Testing the Unbiased Forward Exchange Rate Hypothesis Using a Markov Switching Model and Instrumental Variables*

FABIO SPAGNOLO[a], ZACHARIAS PSARADAKIS[b] AND MARTIN SOLA[b,c]

[a]CARISMA and Department of Economics and Finance, Brunel University, Uxbridge, UK

[b]School of Economics, Mathematics & Statistics, Birkbeck College, London, UK

[c]Department of Economics, Universidad Torcuato Di Tella, Buenos Aires, Argentina

This version: February 2003

Abstract

This paper develops a model for the forward and spot exchange rate which allows for the presence of a Markov switching risk premium in the forward market and considers the issue of testing for the unbiased forward exchange rate (UFER) hypothesis. Using US/UK data, it is shown that the UFER hypothesis cannot be rejected provided that instrumental variables are used to account for within-regime correlation between explanatory variables and disturbances in the Markov switching model on which the test is based.

Key Words: Instrumental variables; Forward exchange rate; Markov chain; Maximum likelihood; Regime switching.

JEL Classification: C13; C22; F31.

*The authors are grateful to Jerry Coakley, John Geweke and two anonymous referees for their helpful and constructive comments.

Address correspondence to: Fabio Spagnolo, Department of Mathematical Science, Brunel University, Uxbridge, UB8 3PH, UK; e-mail: `fabio.spagnolo@brunel.ac.uk`; fax: +44 1895 203303.

1 Introduction

Testing the unbiased forward exchange rate (UFER) hypothesis, that is the hypothesis that the forward foreign exchange rate is an unbiased predictor of the corresponding spot exchange rate, has attracted considerable amount of interest in the literature. The results concerning its empirical validity have however been rather mixed. For example, Engel (1996) surveyed studies which have assumed rational expectations and attempted to attribute the forward rate bias to a foreign-exchange risk premium. His main conclusion was that standard general equilibrium models are unsuccessful in explaining the magnitude of the risk premium and, also, the empirical failure of the unbiasedness hypothesis.

The aim of the present paper is to offer a possible explanation for the rejection of the UFER hypothesis that is often reported in the empirical literature. In particular, we exploit an implication of the consumption capital asset pricing model (CAPM) under structural changes in consumption to reconcile this empirical evidence with general equilibrium models. The motivation for this approach is the empirical finding that consumption dynamics can be characterised successfully by models that allow for structural changes that are driven by a Markov process (see, e.g., Cecchetti et al., 1990; Hall et al., 1997). When combined with the hypothesis of time-varying risk premium, such dynamic behaviour for consumption implies that the risk premium itself is subject to Markov changes in regime.

Allowing for the presence of a Markov switching risk premium leads to a model for the spot rate and the forward premium whose parameters switch stochastically between regimes. A difficulty with such a model, however, that the right-hand side variables are correlated with the disturbances within each regime. It is, therefore, a plausible conjecture that the standard pseudo-maximum likelihood (S–PML) estimator for Markov switching models (see, e.g., Hamilton, 1994, ch. 22), which ignores such correlation, is likely be inconsistent. This inconsistency of the S–PML estimator, combined with the often poor quality of conventional asymptotic inference procedures even in situations where the estimator is consistent (cf. Psaradakis and Sola, 1998), increases considerably the probability of misleading inferences being drawn form the fitted model.

As argued in Psaradakis et al. (2002), the difficulties associated with within-regime orthogonality failures in Markov-switching models can be overcome by using instrumental variables (IV), much in the same way as in single-regime models. Using this approach, we consider the problem of testing the UFER hypothesis using monthly data for the Sterling/Dollar exchange rate. It is demonstrated that, in the context of a model that allows for a time-varying risk premium with Markov regimes is used, the UFER hypothesis cannot be rejected provided that instrumental variables are used to account for the failure of independence between right-hand side variables and disturbances. Moreover, such a model outperforms the forward rate as a mean of forecasting the spot exchange rate.

To fix ideas and notation, the next section of the paper outlines a simple theoretical model for forward exchange pricing and explains how a Markov switching foreign-exchange risk premium can arise. Section 3 discusses the results of standard regression-based tests of the UFER hypothesis and examines the stochastic properties of the risk premium. Section 4 presents our empirical Markov switching model and a small-scale simulation study of the properties of the tests of the UFER hypothesis that are used in our analysis. It also reports and discusses the results from a post-sample analysis that examines whether the model proposed in the paper could be used to improve forecasts of the spot exchange rate. Section 5 summarises and concludes.

2 Modelling the Forward Exchange Rate

Consider a standard infinite-horizon consumption CAPM model in which the behaviour of the equilibrium real return r on any asset is governed by the Euler condition

$$u'(C_t) = \beta E_t[(1 + r_{t+1})u'(C_{t+1})], \tag{1}$$

where $u(\cdot)$ is the utility function, C_t is consumption at date t, β is the discount factor, and $E_t[\cdot]$ denotes mathematical expectation conditional on information available at date t. Using a similar condition for the real ex post return on a foreign asset, and assuming that purchasing power parity and covered interest parity hold, it is easily deduced that

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

where F_t is the one-period forward exchange rate, S_t is the spot exchange rate, and P_t is the price level. Hence, under the assumption that all the variables in (2) are jointly lognormally distributed, the equation may be rewritten as

$$f_t = E_t[s_{t+1}] + \tfrac{1}{2}\text{Var}_t[s_{t+1}] + \text{Cov}_t[R_{t+1}, s_{t+1}], \tag{3}$$

where $f_t = \ln F_t$, $s_t = \ln S_t$, $R_{t+1} = \ln(u'(C_{t+1})/P_{t+1})$, and $\text{Var}_t[\cdot]$ and $\text{Cov}_t[\cdot, \cdot]$ respectively denote variance and covariance conditional on information available at date t. The last two terms on the right-hand side of (3) are typically interpreted in the literature as the sum of the foreign-exchange risk premium and two Jensen inequality terms associated with Siegel's paradox (see, e.g., Obstfeld and Rogoff, 1996, § 8.7.5).

