A RECONSIDERATION OF UNCOVERED INTEREST RATE PARITY
UNDER SWITCHING POLICY REGIMES

Dimitris G. Kirikos *
Technological Educational Institute of Crete
and Hellenic Open University, Greece

* Postal address: Department of Accounting, Technological Educational Institute of Crete, 710 04
Heraklion, Crete, Greece. Tel. +30 2810 379601, Fax +30 2810 254237
E-mail: kirikosd@sdo.teiher.gr
ABSTRACT

It is shown that when monetary authorities manage the interest rate through anti-inflationary policy rules which allow for occasional discrete shifts, then a Markov switching regimes representation is appropriate for the exchange rate and the interest rate differential series. In this setting, tests of the uncovered interest rate parity (UIP) hypothesis, based on the cross-equation restrictions on the parameters of a Markov switching regimes representation for the underlying variables, show that the interest parity relationship is not rejected for the currencies of Germany and the U.K. vis-à-vis the U.S. dollar over the period 1973-1997. These results suggest that previously reported failures of UIP may be the outcome of rational forecast errors induced by central bank interventions.

JEL classification: F31

Keywords: Policy Rules; Switching Regimes; Markov Process; Rational Forecast Errors

1. Introduction

The interest of financial and macro-economists in the empirical validity of the uncovered interest parity (UIP) relationship stems from its importance as a building block of asset market models of exchange rate determination and of many large-scale macroeconometric models. It is thus natural that this long standing research project has generated a voluminous literature, over the post-1973 free floating period, which nevertheless seems not to have dealt with all theoretical and empirical aspects of the hypothesis.

The received evidence suggests that there are predictable excess returns in foreign exchange markets, but it should not escape notice that empirical tests have almost invariably focused on the implied relation between spot and forward exchange rates. In particular, pure arbitrage in foreign exchange markets implies that the forward exchange rate is the rationally expected future spot rate and, therefore, early empirical studies relied on parameter tests of regressions of the *ex post* spot exchange
rate on the lagged forward rate (e.g. Levich, 1979; Frenkel, 1981; Bilson, 1981; Baillie et al., 1983; Bailey et al., 1984). Even though the first results were favorable to the UIP hypothesis, many analysts expressed their skepticism about the usefulness of the particular regression methodology and turned to a more appealing specification whereby the ex post change in the exchange rate is projected on the lagged forward premium or discount.\footnote{Under the latter formulation, the UIP relationship is again identified with the unbiasedness of the forward exchange rate as a predictor of the future spot rate which requires that the coefficient of the forward premium not be significantly different from unity and that the error term be serially uncorrelated.} Such studies (e.g. Levich, 1979; Bilson, 1981; Longworth, 1981; Fama, 1984; Huang, 1984) have produced mixed results for different currencies and time periods, but, overall, the evidence is rather against the unbiasedness of forward market forecasts since ordinary least square (OLS) estimates of the coefficient of the forward premium fall consistently in the vicinity of -3.

Fama (1984) argued that the deviation of the slope of the above regression from unity is a direct measure of the contribution of a risk premium in the variation the forward exchange rate. In particular, under rational expectations, values of this coefficient below $\frac{1}{2}$ imply that the risk premium is more variable than expected exchange rate changes. However, neither static nor general equilibrium models can explain the observed variability of the risk premium under admissible measures of risk aversion (see Lewis, 1995).

Froot and Frankel (1989) used survey data for exchange rate expectations and obtained evidence that the deviation of the above regression slope from unity is mainly due to the correlation between the forward premium and forecast errors rather than the high variability of the risk premium. They attributed this finding to irrational
behavior arising from the presence of heterogeneous traders in the market. However, even when expectations are rational, the forward premium (or the interest rate differential) may be correlated with forecast errors over short samples if agents expect a major policy shift that does not occur in the sample. This is the so-called “peso problem” (see Krasker, 1980). Indeed, Evans and Lewis (1995) estimated a switching regimes model, which allows for potential “peso problem” effects, and did not reject the hypothesis that the forward premium plus an estimated risk premium is equal to the expected exchange rate for three major U.S. dollar exchange rates. Also, they provided evidence that in the presence of policy shifts the Fama regression coefficient is biased downwards.

In recent work, Baillie and Bollerslev (2000) report further evidence that negative estimates of the slope in the Fama regression may be a statistical artefact. Based on the observation that estimates of this coefficient are widely dispersed in short samples, and are even positive and significantly greater than unity over some periods, they conjecture that convergence to the true parameter value of one is very slow. This conjecture is verified by a simulation exercise whereby data generated under the validity of UIP reproduce the results reported in the empirical literature, namely, the persistent autocorrelation of the forward premium and the extreme dispersion of the coefficient in discussion. Thus, there is mounting evidence that empirical findings against UIP may be due to wrong statistical modelling.

McCallum (1994) suggested that the rejection of forward rate unbiasedness does not entail rejection of UIP as long as monetary authorities systematically smooth interest rates and exchange rates by managing interest rate differentials. Such behavior implies that the UIP relation, along with a policy response equation for the interest rate differential, is compatible with estimates around -3 for the coefficient of
the forward premium in the Fama regression. More recently, Anker (1999) reached similar conclusions in the context of a monetary model for a small open economy in which the central bank reacts to exogenous risk premium shocks. In particular, he showed that interest rate smoothing through the manipulation of exchange rate expectations increases the negative correlation between the risk premium and expected exchange rate changes and, thus, introduces additional downward bias in the forward premium regression coefficient.

