# Rethinking the forward premium puzzle in a nonlinear framework

JERRY COAKLEY[a,b] and ANA-MARIA FUERTES[c*]

[a]Department of Accounting, Finance and Management, University of Essex

[b]Department of Economics, Birkbeck College, University of London

[c]Department of Accounting and Finance, City University Business School

May 2001

### Abstract

The forward premium puzzle needs to be reformulated since extant studies address the negative slopes associated with the long dollar swings of the 1980s. By contrast the insignificant coefficients from recent data spans can be explained by an unbalanced regression problem caused by asymmetries in spot returns. These stem from market frictions such as transaction costs and are associated with overshooting of spot rates. Monte Carlo experiments show that asymmetries and overshooting effects produce widely dispersed and statistically insignificant slope coefficients whose small sample mean is close to zero.

JEL classification: C22; F31

Keywords: Asymmetries; Exchange rate overshooting; Monte Carlo

*We thank Mario Levis, Haris Psaradakis, James B. Ramsey, Ron Smith, and participants at the 2000 Annual Conferences of the Society for Nonlinear Dynamics and Econometrics, Computational Finance/Forecasting Financial Markets, and Money, Macro and Finance Group, and at the 2001 Royal Economic Society Annual Conference for helpful comments and discusssions. Responsibility for any errors or omissions is our own.
Corresponding Author: City University Business School, Department of Accounting and Finance, Frobisher Crescent, Barbican Centre, London EC2Y 8HB. Tel +44 207 477 0221. Fax: +44 207 477 8648; E-mail: a.fuertes@city.ac.uk

# 1 Introduction

The forward premium anomaly persists as one of the great puzzles in international finance notwithstanding numerous efforts to resolve it. In a log regression of the spot return — or change in the spot exchange rate — on the forward premium, the reported slope coefficient estimates are typically negative, and often significantly so, instead of unity as implied by the forward rate unbiasedness hypothesis (FRUH). The practical implication is that the spot exchange rate next period moves in the opposite direction to that currently predicted by the forward premium.[1] Understanding the puzzle matters. On one hand, failure to resolve the puzzle suggests researchers still do not fully comprehend the operation of one of the largest and most liquid global financial market in which the leading players are the big international banks. On the other, the puzzle implies that behaviour in foreign (including forward) exchange markets appears inconsistent with intertemporal asset pricing models with plausible levels of risk aversion.

This paper has a number of distinctive features. First it suggests a new definition of the puzzle by reevaluating the empirical evidence from data series which span the 1990s. In so doing, it focuses on the significance as well as the sign and magnitude of the reported regression coefficients. The conclusion is that the traditional interpretation in terms of negative coefficients is associated with data spans dominated by the long dollar swings of the 1980s while longer or more recent data spans point to insignificant slope coefficients. This interpretation is borne out by the results from regressions for the 1976-99 and 1990-99 periods and from rolling five-year regressions. A reconsideration of the findings from other recent studies is also consistent with such an interpretation. The new definition is consistent with the innovative claim of Baillie and Bollerslev (2000) and Maynard and Phillips (1998) that the forward premium bias may be exaggerated even if they arrive at their conclusion via a different route.

Second, it explains the insignificant slope coefficients in terms of an unbalanced regression due to asymmetries in spot returns. This novel rationale is consistent with both empiricl evidence and theoretical considerations.[2]

[1]See Engel (1996) for an interesting survey.

[2]Note that Baillie and Bollerslev (2000) and Maynard and Phillips (1998) also relate their explanations to an unbalanced regression problem but theirs is based on the long memory properties in the forward pemium. We are grateful to Richard Baillie and Peter Phillips for kindly making available unpublished versions of their papers.

Asymmetries are a consequence of transaction costs and market frictions such as short sale constraints which make arbitrage less than perfect. They are supported by evidence of conditional-mean nonlinearities in exchange rates (Clements and Smith, 2001; Kräger and Kugler, 1993), excess returns (Clarida and Taylor, 1997; Coakley and Fuertes, 2001a) and risk premia (Evans and Lewis, 1995; Psaradakis, Sola and Spagnolo, 2000).[3] It follows that deviations from the FRUH or, by implication, uncovered interest parity (UIP) may be more widespread than hitherto realised.

Third, the new formulation is more plausible from a finance viewpoint. It implies that, on average, the forward premium contains no unpriced information on the future spot return. Such an implication dovetails neatly with the notorious difficulty in forecasting nominal exchange rates and beating the random walk model in the econometric literature.[4] Moreover asymmetries of the type uncovered are consistent with the overshooting behaviour of spot exchange rates (Dornbusch, 1976) and other financial series. Such overshooting can plausibly explain the puzzling aspects of deviations from UIP such as sign changes and excess volatility alluded to by Lewis (1995).

Finally Monte Carlo experiments confirm that our proposed rationale can capture the major features of the anomalous regression coefficients. The simulations show that the direct effect of asymmetries in spot returns is to shift the mean of the slope coefficient kernel density close to zero, increase its dispersion by more than 50% and to lead to more than 80% of the coefficients being insignificant. The statistical reason for this downward shift is akin to that for omitted variable bias as in Barnhart, McNown and Wallace (1999) or that for correlation between the explanatory variables and the disturbances as in Psaradakis et al. (2000). However a distinctive aspect of our explanation is that asymmetries of the type observed in spot returns are consistent with overshooting behaviour and excess volatility while the forward premia exhibit no such evidence. Thus, when we add overshooting-induced innovation scale differences between these two sereies to the direct effect of asymmetries, the simulations are able replicate the biased, widely dispersed and insignificant slope coefficients found from spot returns regressions using data from the 1990s.

The remainder of the paper is organised as follows. In §2 the framework

[3]Note that the spot return can be decomposed into the forward premium and the *ex post* excess return.

[4]For an interesting recent contribution see Kilian and Taylor (2000).

for testing the FRUH is outlined and the explanations of recent studies are summarised. In §3 we describe the data and report unit root and symmetry test results as well as estimates for the returns regression. The Monte Carlo experiments are outlined and evaluated in §4 and a final section concludes.

# 2 The forward premium puzzle

## 2.1 Framework

In a CCAPM framework, the Euler equation governing the equilibrium real rate of return — denoted $r_{t+1}$ at time $t+1$ — on any asset is given by:

$$1 = \mathrm{E}_t[(1 + r_{t+1})M_t] \tag{1}$$

where $M_t = \beta u'(C_{t+1})/u'(C_t)$ is the stochastic discount factor or pricing kernel representing the intertemporal marginal rate of substitution, $u'(C)$ denotes the marginal utility of consumption, $\beta$ a discount factor and $\mathrm{E}_t$ is the expectations operator conditional on time $t$ information. The real return from foreign currency speculation is $(F_t - S_{t+1})/P_{t+1}$ where $F_t$ is the forward exchange rate quoted at $t$ for delivery at $t+1$, and $S_{t+1}$ and $P_{t+1}$ are the spot rate and price level, respectively at $t+1$. Since forward speculation requires no initial investment so that its real return should be zero, this with covered interest parity and purchasing power parity imply that (1) can be rearranged as:

$$\mathrm{E}_t\left[\left(\frac{F_t - S_{t+1}}{P_{t+1}}\right)\frac{u'(C_{t+1})}{u'(C_t)}\right] = 0 \tag{2}$$

Assuming risk aversion and that all the variables are jointly lognormally distributed, (2) can be rewritten as:[5]

$$\mathrm{E}_t(s_{t+1}) - f_t = -1/2\ Var_t(s_{t+1}) + Cov_t(s_{t+1}, p_{t+1}) + Cov_t(s_{t+1}, q_{t+1}) \tag{3}$$

where lower case letters denote the natural logarithm of the variable in question, $s_{t+1}$ is the spot exchange rate, $f_t$ the forward rate quoted at $t$ for delivery at $t+1$, $Var_t$ and $Cov_t$ are the variance and covariance, respectively, and $Cov_t(s_{t+1}, q_{t+1})$ is a risk premium where $q_{t+1}$ is the intertemporal

[5] See Obstfeld and Rogoff (1997), footnote 75, page 588 for a derivation.

marginal rate of substitution. Equation (3) suggests is that ex ante excess returns $\mathrm{E}_t(s_{t+1}) - f_t$ may be affected by a risk premium term and this is what asymmetries and associated overshooting may be capturing.