The basis for many empirical investigations of the behaviour of the forward exchange rate is a simplified version of (3) associated with the so-called UFER hypothesis. The latter states that

$$f_t = E_t[s_{t+1}], \tag{4}$$

and

$$s_{t+1} = E_t[s_{t+1}] + \eta_{t+1}, \tag{5}$$

with the public's expectation about s_{t+1} being equal to $E_t[s_{t+1}]$ under rational expectations. Since $\{s_t\}$ and $\{f_t\}$ are typically found in practice to be best described as integrated processes of order one which cointegrate with cointegrating parameter unity, (4)–(5) are often expressed as

$$\Delta s_{t+1} = \alpha + \beta(f_t - s_t) + e_{t+1}, \tag{6}$$

where Δ is the differencing operator defined by $\Delta y_t = y_t - y_{t-1}$. In this representation, the UFER hypothesis is equivalent to $\alpha = 0$ and $\beta = 1$.[1] Under the UFER hypothesis, the logarithm of the forward rate provides an unbiased forecast of the logarithm of the future spot exchange rate.

A common empirical finding, however, is that $\beta < 1$ in (6) (see Engel, 1996). A number of explanations for the empirical failure of this simple model have been offered in the literature leading to the conclusion that the UFER assumption is inappropriate probably because of a breakdown of the rational expectations and/or risk neutrality on the part of economic agents. In such cases, a common strategy is to investigate a "weaker" form of the UFER hypothesis in which the forward rate contains a component which varies randomly over time. Thus, (4) is replaced by

$$f_t = E_t[s_{t+1}] + v_t, \tag{7}$$

in which the forward exchange rate is the sum of the expected future spot rate and a time-varying premium component v_t which may represent compensation for risk-averse speculators in the forward market for holding a net position in foreign exchange.[2]

Many researchers have also shown that the evolution of consumption over time can be described well by dynamic models the parameters of which depend on the state of an (exogenous) finite Markov chain (e.g., Cecchetti et al., 1990; Hall et al., 1997). When combined with the hypothesis of time-varying risk premia, such dynamic behaviour for consumption implies that the risk premium itself is subject to Markov structural changes. To illustrate, let us assume that

$$C_t = \mu_{x_t} + \sum_{j=1}^{h} \varphi_{j,x_t} C_{t-j} + \sigma_{x_t} \zeta_t, \tag{8}$$

where $\{\zeta_t\}$ is white noise and $\{x_t\}$ are regime-indicator variables, independent of $\{\zeta_t\}$. Assume further that nature selects regime at date t with a probability that depends on what regime nature was in at date $t-1$, so that $\{x_t\}$ form a time-homogeneous first-order Markov chain on $\{0, 1\}$ with transition probabilities

$$q = \Pr[x_t = 0 | x_{t-1} = 0], \qquad p = \Pr[x_t = 1 | x_{t-1} = 1]. \tag{9}$$

[1]Since this hypothesis is a consequence of the assumption that market participants use available information efficiently, the hypothesis is sometimes stated in a stronger form in which the forward rate is an unbiased predictor of the future spot rate and the associated prediction errors are serially uncorrelated.

[2]In terms of our theoretical model, $v_t = (1/2)\mathrm{Var}_t[s_{t+1}] + \mathrm{Cov}_t[R_{t+1}, s_{t+1}]$. In the remainder of the paper we shall refer to v_t as the risk premium even though, as mentioned before, it also incorporates two Jensen inequality terms.

Then, it is not difficult to show that, conditionally on $x_t = 0$, the solution for the forward rate is

$$f_t = E_t[s_{t+1}] + \tfrac{1}{2}\text{Var}_t[s_{t+1}] + q\text{Cov}_t[R^{(0)}_{t+1}, s_{t+1}] + (1-q)\text{Cov}_t[R^{(1)}_{t+1}, s_{t+1}], \qquad (10)$$

while, conditionally on $x_t = 1$, the solution is

$$f_t = E_t[s_{t+1}] + \tfrac{1}{2}\text{Var}_t[s_{t+1}] + (1-p)\text{Cov}_t[R^{(0)}_{t+1}, s_{t+1}] + p\text{Cov}_t[R^{(1)}_{t+1}, s_{t+1}], \qquad (11)$$

where $R^{(i)}_t = \ln(u'(C^{(i)}_{t+1})/P_{t+1})$ and $C^{(i)}_{t+1} = \mu_i + \sum_{j=1}^{h} \varphi_{j,i} C_{t+1-j} + \sigma_i \zeta_{t+1}$ for $i \in \{0, 1\}$. This clearly leads to a model with a risk-premium process which is subject to Markov regime-switching. Hence, one may think of the UFER hypothesis as being represented by (7) but with $\{v_t\}$ following a stochastic processes with parameters that are subject to Markov changes. As we shall argue in the sequel, such a characterisation of the foreign-exchange risk premium is consistent with the empirical evidence obtained from our data set.

3 Testing the UFER Hypothesis

The data set used for our empirical analysis consists of 164 end-of-month observations on the natural logarithm of the spot and forward (thirty-day rate) Sterling/Dollar exchange rate for the period from January 1987 to August 2000 (the time series of the spread $f_t - s_t$ is shown in Figure 1). In agreement with other studies, the spot and forward exchange rates are found to be integrated of order one and to cointegrate with cointegrating parameter unity (at least at the 5% significance level). More specifically, the sieve bootstrap t-test for an autoregressive unit root (in a model with a constant term) proposed by Psaradakis (2001) has a P-value of 0.081, 0.079 and 0.041 for $\{s_t\}$, $\{f_t\}$ and $\{f_t - s_t\}$, respectively.[3]

3.1 OLS and IV Results

Starting with the simple model in (6), Table 1 reports the OLS estimates of α and β, along with corresponding Eicker–White standard errors. Evidently, the estimate of β differs markedly from the theoretically correct value of unity and the hypothesis $\beta = 1$ is rejected at the 5% significance level;[4] the Wald test statistic for testing the joint hypothesis $\alpha = \beta - 1 = 0$ has a P-value of 0.106. Moreover, the portmanteau Q statistics for the residuals and their squares indicate the presence of nonlinear temporal dependence in the estimated residuals from (6).

If we allow for a time-varying risk premium, as in (7), then it is well-known that the OLS estimates in (6) will be inconsistent. IV estimates of the parameters in (6) are obtained using a constant, $f_{t-1} - s_{t-1}$, $f_{t-2} - s_{t-2}$ and $f_{t-3} - s_{t-3}$ as instruments.[5] From Table 1, the equation

[3] The P-values for the tests were obtained from 999 bootstrap replications and the order of the autoregressive sieve used was selected from the range $\{0, 1, \ldots, 12\}$ by means of the Akaike information criterion.