While these approaches are indeed successful in reconciling the empirical evidence with UIP, they cannot be tested formally since the assumed policy rules are not informative with regard to the time series properties of the underlying variables. Therefore, in this paper we combine the idea that the UIP puzzle may be due to central bank interventions, with the time series approach of Evans and Lewis (1995) which allows for potential “peso problem” effects. Specifically, it is shown that when monetary authorities manage the interest rate through a standard anti-inflationary policy rule, which is subject to occasional discrete shifts, then a Markov switching regimes representation is appropriate for the exchange rate and the interest rate differential series. Given the evidence in favor of long swings in the exchange rate (see Engel and Hamilton, 1990; Kaminsky, 1993; Evans and Lewis, 1995; Klaassen, 1999; Kirikos, 1998, 2000), it seems that such a policy rule may approximate well the behavior of monetary authorities. In this setting, which allows for rational systematic forecast errors when an expected regime shift does not occur, agents are able to draw probabilistic inferences about the state of the process at any given date and thus form forecasts of the relevant variables based on available information at that date. Such forecasts are non-linear functions of current and past values of the information variables reflecting non-linearities that may arise from random policy changes.
Thus, using this forecast-generating process, the validity of UIP can be tested through tests of the cross-equation restrictions that the parity relation places on the parameters of the joint Markov process.⁴ Also, it is worth noting that this approach does not use the forward exchange rate and is thus independent of forward market forecasts.

The time series implications of a policy rule that allows for discrete intervention shifts are discussed in the next section. The third section contains the forecasting methodology and the derivation of the cross-equation restrictions implied by UIP. Empirical results, based on quarterly data on the currencies of Germany and the U.K. relative to the U.S. dollar over the period 1973 - 1997, are presented in the fourth section, and the final section includes a summary and conclusions.

2. A Model of Discrete Policy Interventions⁵

Consider the uncovered interest rate parity (UIP) equation:

\[ E_t s_{t+1} - s_t = (r_t - r_t^*) + \xi_t, \]

where \( E_t \) denotes expectations conditioned on all information available at time \( t \), \( s_t \) is the natural logarithm of the spot exchange rate measured in units of domestic currency per unit of foreign currency, \( r_t \) and \( r_t^* \) are the domestic and foreign nominal interest rates, respectively, and \( \xi_t \) is a disturbance term which reflects the impact of exogenous shocks.⁶ Given that the change in the exchange rate is a stationary series and that the interest rate differential series has a unit root (see section 4), equation (1) implies that \( \xi_t \) should be non-stationary and cointegrated with \( (r_t - r_t^*) \). Thus, we assume that \( \xi_t \) follows a random walk:

\[ \xi_t = \xi_{t-1} + \zeta_t, \]

where \( \zeta_t \) is white noise.
Also, it will be assumed that the domestic central bank has the following policy rule:

$$\Delta r_t = \beta \pi_t + \delta_t,$$  \hspace{1cm} (3)

where $0 < \beta < 1$, $\pi_t$ is the domestic inflation rate, and $\delta_t$ takes on a given positive value or a given negative value. That is, the domestic monetary authority manages the interest rate so as to counteract inflation. The intervention $\delta_t$ reinforces or loosens the anti-inflationary policy depending on whether the economy is in a high or a low inflation state. Thus, if the rate of inflation is low, the monetary authority might want to reduce the interest rate, and $\delta_t$ takes on a negative value. Alternatively, under a high rate of inflation, $\delta_t$ will take on a positive value to induce a further increase in the interest rate.\(^7\)

Furthermore, let the foreign central bank follow a policy rule that counteracts imported inflation through interest rate targeting:

$$\Delta r_t^* = -\gamma s_t,$$ \hspace{1cm} (4)

where $0 < \gamma < 1$. That is, if there is a depreciation in the foreign currency ($\Delta s_t < 0$), the foreign monetary authority raises the interest rate and vice versa.\(^9\) Then, from (3) and (4) we have: \(^10\)

$$\Delta(r_t - r_t^*) = \beta \pi_t + \gamma \Delta s_t + \delta_t.$$ \hspace{1cm} (5)

Agents realize this pattern of interest rate management but they do not observe the state of policy, that is, they do not know whether $\delta_t$ will take on a positive or a negative value at any given date. However, they will expect a positive intervention if the rate of inflation is high, say $\pi_1$, and a negative intervention if the rate of inflation is low, say $\pi_2$. Thus, agents perceive the following rule for the interest rate differential:
\[ \Delta(r_t - r^*_t) = \beta \pi_t + \gamma \Delta s_t + \delta_{h_t} + \nu_t , \quad \nu_t \sim (0, \sigma^2_{\nu_t}) \]  

(6)

where \( h_t \) is an unobserved state variable that takes on discrete values in the set \{1,2\} and follows a Markov chain with transition probabilities \( p_{ij} = \Pr(h_t = j | h_{t-1} = i) \), \( i,j = 1,2 \), and \( \nu_t \) is an error term. Under this rule, the best that agents can do is to assign probabilities to the current state based on the history of policy interventions.

