OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

The Society for Financial Studies

Do Long-Term Swings in the Dollar Affect Estimates of the Risk Premia?

Author(s): Martin D. D. Evans and Karen K. Lewis
Reviewed work(s):
Source: *The Review of Financial Studies*, Vol. 8, No. 3 (Autumn, 1995), pp. 709-742
Published by: Oxford University Press. Sponsor: The Society for Financial Studies.
Stable URL: http://www.jstor.org/stable/2962237
Accessed: 28/06/2012 17:39

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Oxford University Press and *The Society for Financial Studies* are collaborating with JSTOR to digitize, preserve and extend access to *The Review of Financial Studies.*

STOR®
http://www.jstor.org

Do Long-Term Swings in the Dollar Affect Estimates of the Risk Premia?

Martin D. D. Evans
New York University

Karen K. Lewis
University of Pennsylvania and National Bureau of Economic Research

Foreign exchange returns exhibit behavior difficult to reconcile with standard theoretical models. This article asks whether the recent findings of long swings in exchange rates between appreciating and depreciating periods affect estimates of the foreign exchange risk premium. We demonstrate how the "peso problem" introduced by expected shifts in exchange rate regimes can affect inferences about the risk premium in at least two ways: (1) it can make the foreign exchange risk premium appear to contain a permanent disturbance when it does not; and (2) it can induce bias in the foreign exchange return regressions such as in Fama (1984).

Studies relating the predictable component in excess foreign exchange returns to standard models of the risk premium have proven largely unsuccessful. For example, the results in Fama (1984) suggested that the variability of the risk premium is large and ex-

This article was previously circulated under the title "Are Foreign Exchange Returns Subject to Permanent Shocks?" We are grateful for useful comments from two anonymous referees, Dave Backus, Geert Bekaert, Bob Cumby, Kathryn Dominguez, Charles Engel, Neil Ericsson, Jon Faust, Robert Hodrick, Erzo Luttmer, Andy Lo, and seminar participants at the 1993 Symposium of the Western Finance Association Meetings, the National Bureau of Economic Research Summer Institute, Queen's University, the Federal Reserve Board, the Wharton Macro and Micro Lunch Groups, and New York University. Karen Lewis gratefully acknowledges research support from the National Science Foundation under grant #89-02794. Any errors are our own.

The Review of Financial Studies Fall 1995 Vol. 8, No. 3, pp. 709–742
© 1995 The Review of Financial Studies 0893-9454/95/$1.50

ceeds the variability of the expected change in the exchange rate. Although empirical findings concerning risk premium behavior such as in this case can be reconciled theoretically with foreign exchange models,[1] empirical attempts to do so have not succeeded on the whole.[2]

Recent research has also demonstrated that the dollar appears to undergo periods of persistent appreciation and then depreciation.[3] It therefore seems likely that traders in the market anticipate shifts between these regimes. If so, then this expectation in turn affects the behavior of forward rates relative to observed spot rates. Intuitively, if exchange rates switch infrequently between different regimes, rational anticipations of switches that are not realized over significant periods of time will result in systematic differences between the expected and realized exchange rates through a "peso problem."[4] Since studies identify the foreign exchange risk premium with the systematic component of deviations between forward rates and realized spot rates, then anticipations of long swings in the dollar can potentially affect estimates of this risk premium.

In this article, we examine this possibility. We first provide a new framework for examining switching behavior in exchange rates. In this framework, current exchange rates reflect expectations about future exchange rates as well as other determining variables. We show that when these variables switch between different processes the exchange rate process also switches. Furthermore, when these switches occur, the exchange rate must jump in order to be consistent with the market's new view about the underlying process for the determining variables. The potential for jumps in the exchange rate has not been featured in earlier switching models.

We then estimate the model for the dollar exchange rate against the German mark (DM), the British pound, and the Japanese yen. The exchange rate regimes implied by the model estimates are broadly consistent with conventional views about exchange rates over the

[1] For example, Hodrick and Srivastava (1986) show how the Fama result, that the risk premium is negatively correlated with the expected change in the exchange rate, is consistent with a theoretical model in a complete markets setting.

[2] For example, Bekaert and Hodrick (1992) derive the Hansen and Jaganathan (1991) bounds using foreign money and equity market investments. They find that these bounds are considerably higher than those based upon U.S. equity alone, posing a further challenge to foreign exchange risk premium models. Hodrick (1987) surveys the foreign exchange risk premium literature.

[3] See Engel and Hamilton (1990), for example. As we will describe later, our switching framework differs in important ways from that of Engel and Hamilton. In particular, we allow for a risk premium and the potential for jumps in the exchange rate. Both of these features introduce time variation in predictable excess returns.

[4] Rogoff (1980) first wrote about the "peso problem." See Lewis (1992) for a more detailed description.

floating rate period. We then test a restriction that the actual forward premium equals the foreign exchange risk premium plus the expected change in the exchange rate implied by the model. Strikingly, we cannot reject this restriction. These results suggest that the regime switches identified by our model may have been rationally anticipated *ex ante* by market participants.

We next use our switching model to consider whether long-term swings in the dollar might affect estimates of the risk premium. We first consider the effects upon the observed persistence in the risk premium as reflected in its long-run behavior. We begin by documenting a new feature of excess return behavior that appears to be at odds with standard risk premium models. We test the typical view that the spot rate moves one-for-one with the forward rate in the long run. Interestingly, we find that standard tests would reject this hypothesis. These tests would imply that predictable excess returns on open forward positions have disturbances with the same degree of persistence as the exchange rate itself. Since the exchange rate is often characterized as following a random walk process, this result would say that the process of predictable returns also contains a random walk component.

We then use our switching model as a data-generating process to ask whether rational expectations of potential shifts in the exchange rate process could provide an explanation. Based upon Monte Carlo experiments, we evaluate the likelihood that standard tests would reject the hypothesis that forward rates and spot rates move one for one in the long run when in fact they do. Interestingly, our experiments show this likelihood is quite high. These results show that long swings in the exchange rate can lead to size distortions in the tests of the long-run relationship between spot and forward rates.

We next use our switching model to consider whether long swings in the exchange rate can shed any light on the well-known finding that the forward premium overpredicts the subsequent change in the exchange rate. Fama (1984) provided an interpretation of this finding using regressions of exchange rate changes upon forward premia. Using his identifying assumption that the covariance of the forward premium is uncorrelated with exchange rate forecast errors, typical coefficients in this regression imply that the variance of the risk premium is significantly higher than that of the forward premium. On the other hand, expectations of shifts in the exchange rate process may induce small sample serial correlations between the exchange rate forecast error and the forward premium. If so, the standard identifying assumption used in the Fama decomposition may be violated in standard sample sizes. To examine the effect of this violation upon standard inferences about the risk premium, we use our switching model as a data-generating process to reexamine the Fama regression. We find that anticipated switches bias downward the Fama coefficient and contribute to a higher measured risk premium variance of between 3 percent for the yen and 22 percent for the pound. This finding is consistent with the survey-based evidence in Froot and Frankel (1989) that the Fama regression coefficient is biased by systematic forecast errors.

The structure of the article is as follows. Section 1 develops the switching model for the exchange rate. Section 2 provides new evidence on the risk premium. Based upon standard techniques, the evidence in this section finds that the spot and forward exchange rates do not move together one for one in the long run. Monte Carlo experiments based upon the model show that this result is likely to be an artifact of long swings in the exchange rate. Section 3 describes the effects of these swings upon the high frequency behavior of the exchange rate. Concluding remarks follow.

1. The Switching Model

We begin by analyzing a switching model of the exchange rate that generates long swings in the dollar. In the following sections, we will show that this model is capable of explaining some of the anomalous behavior associated with the observed risk premium.

1.1 A simple model of exchange rate switching

We develop our switching model in the context of a simple but quite general framework that focuses upon the effects of forward-looking behavior by traders. It should be stressed that our goal is not to develop a fundamentals-based model for the exchange rate, but rather to use this framework to consider the general origins and consequences of exchange rate switching. We will then incorporate the basic results from this exercise into our empirical model below.