For tests of the FRUH, the most commonly used specification involves regressing the spot return on the forward premium:[6]

$$s_{t+1} - s_t = \alpha + \beta(f_t - s_t) + \varepsilon_{t+1} \tag{4}$$

where $\varepsilon_{t+1}$ is a random error term and both variables are in logs. Under risk neutrality the FRUH or $E_t s_{t+1} = f_t$ implies $\alpha = 0$ and $\beta = 1$. The spot return regression can also be reparameterised in terms of excess returns or deviations from UIP being a function the forward premium.

$$s_{t+1} - f_t = \alpha + (\beta - 1)(f_t - s_t) + \varepsilon_{t+1} \tag{5}$$

In this context, the FRUH null of $\beta = 1$ implies a theoretical slope coefficient of zero or no deviations from UIP. Lewis (1995) highlighted three implications of the puzzle for excess returns: they are significantly different from zero, they change signs frequently and they are extremely volatile.

Allowing for risk averse agents alters the simple relationship between the forward and expected spot rate as follows:

$$f_t = E_t s_{t+1} + \rho_t \tag{6}$$

where $\rho_t$ is a risk premium. This yields an alternative expression for the forward premium:

$$E_t \Delta s_{t+1} + \rho_t = f_t - s_t \tag{7}$$

Thus the risk premium can be viewed as driving a wedge between the expected spot return and the forward premium leading to deviations from UIP. Fama (1984) analysed the statistical properties of the slope coefficient in (4) in a risk neutral and risk averse world. INn the latter case, the ordinary least squares (OLS) estimate of $\beta$ in this case must satisfy asymptotically:

$$\text{plim}(\hat{\beta}) = \beta = \frac{Cov(E_t \Delta s_{t+1} + \rho_{t,}\ E_t \Delta s_{t+1})}{Var(f_t - s_t)} \tag{8}$$

[6]Maynard and Phillips (1998) provide an interesting discussion of the problems associated with three well known econometric models – levels, spot returns and error correction. See Zivot (2000) also.

where plim denotes the probability limit and $E_t s_{t+1}$ has been substituted for $s_{t+1}$.[7] The numerator in (8) can be expanded to yield:

$$\text{plim}(\hat{\beta}) = \beta = \frac{Var(E_t \Delta s_{t+1}) + Cov(E_t \Delta s_{t+1}, \rho_t)}{Var(f_t - s_t)} \tag{9}$$

Next consider the implications of three hypotheses for the slope coefficient on the basis of this equation:

- $\beta = 1$. This is the traditional FRUH case implying the numerator and denominator of (9) are equal. The existing evidence overwhelmingly rejects this hypothesis.

- $\beta < 0$. This case has attracted most attention in the literature. It requires the covariance of the risk premium and the expected spot return to exceed the variance of the latter, a finding which is difficult to reconcile with theoretical models under plausible levels of risk aversion. The focus on $\beta < 0$ is often based on the average point estimate from 75 published studies of -0.88 reported in Froot and Thaler (1990) and derived from studies using data spans dominated by the 1980s.

- $\beta = 0$. An insignificant slope coefficient represents another potential violation of the FRUH and is supported in recent studies. This involves the weaker condition that the covariance of the risk premium and the expected spot return exactly offsets the variance of the latter.

To our knowledge, no extant solution has tackled the $\beta = 0$ case and so this paper fills an important lacuna in the literature.

## 2.2 Recent explanations

While recent explanations tend to focus on the $\beta < 0$ case — which needs to be addressed to rationalise results using data up to the early 1990s — the reported empirical results from such studies also indicate that the $\beta = 0$ case is relevant. Baillie and Bollerslev (2000) and Maynard and Phillips (1998) focus on misspecification problems in the commonly used returns regression

[7] Since under rational expectations (RE) the difference between the actual and expected spot rate, $s_{t+1} - E_t s_{t+1} = u_{t+1}$, is uncorrelated with any time $t$ information including the forward premium, $E_t s_{t+1}$ can also be substituted for $s_{t+1}$ in (2).

specification. Both these sets of authors attribute the failure to support the FRUH to persistence or long memory in the forward premium. In their view, persistence leads to an unbalanced regression problem in the sense that the dependent and independent variables are integrated of different orders which invalidates standard statistical theory.

Baillie and Bollerslev (2000) employ a stylised UIP model with very persistent (FIGARCH) volatility to provide Monte Carlo evidence for a slope coefficient which is extremely widely dispersed around its true value of unity and has a very slow rate of convergence. While they address the problem of negative slope coefficients found in studies using data up to the early 1990s, their monthly 5-year rolling regressions, 1973:3-1995:11, highlight two issues. On one hand they clearly depict the problem of negative slope coefficients being specific to the 1980s decade. On the other hand the following quote on more recent slope coefficients is revealing for our purposes:

> "Hence, ... the estimated slope coefficient from the anomalous regression exhibits substantial variation, and for the 5-year regressions depicted in Fig. 1, many of the more recent slope coefficients are actually positive, albeit not statistically significant when judged by the usual 95% confidence bands." (pp.476-477)

Maynard and Phillips (1998) explain the puzzling negative slope coefficients by means of endogeneity and serial correlation terms which are associated with nonstandard limiting distributions with long left tails. They employ a more recent daily data span from November 1986 to March 1998 which is not dominated by the 1980s and their returns regression results also furbish support for our version of the puzzzle. No less than six out of seven of the reported slope coefficients are insignificantly different from zero at the conventional 5% level.

Barnhart et al. (1999) build on Fama (1984) to explain the $\beta < 0$ puzzle in terms of simultaneity bias, possibly due to an omitted excess return variable. They decompose the effect of predicted excess returns into movements in the interest rate differential and the spot return. This introduces an extra covariance term between the forward premium and predicted excess returns, leading to simultaneity bias.[8] Nonetheless some four out of seven of their for-

[8]Our approach links to that of Barnhart et al. in that it assumes that the asymmetries in spot returns can be explained by the behavior of excess returns or, under rational expectations, a nonlinear risk premium. On simultaneity bias, see also Liu and Maddala (1992) and Maynard and Phillips (1998).

ward premium regression slope coefficients, 1974:4-1997:3, are insignificantly different from zero.[9] In the next section, five-year rolling regressions suggest that negative slope coefficients are confined to the 1980s and the weight of evience from more recent data spans also points to statistically insignificantly slope coefficients.