In this setting, the UIP relationship (1) is perceived as part of a system that includes the policy rule (6). But from equation (1) we have:

\[ E_t \Delta s_{t+1} = \Delta(r_t - r^*_t) + (r_{t-1} - r^*_{t-1}) + \xi_t \]

and by substitution of (6) into (7):

\[ E_t \Delta s_{t+1} = \beta \pi_t + \gamma \Delta s_t + E_{t-1} \Delta s_t + \delta_{h_t} + \xi_t + \nu_t , \]

(8)

where \( E_{t-1} \Delta s_t = (r_{t-1} - r^*_{t-1}) + \xi_{t-1} \) (from equation 1) and \( \Delta \xi_t = \zeta_t \) (from equation 2) have been taken into account.

Let us conjecture the following solution to (8):

\[ \Delta s_t = \phi_1 \pi_t + \phi_2 \delta_{h_t} + \phi_3 \xi_t + \phi_4 \nu_t \]

(9)

from which we obtain

\[ E_t \Delta s_{t+1} = \phi_1 E_t \pi_{h_t+1} + \phi_2 E_t \delta_{h_t+1} \]

(10)

and

\[ E_{t-1} \Delta s_t = \phi_1 E_{t-1} \pi_t + \phi_2 E_{t-1} \delta_{h_t} \]

(11)

since \( E_t \xi_{t+1} = E_t \nu_{t+1} = E_{t-1} \xi_t = E_{t-1} \nu_t = 0 \).

Now suppose that at the time of forming expectations, agents know that the intervention is \( \delta_t \), that is, the economy is in the state of high inflation \( \pi_t \). Then, from (10) and (11) we have:

\[ E_t \Delta s_{t+1} = E_{t-1} \Delta s_t = \phi_1 (p_{11} \pi_t + p_{12} \pi_{t+1}) + \phi_2 (p_{11} \delta_t + p_{12} \delta_{t+1}) . \]

(12)

Thus, by substituting (9) and (12) into (8), we obtain:
\[
\phi_1(p_{11}\pi_1 + p_{12}\pi_2) + \phi_2(p_{11}\delta_1 + p_{12}\delta_2) = \gamma(\phi_3\pi_h + \phi_4\delta_h + \phi_5\zeta_i + \phi_6v_i) + \\
\phi_1(p_{11}\delta_1 + p_{12}\delta_2) + \phi_2(p_{11}\delta_1 + p_{12}\delta_2) + \beta\pi_h + \delta_h + \zeta_i + v_i,
\]

which holds when
\[
\gamma\phi_1 + \beta = 0 \quad \text{and} \quad \gamma\phi_2 + 1 = \gamma\phi_3 + 1 = \gamma\phi_4 + 1 = 0
\]
or
\[
\phi_1 = -\frac{\beta}{\gamma} \quad \text{and} \quad \phi_2 = \phi_3 = \phi_4 = \frac{1}{\gamma}.
\]

Then, equations (9) and (14) give:
\[
\Delta \epsilon_i = -\frac{1}{\gamma}(\beta\pi_h + \delta_h + \zeta_i + v_i)
\]
that is, the exchange rate change follows a Markov switching regimes process.

Also, by substitution of (15) into (5) we have:
\[
\Delta(r_i - r_i^*) = \beta(\pi_i - \pi_h) + (\delta_i - \delta_h) - \zeta_i - v_i
\]
that is, \(\Delta(r_i - r_i^*)\) has a Markov switching regimes representation as well.12

Furthermore, equation (16) shows that \((r_i - r_i^*)\) and \(\xi_i\) move together over time and, thus, they are cointegrated as they should be, given that the interest rate differential has a unit root (see section 4).13 This means that interest rate management responds to exogenous disturbances too.

To see that policy shifts may introduce the possibility of rational systematic forecast errors, consider the case in which the current policy regime is 1 and does not change in the future. Since agents do not observe the current regime, they form a probabilistic inference that the current regime is 1, say \(\ell_i\). Under these assumptions, the process given in equation (16) implies that the forecast error for changes in the interest rate differential is:
\[
\Delta(r_{i+1} - r_{i+1}^*) - E_i\Delta(r_{i+1} - r_{i+1}^*) = [\ell_i, p_{12} + (1 - \ell_i, p_{22}] \times
\]
\[ x[\beta(\pi_1 - \pi_2) + (\delta_1 - \delta_2)] - \zeta_{t+1} - \nu_{t+1} \]  

(17)

Also, from equation (15), the forecast error for exchange rate changes is:

\[ \Delta_{t+1} - E_t \Delta_{t+1} = -\frac{1}{\gamma} \left\{ [\ell_t p_{12} + (1 - \ell_t) p_{22}] \beta(\pi_1 - \pi_2) + (\delta_1 - \delta_2) + \zeta_{t+1} + \nu_{t+1} \right\} \]  

(18)

Thus, as long as there is uncertainty about the regime (\( \ell_t < 1 \)) and agents attach a positive probability to a change in central bank behavior (\( p_{12} > 0 \)), the forecast error may have a non-zero mean even though agents use the underlying process to form expectations. In addition, given that the inference about the current regime (\( \ell_t \)) depends on past values of the information variables,\(^\text{14}\) equations (17) and (18) imply that forecast errors within a regime will not be independent of the current interest spread (or forward premium). Therefore, if exchange rate changes are represented by a Markov switching process, then the OLS estimate of the slope in the Fama regression, and the corresponding estimates of predicted excess returns, will be biased.