[4] Note that β is not statistically different from zero.

[5] The choice of the instruments is based on the significance of their coefficients in the reduced-form regressions and on the outcome of Sargan's test for instrument validity.

parameters are not significantly different from their theoretically correct values ($\alpha = 0$ and $\beta = 1$) on the basis of a Wald test. Interestingly, while the hypothesis $\beta = 1$ cannot be rejected, the confidence interval for β is very wide and contains negative values of β.[6]

3.2 Stochastic Properties of the Risk Premium

In accordance with the theoretical structure in (10)–(11), we argue that an explanation for these results may lie with the risk premium being subject to discrete Markov shifts. To investigate this possibility, we examine the properties of the series $\{f_t - s_{t+1}\}$, which is plotted in Figure 1. Since, (4)–(5) imply that

$$f_t - s_{t+1} = v_t - \eta_{t+1}, \tag{12}$$

$f_t - s_{t+1}$ must share the same properties with the combined risk premium plus noise process. Hence, any Markov-type nonlinearities in the risk premium are likely to be reflected in the dynamic behaviour of $f_t - s_{t+1}$.

To investigate this possibility, we directly test a single-regime AR(1) model for $f_t - s_{t+1}$ against an alternative in which the autoregressive coefficient and the innovations variance switch between two values according to a (time-homogeneous) Markov transition process. Although the two models are nested, the usual likelihood ratio statistic does not have a chi-squared asymptotic distribution since, under the null hypothesis of a single regime, the transition probabilities are unidentified and the information matrix is singular. To overcome these difficulties, we carry out the test using the standardised likelihood ratio test procedure developed by Hansen (1992). This procedure requires evaluation of the likelihood function across a grid of different values for the transition probabilities and for each state-dependent parameter. Here, the range [0.75, 0.99] in steps of 0.03 (10 gridpoints) is used as a grid for the transition probabilities; for the autoregressive coefficient and the innovations variance, we use the range [0.1, 0.9] and [0.01, 0.17], respectively, in steps of 0.1 and 0.01 (9 gridpoints). The value of the standardised likelihood ratio statistic is 4.281, which has a zero P-value under the null hypothesis.[7] This is strong evidence in favour of Markov regime-switching, especially since Hansen's test is conservative by construction.

It is worth noting that the familiar Akaike information criterion also favours the Markov switching AR(1) model for $f_t - s_{t+1}$ over a single-regime AR(1) model (-686.5 vs -683.2).[8] Moreover, the two-regime models outperform the single-regime ones in terms of the residual diagnostics for non-linear dependence, as can be seen in Table 2.[9]

[6] We also estimated the parameters of (6) using the so-called continuous-updating GMM estimator (Hansen et al., 1996), which is known to be partially robust to weak instruments (see Stock et al., 2002). The GMM estimate of β is -0.724 with standard error 1.315.

[7] The P-value is calculated according to the method described in Hansen (1996), using 1000 random draws from the relevant limiting Gaussian processes and bandwidth parameter $M \in \{0, 1, \ldots, 4\}$.

[8] Psaradakis and Spagnolo (2003) demonstrated the effectiveness of the Akaike information criterion as a mean of selecting the number of regimes in autoregressive models with Markov regime switching.

[9] Diagnostic tests for the switching model are based on standardised residuals computed as in Town (1992).

4 Markov Switching and the UFER Hypothesis

In the light of the empirical evidence reported in the previous section, we proceed to model the forward exchange rate under the assumption that the risk premium process is subject to Markov changes in regime. As explained in Section 2, such an assumption is consistent with a consumption CAPM model where consumption is subject to Markov regime-switching.

4.1 Empirical Model

Letting $\{x_t\}$ be a time-homogeneous first-order Markov chain on $\{0,1\}$ with transition probabilities $p = \Pr[x_t = 1|x_{t-1} = 1]$ and $q = \Pr[x_t = 0|x_{t-1} = 0]$, we assume that $\{v_t\}$ satisfies the difference equation

$$v_t = \rho_{x_t} v_{t-1} + \omega_{x_t}\varepsilon_t, \tag{13}$$

where $\{\varepsilon_t\}$ is a white noise with zero mean and unit variance, independent of $\{x_t\}$.[10] Then, (5), (7) and (13) imply that

$$\Delta s_{t+1} = (f_t - s_t) - \rho_{x_t}(f_{t-1} - s_t) + \varepsilon^*_{t+1}, \tag{14}$$

where $\varepsilon^*_{t+1} = \eta_{t+1} - \omega_{x_t}\varepsilon_t - \rho_{x_t}\eta_t$.

Equation (14) can provide the basis for a test of the UFER hypothesis. Notice, however, that s_t, f_t and f_{t-1} are correlated with the noise term ε^*_{t+1} (in each regime) through η_{t+1} and η_t. It is, therefore, likely that S–PML estimates of the parameters in (14) will be inconsistent. One way of overcoming the difficulties associated with within-regime correlation between the regressors and the disturbance term is to use some form of IV estimation where the reduced-form equations relating the endogenous regressors to the instruments have state-dependent parameters too.

Our analysis is based on the following system of equations:

$$\Delta s_{t+1} = \alpha + \beta z^*_{1t} - \gamma_{x_t} z^*_{2t} + \sigma_{x_t}\xi_{1,t+1}, \tag{15}$$

$$f_t - s_t = \delta^{(1)}_{x_t} + \delta^{(2)}_{x_t}(f_{t-1} - s_{t-1}) + \delta^{(3)}_{x_t}(f_{t-2} - s_{t-2}) + \theta_{x_t}\xi_{2t}, \tag{16}$$

$$f_{t-1} - s_t = \lambda^{(1)}_{x_t} + \lambda^{(2)}_{x_t}(f_{t-2} - s_{t-1}) + \lambda^{(3)}_{x_t}(f_{t-3} - s_{t-2}) + \phi_{x_t}\xi_{3t}, \tag{17}$$

where $z^*_{1t} = E[(f_t - s_t)|x_t, \Omega_t]$, $z^*_{2t} = E[(f_{t-1} - s_t)|x_t, \Omega_t]$, $\Omega_t = \{(s_m, f_m), 1 \leqslant m \leqslant t\}$, and $E[\xi_{it}] = E[\xi^2_{it} - 1] = 0$ for $i \in \{1,2,3\}$. In (15), the UFER hypothesis is equivalent to $\alpha = 0$ and $\beta = 1$.