The latter together with the evidence from the above studies leads to the conclusion that the $\beta = 0$ case is the version of the puzzle that needs to be explained for more recent data spans. The $\beta < 0$ case, while clearly important, appears to be an artifact of the long swings in the US dollar in the 1980s. Psaradakis et al. (2000) and Evans and Lewis (1995) provide a rationale by incorporating these swings via a RE assumption into their analysis of risk premia employing a Markov switching approach. The former use instrumental variables (IV) to account for within-regime correlation between explanatory variables and disturbances to show that the FRUH cannot be rejected using IV. The latter allow for jumps in the exchange rate and use Monte Carlo experiments to show that the "peso" problem has implications for both the low and high frequency behavior of exchange rates.

# 3 Empirical Analysis

## 3.1 Data and unit root tests

The data consist of bilateral US dollar spot and forward exchange rates from Datastream for four currencies in which North America, Europe and Asia are represented: the Canadian dollar, French franc, Deutschmark and Japanese yen. The choice of numeraire is determined by the central role of the US dollar as both reserve currency and as unit of account in international trade. Beginning-of-month (average of bid and ask) spot and forward rates are employed for the duration of the current floating rate regime: 1976:2-1999:5.[10] Both rates are defined as the domestic price of foreign currency. By using monthly frequency and 1-month forward contracts, overlapping data effects are avoided.

The results of ADF and Phillips-Perron [PP] unit root tests are reported

[9]Table 4, p.276.

[10]Both rates are defined as the domestic price of foreign currency. The data commence in 1978:7 for the yen and end in 1998:12 for the franc and mark due to the advent of the single currency in January 1999. The empirical analysis was undertaken in GAUSS 3.2.26.

in Table 1.[11]

[Table 1 around here]

The selected $k$ suggest that the forward premium is highly autocorrelated in contrast with the spot return. Since the large number of parameters implied by a high lag order $k$ in the former may reduce power, insignificant lower order terms were excluded. Both tests clearly reject the unit root null at the 1% level for all the spot return series in accordance with the literature. The ADF test rejects the null for three forward premium series — the Canadian dollar and yen at the 5% level and French franc at the 1% level — and the PP test rejects for all, the Canadian dollar and franc at the 1% level, the yen at the 5% level and the mark at the 10% level.

Our findings are in agreement with most extant studies such as Berben and van Dijk (1999) and Hai, Mark and Wu (1997) which support stationarity in forward premia. A few dissenting voices such as Crowder (1994) and, indirectly, Evans and Lewis (1995) suggest evidence of unit root behavior. Relatedly, Baillie and Bollerslev (1994, 2000), Byers and Peel (1996) and Maynard and Phillips (1998) concur that the forward premium is fractionally integrated.[12]

The results for the returns regression are given in Table 2 for different sample sizes $n$, corresponding to the full sample period and to the pre- and post-1990 samples to separate out the effect of the 1980s.

[Table 2 around here]

In keeping with the literature very low $R^2$ statistics are obtained with an upper bound of just 5.5%. For the full sample period the average slope coefficient is -0.87 which is in line with most existing findings.[13] For the 1976-1989 period, all the coefficients are significantly negative (except that for the French franc) reflecting the dominant effect of the 1980s. Interestingly, for the 1990s no coefficient is significantly different from zero.

[11] The PP truncation lag is chosen by the Newey-West formula. The ADF lag order $k$ is selected following the general-to-specific approach proposed in Ng and Perron (1995), starting from $k_{\mathsf{max}} = 12$.

[12] See Baillie (1996) for an excellent survey of fractional integration.

[13] Although the FRUH cannot be rejected for the French franc, this result does not appear to be stable since the null is readily rejected for both subsamples.

To investigate further the potentially distortionary effects of the 1980s, rolling 5-year returns regressions were run over the entire sample period. The slope coefficients with their standard error bands and the $t$-ratios are depicted in Figures 1 and 2, respectively.

[Figures 1 and around here]

These highlight two issues. First, the slope coefficient estimates using 1980s data are very inefficient as indicated by a marked widening of the two standard error bands. Secondly, while the majority of the estimated coefficients are statistically insignificant at the usual 5% level, the significant coefficients are typically confined to the 1980s. The insignificance of the majority of the slope coefficients from the rolling regressions and from the 1990s regressions is the basis for focusing on the new $\beta = 0$ definition of the puzzle. The existing definition of $\beta < 0$ appears to be an artifact of using 1980s data.

## 3.2 TAR bootstrap tests

A nonlinear framework is employed to explore spot returns and forward premia for potential frictions captured by asymmetries. Our approach is distinctive in opting for the parsimony of a threshold autoregressive (TAR) specification to capture the jumps or overshooting behaviour of exchange rates typical of the Dornbusch (1976) sticky price monetary model and to permit a no-arbitrage, transaction cost band.[14] In particular, consider the following regime-switching dynamic model:

$$\begin{aligned} &\Delta z_t = \left\{ I_{1t}\rho_1(z_{t-1}-\mu) + I_{2t}\rho_2(z_{t-1}-\mu) \right\} + \varepsilon_t \\ &I_{1t} = \begin{cases} 1 & if\ z_{t-1}-\mu \geq \theta_1 \\ 0 & otherwise \end{cases} \quad I_{2t} = \begin{cases} 1 & if\ z_{t-1}-\mu \leq -\theta_2 \\ 0 & otherwise \end{cases} \end{aligned} \tag{10}$$

where $\varepsilon_t \sim NID(0, \sigma_\varepsilon^2)$, and $\theta_i, i = 1, 2$ are the threshold parameters with $(\theta_1, \theta_2) \geq 0$ as an identifying restriction. If the stochastic process $z_t$ is stationary, the parameter $\mu$ represents its long run equilibrium or attractor.

[14] See Coakley and Fuertes (2001b) and O'Connell (1998) for TAR applications to real exchange rates. Dynamics varying by regime could also be modelled by the less parsimonious smooth transition autoregressive (STAR) models. See Tong (1990) and Granger and Teräsvirta (1993) for a comprehensive overview of (S)TAR and other nonlinear time series models.

To deal with serially correlated errors, (10) is augmented by adding lagged differences, $\sum_{i=1}^{k} \beta_i \Delta z_{t-i}$.

This model is a three-regime (discontinuous) generalization of the Enders and Granger (1998) two-regime TAR. The addditional (inner) regime is justified on theoretical grounds as stemming from transaction costs or noise trading activity (Coakley and Fuertes, 2001a). Thus this specification can parsimoniously capture local nonstationarity or persistence in a process which is globally stationary provided $-2 < (\rho_1, \rho_2) < 0$. Linear symmetric AR dynamics is nested in (10) for the joint restrictions $\rho_1 - \rho_2 = 0$ and $\theta_1 + \theta_2 = 0$ and this provides the basis for symmetry tests.

The null of symmetric dynamics is tested against TAR asymmetries using a bootstrap likelihood ratio statistic $[\eta]$ proposed by Coakley and Fuertes (2000).[15] A bootstrap method is employed since the discontinuity implied by the random walk band in the conditional mean of (10) invalidates standard asymptotic inference. The alternative comprises both sign asymmetry — differential adjustment ($\rho_1 - \rho_2 \neq 0$) to positive and negative deviations from the attractor — and amplitude asymmetry or differential adjustment ($\theta_1 + \theta_2 \neq 0$) to large and small deviations.

To implement the bootstrap test, the parameters of TAR model (10) are estimated by conditional least squares. A grid search is conducted to find the threshold values $(\theta_1, \theta_2)$ that minimise the residual sum of squares. Under Gaussian innovations, the latter amounts to maximizing the likelihood function of the model. Two versions of the test are reported: one which proxies the sample mean of $z_{t-1}$ by $\mu$ and another in which the mean is consistently estimated.[16] The augmentation lag $k$ is identified as for the ADF tests.