We posit that the exchange rate depends upon its expected change as well as some current variables:

$$s_t = \alpha(E_t s_{t+1} - s_t) + x_t \tag{1}$$

where s_t is the logarithm of the exchange rate (measured as the foreign currency price of dollars), $\alpha > 0$ is a parameter, E_t is the expectations operator conditional upon time t information, and x_t is the logarithm of the composite effects of variables that affect the exchange rate. This equation can be written more simply as

$$s_t = \phi E_t s_{t+1} + y_t \tag{1a}$$

where $\phi \equiv \alpha/(1+\alpha)$ and $y_t \equiv (1-\phi)x_t$.

Specifications like Equation (1a) have been considered extensively in the literature on exchange rates.[5] As such, Equation (1a) provides a useful benchmark for our current purpose of asking how switches in y_t may affect the exchange rate.[6]

To examine how switches in the process for these y_t variables affect the exchange rate, we assume that they can follow two different processes.[7] Switches between the two processes are governed by a state variable z_t that takes the values of zero or one. Realizations of y_t from the process in state z are denoted by $y_t(z)$. We assume that z_t follows an independent first-order Markov process with transition probabilities: $\lambda_i = \Pr(z_{t+1} = i \mid z_t = i)$ for $i = 0, 1$.

The exchange rate must satisfy Equation (1a) regardless of the current process generating y_t. Expectations of next period's exchange rate, $E_t s_{t+1}$, will depend upon the probability weighted average of each state next period. Assuming market participants know that y_t is currently following the state $z = i$ process at time t, then $E_t s_{t+1} = \lambda_i E_t s_{t+1}(i) + (1 - \lambda_i) E_t s_{t+1}(j)$ for $i \neq j$ where $E_t s_{t+1}(z)$ is the conditional time t expectation of the exchange rate at $t+1$ if y_{t+1} is realized from the state z process at $t + 1$. Substituting these expectations into Equation (1a) gives

$$\begin{bmatrix} s_t(1) \\ s_t(0) \end{bmatrix} = \phi \begin{bmatrix} \lambda_1 & 1 - \lambda_1 \\ 1 - \lambda_0 & \lambda_0 \end{bmatrix} E_t \begin{bmatrix} s_{t+1}(1) \\ s_{t+1}(0) \end{bmatrix} + \begin{bmatrix} y_t(1) \\ y_t(0) \end{bmatrix}.$$

Rewriting this system of equations in matrix form, we have

$$\mathbf{S}_t = \phi \boldsymbol{\Lambda} E_t \mathbf{S}_{t+1} + \mathbf{Y}_t \tag{2}$$

where $\mathbf{S}_t$, $E_t\mathbf{S}_{t+1}$, and $\mathbf{Y}_t$ are the vector of state-contingent spot rates, expected future spot rates, and variables that influence the exchange rate, respectively, and where $\boldsymbol{\Lambda}$ is the transition probability matrix.

Equation (2) describes the dynamics of the exchange rates when y_t is realized from a particular process. In regime z, the current exchange

[5] Empirical studies of (1′) using standard rational expectations assumptions and specific measures for y_t, such as money supplies, have not adequately explained exchange rate behavior. See Meese and Rogoff (1983), for example. More recently, Mark (1995) has found that monetary variables have significant explanatory power for exchange rate movements over long horizons.

[6] Notice that in using (1′) we are not taking a stand on the identity of y_t. For our purposes, y_t could include endogenous as well as exogenous variables, and may follow arbitrarily complicated processes. Moreover, we will show that the empirical implications of (1′) are quite different when there are switches in the process for y_t than when y_t follows a single process, as has been typically assumed in the literature.

[7] In this article we do not model the switching behavior of fundamentals endogenously. However, switches in the fundamentals process may reflect such influences as shifts in the behavior of monetary regimes. For example, Kaminsky and Lewis (1993) find that U.S. monetary policy underwent shifts between contractionary and expansionary regimes roughly consistent with dollar appreciating and depreciating regimes during the late 1980s.

rate $s_t(z)$ depends upon the weighted average of expected future exchange rates from both regimes and the y_t realized from the given state, $y_t(z)$. Over time, the y_t process will switch so that the observed exchange rate, s_t, will be equal to $s_t(1)$ or $s_t(0)$ depending on the value of the state variable z_t:

$$s_t = z_t s_t(1) + (1 - z_t) s_t(0) \tag{3}$$

Equations (2) and (3) provide a framework for examining how switches in y_t affect exchange rates. For tractability, it is useful to focus on an example in which y_t switches between simple processes. For this purpose, we will examine a case where the y_t processes each follow a random walk with drift, a case that allows us to compare our results to those of Engel and Hamilton (1990). The Appendix presents solutions for the exchange rate when y_t switches between more general time series processes.

In our particular case, the vector $\mathbf{Y}_t \equiv [y_t(1) y_t(0)]'$ follows

$$\mathbf{Y}_t = \mathbf{A} + \mathbf{Y}_{t-1} + \epsilon_t \tag{4}$$

where $\mathbf{A}$ and ϵ_t are 2×1 vectors of drift coefficients and innovations (with $\mathbf{E}_{t-1}\epsilon_t = 0$), respectively. With this specification, we can solve Equation (2) for $\mathbf{S}_t$ using the method of undetermined coefficients, giving

$$\mathbf{S}_t = (\mathbf{I} - \phi\mathbf{\Lambda})^{-1}\phi\mathbf{\Lambda}(\mathbf{I} - \phi\mathbf{\Lambda})^{-1}\mathbf{A} + (\mathbf{I} - \phi\mathbf{\Lambda})^{-1}\mathbf{Y}_t \tag{5}$$

where $\mathbf{I}$ is a 2×2 identity matrix.

Equation (5) shows that exchange rates depend upon the discounted present value of the probability-weighted averages of drift terms in all future periods and the present value of the current level of y_t. Note that even if current y_t were the same across states so that $y_t(1) = y_t(0)$, the spot rates in each state will not be equal. In other words, $s_t(1)$ will not equal $s_t(0)$ because the expected future evolution between the two y_t processes will differ across states, as captured by $(\mathbf{I} - \phi\mathbf{\Lambda})^{-1}$. The Appendix shows that this feature of the exchange rate model is not specific to these particular y_t processes.

Taking the first difference of Equation (5) highlights the dynamics of the exchange rate within a regime:

$$\Delta\mathbf{S}_{t+1} = (\mathbf{I} - \phi\mathbf{\Lambda})^{-1}\mathbf{A} + (\mathbf{I} - \phi\mathbf{\Lambda})^{-1}\epsilon_{t+1} \tag{5a}$$

Changes in the exchange rate are equal to the change in y_t weighted by their discounted present value effect in all future periods, including the transition probabilities of switching between states in all future periods summarized in $\mathbf{\Lambda}$.

Our example solution for the equilibrium exchange rate incorporates the probability of switching between two different y_t processes. Comparing this solution to standard present value solutions makes this factor transparent. To do this, we rewrite Equation (2) as

$$\begin{aligned}\mathbf{S}_t &= \phi E_t \mathbf{S}_{t+1} + \mathbf{Y}_t + \phi(\boldsymbol{\Lambda} - \mathbf{I})E_t\mathbf{S}_{t+1} \\ &= \phi E_t \mathbf{S}_{t+1} + \mathbf{Y}_t + \phi \nabla \mathbf{S}_t. \end{aligned} \tag{6}$$

The first two terms on the right-hand side of Equation (6) correspond to the standard exchange rate equation for each state as in Equation (2). The last term, $\nabla \mathbf{S}'_t \equiv [\lambda_1 - 1, \lambda_0 - 1][E_t s_{t+1}(1) - E_t s_{t+1}(0)]$, is the expected jump in the exchange rate resulting from switches in the y_t process. To see the implications of this term on the exchange rate within a regime, we iterate Equation (6) forward to obtain

$$\mathbf{S}_t = E_t \sum_{i=0}^{\infty} \phi^i [\mathbf{Y}_{t+i} + \phi \nabla \mathbf{S}_{t+i}]. \tag{7}$$

Recall that $\mathbf{S}_t \equiv [s_t(1)s_t(0)]'$ and $\mathbf{Y}_t \equiv [y_t(1)y_t(0)]'$. Hence Equation (7) shows that the exchange rate $s_t(z)$ is equal to the present value of y_t in the current process z, and the present value of expected future changes in the exchange rate caused by shifts in this process.[8]

We can identify two different factors affecting exchange rates from Equation (7). First, innovations to the y_t process affect the market's perception of the present value of future y_{t+j}, $j > 0$. Second, new information affects the expected future size of the jump in exchange rates when switches in the process occur. In our solution for the exchange rate in Equation (5a), these effects are identified by the diagonal elements of $(\mathbf{I} - \phi\boldsymbol{\Lambda})^{-1}$.