Estimation and testing in the context of the Markov switching model in (15)–(17) – where neither the error terms $\xi_{1,t+1}$, ξ_{2t} and ξ_{3t} nor the Markov chain x_t are observed – can be carried

[10] To ensure covariance stationarity of $\{v_t\}$, it is also assumed that $q\rho_0^2 + p\rho_1^2 + (1-p-q)\rho_0^2\rho_1^2 < 1$ and $q\rho_0^2 + p\rho_1^2 < 2$ (cf. Francq and Zakoïan, 2001).

out by using the a variant of the recursive algorithm discussed in Hamilton (1994, ch. 22), which is presented in the Appendix. This gives as a by-product the sample likelihood function which can be maximised numerically with respect to (α, β, γ_0, γ_1, σ_0, σ_1, $\delta_0^{(1)}$, $\delta_1^{(1)}$, $\delta_0^{(2)}$, $\delta_1^{(2)}$, $\delta_0^{(3)}$, $\delta_1^{(3)}$, θ_0, θ_1, $\lambda_0^{(1)}$, $\lambda_1^{(1)}$, $\lambda_0^{(2)}$, $\lambda_1^{(2)}$, $\lambda_0^{(3)}$, $\lambda_1^{(3)}$, ϕ_0, ϕ_1, p, q), subject to the constraint that p and q lie in the open unit interval. We refer to estimates obtained from this procedure as IV–PML estimates. In general, the structure of the instrumenting equations (16) and (17) has to be determined empirically. In our case, two lags of $f_t - s_t$ and $f_{t-1} - s_t$ were found to be highly significant and enough to produce well-behaved (standardised) residuals.[11]

In Table 3, we report Gaussian IV–PML estimates of the parameters in (15), along with corresponding standard errors.[12] The fitted model is well-specified, having standardised residuals which exhibit no signs of linear or nonlinear temporal dependence. The IV–PML point estimate of β is very close to the value implied by the UFER hypothesis and the hypotheses $\beta = 1$ and $\alpha = \beta - 1 = 0$ cannot be rejected at any significance level. Finally, the estimated transition probabilities suggest that the Markov chain that drives changes in regime is highly persistent, so if the system is in either of the two regimes, it is likely to remain in that regime.

Figure 2 shows plots of the inferred probabilities of being in the regime represented by $x_t = 1$ at each point in the sample, along with the risk-premium $f_{t-1} - s_t$.[13] It is clear that there were several changes in regime. For example, the filter allocates a high probability of state $x_t = 0$ in the early 1990s, probably caused by the UK's exit from the ERM following disruption on the financial markets.

To evaluate the empirical relevance of the use of the IV–PML estimator, we also estimate the parameters in (14) using the standard algorithm (Hamilton, 1994, ch. 22). Thus, in Table 3 we report S–PML parameter estimates for the model

$$\Delta s_{t+1} = \alpha + \beta(f_t - s_t) - \gamma_{x_t}(f_{t-1} - s_t) + \sigma_{x_t}\xi_{t+1}. \tag{18}$$

As in the IV–PML case, the fitted model appears to be well-defined. The S–PML estimate of β, however, differs significantly from the value that is consistent with the UFER hypothesis and the hypotheses $\beta = 1$ and $\alpha = \beta - 1 = 0$ are rejected at any significance level. We speculate that the most likely explanation of this result is the inconsistency of the S–PML estimator due to orthogonality failure. In the following section we carry out some sampling experiments to investigate the validity of this claim.

[11] In the following section we carry out some sampling experiments to investigate how weak instruments can affect the power properties of our test.

[12] The likelihood function was maximized by using the Broyden–Fletcher–Goldfarb–Shanno variable-metric algorithm with numerically computed derivatives. Estimated standard errors are obtained using the Huber–White sandwich estimator.

[13] These are the filter probabilities $\Pr[x_t = 1|\Omega_t; \widehat{\boldsymbol{\vartheta}}]$, where $\widehat{\boldsymbol{\vartheta}}$ is the IV–PML estimate of the structural parameters in (15)–(17).

4.2 A Simulation Study

As mentioned earlier, the presence of f_t in the right-hand side of (14) and the replacement of the latent term $f_{t-1} - E_{t-1}[s_t]$ by the observed $f_{t-1} - s_t$ mean that the S–PML estimator of the parameters in (18) is likely to be inconsistent. As a means of overcoming the problem, we proposed to use an IV–PML estimator. In this subsection of the paper, we investigate the properties of different estimators using a few sampling experiments. The data-generating process in our simulation study is defined by the equations

$$\begin{aligned} f_t &= E_t[s_{t+1}] + v_t, \\ v_t &= \rho_{x_t} v_{t-1} + \varepsilon_t, \\ s_t &= s_{t-1} + \eta_t, \end{aligned} \tag{19}$$

where, conditionally on x_t,

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim \text{NID}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{x_t}^{(11)} & \sigma_{x_t}^{(12)} \\ \sigma_{x_t}^{(12)} & \sigma_{22} \end{bmatrix} \right). \tag{20}$$

As before, $\{x_t\}$ is a time-homogeneous first-order Markov chain on $\{0, 1\}$ with transition probabilities $p = \Pr[x_t = 1 | x_{t-1} = 1]$ and $q = \Pr[x_t = 0 | x_{t-1} = 0]$, so our data-generating mechanism is consistent with the UFER hypothesis with a Markov switching risk premium. To ensure the relevance of our simulations, the values of the parameters $(\rho_{x_t}, \sigma_{x_t}^{(11)}, \sigma_{x_t}^{(12)}, \sigma_{22}, p, q)$ are chosen on the basis of empirical results obtained with the Sterling/Dollar exchange-rate data. The initial values for s_t and f_t are historical values from our sample, while the initial value for x_t is drawn from its estimated stationary distribution. The experiments proceed by generating artificial time series for s_t and f_t of length $T \in \{150, 300, 500, 1000\}$. Then, in each of 1000 Monte Carlo replications, the OLS and IV estimates of the parameters of (6) and the Gaussian S–PML and IV–PML estimates of the parameters of (18) and (15)–(17), respectively, are computed, as well as a t-statistic and a Wald statistic for respectively testing the null hypotheses $\beta = 1$ and $\alpha = \beta - 1 = 0$.