3. Empirical Modeling and Forecasting

As it stands, equation (1) has little empirical content because it involves unobserved expectations of exchange rate depreciation conditioned on an information set which is in general unavailable to an econometrician. Thus, to assess the validity of UIP empirically, one has to work with a subset of the full information set and to replace expectations of future exchange rate changes with observed variables. In addition, one has to account for non-stationarity of the variables \( s_t \) and \( r_t - r_t^* \) if biased test procedures are to be avoided. While there is compelling evidence that the natural logarithm of the exchange rate is a difference stationary series, there is also a wide-spread belief that the interest rate differential is borderline stationary, at least for the industrial countries.
However, the results of the next section strongly suggest that the differential yields on eurocurrency deposits are non-stationary for the currencies of Germany, the U.K. and the U.S. Thus, in this section, we will consider the implications of the UIP hypothesis under both stationary and non-stationary interest differentials.

It will be assumed that monetary authorities manage the interest rate differential through occasional discrete shifts in the underlying policy regime, as discussed in the previous section. That is, instead of assuming that the vector time series $x_t = [(r_t-r_t^*) \Delta s_t]'$ or $x_t = [\Delta(r_t-r_t^*) \Delta s_t]'$ clusters around a constant mean $\mu$, we can use the more general stochastic representation:

$$x_t = \mu_h + u_t, \quad u_t \sim N(0, \Omega_h)$$

where $h_t$ is an unobserved state variable that takes on values in the set {1,2} and follows an irreducible Markov chain with stationary transition probability matrix:

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

where $p_{ij} = Pr(h_t=j|h_{t-1}=i), i,j=1,2$ and the sum of each row is unity.

State 1 will be referred to as the depreciation state while state 2 will be associated with an appreciation of the domestic currency. Thus, when the process is in state 1 ($h_t=1$), $x_t$ is drawn from a $N(\mu_1, \Omega_1)$ distribution, whereas for $h_t=2$, $x_t$ is drawn from a $N(\mu_2, \Omega_2)$ distribution.

The above specification is chosen for several reasons. First, there is convincing evidence that exchange rate changes follow a univariate Markov switching process (see Engel and Hamilton, 1990; Kaminsky, 1993; Klaassen, 1999; Kirikos, 1996, 1998, 2000). Second, the assumption of different variances across regimes permits the construction of a test of the Markovian dynamics, which otherwise would be impossible because, under the null hypothesis $H_0: \mu = \mu$, and $\Omega_1 = \Omega_2$, the parameters $p_{11}$ and $p_{22}$ are
not identified. Third, the computations involved in estimating the model are less burdensome and are carried out by a GAUSS program.

Maximum likelihood estimates (MLE) of the parameters $\mu_1, \mu_2, \Omega_1, \Omega_2, p_{11}, p_{22}$ can be obtained through the EM algorithm which is a method of maximising the sample likelihood function by iterating on the normal equations (see Hamilton, 1990). One starts with a guess of the parameter vector $\theta$, calculates the smoothed inferences $Pr(h_t=j|x_1,x_2,...,x_T;\theta)$, $t=1,...,T$, $j=1,2$, i.e. the probabilities that the process is in a particular regime at any given date based on the full sample of observations ($T$), and uses them in the normal equations to obtain new estimates of the parameters. The latter are then used to recalculate the smoothed inferences and the parameter vector, and the process continues until a convergence criterion is satisfied. The final parameter vector corresponds to a local maximum of the likelihood surface, and, therefore, one should try several different starting values for $\theta$ and select the estimated parameter vector that gives the greatest value of the likelihood function.

Forecasts of future values of $x$, based on currently available information, can be obtained as follows. The probability that the process will be in state $j$ ($j=1,2$) $m$ periods from now ($m>0$), given that it is currently in state $i$, is the element of the $i$th row and the $j$th column of the matrix $P^m$ (see e.g Bhat, 1972). Also, let $\chi_t=(x_1,x_2,...,x_t)$ and $\eta_t=(h_1,h_2,...,h_t)$ be the histories of $x$ and $h$, respectively, up to time $t$. Then, assuming that the state variable $h_t$ is observed, we have:

$$E(x_{t+m} | \chi_t, \eta_t) = \omega P^m \begin{bmatrix} \mu_1' \\ \mu_2' \end{bmatrix}$$

where $\omega=[1 \ 0]$ for $h_t=1$, $\omega=[0 \ 1]$ for $h_t=2$, and $\mu_1'$ and $\mu_2'$ are 1x2 vectors of means for states 1 and 2, respectively. However, the state variable $h_t$ is in fact unobserved, and thus forecasts of future values of $x$ must be based on probabilistic inferences about the
state at date $t$ conditioned on the history of $x$ through that date. These probabilities are also called filter inferences because they are obtained from a non-linear filter. Thus, letting $a'_t = [Pr(h_t=1|\mathcal{X}_t) \; Pr(h_t=2|\mathcal{X}_t)]$ be the vector of filter inferences, and defining $\mu_t = [\mu_1 \; \mu_2]'$, we have the following non-linear forecasts based on the history of $x$ only:

$$E(x'_{t+m} | \mathcal{X}_t) = a'_t \; P^m \; \mu_t.$$  \hspace{1cm} (22)