We have shown that shifts in the behavior of the underlying determinants of the exchange rate will lead to accompanying jumps in the exchange rate. Standard switching models of the exchange rate have ignored these jumps, positing instead that switches in the exchange rate simply involve changes in its dynamic behavior. Later, we allow for the presence of these jumps in estimating a new switching model.

1.2 The econometric model

Our empirical switching model is based upon the example presented above. Combining the process for the exchange rate in Equation (3)

[8] Notice that standard monetary models of the exchange rate do not take account of both these effects. It is therefore possible that these models would be more successful econometrically if they incorporated the effects of switching. It is also likely that the probability of a shift in regimes is endogenous to policy variables. We leave these issues for future research.

with the exchange rate solution in Equation (5a), the equilibrium process of the exchange rate can be written as

$$s_{t+1} = z_{t+1}s_{t+1}(1) + (1 - z_{t+1})s_{t+1}(0) \tag{8}$$

where

$$s_{t+1}(1) = s_t(1) + \mu_1 + \eta_{t+1}(1)$$

and

$$s_{t+1}(0) = s_t(0) + \mu_0 + \eta_{t+1}(0).$$

Note from Equation (5a) that the drift parameters μ_i and the innovations $\eta_t(z)$ depend upon the underlying drift and innovations to the y_t processes, $\mathbf{A}$ and ϵ_t. Switches between regimes are governed by the discrete state variable, z_t, that takes on values of 1 or 0. This state variable in turn follows a first-order Markov process with transition probability matrix $\boldsymbol{\Lambda}$.

The specification in Equation (8) differs from the exchange rate switching model developed by Engel and Hamilton (1990). To see how, we rewrite Equation (8) as

$$\begin{aligned}\Delta s_{t+1} &= z_{t+1}\Delta s_{t+1}(1) + (1 - z_{t+1})\Delta s_{t+1}(0) + \Delta z_{t+1}[s_t(1) - s_t(0)] \\ &= z_{t+1}\mu_1 + (1 - z_{t+1})\mu_0 + z_{t+1}\eta_{t+1}(1) + (1 - z_{t+1})\eta_{t+1}(0) \\ &\quad + \Delta z_{t+1}[s_t(1) - s_t(0)].\end{aligned} \tag{9}$$

The first four terms on the right-hand side show that regime switches affect the dynamics of the exchange rate as in the regime-switching model of Engel and Hamilton (1990). However, our model also allows the exchange rate to jump when the regime switches (i.e., when $\Delta z \neq 0$) through the last term.[9] As explained above, such jumps are likely when market participants revise discretely their expectations of the future path of y_t even if current y_t are relatively unchanged.

Estimating this model requires identifying the jumps in the exchange rate associated with a change in regime. For this purpose, we use the information in forward exchange rates through the identity that the conditionally expected exchange rate equals the forward premium plus the foreign exchange risk premium; that is, $E_t\Delta s_{t+1} \equiv f_t - s_t + \theta_t$, where f_t is the logarithm of the time t forward rate on a contract to buy or sell dollars next period and θ_t is the risk premium on this position. Note that $E_t\Delta s_{t+1}$ can be written as the probability weighted

[9] In principle, we could allow the exchange rate to jump according to a different process. In the absence of a model to econometrically identify these jumps on theoretical grounds, we allow them to be determined by the difference between the two potential processes.

average conditional upon each regime:

$$E_t\Delta s_{t+1} = z_t[\lambda_1\mu_1 + (1-\lambda_1)\mu_0 - (1-\lambda_1)(s_t(1) - s_t(0))] \qquad (10)$$
$$+ (1-z_t)[(1-\lambda_0)\mu_1 + \lambda_0\mu_0 + (1-\lambda_0)(s_t(1) - s_t(0))]$$

where λ_i is the transition probability of remaining in regime i from one period to the next. Setting Equation (10) equal to $f_t - s_t + \theta_t$, and solving for $s_t(1) - s_t(0)$ implies

$$s_t(1) - s_t(0) = [(f_t - s_t) - \psi(z_t) + \theta_t]/\Gamma(z_t) \qquad (11)$$

where

$$\psi(1) = \lambda_1\mu_1 + (1-\lambda_1)\mu_0$$

$$\psi(0) = \lambda_0\mu_0 + (1-\lambda_0)\mu_1$$

$$\Gamma(1) = \lambda_1 - 1$$

$$\Gamma(0) = 1 - \lambda_0.$$

Within a regime, the expected size of the jump varies in response to changes in both the risk premium θ_t and the forward premium $f_t - s_t$. Thus, given an assumption about the behavior of the risk premium, we can use the forward premium and the parameters of the model to identify $s_t(1) - s_t(0)$. Since this term only affects the dynamics of the exchange rate in Equation (9) when $\Delta z_t \neq 0$, we only need to know the risk premium when a change in regime occurs. We therefore make the identifying assumption that the risk premium $\theta_t = \theta(1)$ during switches from regime 1 to 0, and $\theta_t = \theta(0)$ during switches from regime 0 to 1. With these restrictions, we use Equation (11) to identify the magnitude of the jump, $[s_t(1) - s_t(0)]$, when $\Delta z_t \neq 0$ in our estimation Equation (9).[10] Note that we do not restrict the time-varying behavior of the risk premium within regimes, so this identifying assumption is fairly weak. Below we examine the sensitivity of our results to this assumption.

1.3 Results

We estimate our model using the spot and forward exchange rates sampled at the end of the month from Citicorp Database Services for

[10] Since $s_t(1)$ and $s_t(0)$ are nonstationary variables, it may seem as though the magnitude of the jump can become arbitrarily large over time. Note, however, that through Equation (11) we have bounded the size of the jump through the forward premium. As long as the forward premium is stationary, then the size of the jump will be bounded as well. To verify that the size of the jumps were not excessive, we used Equation (10) to back out the size of the jump in our sample. The estimates as a proportion of the standard deviation of exchange rate changes ranged from a minimum of zero to a maximum of 1.6 for the yen, 3 for the pound, and 3.6 for the DM.

the period 1975 to 1989.[11] In particular, we examine the exchange rate for the U.S. dollar against the Deutschemark, the British pound, and the Japanese yen.

Table 1 shows the results of estimating the model given by Equations (9) and (11) with maximum likelihood, using a modified version of the Hamilton (1989) filter. We provide estimates at the monthly frequency as well as the quarterly frequency for comparison with the quarterly estimates in Engel and Hamilton (1990). The first two columns show the estimates of the transition probabilities of staying in each regime. The point estimates are mostly above 0.50, particularly for the pound and the DM. The estimates on the drift terms, μ_1 and μ_0, indicate that these states reflect dollar appreciating and dollar depreciating states, respectively. The estimates of μ_1 are positive, while those of μ_0 are negative.

Columns (5) and (6) report the estimates of the risk premium during the period when the process switches from one regime to the other. In column (5) the risk premium estimates of $\theta(1)$ are negative, while in column (6) the estimates of $\theta(0)$ are positive, except for the statistically insignificant quarterly pound estimates. These estimates have an interesting interpretation. During periods before a switch to a depreciating regime, traders willing to buy dollars forward required a lower forward rate, f_t, relative to the expected spot rate, $E_t s_{t+1}$, to compensate for the risk that the price of dollars at the spot rate may fall. Therefore, $\theta(1)$ is negative. By analogous reasoning, $\theta(0)$ is positive.

Columns (7) and (8) give the estimates of the variances in each state. Consistent with Engel and Hamilton (1990), the variance of the appreciating dollar state is higher than the depreciating state, except for the pound where the pattern is reversed.