4.2.1 OLS, IV, S–PML

Table 4 reports the mean and standard deviation of the empirical distribution of the OLS, IV, and S–PML estimators of β, along with the rejection frequencies of the corresponding 5%-level t-type tests for $\beta = 1$ and Wald tests for the joint hypothesis $\alpha = \beta - 1 = 0$. The OLS and S–PML estimators are evidently severely biased away from unity, a bias which remains substantial even in samples of 1000 observations. As a consequence, the corresponding t-tests and Wald tests falsely reject the null hypotheses $\beta = 1$ and $\alpha = \beta - 1 = 0$ in almost all Monte Carlo replications. Using the standard IV estimator does not rescue the UFER hypothesis, with the bias remaining substantial (probably due to the omitted variable problem) and the test rejection frequencies only slowly decreasing with the sample size.[14]

4.2.2 IV–PML

In addition to providing evidence on the finite-sample properties of the IV–PML estimator, our Monte Carlo experiments also assess the relevance of the instruments used in our application. The latter is a worthwhile exercise since standard statistical inference procedures based on IV-type estimators can be considerably unreliable when marginally relevant (or weak) instruments are used. The difficulty is that the various procedures that have been proposed in the literature for detecting weak instruments and carrying out robust hypothesis tests (see Stock et al., 2002) cannot be used in the context of a Markov switching model.

Building on the suggestion of Stock and Yogo (2002) that the definition of weak instruments should depend on the purpose to which the instruments are put to, we consider instruments to be relevant or strong from the perspective of a test of the UFER hypothesis if the actual probability of Type I error of the test is close to its nominal level. Alternatively, instruments may be considered to be strong if the bias of the IV–PML estimator (relative to the S–PML estimator or an IV–PML estimator based on a different set of instruments) is fairly small.

To address these issues, IV–PML estimates of the parameters of (15) are computed in each Monte Carlo replication using the following sets of instruments:[15]

(S1) $(1, f_{t-1} - s_{t-1}, f_{t-2} - s_{t-2})$ for $f_t - s_t$ and $(1, f_{t-2} - s_{t-1}, f_{t-3} - s_{t-2})$ for $f_{t-1} - s_t$ (these are the instruments used in Section 4.1);

(S2) $(1, f_{t-1} - s_{t-1})$ for $f_t - s_t$ and $(1, f_{t-2} - s_{t-1})$ for $f_{t-1} - s_t$;

(S3) $(1, f_{t-1} - s_{t-1}, f_{t-2} - s_{t-1})$ for both $f_t - s_t$ and $f_{t-1} - s_t$.

In each case, the Monte Carlo mean and standard deviation of the distribution of the IV–PML estimator of β are computed, along with the rejection frequencies of tests of the hypotheses $\beta = 1$ and $\alpha = \beta - 1 = 0$. We also compute the rejection frequencies of tests of the hypotheses $\beta = 0$ and $\delta_i^{(j)} = \lambda_i^{(j)} = 0$ ($i \in \{0, 1\}$, $j \in \{1, 2, 3\}$); the latter may be thought of as the equivalent of the F-test for weak instruments that is often used in linear IV regressions.

From the results reported in Table 5, it is clear that the IV–PML estimator outperforms all other estimators. Although it is slightly biased in the smaller sample (when the instruments S2 and S3 are used), the bias is a decreasing function of the sample size and the empirical size of the t-test and Wald test of the UFER hypothesis converges to the nominal 5% value.

[14] As in the empirical study, estimates of the parameters are obtained using a constant, $f_{t-1} - s_{t-1}$, $f_{t-2} - s_{t-2}$ and $f_{t-3} - s_{t-3}$ as instruments. We also experimented with alternative sets of instruments with little change in the simulation results.

[15] Although we experimented with several alternative sets of instruments, we shall only report results for a few sets, which should, however, be sufficient to illustrate the main point.

Turning to tests of the significance of β, $\delta_i^{(j)}$ and $\lambda_i^{(j)}$, when the instruments S1 are used the hypothesis $\beta = 0$ is rejected in almost all Monte Carlo replications, while rejection rates are not as high when the instruments S2 and S3 are used. A possible explanation may be that instruments S2 and S3 are weak compared to the instruments S1. This explanation is further corroborated by the fact that the instruments S1 appear to be jointly significant much more often than the instruments S2 and S3. In addition, the bias of the IV–PML estimator is smaller when the instruments S1 are used (especially in smaller samples) and so is the size distortions of tests of the UFER hypothesis.

In summary, it appears that failure to take into account within-regime correlation between a regressor and the disturbance term in the Markov switching model will result in almost certain rejection of the UFER hypothesis, a problem which can be overcome, to a large extent, by the use of the IV–PML estimator, provided that relevant instruments are used.

4.3 A Forecast Exercise

In this subsection we examine how a model that allows for the presence of a time-varying foreign-exchange risk premium can be exploited for forecasting purposes. More specifically, assuming that the UFER hypothesis holds (i.e., $\alpha = \beta - 1 = 0$), we investigate whether taking into account the information contained in the autocorrelation structure of the risk premium yields improvements in the accuracy of out-of-sample forecasts of the forward exchange rate.

The competing forecasting models under consideration are the following:

(M1) The standard model (6) with no risk premium, which implies that

$$\widehat{s}_{T+1} = f_T, \tag{21}$$

where $\widehat{s}_{T+1}$ denotes the one-step-ahead forecast at forecast origin T.

(M2) A model with a linear AR(1) risk premium, i.e. $s_{t+1} = f_t + \gamma(f_{t-1} - s_t) + e_t$, for which

$$\widehat{s}_{T+1} = f_T + \widehat{\gamma}(f_{T-1} - s_T) \tag{22}$$

where $\widehat{\gamma}$ is an IV estimate of γ.