We can now consider the restrictions that the UIP relationship, given in equation (1), places on the parameters of the stochastic representation (19). First, if the interest rate differential is a stationary series, then the disturbance term $\xi_t$ in (1) is assumed to be white noise. Thus, by updating equation (1) and by taking expectations as of time $t$, we have:

$$E_t \Delta s_{r+1} = E_t (r_{t+1} - r^*_{t+1}).$$  \hspace{1cm} (23)

Assuming that the vector $x_t = [\Delta(r_t - r^*_t) \; \Delta s_t]'$ has the stochastic representation given in equation (19) and letting $e_t = [1 \; 0]'$ and $e_2 = [0 \; 1]'$, equations (22) and (23) give:

$$a'_t P^2 \mu_t e_2 = a'_t P \mu_t e_1.$$  \hspace{1cm} (24)

If the UIP hypothesis is correct, equation (24) must hold for general probability vector $a'_t$ and, thus, the validity of UIP implies the following cross-equation restrictions for the parameters:

$$P^2 \mu_t e_2 = P \mu_t e_1.$$  \hspace{1cm} (25)

However, if the interest rate differential series $(r-r^*)$ is not stationary, then the vector time series should be $x_t = [\Delta(r_t - r^*_t) \; \Delta s_t]$ and the disturbance term $\xi_t$ in equation (1) is assumed to follow a random walk as in (2). Thus, by updating equation (7) and taking expectations as of time $t$, we have:

$$E_t \Delta s_{r+1} = E_t \Delta((r_{t+1} - r^*_{t+1}) + (r_t - r^*_t) + \xi_t)$$  \hspace{1cm} (26)
since \( E_t \bar{\xi}_{t+1} = \bar{\xi}_t \) from equation (2). Then, given the representation (19) for the vector \( \mathbf{x}_t = [\Delta(r_t - r_t^*) \ \Delta s_t] \), equations (1), (22), and (26) imply:

\[
a_t' P^2 \mu_n e_2 - a_t' P \mu_n e_2 = a_t' P \mu_n e_1. \tag{27}
\]

Again, if the UIP hypothesis is valid, equation (27) must hold for general probability vector \( a_t' \) and, therefore, the restrictions placed on the parameters of the Markov switching process are the following:

\[
P(P-I)\mu_n e_2 = P\mu_n e_1, \tag{28}
\]

These can be tested by means of a Wald test. More precisely, letting \( \hat{\lambda} = P(P-I)\mu_n e_2 - P\mu_n e_1 \) denote deviations from equation (28) and \( \mathbf{b} \) the vector of parameters with variance-covariance matrix \( \text{Var}(\mathbf{b}) \), we have the following convergence in distribution (see Amemiya, 1985, pp. 141-146):

\[
W = \hat{\lambda}' \left[ \frac{\partial \hat{\lambda}}{\partial \mathbf{b}} \text{Var}(\mathbf{b}) \frac{\partial \hat{\lambda}}{\partial \mathbf{b}} \right]^{-1} \hat{\lambda} \xrightarrow{d} \chi^2(2). \tag{29}
\]

where \( W \) is the Wald statistic and \( \chi^2(2) \) denotes a \( \chi^2 \) distribution with two degrees of freedom. The partial derivatives in \( W \) involve highly non-linear expressions in \( \mathbf{b} \), but they can be computed numerically.

4. Empirical Results

The UIP hypothesis is assessed here under the assumption that the joint dynamic behavior of the information variables is described by the Markov switching process discussed earlier. The quarterly data for Germany, the U.K., and the U.S. cover the period from 1973:III to 1997:III (97 observations).17

Since the presence of unit roots in the underlying series makes a difference for the cross-equation restrictions implied by the UIP hypothesis (see equations 25 and 28), it is important to conduct some preliminary unit root tests. The Phillips-Perron (1988)
test, reported in Table 1, suggests that the exchange rate series \((s)\) and the interest rate differential series \((r_t-r_t^*)\) are difference stationary. Indeed, the value of the \(Z_t\) statistic \(^{18}\) provides strong evidence in favor of a unit root in the *levels* of the exchange rate series, while the hypothesis of a unit root is rejected for the first differences. Similar results are obtained for the interest rate differential between Germany and the U.S.\(^{19}\) Thus, the interest rate differential series appear to be difference stationary, \(^{20}\) and this implies that the specification given in equations (19) and (20) would be meaningful for the vector series \(x_t = [\Delta(r_t-r_t^*) \Delta s_t]'\). Under this specification, the *levels* of the interest differential and of the exchange rate follow a process with stationary errors around stochastically shifting trends. Then, the cross-equation restrictions implied by UIP are given by equation (28).\(^{21}\)

Maximum likelihood estimates (MLE) of the parameters are reported in Table 2. They were obtained via the EM algorithm, and their asymptotic standard errors were based on numeric second derivatives of the logarithm of the likelihood function. In each case, several different starting values of the parameter vector were tried and the selected estimates are those that gave the greatest value of the sample likelihood function.