Columns (9) through (13) report the marginal significance levels for a series of diagnostic tests. The columns labeled T_1 and T_2 report those levels for the hypothesis that the first-order serial correlation of the residuals is zero for states 1 and 0, respectively. None of these hypotheses are rejected at the 95 percent confidence level, and the restriction is rejected at the 90 percent confidence level only for the quarterly pound in state 1. The columns with headings T_3 and T_4 report the marginal significance levels for LM tests that the first-order ARCH coefficients are zero in the residuals, $\eta_t(z)$, for states 1 and 0,

[11] These data were kindly provided by Geert Bekaert and Robert Hodrick and are described in Bekaert and Hodrick (1993). These series match up the current forward rate with the appropriate spot rate on the actual corresponding delivery date. As such, this series treats complications such as holidays and contracts expiring at the end of the month in the same way as they are settled in the foreign exchange market.

Table 1
Maximum likelihood estimates of switching model

	Parameters								Diagnostic tests				
	λ_1 (1)	λ_0 (2)	μ_1 (3)	μ_0 (4)	$\theta(1)$ (5)	$\theta(0)$ (6)	σ_1^2 (7)	σ_0^2 (8)	T_1 (9)	T_2 (10)	T_3 (11)	T_4 (12)	T_5 (13)
Monthly data													
Pound	0.936	0.990	1.715	−0.623	−0.256	0.183	7.463	9.735	0.955	0.735	0.461	0.611	0.184
	(0.015)	(0.010)	(0.333)	(0.303)	(0.390)	(0.306)	(1.276)	(1.331)					
Mark	0.979	0.924	0.037	−0.827	−0.428	1.040	11.313	3.862	0.745	0.889	0.206	0.379	0.876
	(0.010)	(0.044)	(0.354)	(0.448)	(0.350)	(0.645)	(1.646)	(1.161)					
Yen	0.933	0.614	0.550	−2.067	−0.428	3.111	7.005	4.711	0.159	0.621	0.895	0.432	0.278
	(0.018)	(0.246)	(0.281)	(1.076)	(0.292)	(0.900)	(0.971)	(2.987)					
Quarterly data													
Pound	0.836	0.964	4.104	−1.080	0.880	−0.571	21.285	33.474	0.056	0.121	0.210	0.319	0.288
	(0.049)	(0.080)	(1.127)	(1.094)	(1.404)	(1.168)	(8.491)	(9.443)					
Mark	0.941	0.808	1.493	−4.310	−1.586	5.335	20.407	10.663	0.410	0.432	0 680	0.728	0.811
	(0.028)	(0.105)	(0.723)	(1.187)	(0.748)	(1.520)	(4.413)	(4.737)					
Yen	0.837	0.337	1.943	−10.557	−0.977	6.214	19.357	9.010	0.925	0.709	0.187	0.489	0.834
	(0.057)	(0.193)	(0.982)	(2.738)	(0.851)	(1.705)	(5.397)	(5.174)					

$$s_t = z_t s_t(1) + (1 - z_t)s_t(0) \qquad f_t = E_t s_{t+1} - z_t\theta(1) - (1 - z_t)\theta(0)$$

$$s_t(i) = \mu_1 + s_{t-1}(i) + \eta_t(i) \qquad \eta_t(i) \sim N(0, \sigma_i^2) \qquad \mathrm{Prob}(z_{t+1} = i \mid z_t = i) = \lambda_i\ i = 0, 1$$

The model is estimated with the ask spot rate s_t and one-period bid forward rate f_t. Columns (1) to (8) report the maximum likelihood estimates with the asymptotic standard errors in parenthesis. Marginal significance levels for the diagnostic tests are shown in columns (9) to (13). T_1 and T_2 are the MSLs for the tests of no serial correlation in the errors $\eta_t(i)$ for the state 1 and state 0 process, while T_3 and T_4 are the MSLs for the tests of first-order ARCH in $\eta_t(i)$ in states 1 and 0. T_5 is the MSL for the test of the cross-equation restrictions.

respectively. Again, none of these statistics reject the hypothesis at standard confidence levels.

The model restricts the jumps between regimes, $s_t(1) - s_t(0)$, to depend upon the forward premium as defined in Equation (11). In the last column of Table 1 we provide a diagnostic test of this restriction. We consider periods in which shifts occur and test whether changes in the spot rate process over this period are consistent with the forward rate restriction imposed upon the full model.

Combining Equations (9) and (11) we can write the restrictions between the spot and forward rates as

$$\Delta s_{t+1} = \begin{cases} b_0[f_t - s_t + \theta(0)] - b_0\lambda_0\mu_0 + \eta_{t+1}(1) & \text{when } \Delta z_{t+1} = 1 \\ b_1[f_t - s_t + \theta(1)] - b_1\lambda_1\mu_1 + \eta_{t+1}(0) & \text{when } \Delta z_{t+1} = -1 \end{cases} \tag{12}$$

where, under the null hypothesis, $b_i = (1 - \lambda_i)^{-1}$ where λ_i are the transition probabilities from the Markov process for z_t.

Column (13) of Table 1 reports the marginal significance levels of the LM tests of this hypothesis. As the numbers show, none of the LM statistics are significant at the 5 percent level. Thus, we cannot reject the hypothesis that market participants rationally included the effects of regime switches when forecasting future exchange rates.[12] Engel and Hamilton (1990) also reported a test of the relationship between the spot rate process implied by their Markov switching model and the forward premium. They tested and rejected the hypothesis that the expected future exchange rate from their model was equal to a forward premium up to a white noise error term. In our case, we cannot reject the hypothesis that the expected future exchange rate generated by our model is equal to the forward premium plus our estimates of the risk premium.

Figures 1, 2, and 3 show some of the implications of the estimated models using quarterly data. The upper panel of each figure plots the logarithm of the exchange rate. The lower panel shows the probability of being in regime 1, the appreciating regime, estimated from the whole data sample (i.e., the smoothed probabilities). In the case of the pound and mark, the long swings in the exchange rates closely correspond to the estimated regimes. By contrast, the yen appears to have been in a mildly appreciating regime for most of the sample,

[12] To test the sensitivity of our estimates to the assumption about the risk premia during regime switches, we also estimated models where $\theta_t = \theta(1) + \omega_t(1)$ during switches from regime 1 to 0, and $\theta_t = \theta(0) + \omega_t(0)$ during switches from 0 to 1 where $\omega_t(i)$ are normally distributed errors with zero means. The results from estimating these models are almost identical to those reported in Table 1.

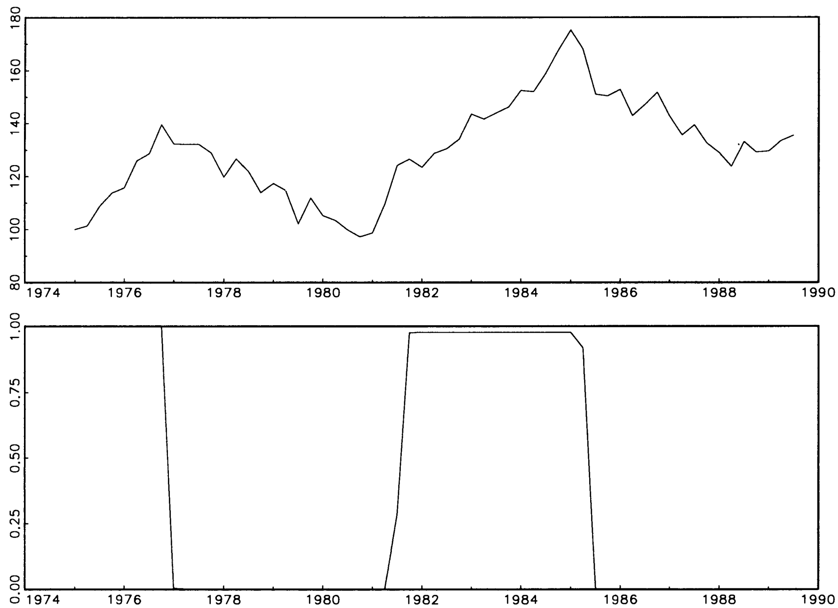

Figure 1
The upper panel plots the log of the pound/dollar exchange rate (1975 = 100). The lower panel plots the probability of being in regime 1, the appreciating regime. This is the smoothed probability based on the whole data sample using the estimates from the quarterly pound/dollar model reported in Table 1.

punctuated by short periods when the exchange rate switched to a sharply depreciating regime.