(M3) A model with a Markov switching AR(1) risk premium, as in (14), for which

$$\widehat{s}_{T+1} = f_T + (1 - \widehat{\pi}_T)\{\widehat{q}\widehat{\gamma}_0 + \widehat{\gamma}_1(1 - \widehat{q})\} + \widehat{\pi}_T\{(1 - \widehat{p})\widehat{\gamma}_0 + \widehat{\gamma}_1\widehat{p}\}(f_{T-1} - s_T), \tag{23}$$

where $(\widehat{q}, \widehat{p}, \widehat{\gamma}_0, \widehat{\gamma}_1)$ is the IV–PML estimate of $(q, p, \gamma_0, \gamma_1)$ and $\widehat{\pi}_t = \Pr[x_t = 1|\Omega_t; \widehat{\boldsymbol{\vartheta}}]$.[16]

Comparison of M2 and M3 with M1 reveals what the loss in terms of forecast performance is from using the simple forward rate as a forecast of the future spot rate in the presence of a time-varying risk premium in the forward market.

[16] Here we assume that the public's information consists only of observations on the exchange rates.

Forecast comparisons are carried out as follows. For each of 12 subseries $\{(s_1, f_1), \ldots, (s_\ell, f_\ell)\}$ of the data, with $\ell \in \{152, \ldots, 163\}$, we estimate the parameters of the forecasting models and compute the one-period-ahead forecast and the associated forecast error. Based on these 12 forecast errors, we then compute traditional loss measures, which include the mean squared error (MSE), the root mean squared error (RMSE), the mean absolute error (MAE), the mean squared percent error (MSPE), the root mean squared percent error (RMSPE), and the mean absolute percentage error (MAPE). In addition, we also calculate Theil's inequality coefficient.

Table 6 summarises the results of the forecast comparisons in terms of ratios of the individual loss measures. The results are all in favour of the M3. Based on the MSE and the MSPE criteria, there is a loss of 16% when M2 is used against the M3 specification and 5% against the forward rate (M1). The other criteria tell the same story. Turning to M3, the results indicate a gain of 10% against the forward rate using the MSE and MSPE, and similar results are obtained using the other accuracy criteria. In summary, it appears that using linear models as an approximation to nonlinear one could result in substantially inferior forecast performance.

5 Summary

In this paper, we have considered the problem of testing the unbiased forward exchange rate (UFER) hypothesis in the presence of structural instability. Our analysis has been based on a theoretical model which allows for the presence of a time-varying risk premium in the forward market and Markov regime-switching behaviour in consumption. Under these conditions, the risk-premium process itself is subject to Markov changes in regime. As a consequence, a model for the spot rate and premium has parameters that depend on the state of nature and explanatory variables which are correlated with the disturbances. It is, therefore, reasonable to expect the standard pseudo-maximum likelihood estimator for Markov-switching models to be inconsistent, and have argued that valid inference requires the use of an estimator like the instrumental variables pseudo-maximum likelihood estimator discussed in this paper and in Psaradakis et al. (2002). Using such inference procedures, we have shown that the UFER hypothesis cannot be rejected for the Sterling/Dollar forward and spot exchange rates. We have also carried out a few Monte Carlo experiments which have demonstrated that, even if we account for structural breaks, failure to take into account the orthogonality failure within regimes will almost certainly result in the rejection of the UFER hypothesis. Finally, we have shown that the proposed Markov switching model provides more accurate forecasts of the spot exchange rate than the forward rate.

6 Appendix: IV–PML Estimation

Estimates of the parameters of regime-switching models are obtained using procedures which are similar to those described in Hamilton (1994, ch. 22). In the case of IV–PML estimation, the regressors are instrumented and the reduced-form regressions for the instruments have state-dependent parameters. The conditional probability density function of the data $\mathbf{w}_t = (\Delta s_{t+1}, f_t - s_t, f_{t-1} - s_t)'$ given the state x_t and the history of the system can thus be written as

$$\begin{aligned}
\mathrm{pdf}(\mathbf{w}_t | x_t, \mathbf{w}_{t-1}, ..., \mathbf{w}_1; \boldsymbol{\vartheta}) \quad &= \quad \frac{1}{\sqrt{2\pi\sigma_{x_t}^2}} \exp\left\{ -\frac{1}{2} \left(\frac{\Delta s_{t+1} - \alpha - \beta \widetilde{z}_{1t} - \gamma_{x_t} \widetilde{z}_{2t}}{\sigma_{x_t}} \right)^2 \right\} \\
&\times \frac{1}{\sqrt{2\pi\theta_{x_t}^2}} \exp\left\{ -\frac{1}{2} \left(\frac{z_{1t} - \delta_{x_t}^{(1)} - \delta_{x_t}^{(2)} z_{1,t-1} - \delta_{x_t}^{(3)} z_{1,t-2}}{\theta_{x_t}} \right)^2 \right\} \\
&\times \frac{1}{\sqrt{2\pi\phi_{x_t}^2}} \exp\left\{ -\frac{1}{2} \left(\frac{z_{2t} - \lambda_{x_t}^{(1)} - \lambda_{x_t}^{(2)} z_{2,t-1} - \lambda_{x_t}^{(3)} z_{2,t-2}}{\phi_{x_t}} \right)^2 \right\}.
\end{aligned}$$

Here, $z_{1t} = f_t - s_t$, $z_{2t} = f_{t-1} - s_t$, $\widetilde{z}_{1t} = \delta_{x_t}^{(1)} + \delta_{x_t}^{(2)} z_{1,t-1} + \delta_{x_t}^{(3)} z_{1,t-2}$, $\widetilde{z}_{2t} = \lambda_{x_t}^{(1)} + \lambda_{x_t}^{(2)} z_{2,t-1} + \lambda_{x_t}^{(3)} z_{2,t-2}$, and $\boldsymbol{\vartheta} = (\alpha, \beta, \gamma_0, \gamma_1, \sigma_0, \sigma_1, \delta_0^{(1)}, \delta_1^{(1)}, \delta_0^{(2)}, \delta_1^{(2)}, \delta_0^{(3)}, \delta_1^{(3)}, \theta_0, \theta_1, \lambda_0^{(1)}, \lambda_1^{(1)}, \lambda_0^{(2)}, \lambda_1^{(2)}, \lambda_0^{(3)}, \lambda_1^{(3)}, \phi_0, \phi_1, p, q)'$.