Wald tests of the Markovian dynamics are reported in Table 3. In particular, the Table contains the value of the Wald test statistics for testing the null hypotheses \(H_0: \mu_1=\mu_2\) and \(H_0: p_{11}=1-p_{22}\), respectively. The latter hypothesis implies that the probability that the process is currently in state 1 does not depend on the state of the previous period, that is, the state variable \(h_t\) is not characterized by the Markov property. Also, under the hypothesis \(H_0: \mu_1=\mu_2\) the process has a non-shifting mean across regimes. In both cases, however, regularity conditions hold, and thus standard asymptotic theory applies, only when \(\sigma_1^2 \neq \sigma_2^2\) (see Engel and Hamilton, 1990; Hamilton, 1993). The test statistics are, respectively:
\[(3' \hat{\theta})'[Z' \cdot Var(\hat{\theta}) \cdot Z]^{-1}(3' \hat{\theta}) \sim \chi^2(2) \quad (30)\]

\[
\frac{(\hat{p}_{11} + \hat{p}_{22} - I)^2}{Var(\hat{p}_{11}) + Var(\hat{p}_{22}) + 2Cov(\hat{p}_{11}, \hat{p}_{22})} \sim \chi^2(I) \quad (31)
\]

where

\[
\hat{\rho} = [\hat{\mu}_{1[\Delta(r-r^*)]}, \hat{\mu}_{2[\Delta(r-r^*)]}, \hat{\sigma}_{2[\Delta(r-r^*)]}^2, \hat{p}_{11}, \hat{p}_{22}, \sigma_{2[\Delta(r-r^*)]}^2]
\]

\[
Cov_{\Delta}(r-r^*, \Delta s) \sigma_{2[\Delta(r-r^*)]}^2, Cov_{\Delta}(\Delta r, \Delta s) \sigma_{2[\Delta(r-r^*)]}^2\]

is the vector of estimated parameters, and

\[
\mathbf{Z}' = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
\]

The results of these tests are favorable to the proposed specification since the Markov property is not rejected for both data sets (i.e. the null \(H_0: p_{11} = 1 - p_{22}\) is strongly rejected). Also, the hypothesis of non-shifting means across regimes is rejected in the case of the U.K. but not for the German data.

The Wald test statistics (equation 29), given in Table 4, are based on the estimates of the bivariate model and provide compelling evidence in favor of the validity of the cross-equation restrictions (equation 28) imposed by UIP on the parameters of the Markov process. These results suggest that the uncovered interest parity relationship was maintained for the currencies of Germany and the U.K. vis-à-vis the U.S. dollar, over our sample period.

### 5. Concluding Remarks

Theoretical attempts to reconcile the empirical findings on the predictability of excess returns in foreign exchange markets with central bank interventions (McCallum, 1994; Anker, 1999) have proven largely successful. However, these approaches cannot be tested formally because they are not informative with regard to the time
series properties of the underlying variables. Furthermore, a burgeoning empirical literature has provided compelling evidence that exchange rates are characterized by long swings whose dynamics are well represented by Markov switching regimes processes (Engel and Hamilton, 1990; Kaminsky, 1993; Evans and Lewis, 1995; Klaassen, 1999; Kirikos, 1998, 2000). Therefore, in this paper we have combined the policy response explanation of the UIP puzzle with the time series dynamics implied by a Markov switching regimes representation.

By postulating an anti-inflationary policy rule, which allows for discrete shifts in managing the interest rate, we are able to conclude that the dynamic behavior of the exchange rate and the interest rate differential is described by a Markov switching regimes process. This model implies that forecast errors may be orthogonal to, but not independent of, past values of the information variables, even though agents form expectations rationally. That is, whenever agents expect a shift in policy that does not occur, we have “peso problem” effects that make the behavior of investors look irrational. Thus, forecasts of exchange rate changes and excess returns can be improved if the possibility of policy regime shifts is accounted for. Obviously, when this possibility exists, the coefficient of the forward premium in the Fama regression will be biased, as shown by Evans and Lewis (1995).

In this setting, the approximate validity of the proposed policy rule can be checked by testing whether the joint dynamics of exchange rate and interest differential changes are characterized by the Markov property. In addition, we can formally test the empirical relevance of the remaining part of our system, namely the UIP hypothesis, by means of a statistical test of the cross-equation restrictions that it imposes on the parameters of the Markov switching regimes representation.
Indeed, estimates of the stochastic segmented trends of Engel and Hamilton (1990) for the exchange rate and the interest differential, based on quarterly data on the currencies of Germany and the U.K. relative to the U.S. dollar over the period 1973-1997, suggest that the joint behavior of these series is driven by Markovian dynamics. This finding lends support to the postulated central bank behavior. Also, tests of the UIP-implied cross-equation restrictions show that the parity relationship is not rejected, thus implying the absence of predictable excess returns over our sample period. Similar results for the currencies of Greece, Italy, and Portugal relative to the U.S. dollar, over the post-1973 free floating period, are reported in Kirikos (2002). Therefore, our evidence suggests that previous empirical findings in favor of predictable excess returns in foreign exchange markets may well be the outcome of the presence of rational forecast errors induced by central bank interventions.
REFERENCES