2. Long-Run Behavior in the Risk Premium: A New Anomaly?

Past studies of foreign exchange returns have uncovered significant anomalies in the high frequency behavior of returns. In this section, we first show that an anomaly also appears to be present in the low frequency behavior of returns.[13] Based upon our model above, we will then demonstrate that rational anticipations of long-term swings in the exchange rate process can generate this low frequency behavior.

[13] Recent studies that have examined the long-run relationship in spot and forward rates include Hakkio and Rush (1989) and McCallum (1992).

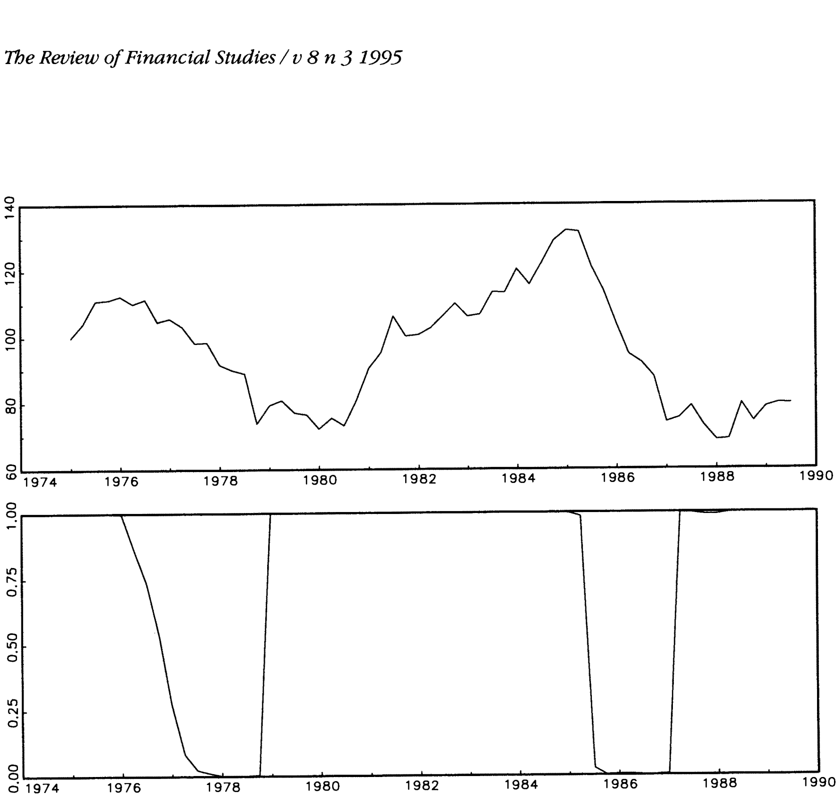

Figure 2
The upper panel plots the log of the mark/dollar exchange rate (1975 = 100). The lower panel plots the probability of being in regime 1, the appreciating regime. This is the smoothed probability based on the whole data sample using the estimates from the quarterly mark/dollar model reported in Table 1.

2.1 The long-run relationship between spot and forward rates

From the definitions of spot and forward exchange rates, we can write the speculative excess return on a forward contract in period t to buy dollars in period $t+1$ as

$$s_{t+1} - f_t = \theta_t + e_{t+1} \tag{13}$$

where θ_t is the risk premium and e_{t+1} is the market's exchange rate forecast error.

Equation (13) illustrates why empirical studies typically treat excess returns as covariance stationary processes, called "$I(0)$" in the literature. The equation shows that excess returns are comprised of a risk premium, θ_t, and a forecast error, e_{t+1}. Risk premia have been considered stationary on theoretical grounds.[14] And under rational ex-

[14] For example, standard models of time-varying risk premia imply that risk premia are stationary since they depend upon the time-series properties of the change in consumption. See Grossman and Shiller (1981) and Backus, Gregory, and Zin (1989) for some applications.

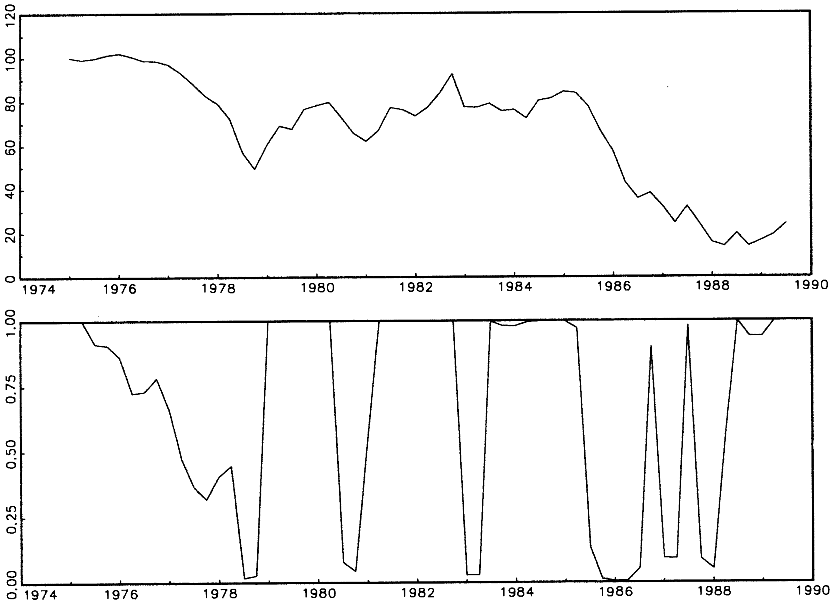

Figure 3
The upper panel plots the log of the yen/dollar exchange rate (1975 = 100). The lower panel plots the probability of being in regime 1, the appreciating regime. This is the smoothed probability based on the whole data sample using the estimates from the quarterly yen/dollar model reported in Table 1.

pectations, forecast errors follow a stationary process. Since the sum of two stationary variables must be stationary, the sum of the risk premium and the forecast error must also be stationary under the assumptions above. By contrast, the levels of spot and forward rates have been found to follow processes with very persistent shocks, well-approximated as permanent disturbances. These disturbances are covariance stationary only after differencing.[15] Processes with these types of shocks have been denoted "$I(1)$" in the literature.

The requirement that both sides of Equation (13) must be stationary places restrictions upon the relationship between spot and forward rates. Specifically, if spot and forward rates are $I(1)$, excess returns will only be $I(0)$ stationary when permanent shocks to s_t and f_t cancel out. This restriction has two implications for the long-run relationship

[15] Meese and Rogoff (1983) and Meese and Singleton (1982) found that exchange rates follow a random walk. More recently, Baillie and Bollerslev (1989) test directly for unit roots in exchange rates and find that exchange rates and forward rates are cointegrated.

between spot and forward rates. First, it requires that the variables in the vector defined as $\mathbf{X}_t \equiv [s_{t+k}, f_t]'\forall t$ be cointegrated. And second, it places restrictions on the cointegrating relationship between these variables. Specifically, a cointegrating vector is the linear combination of s_t and f_t that is stationary. In other words, if the cointegrating vector is defined as $\boldsymbol{\alpha}$, then $\boldsymbol{\alpha}'\mathbf{X}_t$ must be $I(0)$ stationary. Clearly, if excess returns are stationary, then the cointegrating vector, α, must be equal to $[1, -1]$ since premultiplying by this vector, $\boldsymbol{\alpha}'\mathbf{X}_t$, gives excess returns.[16]

Therefore, we can assess whether excess returns are stationary by estimating the cointegrating regression of spot rates on forward rates and testing whether the coefficient on forward rates equals one. In other words, in the cointegrating regression,

$$s_{t+1} = a_0 + a_1 f_t + v_{t+1}, \tag{14}$$

a_1 must be equal to 1 under the null hypothesis of stationary excess returns.