The bootstrap p-values for the $\eta$ statistic are computed by resampling using the null model(s) estimates and normally distributed random distur-

[15] Their small-scale Monte Carlo analysis of the bootstrap $\eta$ test suggests that it is correctly sized and has reasonable power for a sample span such as ours. Note that Wu and Zhang (1996) uncover a different type of asymmetry depending on whether the US dollar forward rate is at a premium or discount against the yen and Deutschmark.

[16] The latter is motivated by the argument that the sample mean of an asymmetric series is a biased estimator of $\mu$ (Tong, 1990). This second version involves a grid search for $\mu$ and the addition of a constant in (10) since otherwise the residuals might have a nonzero mean. The grids for $\mu$ and $(\theta_1, -\theta_2)$ are confined to between the $15^{\text{th}}$ and $85^{\text{th}}$ percentiles of $\{z_{t-1}\}$ and $(\{z_{t-1} - \mu\}^+, \{z_{t-1} - \mu\}^-)$ respectively. As Chan and Tsay (1998) show, this yields $\sqrt{n}$- and $n$-consistent estimates for $\mu$ and $\theta_j, j = 1, 2$, respectively.

bances whose variance equals the estimated residual variance.[17] We implement the $m$-out-of-$n$ or rescaled bootstrap formally introduced by Bickel, Götze and van Zwet (1997) to safeguard against non-smoothness-induced failure of first order (consistency) of the bootstrap. Table 3 reports the results.

[Table 3 around here]

The symmetric adjustment null is rejected for all spot return series in favour of TAR dynamics at the 5% level or better, supporting the conclusions of Kräger and Kugler (1993) and Clements and Smith (2001). However the forward premium show no evidence of asymmetries.[18] An interesting feature of the test results is that $|1+\rho_1| = |1+\rho_2|$ for all four spot return series which is consistent with overshooting spot exchange rate behaviour. This is important both because it links asymmetries with overshooting models of exchange rates theoretical and because it can explain sign changes in spot returns and their excess volatility as compared with the forward premium series. The more stable behaviour of the forward premium series is consistent with the notion that jumps in the spot rate in response to time $t$ innovations will be approximately offset by those in the forward rate.

Repeating the estimation and testing for the 1990s the evidence for the latter is even stronger. Table 4 reports the estimates of sign asymmetry measured by $|\hat{\rho}_1 - \hat{\rho}_2|$ and amplitude asymmetry measured by $\hat{\theta}_1 + \hat{\theta}_2$ for both the full sample period, 1976-99, and the 1990-99 period.

[Table 4 around here]

It is evident that average sign asymmetry in the 1990s is more than double that for the full sample period while average amplitude asymmetry is more or less unchanged as may be expected if it is capturing transaction costs. Since the range of asymmetries in spot returns is increased in the 1990s, the rejection of the symmetry null is even stronger.

[17] The first 100 observations of each of $B = 599$ bootstrap samples are discarded. The latter is chosen, following Davidson and MacKinnon (1998), to satisfy $B > 399$ and so that $\alpha(1+B)$ is an integer at typical significance levels such as $\alpha = 0.05$ to achieve an exact test. A less restrictive but computationally more intensive approach would involve resampling from the empirical distribution of the estimated residuals.

[18] Berben and van Dijk (1999) find evidence of sign asymmetries in the forward premia for the UK pound, French franc and yen 1976:1-1992:8 using a Wald test based on a different (continuous) TAR specification.

# 4 Monte Carlo experiments

Monte Carlo experiments are conducted to isolate the impact of asymmetries and of overshooting effects and assess their separate and combined contribution to the slope coefficient bias.[19] Accordingly, artificial time series are generated for two variables, $y_t$ and $x_t$, which — while linearly related according to the FRUH in equation (4) — abstract from the other characteristics of the spot return and forward premium variables. We set up a simple linear regression model of $y_t$ on $x_t$ and then study the consequences of adding asymmetries and overshooting effects to $y_t$.

Let $x_t$ — below associated with the forward premium — have the following linear AR(1) dynamics:

$$\Delta x_t = \rho x_{t-1} + u_t, \qquad u_t \sim NID(0, \sigma_u^2) \tag{11}$$

Another variable $y_t$ — below associated with the spot return — is generated by assuming the following linear (regression) relation with $x_t$:

$$y_t = \alpha + \beta x_t + \nu_t, \qquad \nu_t \sim NID(0, \sigma_\nu^2)$$

where $\alpha = 0, \beta = 1$ and $\nu_t$ is independent from $u_t$. Thus the generating mechanism of $y_t$ in AR form is:

$$\Delta y_t = \rho y_{t-1} + \zeta_t \tag{12}$$

where the innovations are Gaussian MA(1), $\zeta_t = u_t + \nu_t - \tilde{\rho}\nu_{t-1}$ with $\tilde{\rho} = \rho + 1$. Next, let $y_t$ be affected by an unobservable random component introducing asymmetries, $y_t^a = y_t + w_t^a$. We allow for both sign and amplitude asymmetries, yielding the following dynamics:

$$\begin{aligned} &\Delta y_t^a = I_{1t}(\rho - \tau) y_{t-1}^a + I_{2t}(\rho + \tau) y_{t-1}^a + \zeta_t \\ &I_{1t} = \begin{cases} 1 & if\ y_{t-1}^a \geq \delta_1 \\ 0 & otherwise \end{cases} \quad I_{2t} = \begin{cases} 1 & if\ y_{t-1}^a \leq -\delta_2 \\ 0 & otherwise \end{cases} \end{aligned} \tag{13}$$

where $(\tau, \delta_1, \delta_2) \geq 0$. For $\delta_1 = \delta_2 = \tau = 0$ the stochastic behavior in (13) reduces to (12).

To investigate the effects of the asymmetries on the distribution of the slope coefficient, we generate 10,000 artificial time series of length $n + 200$

[19] We are indebted to Haris Psaradakis for this suggestion.

using (12) for $y_t$, (13) for $y_t$ with asymmetries denoted $y_t^a$, and (11) for $x_t$. The actual parameter values used in the simulations are $\rho = -0.9$, $\sigma_u = \sigma_\nu = 1$, and $n = 250$. The range of asymmetries in (13) is calibrated to a set of realistic values $\tau$, $\delta_1$ and $\delta_2$ based on the estimated TAR parameters for our sample spot return series as reported in Table 4. Accordingly, since the estimated average $|\hat{\rho}_1 - \hat{\rho}_2|$ for the 1990s is 0.645, we set $2\tau = 0.645$. Similarly, since the 1990s average of $\hat{\theta}_1/\hat{\sigma}_\varepsilon$ and $\hat{\theta}_2)/\hat{\sigma}_\varepsilon$ in (6) are 0.851 and 0.419, respectively, we set $\delta_1 = 0.851\sigma_\zeta$ and $\delta_2 = 0.0419\sigma_\zeta$.

Figure 3 depicts the small sample distribution of the slope coefficient from the regressions of $\{y_t\}$ and $\{y_t^a\}$ on $\{x_t\}$, denoted $\beta$ and $\beta^a$, respectively, and their corresponding $t$-ratios.[20]

[Figure 3 around here]

Interestingly, while the kernel density of $\beta$ is centred on its theoretical value of unity, that for $\beta^a$ is shifted sharply to the left and centred close to zero. Asymmetries lead to a sharp decrease in efficiency as indicated by the 53.5 per cent increase in the standard deviation of the $\beta^a$ over the $\beta$ estimates. Some 83 per cent of the biased coefficients are statistically insignificant as suggested by the associated $t$-statistics for $\beta = 0$. Qualitatively similar results on downward bias were obtained from simulations using the average range of asymmetries for the full sample period.