### TABLE 1. Phillips-Perron test for unit roots

<table>
<thead>
<tr>
<th>Truncation Lag</th>
<th>With Trend</th>
<th>Without Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>12</td>
<td>4</td>
</tr>
</tbody>
</table>

**A. Exchange Rates**

| | With Trend | | Without Trend | |
|----------------|------------|----------------|----------------|
| ln(DM/$) | -2.09 | -2.13 | -1.57 | -1.60 |
| ln($/BP) | -2.34 | -2.39 | -2.38 | -2.40 |
| Δln(DM/$) | -7.24 | -7.17 | -7.27 | -7.19 |
| Δln($/BP) | -7.89 | -7.83 | -7.89 | -7.86 |

**B. Interest Rate Differentials**

| | With Trend | | Without Trend | |
|----------------|------------|----------------|----------------|
| r\(_{GER}\)-r\(_{US}\) | -2.03 | -2.29 | -1.90 | -2.12 |
| r\(_{US}\)-r\(_{UK}\) | -3.00 | -2.96 | -2.98 | -2.93 |
| Δ(r\(_{US}\)-r\(_{UK}\)) | -8.44 | -9.45 | -8.50 | -9.52 |

**Notes:** The truncation lag corresponds to the maximum order of non-zero autocorrelations of the differenced series (see Phillips, 1987, p. 285) and is selected according to the suggestion of Schwert (1989) that the maximum length should be \(\text{int}\{12(T/100)^{1/4}\}\) and the minimum \(\text{int}\{4(T/100)^{1/4}\}\), where \(T\) is the sample size. Asymptotic critical values for testing the null hypothesis of a unit root are as follows (see Fuller, 1976, p. 373). With trend: -4.04, -3.45, -3.15 at the 1%, 5%, and 10% significance levels, respectively. Without trend: -3.51, -2.89, -2.58 at the 1%, 5%, and 10% significance levels, respectively. The test is a left-hand tail test.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\Delta (r_{\text{GER}} - r_{\text{US}})$</th>
<th>$\Delta \ln (\text{DM/$US)}$</th>
<th>$\Delta (r_{\text{US}} - r_{\text{UK}})$</th>
<th>$\Delta \ln (\text{US/BP})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>0.0322</td>
<td>1.6626</td>
<td>-0.0069</td>
<td>1.6063</td>
</tr>
<tr>
<td></td>
<td>(0.0916)</td>
<td>(1.1709)</td>
<td>(0.00067)</td>
<td>(0.6564)</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>-0.0103</td>
<td>-1.0937</td>
<td>0.0195</td>
<td>-2.6302</td>
</tr>
<tr>
<td></td>
<td>(0.0227)</td>
<td>(0.7122)</td>
<td>(0.0814)</td>
<td>(0.9869)</td>
</tr>
<tr>
<td>$\sigma_1^2$</td>
<td>0.2136</td>
<td>25.1469</td>
<td>0.0332</td>
<td>15.3482</td>
</tr>
<tr>
<td></td>
<td>(0.0816)</td>
<td>(7.7001)</td>
<td>(0.0079)</td>
<td>(4.2555)</td>
</tr>
<tr>
<td>$\sigma_2^2$</td>
<td>0.0307</td>
<td>23.3980</td>
<td>0.3002</td>
<td>29.3741</td>
</tr>
<tr>
<td></td>
<td>(0.0090)</td>
<td>(4.1719)</td>
<td>(0.0678)</td>
<td>(7.8554)</td>
</tr>
<tr>
<td>$Cov_1$</td>
<td>0.1461</td>
<td></td>
<td>0.3025</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.4743)</td>
<td></td>
<td>(0.1472)</td>
<td></td>
</tr>
<tr>
<td>$Cov_2$</td>
<td>0.0523</td>
<td></td>
<td>0.2918</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1273)</td>
<td></td>
<td>(0.4519)</td>
<td></td>
</tr>
<tr>
<td>$p_{11}$</td>
<td>0.9305</td>
<td></td>
<td>0.9139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0656)</td>
<td></td>
<td>(0.0482)</td>
<td></td>
</tr>
<tr>
<td>$p_{22}$</td>
<td>0.9663</td>
<td></td>
<td>0.9068</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0263)</td>
<td></td>
<td>(0.0514)</td>
<td></td>
</tr>
<tr>
<td>$L^*$</td>
<td>116.007</td>
<td>142.406</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Numbers in parentheses below the estimated parameters are standard errors. 