It should be noted that Equation (14) differs from ordinary regressions. Under the null hypothesis, the residual in the cointegrating regression equals the sum of the time-varying component of the risk premium and the forecast error (i.e., $v_{t+1} = \theta_t - a_0 + e_{t+1}$). If both are stationary, then the sum must be asymptotically independent of all permanent shock components so that θ_t and e_{t+1} cannot contaminate the estimate of a_1.[17] In this case, we should find that $a_1 = 1$.[18]

On the other hand, rejecting $a_1 = 1$ implies that the process followed by predictable excess returns shares the same persistence as the forward rate process. To see why, subtract the forward rate from both sides of the cointegrating regression of Equation (14) to yield

$$s_{t+1} - f_t = \theta_t + e_{t+1} = a_0 + (a_1 - 1)f_t + v_{t+1}. \tag{15}$$

Since v_t is stationary by construction, the effect of permanent shocks in the forward rate upon excess returns will depend upon the coefficient $(a_1 - 1)$. Clearly, if a_1 equals one, excess returns depend only upon the stationary term, v_t. If a_1 does not equal one, however, excess returns will depend upon the stationary variable v_t, but will also inherit the permanent component in the forward rate.

---

[16] Technically, α need only be proportional to $[1, -1]$ for $\alpha' X_t$ to be stationary. In the cointegrating regression, we will normalize the coefficient on s_{t+1} to equal one so that the proportionality coefficient also equals one.

[17] See Stock (1987) or the surveys in Campbell and Perron (1991) and Diebold and Nerlove (1990).

[18] This test is valid only if s_t and f_t are cointegrated and share a unique common trend. We test for cointegration later.

2.2 Cointegration test results

In order to test whether the estimates of a_1 are significantly different from one, two econometric issues need to be addressed. First, the asymptotic distribution of a_1 is nonnormal, invalidating OLS standard errors. Second, the OLS estimates of the parameters in cointegrating regressions are known to be biased in finite samples. For these two reasons, we apply the bias adjustment and inference techniques described in Stock and Watson (1989).[19] For efficiency, we estimate the system of three equations jointly. Details of the techniques are provided in the appendix.

The first two columns of Table 2 report the results from estimating the cointegrating regression in Equation (14) using the spot bid rate and the forward ask rate at horizons of 1 and 3 months using monthly frequency data. Following Bekaert and Hodrick (1993), we construct the excess returns from buying dollars forward at the market ask rate and selling these dollars at the future bid spot rate.

Column (1) shows the coefficient estimates based upon OLS. As these estimates show, the coefficients are all less than one. Column (2) provides p-values for the hypothesis that $a_1 = 1$ under different assumptions about the structure of the residual process in each equation. The upper p-value corrects for conditional heteroskedasticity in the regression residuals, while the lower p-value corrects for both heteroskedasticity and serial correlation (an MA(2) process).[20] These p-values show that the null hypothesis is strongly rejected for the Japanese yen and less strongly rejected for the German mark. On the other hand, the results for the pound are sensitive to serial correlation assumptions. The row labeled "Joint" strongly rejects the hypothesis that the coefficients in the three equations are jointly equal to one at both horizons.

All of these results are based upon equations that use bid spot rates and ask forward rates. However, as noted by Stambaugh (1988), the true value of the spot and forward rates may lie somewhere in between the bid and the ask rate. If so, our use of the ask forward rate may introduce a measurement error into the right-hand side of Equation (14) that might bias the tests toward rejection. To consider this possibility, we reestimated the equations using the average of the bid and ask rates for both the forward and spot rates. If measurement error has important effects upon our estimates, then we should observe

---

[19] Before running the regressions, we used the Johansen (1988) procedure to check that the pairs of individual spot and forward rates were cointegrated. The same procedure was also used to test for the number of trends in the vectors of three spot and three forward rates used in the system estimation. We could not reject the hypothesis that all these vectors contained three trends.

[20] We obtain similar results allowing for higher order MA processes in the residuals.

Table 2
Cointegration results

Currency	Bid spot and ask forward		Averages of bid and ask	
	a_1 (1)	Asy p-value H_0: $a_1 = 1$ (2)	a_1 (3)	Asy p-value H_0: $a_1 = 1$ (4)
$k = 1$				
Pound	0.970	0.086 0.223	0.971	0.209 0.463
Mark	0.986	0.028 0.038	0.986	0.010 0.013
Yen	0.992	< 0.001 < 0.001	0.991	< 0.001 < 0.001
Joint		< 0.001 0.002		< 0.001 < 0.001
$k = 3$				
Pound	0.909	0.013 0.204	0.909	0.021 0.297
Mark	0.946	0.074 0.074	0.946	0.052 0.057
Yen	0.967	< 0.001 0.074	0.967	< 0.001 0.052
Joint		< 0.001 < 0.001		< 0.001 < 0.001

$$s_{t+k} = a_0 + a_1 f_t + v_t \tag{14}$$

Columns (1) and (3) report the OLS estimates of a_1 in the regression, where s_t is the spot exchange rate and f_t is the k-month ahead forward rate. The regressions are estimated with monthly data from January 1975 to December 1989. Columns (2) and (4) show the p-values from Wald tests of H_0: $a_1 = 1$. The methods developed by Stock and Watson (1989) are used to account for the finite sample bias in the OLS estimates of a_1 when calculating these statistics. The upper p-value allows for heteroskedasticity in the residuals, the lower value allows for heteroskedasticity and serial correlation.

different results using the bid-ask average instead of the bid and ask rates individually.

Column 3 of Table 2 reports the OLS estimates of a_1 using the bid-ask averages. These results are very similar to those in column (1) and are even identical in most cases.

The similarity between the results using average bid-ask rates and individual bid-ask rates indicate that our findings are robust to the presence of measurement error in forward rates. On the basis of this evidence, one might conclude that spot and forward rates are not cointegrated one for one, even though this is the implicit assumption in studies of foreign exchange returns. Taken literally, this result implies that fluctuations in excess returns have a permanent component. Since conventional assumptions require forecast errors to be stationary and uncorrelated with all current information, one interpretation is that the risk premium is subject to permanent shocks.

2.3 Switching and cointegrating regressions

Our switching framework provides an alternative explanation to this striking result. Intuitively when market participants anticipate a switch in exchange rate regime, this anticipation creates a "peso problem" in the expected exchange rate during periods when switches do not materialize. During these periods, the "peso problem" may induce a great deal of persistence in the deviation between the future realized spot rate and the expected spot rate. While rational expectations implies that forecast errors are serially uncorrelated over long samples, the forecast errors during samples with few regime changes, as found in Figures 1, 2, and 3, may be sufficiently persistent to account for the cointegrating results in Table 2.

To see why, note that the forecast errors based upon the switching model in Equation (9) can be written

$$\begin{aligned} s_{t+1} - E_t s_{t+1} &= z_{t+1}\eta_{t+1}(1) + (1 - z_{t+1})\eta_{t+1}(0) \\ &\quad + [\Delta z_{t+1} - \Gamma(z_t)]E_t[s_{t+1}(1) - s_{t+1}(0)], \end{aligned} \tag{16}$$

where $\Gamma(1) = \lambda_1 - 1$ and $\Gamma(0) = 1 - \lambda_0$. During periods when there is no change in regime so that $\Delta z_{t+1} = 0$, forecast errors include the "peso problem" term $-\Gamma(z_t)E_t[s_{t+1}(1) - s_{t+1}(0)]$. This component is highly serially correlated as long as $\lambda_i < 1$ because the expected future exchange rates within each regime differ. In particular, $E_t[s_{t+1}(1) - s_{t+1}(0)]$ contains two parts: first, the difference between the two drift terms as in Engel and Hamilton (1990), $\mu_1 - \mu_0$, and second, the expected size of the jump component, $E_t(s_t(1) - s_t(0))$. Of course, over sufficiently long periods, the sample will include many regime changes so that errors in forecasting Δz_t will be serially uncorrelated. Since the expectation of the change in the regime state variable, Δz_{t+1}, equals its probability, $\Gamma(z_t)$, in large samples, $[\Delta z_{t+1} - \Gamma(z_t)]E_t[s_{t+1}(1) - s_{t+1}(0)]$ will be serially uncorrelated and have an unconditional mean of zero.