References

[1] Cecchetti, S. G., Lam, P., and Mark, N. C. (1990), Mean Reversion in Equilibrium Asset Prices, *American Economic Review*, 80, 398–418.

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

[3] Francq, C., and Zakoïan, J.-M. (2001), Stationarity of Multivariate Markov-Switching ARMA Models, *Journal of Econometrics*, 102, 339–364.

[4] Hall, S. G., Psaradakis, Z., and Sola, M. (1997), Cointegration and Changes in Regime: The Japanese Consumption Function, *Journal of Applied Econometrics*, 12, 151–168.

[5] Hamilton, J. D. (1994), *Time Series Analysis*, Princeton: Princeton University Press.

[6] Hansen, B. E. (1992), The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP, *Journal of Applied Econometrics*, 7, S61–S82.

[7] Hansen, B. E. (1996), Erratum: The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP, *Journal of Applied Econometrics*, 11, 195–198.

[8] Hansen, L. P., Heaton, J., and Yaron, A. (1996), Finite Sample Properties of Some Alternative GMM Estimators, *Journal of Business and Economic Statistics*, 14, 262–280.

[9] Obstfeld, M., and Rogoff, K. (1996), *Foundations of International Macroeconomics*, Cambridge, Massachusetts: MIT Press.

[10] Psaradakis, Z. (2001), Bootstrap Tests for an Autoregressive Unit Root in the Presence of Weakly Dependent Errors, *Journal of Time Series Analysis*, 22, 577–594.

[11] Psaradakis, Z., and Sola, M. (1998), Finite-Sample Properties of the Maximum Likelihood Estimator in Autoregressive Models with Markov Switching, *Journal of Econometrics*, 86, 369–386.

[12] Psaradakis, Z., Sola, M., and Spagnolo, F. (2002), Estimating Fama–Campbell–Shiller Term Structure Equations in the Presence of Risk Premia and Markov Regime Shifts, mimeo, School of Economics, Mathematics and Statistics, Birkbeck College, University of London.

[13] Psaradakis, Z., and Spagnolo, N. (2003), On the Determination of the Number of Regimes in Markov-Switching Autoregressive Models, *Journal of Time Series Analysis*, forthcoming.

[14] Stock, J. H., Wright, J. H., and Yogo, M. (2002), A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments, *Journal of Business and Economic Statistics*, 20, 518–529.

[15] Stock, J. H., and Yogo, M. (2002), Testing for Weak Instruments in Linear IV Regression, mimeo, Department of Economics, Harvard University.

[16] Town, R. J. (1992), Merger Waves and the Structure of Merger and Acquisition Time-Series, *Journal of Applied Econometrics*, 7, S83–S100.

Table 1. Estimation Results for Equation (6)

	OLS		IV	
α	0.000	(0.003)	-0.003	(0.003)
β	-0.256	(0.623)	-1.142	(1.352)
Std. Err.	0.029		0.030	
$t(\beta = 1)$	-2.016	[0.044]	-1.584	[0.113]
Wald $(\alpha=\beta-1=0)$	4.488	[0.106]	2.598	[0.273]
$Q(12)$	5.200	[0.951]	5.799	[0.926]
$Q(18)$	10.206	[0.925]	9.850	[0.937]
$Q_2(12)$	24.988	[0.015]	23.782	[0.022]
$Q_2(18)$	26.181	[0.096]	26.010	[0.100]

NOTES: Estimated standard errors are in parentheses and P-values for test statistics are in square brackets. $Q(k)$ is the residual Ljung–Box statistic at lag k, $Q_2(k)$ is the squared-residual Ljung–Box statistic at lag k, and $t(\beta = 1)$ is the t-statistic for the hypothesis $\beta = 1$. Wald is the Wald statistic for testing the hypothesis $\alpha = \beta - 1 = 0$.

Table 2. Diagnostics for Models of the Risk Premium

	AR(1)		Switching AR(1)	
$Q(12)$	5.122	[0.953]	5.643	[0.933]
$Q(18)$	11.364	[0.878]	10.782	[0.903]
$Q_2(12)$	29.152	[0.003]	5.355	[0.945]
$Q_2(18)$	30.801	[0.030]	8.894	[0.962]

NOTE: $Q(k)$ is the residual Ljung–Box statistic at lag k, and $Q_2(k)$ is the squared-residual Ljung–Box statistic at lag k. P-values for test statistics are in square brackets.

Table 3. Estimation Results for Equation (18)

	S–PML		IV–PML	
α	0.000	(0.002)	0.000	(0.002)
β	0.030	(0.362)	0.822	(0.167)
γ_0	0.217	(0.107)	−0.923	(0.003)
γ_1	−0.162	(0.126)	−0.358	(0.012)
σ_0	0.020	(0.002)	0.018	(0.002)
σ_1	0.036	(0.002)	0.036	(0.001)
q	0.988	(0.011)	0.972	(0.026)
p	0.994	(0.023)	0.986	(0.025)
logL	358.14		1450.9	
$t(\beta = 1)$	−2.679	[0.007]	− 1.066	[0.286]
Wald ($\alpha{=}\beta{-}1{=}0$)	25.740	[0.000]	0.060	[0.970]
$Q(12)$	5.374	[0.944]	7.088	[0.852]
$Q(18)$	8.407	[0.972]	10.378	[0.919]
$Q_2(12)$	9.163	[0.689]	13.665	[0.322]
$Q_2(18)$	11.743	[0.860]	16.099	[0.586]

NOTE: See notes to Table 1.

