Bilson, J.F.O. and Marston, R.C (eds.). *Exchange Rate Theory and Practice*, National Bureau of Economic Research Conference Report Chicago and London: University of Chicago Press, 121-51.

[18] Domowitz, I. and Hakkio, C. (1985), "Conditional Variance and the Risk Premium in the Foreign Exchange Market," *Journal of International Economics*, 19, 47-66.

[19] Dooley, M.P. and Isard, P. (1980), "Capital Controls, Political Risk, and Deviations from Interest-Rate Parity," *Journal of Political Economy*, 88, 370-84.

[20] Dumas, B. (1992), "Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World," *Review of Financial Studies*, 5, 153-80.

[21] Dumas, B. (1994), "Partial Equilibrium versus General Equilibrium Models of International Capital Market," in van der Ploeg, F. (ed.) *The Handbook of International Macroeconomics*, Oxford: Blackwell, Ch. 10.

[22] Eitrheim, O, and Teräsvirta, T. (1996), "Testing the Adequacy of Smooth Transition Autoregressive Models," *Journal of Econometrics* 74, 59-75.

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

[24] Flood, R.B. and Rose, A.K. (1996), "Fixes: Of the Forward Discount Puzzle," *Review of Economics and Statistics*, 78, 748-52.

[25] Flood, R.B. and Rose, A.K. (2002), "Uncovered Interest Parity in Crisis," *International Monetary Fund Staff Papers*, 49, 252-66.

[26] Flood, R.B. and Taylor, M.P. (1996), "Exchange Rate Economics: What's Wrong with the Conventional Macro Approach?", in Frankel, J.A. and Galli, G. and Giovannini, A. (eds.), *The Microstructure of Foreign Exchange Markets*, National Bureau of Economic Research Conference Report series. Chicago and London: University of Chicago Press, 261-94.

[27] Frankel, J.A. (1980), "Tests of Rational Expectations in the Forward Exchange Market," *Southern Economic Journal*, 46, 1083-1101.

[28] Frankel, J.A. and Engel, C. (1984), "Do Asset Demand Functions Optimize Over the Mean and Variance of Real Returns? A Six Currency Test," *Journal of International Economics*, 17, 309-323.

[29] Frankel, J.A. and MacArthur, A.T. (1988), "Political vs. Currency Premia in International Real Interest Rate Differentials: A Study of Forward Rates for 24 Countries," *European Economic Review*, 32, 1083-114.

[30] Frenkel, J.A. and Levich, R.M. (1975), "Covered Interest Arbitrage: Unexploited Profits?", *Journal of Political Economy*, 83, 325-38.

[31] Frenkel, J.A. and Levich, R.M. (1977), "Transaction Costs and Interest Arbitrage: Tranquil versus Turbulent Periods," *Journal of Political Economy*, 85, 1209-26.

[32] Froot, K.A. and Thaler, R.H. (1990), "Foreign Exchange," *Journal of Economic Perspectives*, 4, 179-92.

[33] Granger, C.W.J. and Teräsvirta, T. (1993), *Modelling Nonlinear Economic Relationships*, Oxford; New York; Toronto and Melbourne: Oxford University Press.

[34] Granger, C.W.J. and Lee, T-H. (1999), "The Effect of Aggregation on Nonlinearity," *Econometric Reviews*, 18, 259-69.

[35] Grossman, S. and Weiss, L. (1983), "A Transactions-Based Model of the Monetary Transition Mechanism," *American Economic Review*, 73, 871-80.

[36] Hansen, L.P. and Hodrick, R.J. (1980), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829-53.

[37] Hirshleifer, D. (1988), "Residual Risk, Trading Costs, and Commodity Futures Risk Premia," Review of Financial Studies, 1, 173-93.

[38] Hollifield, B., and Uppal, R. (1997), "An Examination of Uncovered Interest Parity in Segmented International Commodity Markets," *Journal of Finance*, 52, 2145-2170.

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

[40] Huisman, R., Koedijk, K., Kool, C. and Nissen, F. (1998), "Extreme Support for Uncovered Interest Parity," *Journal of International Money and Finance*, 17, 211-28.

[41] Levich, R.M. (1985), "Empirical Studies of Exchange Rates: Price Behavior, Rate Determination, and Market Efficiency," in Jones, R.W. and Kenen, P.B. (eds.) *The Handbook of International Economics*, vol. II, Amsterdam: North-Holland, 979-1040.

[42] Lewis, K.K. (1995), "Puzzles in International Financial Markets," Grossman, G. M.; Rogoff, K. (eds.) *The Handbook of International Economics*, vol. 3. Amsterdam; New York and Oxford: Elsevier, North-Holland, 1913-71.

[43] Liu, J. and Longstaff, F. (2003), "Losing Money on Arbitrages: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities," *Review of Financial Studies*, forthcoming.

[44] Lyons, R.K. (2001), *The Microstructure Approach to Exchange Rates*, Cambridge and London: MIT Press.

[45] Mark, N.C. (1988), "Time Varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts," *Journal of Financial Economics*, 22, 335-54.

[46] Mark, N.C. and Moh, Y.-K. (2002), "Continuous-Time Market Dynamics, ARCH Effects, and the Forward Premium Anomaly," University of Notre Dame and Tulane University, *mimeo*.

[47] Moh, Y-K. (2002), "Exchange Rate Dynamics in a Continuous-Time Model of Uncovered Interest Parity and the Regulated Jump-Diffusion Interest Differential," Tulane University, *mimeo.*

[48] Neal, R. (1996), "Direct Tests of Index Arbitrage Models," *Journal of Financial and Quantitative Analysis*, 31, 541-62.

[49] Newey, W.K. and West, K.D. (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-8.