To examine whether "peso problems" in the forecast errors could have affected our cointegration results in Table 2, we conducted a series of Monte Carlo experiments. In the first set of experiments, we used the switching model estimates to generate an empirical distribution for the estimates of a_1 in the cointegrating regression in Equation (14). First, we used Equation (8) and our estimates of the drift terms and the variances in each state, μ_i and σ_i, to generate a sequence for $s_t(z)$ equal to the length of our sample. We then drew realizations for z_t based upon the estimated transition probability matrix, and from these realizations generated a series for the observed exchange rate process. The forward rate was constructed from these series using Equation (11) under the assumption that the risk premia were equal to our estimates of $\theta(z_t)$ within each regime.[21] The cointegrating regression in Equation (14) was then run with the generated spot and forward rates and the OLS estimates of a_1 were saved. This whole procedure was repeated 1000 times to generate an empirical distribution for a_1.

Column (1) of Table 3 reports the OLS estimates of a_1 calculated from the data.[22] Column (2) shows the probability of observing the OLS estimates of a_1 when the data is generated by the switching model. As the table shows, all the probabilities are above 30 percent. Thus, if market participants rationally anticipate switches in the exchange rate regime and the risk premium is stationary, a researcher would quite likely find estimates of a_1 as low as we find in the data.[23]

Since "peso problems" occur only in small samples, it is natural to ask whether the results in Table 2 could also be due to other small sample effects. Columns (3) and (4) of Table 3 address this question. These columns present p-values for the hypothesis that the OLS estimates of a_1 are equal to one based on the small sample distribution of the Stock and Watson (1989) test statistics calculated from Monte Carlo experiments. These experiments exclude the possibility of regime switching and are described in the Appendix. The p-values in column (3) are based on experiments that assume conditional homoskedasticity in the data generation process, while those in column (4) are based on experiments that allow for conditional heteroskedasticity. As the table shows, the p-values based on the small sample distributions of the Stock and Watson test statistics are typically smaller than those reported in Table 2. Thus, the cointegration results do not appear to be explained by generic small sample problems.

Overall, the results in Table 3 show that standard cointegrating regression tests are likely to find that spot rates and forward rates

[21] In these experiments, we treat the risk premium as constant in each regime in order to understate as much as possible the serial correlation in excess returns. Thus, a finding that exchange rate switching biases estimates of the cointegrating regressions in these experiments can only be attributed to the "peso problem" effects. Serial correlation in risk premia can only add further bias than we find below.

[22] The estimates in the lower panel differ from those in Table 2 because we estimated the cointegrating regression with quarterly rather than monthly data to conform to the switching model estimates.

[23] In order to consider whether the evidence in Table 3 was sensitive to our particular characterization of jumps, we also conducted Monte Carlo experiments without allowing for jumps, as in Engel and Hamilton (1990), and found similar results. However, the model without jumps cannot generate time variation in expected exchange rates, a feature that will be necessary to explain the high frequency behavior of returns.

Table 3
Monte Carlo cointegration regressions

		p-values for H_0: $a = 1$		
		Switching	No switching	
Currency horizon	a_1 (1)	(2)	Homoskedasticity (3)	Heteroskedasticity (4)
$k = 1$				
Pound	0.970	0.350	0.429	0.420
Mark	0.986	0.494	0.011	0.008
Yen	0.992	0.479	< 0.001	0.001
$k = 3$				
Pound	0.907	0.382	0.033	0.024
Mark	0.949	0.400	0.031	0.035
Yen	0.966	0.566	0.001	< 0.001

$$s_{t+k} = a_0 + a_1 f_t + v_t \tag{14}$$

Column (1) reports the OLS estimates of a_1 in the regression, where s_t is the spot exchange rate and f_t is the k-month ahead forward rate. The regressions are estimated with monthly data for $k = 1$ and quarterly data for $k = 3$. Column (2) reports the p-value for H_0: $a_1 = 1$ based on the empirical distribution for a_1 calculated by Monte Carlo simulation of the switching models. Columns 3 and 4 report p-values for H_0: $a_1 = 1$ based on the small sample distribution of the test statistic. These distributions are generated by Monte Carlo simulation and account for the finite sample bias in the estimates of a_1 but ignore the presence of switching. The p-values in columns (3) and (4) are calculated from simulations that assume the residuals in Equation (14) are homoskedastic and heteroskedastic, respectively.

do not move together one for one in the long run if the exchange rate switches between processes of appreciation and depreciation. Intuitively the expectation of a switch from an appreciating to a depreciating dollar regime implies that forecast errors will be serially correlated for periods in which this shift is not realized. This "peso problem" induces a small sample serial correlation in forecast errors that has a great deal of persistence. Our experiments show that this serial correlation biases the estimates of the cointegrating regression coefficients.

3. High Frequency Behavior of the Risk Premium: How Much Is a "Peso Problem?"

We will now examine whether "peso problems" due to long swings in the exchange rate can help explain some of the anomalous behavior in excess returns at high frequencies.

3.1 Switching and higher frequency movements in the risk premium

The empirical exchange rate literature has documented that the forward premium, $f_t - s_t$, is a biased predictor of the future change in the exchange rate, Δs_{t+1}. The nature of this finding has been interpreted as evidence that the time varying risk premium is correlated with the forward discount. To see why, consider the regression due to Fama (1984):

$$\Delta s_{t+1} = \beta_0 + \beta(f_t - s_t) + w_{t+1} \tag{17}$$

Using the identity, $\Delta s_{t+1} \equiv f_t - s_t + \theta_t + s_{t+1} - E_t s_{t+1}$, and the standard rational expectations assumption that the covariance between the forward premium and the forecast error is zero, that is, $\text{Cov}(s_{t+1} - E_t s_{t+1}, f_t - s_t) = 0$, least squares theory implies that the estimate of β is

$$\hat{\beta} = 1 + \text{Cov}(\theta_t, f_t - s_t)/\text{Var}(f_t - s_t). \tag{18}$$

Under this assumption, an estimate of β different from one implies covariation between the risk premium θ_t and the forward premium.

In practice, the coefficients in this regression tend to be significantly negative. Based upon Fama's decomposition of the β coefficient in Equation (18), negative estimates imply that the variance of the risk premium exceeds the variance of the forward premium.[24] While negative estimates such as these can be reconciled with the theoretical predictions of the asset pricing model for foreign exchange, empirical attempts to do so have been unsuccessful, as described in the introduction. Thus, from the perspective of standard rational expectations, the regression results represent a puzzle.

The importance of the "peso problem" in explaining the anomalous low frequency behavior of foreign exchange returns in Table 2 suggests that this problem may contribute to the high frequency puzzle as well. When market participants anticipate shifts in the exchange rate process, forecast errors may be serially correlated during periods when the switches do not occur. Therefore, the condition that $\text{Cov}(s_{t+1} - E_t s_{t+1}, f_t - s_t) = 0$ used to derive the Fama decomposition in Equation (18) may not hold within the sample.

To see the effects of switching more clearly, consider a projection of the risk premium upon the forward premium,

$$\theta_t = \theta(f_t - s_t) + u_t, \tag{19}$$

where u_t is the component of the risk premium that is uncorrelated with the forward premium.[25] Thus, the decomposition in Equation (18) implies that $\beta = 1 + \theta$. In the presence of exchange rate

[24] In particular, since $\hat{\beta} < 0$, $\text{Cov}(\theta_t, f_t - s_t)/\text{Var}(f_t - s_t) < -1$. In turn, this inequality implies that $|\text{Cov}(\theta_t, f_t - s_t)| > \text{Var}(f_t - s_t)$. Finally, this last inequality can only be true if $\text{Var}(\theta_t) > \text{Var}(f_t - s_t)$.

[25] In Equation (19), we have subsumed the constant term for expositional clarity. However, none of our main conclusions would be altered by including this term.

switching, this decomposition may not hold except in very large samples. To see why, we use Equation (19) to substitute for θ_t in Equation (11). Substituting the resulting expression for the difference in spot rates for $s_t(1) - s_t(0)$ in Equation (9), we obtain

$$\begin{aligned}\Delta s_{t+1} = & \left[\frac{\Gamma(z_t) - \Delta z_{t+1}}{\Gamma(z_t)}\right](z_t\mu_1 + (1 - z_t)\mu_0) \\ & + \left[\frac{\Delta z_{t+1}}{\Gamma(z_t)}\right](1 + \theta)(f_t - s_t) \\ & - \left[\frac{\Delta z_{t+1}}{\Gamma(z_t)}\right] u_t + z_{t+1}\eta_{t+1}(1) + (1 - z_{t+1})\eta_{t+1}(0). \quad (20)\end{aligned}$$

This equation expresses the exchange rate change in terms of the forward premium and the components of the switching model.