The observed range of asymmetries for the 1990s was reduced by a factor of 40 — this sufficed to produce a no-arbitrage band of $\mp 0.0006$ or six basis points in line with actual transaction costs — and the simulations were repeated. Reassuringly they still resulted in biased coefficients centred on 0.095 as in the original experiment depicted in Figure 3 and an increase in dispersion of 43 per cent.[21]

The direct effect of asymmetries is statistically analogous to that of an omitted variable or simultaneity bias. Allowing $\Delta s_{t+1}$ to be influenced by $w_{t+1}^a$, where $\Delta s_{t+1}^a = \Delta s_{t+1} + w_{t+1}^a$, leads to the following bias in the slope coefficient:

$$\frac{Cov(f_t - s_t, w_{t+1}^a)}{Var(f_t - s_t)} \tag{14}$$

[20] The smooth densities were calculated by an Epanechnikov kernel with bandwitdth ($h$) selected as in Silverman (1986).

[21] Qualitatively similar results were obtained from an alternative Monte Carlo experiment where we introduced TAR(1) asymmetries via $\nu_t$ in $y_t = \alpha + \beta x_t + \nu_t$.

The similarity between this and the corresponding bias term in Barnhart et al. (1999) is straightforward. The effect in both cases, is to shift the slope coefficient to the left.

From Figure 3 it is evident that the empirical distribution of the slope coefficient $\beta^a$ is not as dispersed as might be expected since it does not produce absolutely large negative values and some 17 per cent of the coefficients are still significant on average. However, one of the stylised facts of the forward premium puzzle is the markedly larger scale of innovations to spot returns relative to the forward premium as Maynard and Philips (1998) inter alia have pointed out. This scale differnece may be attributed to the overshooting behaviour of spot rates associated with sign asymmetries. Accordingly, new artificial data are generated using the innovation scaling relationship $\sigma_\nu = 12\sigma_u$ (and denoted $y_t^s$) which reflects the observed ratio in the 1990s and will be described as overshooting efects.

Figure 4 graphs the small sample distribution of the slope coefficient from the new regressions of $\{y_t^s\}$ and $\{y_t^{a,s}\}$ on $\{x_t\}$, denoted $\beta^s$ and $\beta^{a,s}$, respectively, and their corresponding $t$-ratios.

[Figure 4 around here]

Two effects are evident. First, while overshooting effects on theie own have no impact on the mean (of unity) of the slope kernel density, almost three quarters of the coefficients are statistically insignificant. Second, the returns regressions combining both asymmetries and e overshooting effects produce a much more dispersed distribution of slope coefficients, $\beta^{a,s}$, whose sample standard deviation is almost nine times larger than that of $\beta^a$ from the asymmetries only DGP. More importantly, while still centred on 0.1, virtually all (in excess of 94 per cent) of the coefficients are now statistically insignificant at the usual 95 per cent confidence level.

Our results imply that the direct effect of asymmetries is sufficient to shift the mean of the slope kernel density to zero and produce a large proportion of insignificant slope coefficients. Moreover, when supplemented by overshooting effects or innovation scale differences, our simulations can replicate the widely dispersed and insignificant coefficients for the 1990s reported above and by a number of extant studies.

# 5 Conclusions

This paper addresses the forward premium puzzle for four major US dollar exchange rates using monthly data 1976-1999. It is argued that the forward premium puzzle is exaggerated and needs to be reformulated. This is necessary since extant studies address the negative slopes associated with the long dollar swings of the 1980s. By contrast, the insignificant coefficients from recent data spans can be explained by an unbalanced regression problem caused by asymmetries in spot returns. These stem from market frictions such as transaction costs and noise trading and are also associated with overshooting in spot rates. This new formulation is more plausible than the traditional versions since it simply implies that, on average, the forward premium contains no unpriced information on the future spot return.

Bootstrap likelihood ratio tests confirm asymmetries in the spot returns but not the forward premium series. Monte Carlo experiments show that asymmetries and associated overshoting effects produce widely dispersed and statistically insignificant slope coefficients whose kernel density is centred close to zero. This dovetails both with the extant econometric evidence on the non-forecastability of spot exchange rates and is consistent with the existence of markets frictions leading deviations from UIP.

# References

[1] Baillie R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5-59.

[2] Baillie, R.T., Bollerslev, T., 1994. The long memory of the forward premium. Journal of International Money and Finance 13, 565-571.

[3] Baillie, R.T., Bollerslev, T., 1996. The forward premium anomaly is not as bad as you think. Econometrics and Economic Theory Paper No. 9606, Michigan State University.

[4] Balke, N.S. and Wohar, M.E., 1998. Nonlinear dynamics and covered interest arity. Empirical Economics 23, 535-559.

[5] Barnhart, S.W., McNown, R., Wallace M.S., 1999. Noninformative tests of the unbiased forward exchange rate, Journal of Financial and Quantitative Analysis 34, 265-291.

[6] Berben, R.-P., van Dijk, D. 1999. Unit root tests and asymmetric adjustment: A reassessment, Mimeo, Department of Appled Economics, University of Nijmegen.

[7] Bickel, P. J., Götze, F., van Zwet, W. R., 1997. Resampling fewer than n observations: Gains, losses and remedies for losses. Statistical Sinica, 7, 1-31.

[8] Byers, J.D., Peel D.A., 1996. Long-memory risk premia in exchange rates. Manchester School 64, 421-438.

[9] Chan, K.S., Tsay, R.S., 1998. Limiting properties of the least squares estimator of a threshold autoregressive model. Annals of Statistics 21, 420-433.

[10] Clarida, R.H., Taylor, M.P., 1997. The term structure of forward exchange premiums and the forecastability of spot exchange rates: Correcting the errors. Review of Economics and Statistics 79, 353-361.

[11] Clements, M.P., Smith, J., 2001. Evaluating forecasts from SETAR models of exchange rates. Journal of International Money and Finance 20, 133-148.

[12] Coakley, J., Fuertes, A.M., 2000. Bootstrap LR tests for sign and amplitude asymmetries. Birkbeck College Discussion Paper No. 4/2000

[13] Coakley, J., Fuertes, A.M., 2001a. A nonlinear analysis of excess foreign exchange returns. The Manchester School, Forthcoming.

[14] Coakley, J., Fuertes, A.M., 2001b. Border costs and real exchange rate dynamics in Europe, Journal of Policy Modeling, Forthcoming.

[15] Crowder, W.J., 1994. Foreign exchange market efficiency and common stochastic trends. Journal of International Money and Finance 13, 551-564.

[16] Davidson, R., MacKinnon, J.G. 1998. The power of bootstrap tests. Queen's Institute for Economic Research Discussion Paper No. 937.

[17] Dornbusch, R. 1976. Expectations and exchange rate dynamics. Journal of Political Economy 84, 1161-1176.

[18] Enders, W., Granger, C.W.J., 1998. Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates. Journal of Business and Economic Statistics 16, 304-311.

[19] Engel, C., 1996. The forward discount anomaly and the risk premium: A survey of recent evidence. Journal of Empirical Finance 3, 123-192.

[20] Evans, M.D.D., Lewis, K.K. 1995. Do long-term swings in the dollar affect estimates of the risk premia? Review of Financial Studies 8, 709-742.