$L^*$ is the value of the log likelihood function. MLE are obtained via the EM algorithm, and standard errors are based on numerically computed second derivatives.
TABLE 3. Wald tests of Markovian dynamics

<table>
<thead>
<tr>
<th></th>
<th>( H_0: \mu_1 = \mu_2 ) ( \chi^2(2) )</th>
<th>( H_0: p_{11} = 1 - p_{22} ) ( \chi^2(1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GERMANY/US</td>
<td>3.915 (0.1412)</td>
<td>130.949 (0.0000)</td>
</tr>
<tr>
<td>US/UK</td>
<td>11.127 (0.0038)</td>
<td>103.571 (0.0000)</td>
</tr>
</tbody>
</table>

*Notes:* Numbers in parentheses next to the estimated statistics are marginal significance levels or \( p \)-values.
## TABLE 4. Tests of cross-equation restrictions

<table>
<thead>
<tr>
<th></th>
<th>Wald test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>GERMANY/US</td>
<td>1.607 (0.4477)</td>
</tr>
<tr>
<td>US/UK</td>
<td>3.892 (0.1428)</td>
</tr>
</tbody>
</table>

*Notes:* Numbers in parentheses next to the estimated statistics are marginal significance levels or *p*-values.
ENDNOTES

1 For a discussion of the implications of the two alternative specifications see McCallum (1994). A survey of evidence on forward market forecasts is given in Engel (1996).

2 Tests of UIP based on linear Vector Autoregressive (VAR) forecasts are reported in Hakkio (1981), Baillie et al. (1983), Ito (1988), and Kirikos (1993). However, these specifications do not capture the dynamics implied by changes in policy regimes.

3 Nonlinear behaviour may also characterise the reversion of the exchange rate towards its fundamental value (Kilian and Taylor, 2001).

4 Engel and Hamilton (1990) have provided some evidence against UIP, using a Markov switching regimes specification with quarterly data, but their results do not include tests of the statistical significance of the implied cross-equation restrictions.

5 Sections 2 and 3 draw on Kirikos (2002).

6 As argued by McCallum (1994), the term $\xi$, may represent time-varying aggregation and other effects, as well as risk premia. That is, the possibility of a time-varying risk premium is not the only reason for including a disturbance term in equation (1) as is widely held.

7 It is reasonable to assume that the coefficient $\beta$ is less than unity because, otherwise, this policy rule may cause the interest rate to vary excessively.

8 Hamilton (1988) has argued that similar discrete changes in the Fed’s operating procedures in the ‘80s caused the U.S. interest rate to follow a Markov switching regimes process.

9 Some high inflation EU countries have followed similar rules in their efforts to fulfill the inflation criterion for the monetary union.
Note that a policy rule that targets the interest rate differential, $\Delta(r_t - r_t^*) = \gamma \Delta s_t + \delta_t$, followed by either country, would suffice for the solution drawn below for the change in the exchange rate. Also, a similar solution obtains if the foreign central bank targets inflation through interest rate management, that is, $\Delta r_t^* = \gamma \pi_t^*$ instead of (4). Then, taking into account that $\Delta s_t = \pi_t - \pi_t^* + \Delta q_t$, where $q_t$ is the real exchange rate, we obtain an extended form of equation (5) for the change in the interest rate differential. The same solution obtains if we base expectations on intervention $\delta_2$.

Note that even when realized and perceived variables coincide, the interest rate differential will still have a Markov switching regimes representation since the error term $\nu_t$ has state dependent variance.

It should be noted that even McCallum’s (1994) solution for $\Delta s_t$, which is derived under a different policy rule, is valid only if $(r_t - r_t^*)$ and $\xi_t$ are cointegrated, when $(r_t - r_t^*)$ has a unit root.

In fact, these inferences depend non-linearly on the information variables (see section 3).

Kirikos (2002) reports similar results for the currencies of Greece, Italy, and Portugal relative to the U.S. dollar.

The specification in equations (19) and (20) is the same as in Engel and Hamilton (1990). Evans and Lewis (1995) extended this specification by including a term which reflects the jump in the exchange rate process when a change in the regime takes place. This jump reflects the *ad hoc* assumption of switches in the process of the underlying determinants of the exchange rate and is restricted to depend on the forward premium.
The exchange rate series are averages of monthly data available by the Federal Reserve via the Web site http://www.stls.frb.org/fred. The interest rates are averages of daily returns on euro-currency deposits and they were obtained from the DataResources Inc. (DRI) database. These returns are quoted at an annual rate and are transformed to quarterly by the formula \((1+r/100)^{1/4}-1\), where \(r\) is the annual rate. The first differences of the exchange rates and the interest rate differentials are multiplied by 100 to express changes in percentage terms.

This is a transformation of the simple Dickey-Fuller (1979) \(t\)-statistic to account for weakly dependent and heterogeneously distributed innovations in the stochastic representation of the series (see Phillips, 1987; Phillips and Perron, 1988; Hamilton, 1994, pp. 506-514).

The evidence of a unit root in the \textit{levels} of the U.S./U.K. interest differential is not as strong, since the hypothesis of a unit root is not rejected only at the 1% significance level, while at the 5% level we have a borderline rejection (case without trend). Given this evidence and that the autocorrelation function of the series in \textit{levels} dies out very slowly, it is reasonable to treat the series as difference stationary too.

Similar results obtain from the Augmented Dickey-Fuller (ADF) test which is appropriate when the true process of the series is an autoregression (see Hamilton, 1994, pp. 516-529).

The results concerning the empirical relevance of UIP do not change when the levels of interest rate differentials are treated as stationary series (Kirikos and Terzakis, 1999).