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

[51] Sarno, L. and Taylor, M.P. (2002), *The Economics of Exchange Rates*, Cambridge: Cambridge University Press.

[52] Sercu, P. and Wu, X., (2000), "Uncovered Interest Arbitrage and Transaction Costs: Errors-in-Variables versus Hysteresis Effects," University of Leuven and City University of Hong Kong, *mimeo.*

[53] Sharpe, W.F. (1994), "The Sharpe Ratio," *Journal of Portfolio Management*, 21, 49-58.

[54] Shleifer, A. and Vishny, R. (1997), "The Limits to Arbitrage," *Journal of Finance*, 52, 35-56.

[55] Skalin, J. (1998), "Testing Linearity Against Smooth Transition Autoregression Using Parametric Bootstrap," Stockholm School of Economics, *mimeo.*

[56] Sofianos, G. (1993), "Index Arbitrage Profitability," *Journal of Derivatives*, 1, 6-20.

[57] Taylor, M.P. (1987), "Covered Interest Parity: A High-Frequency, High-Quality Data Study," *Economica*, 54, 429-38.

[58] Taylor, M.P. (1989), "Covered Interest Arbitrage and Market Turbulence," *Economic Journal*, 99, 376-91.

[59] Taylor, M.P. (1995), "The Economics of Exchange Rates," *Journal of Economic Literature*, 33, 13-47.

[60] Teräsvirta, T. (1994), "Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89, 208-18.

[61] Teräsvirta, T. (1998), "Modelling Economic Relationship with Smooth Transition Regressions," in Giles, D.E.A. and Ullah, A. (eds.), *Handbook of Applied Economic Statistics*, New York: Marcel Dekker, 507-52.

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

Figure 1. Estimated transition functions

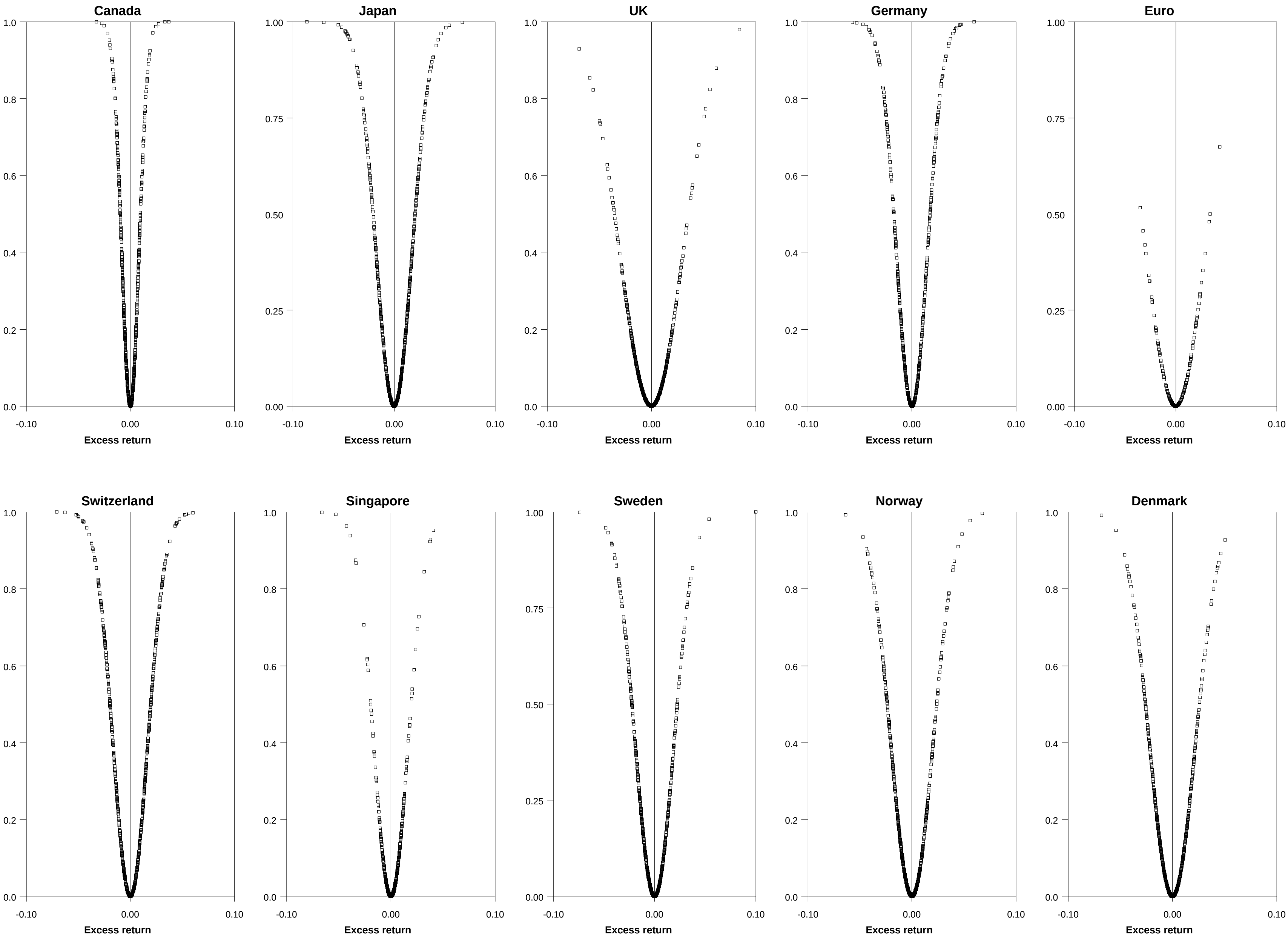

Figure 2. Sharpe ratios and estimated transition functions