Table 4. Results of Monte Carlo Experiments: OLS, IV, S–PML

				Rej. Freq.	
	T	Bias	Std. Dev.	t-test ($\beta{=}1$)	Wald ($\alpha{=}\beta{-}1{=}0$)
OLS	150	−0.994	0.085	1.000	1.000
	300	−0.995	0.059	0.998	1.000
	500	−0.995	0.045	1.000	1.000
	1000	−0.995	0.032	0.995	1.000
IV	150	−1.025	0.378	0.540	0.432
	300	−0.995	0.515	0.362	0.401
	500	−0.965	0.464	0.425	0.429
	1000	−0.993	0.452	0.402	0.411
S–PML	150	−0.979	0.115	1.000	1.000
	300	−0.975	0.080	0.999	1.000
	500	−0.975	0.060	0.994	0.999
	1000	−0.973	0.042	0.991	1.000

Table 5. Results of Monte Carlo Experiments: IV–PML

				Rej. Freq.			
IV–PML	T	Bias	Std. Dev.	t-test (β=1)	Wald (α=β−1=0)	t-test (β=0)	Wald ($\delta_i^{(j)}$=$\lambda_i^{(j)}$=0)
S1	150	−0.081	0.127	0.073	0.068	0.990	0.981
	300	−0.065	0.081	0.057	0.060	1.000	0.990
	500	−0.006	0.042	0.053	0.058	1.000	1.000
	1000	−0.001	0.004	0.049	0.054	1.000	1.000
S2	150	−0.242	0.512	0.129	0.137	0.632	0.851
	300	−0.137	0.313	0.088	0.090	0.689	0.883
	500	−0.113	0.112	0.070	0.079	0.798	0.892
	1000	−0.051	0.031	0.061	0.065	0.851	0.930
S3	150	−0.231	0.459	0.106	0.112	0.723	0.691
	300	−0.089	0.258	0.060	0.081	0.738	0.701
	500	−0.044	0.154	0.053	0.061	0.813	0.835
	1000	−0.009	0.018	0.046	0.057	0.872	0.891

Table 6. Out-of-Sample Forecast Comparisons

	MSE	MAE	RMSE	MSPE	MAPE	RMSPE	Theil
M1	0.00101	0.02342	0.03180	0.00765	0.06148	0.24796	0.03840
M2	0.00106	0.02497	0.03265	0.00805	0.06563	0.25619	0.03939
M3	0.00092	0.02251	0.03042	0.00688	0.05798	0.24080	0.03622

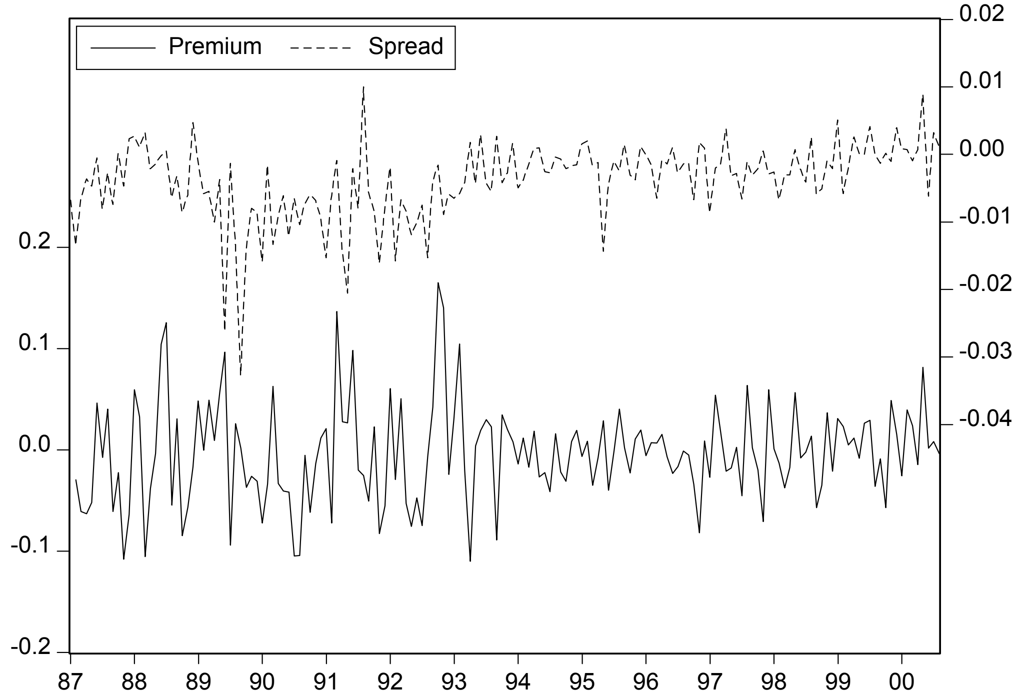

Figure 1: Premium ($f_t - s_{t+1}$) and Spread ($f_t - s_t$).

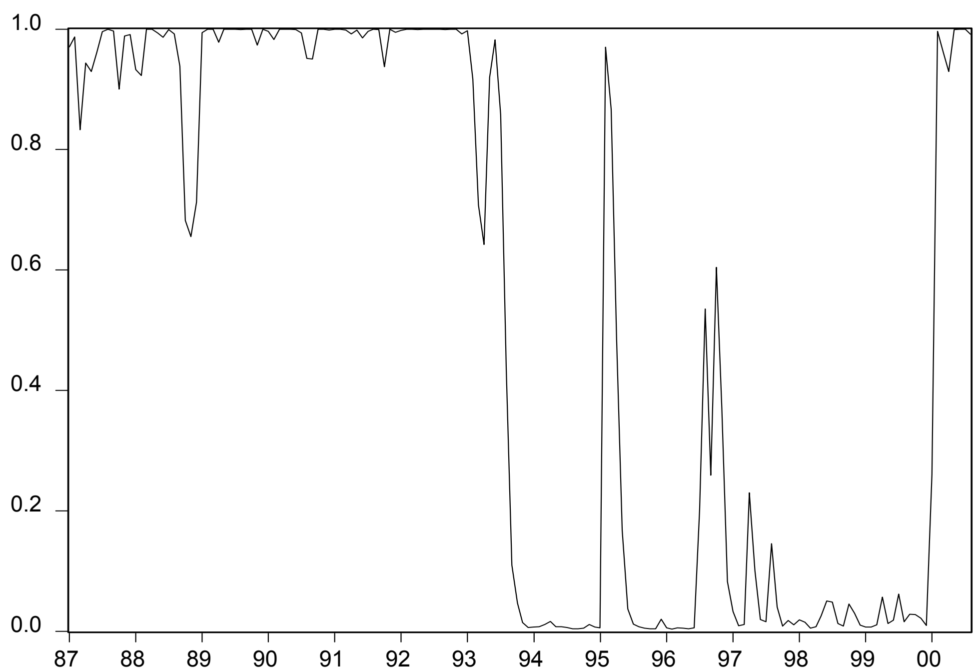

Figure 2: Inferred probability of being in the regime characterized by $x_t = 1$.