[21] Fama, E., 1984. Forward and spot exchange rates. Journal of Monetary Economics 14, 319-338.

[22] Froot, K.A., Thaler, R. H., 1990. Anomalies: Foreign exchange. Journal of Economic Perspectives 4, 179-192.

[23] Granger, C. W. J., Teräsvirta, T., 1993. Modelling nonlinear economic relationships. Oxford University Press: Oxford.

[24] Hai, W., Mark, N. C., Wu, Y., 1997. Understanding spot and forward exchange rate regressions. Journal of Applied Econometrics, 12, 715-734.

[25] Kilian, L., Taylor M. P., 2000. Why is it so difficult to beat the random walk forecast of exchange rates? Mimeo, University of Warwick.

[26] Kräger, H., Kugler, P., 1993. Nonlinearities in foreign exchange markets: a different perspective. *Journal of International Money and Finance* 12, 195-208.

[27] Lewis, K.K., 1995. Puzzles in international financial markets. In *Handbook of international economics*, Grossman, G.M. and Rogoff, K. eds, Vol. 3, 1913-1971, North-Holland, Amsterdam.

[28] Liu, P. C., Maddala, G. S., 1992. Rationality of survey data and tests for market efficiency in the foreign exchange markets. *Journal of International Money and Finance* 11, 366-381.

[29] Maynard, A., Phillips, P.C.B., 1998. Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. Cowles Foundation for Research in Economics, Yale University.

[30] Obstfeld, M., Rogoff, K., 1997. *Foundations of International Macroeconomics*, MIT Press: Cambridge Mass.

[31] O'Connell, P.G.J., 1998. Market frictions and real exchange rates. *Journal of International Money and Finance* 17: 1-19.

[32] Ramsey, J. B., 1969. Tests for specification error in classical linear least squares regression analysis. *Journal of the Royal statistical Society* B31, 250-271.

[33] Silverman, B.W., 1986. *Density estimation for statistics and data analysis*. Chapman Hall: London.

[34] Spagnolo, F., Psaradakis, H., Sola, M., 2000. Instrumental-variables estimation in Markov switching models with an application to testing the unbiased forward exchange rate hypothesis. Birkbeck College Discussion Paper No. 26/2000.

[35] Tong, H., 1990. *Non-linear time series: A dynamical system approach*. Clarendon Press: Oxford.

[36] Wu, Y., Zhang, H., 1996. Asymmetry in forward exchange rate bias: A puzzling

Table 1 Unit root tests

| | Spot returns ($s_{t+1} - s_t$) | | | Forward premia ($f_t - s_t$) | | |
|---|---|---|---|---|---|---|
| | PP | ADF | $k$ | PP | ADF | $k$ |
| C$ | -18.03* | 17.97* | 0 | -4.44* | -3.38** | 10 |
| FFr | -15.92* | -16.62* | 0 | -5.75* | -3.86* | 9 |
| DM | -16.69* | -10.47* | 1 | -2.63*** | -1.99 | 9 |
| Yen | -16.67* | -7.05* | 5 | -3.30** | -2.77** | 11 |

Notes:
[a]Phillips-Perron test. Truncation lag selected by $q = \text{int}[4(T/100)^{2/9}]$.
[b]Standard augmented Dickey-Fuller test.
[c]$k$ is the maximum lag order selected for the ADF (and TAR) equations.
* Significant at 1% level, ** Significant at 5% level, *** Significant at 10% level.

Table 2 Spot return regression estimates. Eq. (4)

| Currency | Sample | $\hat{\beta}$ $(s.e.)$ | $R^2$ |
|---|---|---|---|
| C$ | 1976-99 | -1.19 (0.511) | 0.019 |
| | 1976-89 | -1.64 (0.752) | 0.028 |
| | 1990-99 | -0.72 (0.673) | 0.010 |
| FFr | 1976-98 | 1.09 (0.527) | 0.015 |
| | 1976-89 | 0.97 (0.595) | 0.016 |
| | 1990-98 | 1.53 (1.308) | 0.013 |
| DM | 1976-98 | -0.68 (0.716) | 0.003 |
| | 1976-89 | -3.78 (1.494) | 0.037 |
| | 1990-98 | 0.17 (1.214) | 0.000 |
| Yen | 1978-99 | -2.71 (0.906) | 0.035 |
| | 1978-89 | -3.44 (1.233) | 0.055 |
| | 1990-99 | -2.58 (1.554) | 0.024 |

Table 3 Symmetry tests

| | Spot returns ($s_{t+1} - s_t$) | | Forward premia ($f_t - s_t$) | |
|---|---|---|---|---|
| | $\eta$ | $\eta_{\hat{\mu}}$ | $\eta$ | $\eta_{\hat{\mu}}$ |
| C$ | 17.19 | 18.07 | 0.74 | 0.72 |
| | [.006] | [.001] | [.938] | [.984] |
| FFr | 12.36 | 13.19 | 8.27 | 10.37 |
| | [.017] | [.008] | [.139] | [.092] |
| DM | 14.44 | 14.35 | 7.75 | 7.50 |
| | [.012] | [.009] | [.141] | [.133] |
| Yen | 12.30 | 12.27 | 2.82 | 2.94 |
| | [.043] | [.040] | [.750] | [.608] |

Notes:
[a] $\eta$ denotes threshold symmetry LR statistics with $p$-values in brackets. The subscript $\hat{\mu}$ indicates that $\mu$ is estimated instead of proxied by the sample mean.
[b] $k$ is the maximum lag order selected for the ADF and TAR equations.
[c] $m$-out-of-$n$ bootstrap empirical values in brackets, $m = n^{.97}$.

Table 4 Estimated aggregate range of asymmetries

| | 1976-1999 | | | 1990-1999 | | |
|---|---|---|---|---|---|---|
| | $\lvert\hat{\rho}_1 - \hat{\rho}_2\rvert$ | $\hat{\theta}_1 + \hat{\theta}_2$ | $\frac{(\hat{\theta}_1+\hat{\theta}_2)}{\hat{\sigma}_\varepsilon}$ | $\lvert\hat{\rho}_1 - \hat{\rho}_2\rvert$ | $\hat{\theta}_1 + \hat{\theta}_2$ | $\frac{(\hat{\theta}_1+\hat{\theta}_2)}{\hat{\sigma}_\varepsilon}$ |
| DM | .3999 | .0328 | 1.00 | .8324 | .0325 | 1.0926 |
| FFr | .2190 | .0098 | .3036 | .8740 | .0353 | 1.1806 |
| C$ | .4194 | .0111 | .8197 | .5899 | .0110 | .8679 |
| Yen | .1904 | .0292 | .8181 | .2830 | .0147 | .4316 |
| Average | .3072 | .0207 | .7354 | .6448 | .0234 | .8932 |

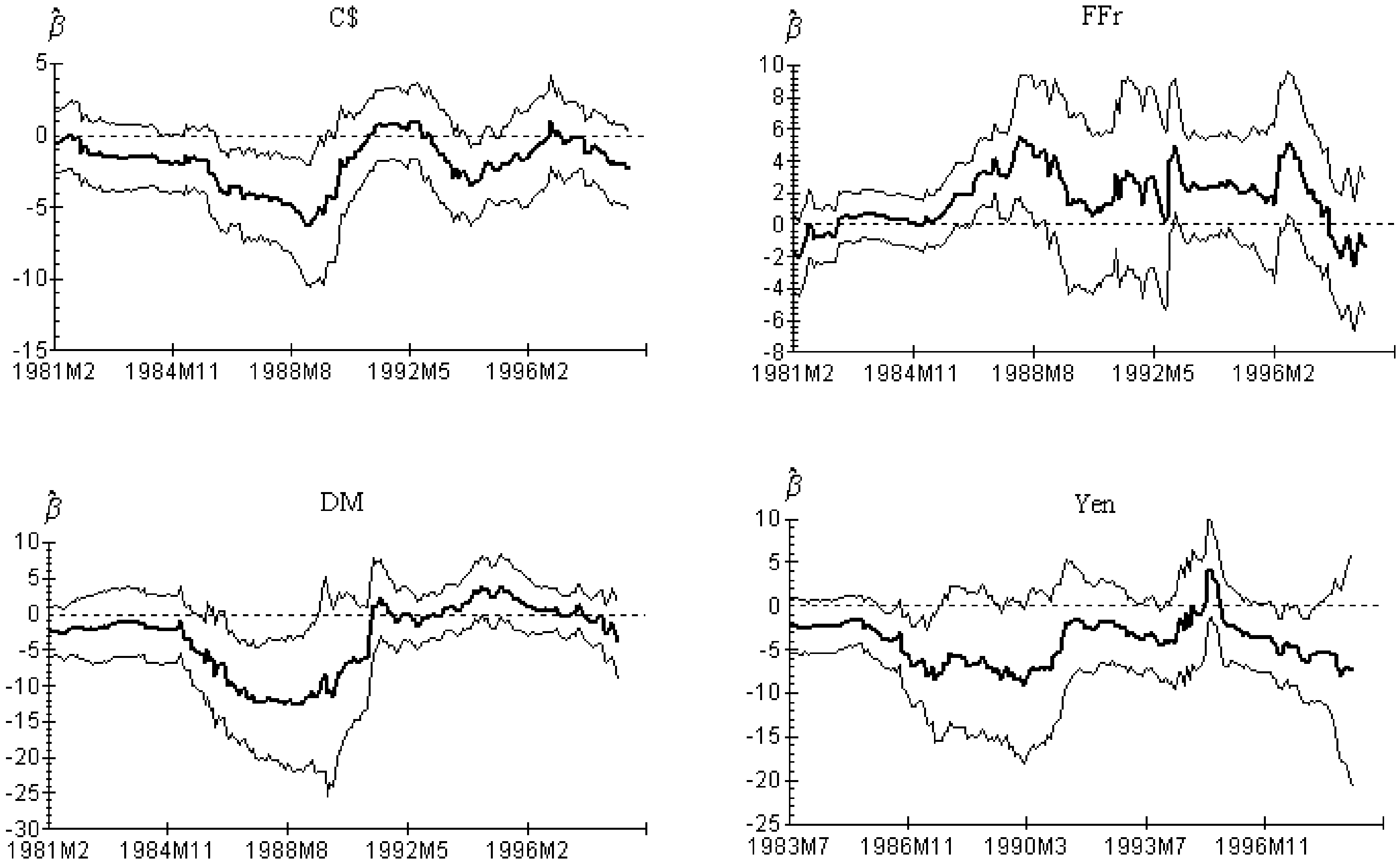


Figure 1. Coefficient of forward premium and its two standard error bands based on 5-year rolling OLS

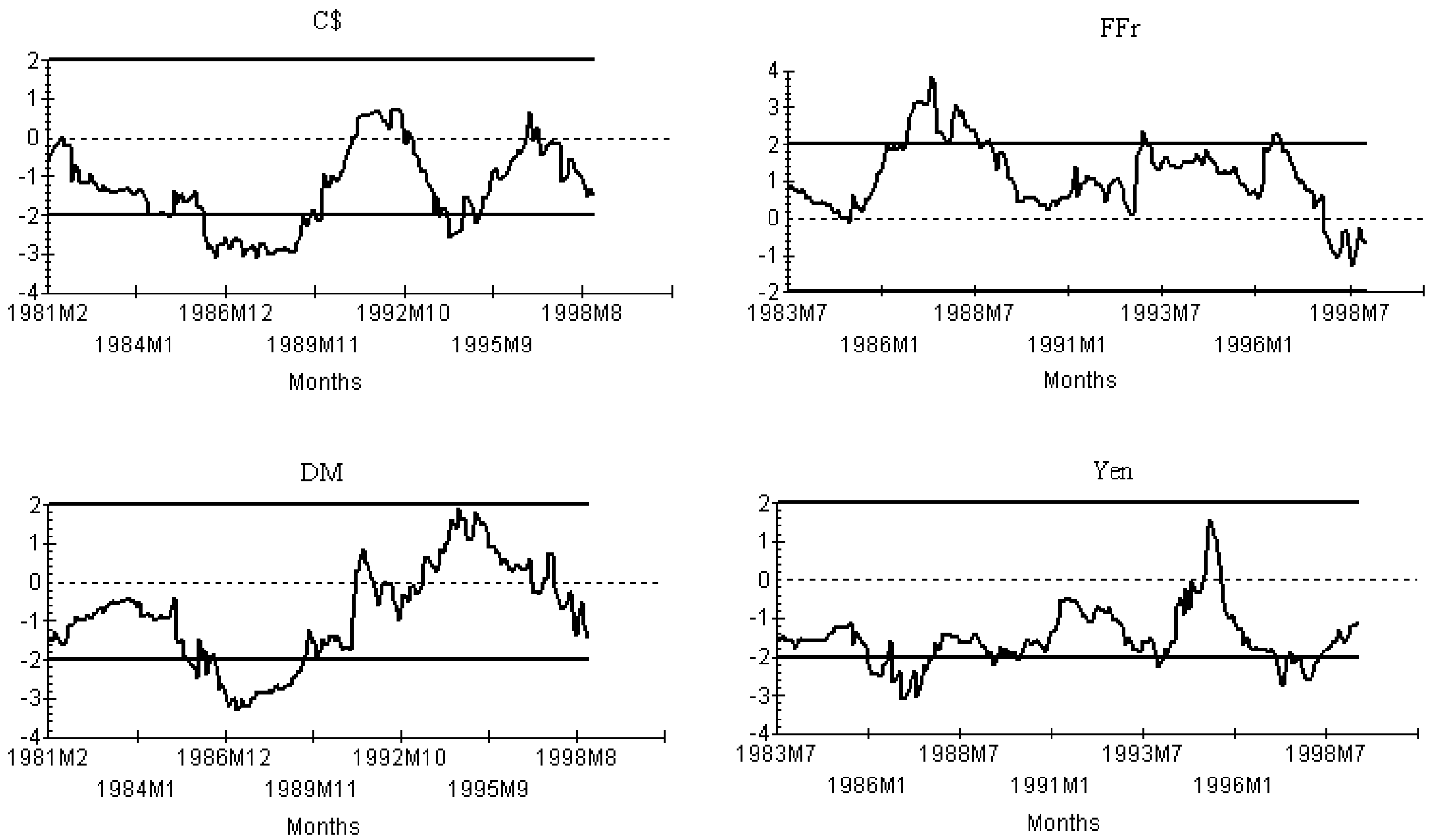


Figure 2 Ratio of coefficient estimate to its standard error from 5-year rolling OLS and asymptotic ±2 bands.

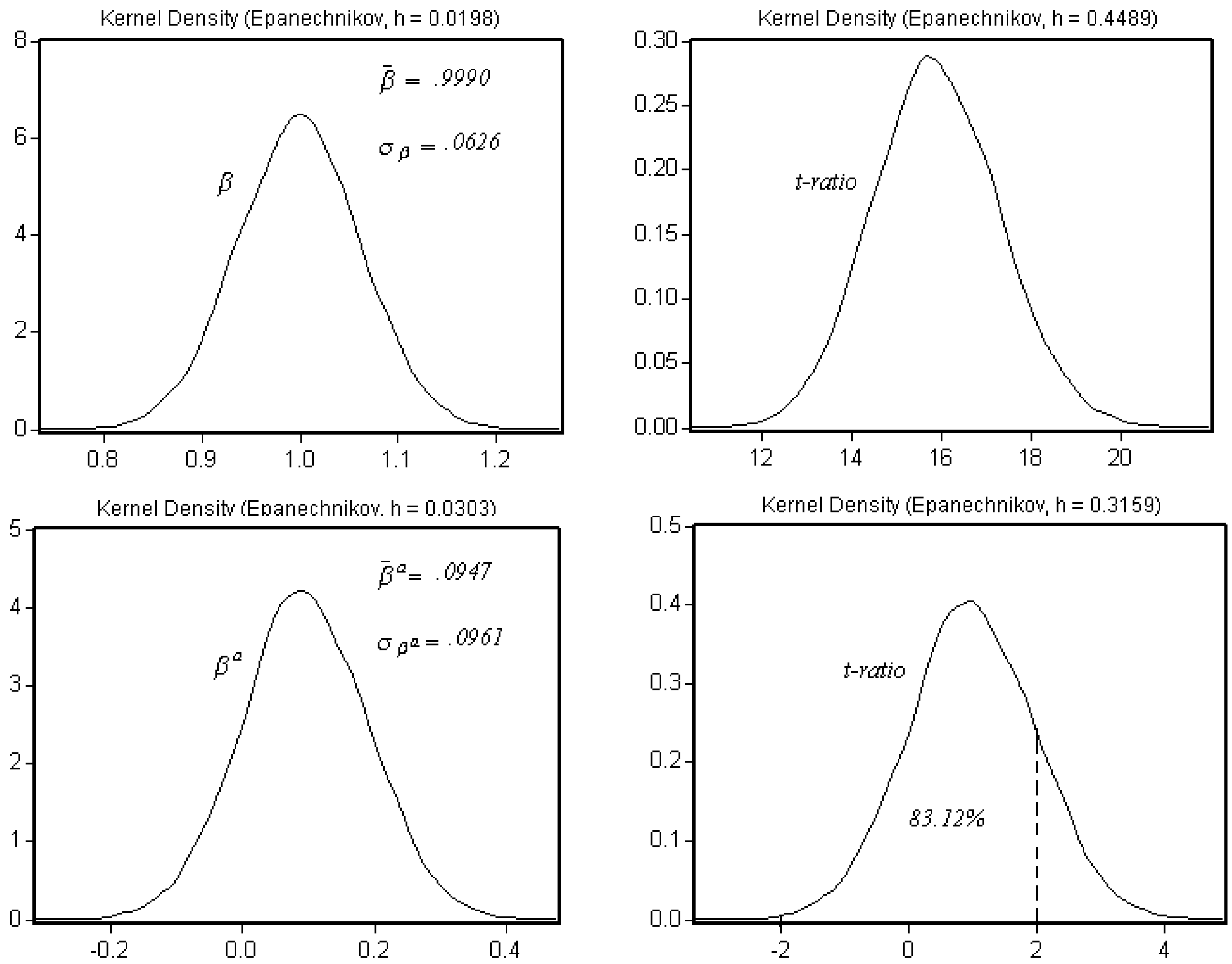


Figure 3. Simulated effect of asymmetries.

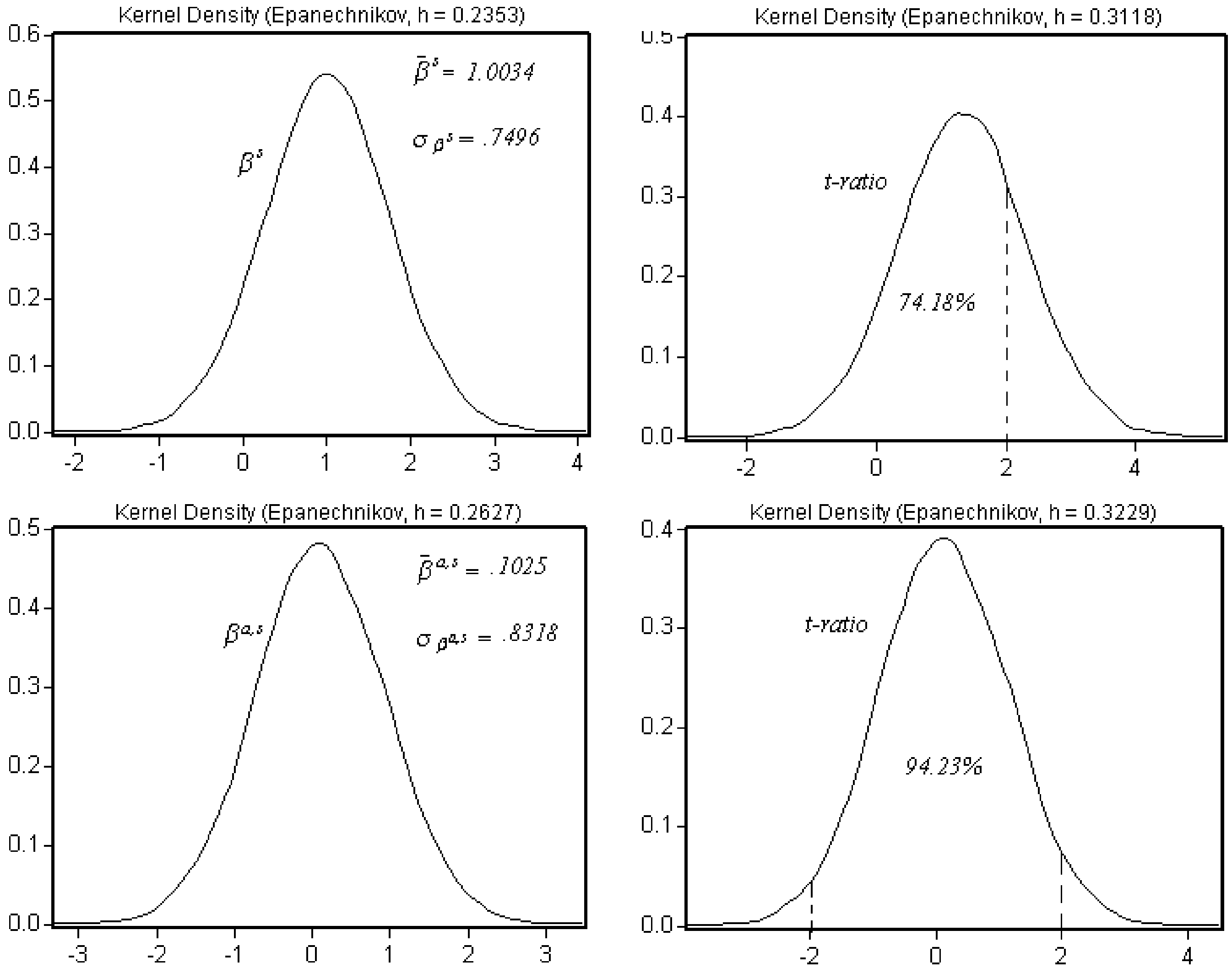


Figure 4. Simulated effect of asymmetries and scale of innovations.