In the presence of exchange rate switching, the estimates of β depend upon the covariance between the forward premium and each of the components on the right-hand side of Equation (20). The potential biases to $\hat{\beta}$ in small samples can be evaluated by considering these covariances. For this purpose, note first that the last three terms depend upon errors uncorrelated with the forward premium. In particular, u_t is uncorrelated with $f_t - s_t$ through Equation (19) and $\eta_{t+1}(i)$ are innovations uncorrelated with all variables known at time t. Therefore, the estimates of β can only be affected by the first two terms on the right-hand side of Equation (20).

The first term on the right-hand side of Equation (20) incorporates the effects of the expected change in the exchange rate. The effects of the current drift in the exchange rate, $z_t\mu_1 + (1 - z_t)\mu_0$, depends upon the deviation of the change in the switching process, Δz_{t+1}, from its expected change, $E_t\Delta \mathbf{z}_{t+1} = \Gamma(z_t)$. Thus, the deviation of $\Gamma(z_t)$ from Δz_{t+1} represents an expected switch in regimes not realized during a particular period. In large samples this deviation will be serially uncorrelated and will have an unconditional mean of zero.

The second term captures the effects of the covariation between the risk premium and the forward rate. In samples with many switches in regime, the unconditional mean of $\Delta z_{t+1}/\Gamma(z_t)$ will equal one so that $\hat{\beta} = (1 + \theta)$ as in the Fama decomposition. In small samples, by contrast, the deviation between Δz_{t+1} and $\Gamma(z_t)$ may introduce additional small sample bias into the estimate of β.

3.2 Empirical evidence

To evaluate the potential biases introduced by the long swings in the exchange rate, we re-estimated the standard Fama regression and then asked how much the "peso problem" generated by the switching

Table 4
Monte Carlo Fama regressions

	Currency	p-value	Monte Carlo experiments	
	$\theta = \beta - 1$ (1)	H_0: $\theta = 0$ (2)	Bias (3)	Ratio (4)
Monthly data				
Pound	−3.266	< 0.001	−0.726 (3.438)	1.222 (1.053)
Mark	−4.502	0.001	−1.068 (3.253)	1.237 (0.722)
Yen	−3.022	< 0.001	−0.107 (0.607)	1.035 (0.201)
Quarterly data				
Pound	−3.347	0.001	−0.724 (2.691)	1.216 (0.804)
Mark	−4.448	0.004	−0.720 (2.735)	1.162 (0.615)
Yen	−3.955	< 0.001	−0.124 (0.700)	1.031 (0.177)

$$\Delta s_{t+1} = \beta_0 + \beta(f_t - s_t) + w_{t+1} \tag{17}$$

Column (1) reports the value for θ implied by the OLS estimates of β under the hypothesis that the risk premium is related to the forward discount by

$$\theta_t = \theta(f_t - s_t) + u_t \tag{19}$$

where u_t is uncorrelated with $f_t - s_t$. s_t is the ask spot rate and f_t is the one-period bid forward rate. Column (2) reports the p-value for H_0: $\theta = 0$, based on Wald tests that allow for heteroskedasticity in the residuals u_{t+1}. Column (3) reports the bias in the estimate θ, measured as $\theta^* - \theta$ where θ^* is the value of θ implied from the regression in Equation (17) based on simulated data from the switching model and Equation (19). The table reports the mean bias with the standard deviation in parenthesis of the empirical distribution based on 1000 simulations. Column (4) reports the mean and standard deviation of the ratio θ^*/θ.

model contributed to the typical findings. Specifically, we estimated Equation (17) by OLS and constructed the estimate of $\theta = \hat{\beta} - 1$. Column (1) of Table 4 shows the implied estimates of θ from these regressions. Notably, all of the estimates are negative and exceed three in absolute value. Column (2) gives the p-values based on t tests (adjusted for heteroskedasticity) for the hypothesis that $\theta = 0$. The marginal significance levels on all the statistics are much less than 1%, indicating a very strong rejection of the hypothesis.

To study how the estimates of θ may be affected by "peso problems," we then conducted the following Monte Carlo experiment. First, we generated data on spot rates from Equation (8) using the estimates of λ_i, μ_i, and σ_i^2 in Table 1. Next, we substituted the expression for the risk premium in Equation (19) into Equation (11) and solved for the forward rate, setting θ equal to $\hat{\beta} - 1$. Based upon the generated forward and spot rates, we then ran the Fama regression of Equation (17) and saved the OLS estimates of θ as θ^*. We repeated these steps 1000 times to form the empirical distribution for θ^*. Comparing this distribution against θ allows us to examine the bias induced by the "peso problem" relative to the standard assumption that the predictable excess return is entirely due to the risk premium.

Column (3) of Table 4 reports the mean of the bias, $\theta^* - \theta$, based upon the Monte Carlo distribution (the median values were essentially the same). While the Fama coefficient is biased downward in all cases, the mean bias is a good deal larger for the pound and the mark than for the yen. The standard deviations of the bias shown in parentheses demonstrate that the distributions of the bias are quite spread out. This observation further emphasizes the problems in using standard asymptotic theory as in column (2) to draw implications about the risk premium.

We also used the Monte Carlo experiments to examine how much "peso problems" contribute to the measured variability in the risk premium from the Fama regression. To answer this question, we generated the empirical distribution of θ^*/θ. This ratio is equal to the standard deviation of the measured risk premium divided by the standard deviation of the risk premium implied by the Fama interpretation.[26] Column (4) reports the mean and standard deviation of θ^*/θ. In all cases, the mean value for θ^*/θ implies that the standard deviation of the measured risk premium exceeds the true risk premium imposed on the experiment. For the pound and the mark, the standard deviation of the measured risk premium are about 20 percent higher than the actual standard deviation (i.e., the estimates are near 1.2), while the values for the yen are about 3 percent higher. Column (4) also shows that the empirical distribution of these measures are quite disperse, again underscoring the difficulty of using standard inference techniques to interpret the behavior of the risk premium when "peso problems" are present. These results are consistent with recent evidence using options data in Bates (1994) and Baily and Kropywiansky's (1994) finding that the "peso problem" can partially, but not fully, explain the forward discount bias.

Overall, these results place some new perspective on the interpretation of the Fama regression results. The conventional interpretation suggests that the risk premium covaries strongly in the opposite direction of the forward premium. Our findings suggest that anticipations of

[26] Strictly speaking, θ^*/θ measures the ratios of lower bounds on the standard deviations of the risk premia because, as Equation (19) shows, the risk premia may also vary independently of the forward premium through the u_t terms.

exchange rate regime shifts biases the regression coefficients toward this finding. Also, the presence of this "peso problem" contributes to an upward bias in the measured variability of the risk premium.

4. Concluding Remarks

The behavior of foreign exchange returns has been difficult to reconcile with standard theoretical models. In this article, we have shown that some of this anomalous behavior can be explained by rational expectations about shifts in the exchange rate between appreciating and depreciating processes. In this case, predictable excess foreign exchange returns over periods with infrequent switches are affected by small sample serial correlation in forecast errors. The "peso problem" induced by anticipations of future switches in exchange rate regimes appear to explain an apparent permanent shock to the risk premium as well as an important fraction of the risk premium variability. This evidence suggests that long swings in the exchange rate can have important effects upon inferences about the risk premium.

Appendix

A.1 More general exchange rate solutions

In this appendix we describe how the exchange rate switching model presented in Section 1 can be solved with more general fundamentals processes. We also present an example in which the exchange rate jumps when there is a switch in regime even though there is no jump in the fundamentals process.

We consider solutions to the switching model in which the vector of fundamentals $\mathbf{Y}_t \equiv [y_t(1), y_t(0)]'$ follows a vector ARIMA process. In order to solve for the equilibrium vector of possible exchange rates, $\mathbf{S}_t \equiv [s_t(1), s_t(0)]'$, we first write the ARIMA process in companion form:

$$\mathbf{W}_t = \mathbf{B}_0 + \mathbf{B}\mathbf{W}_{t-1} + \mathbf{C}\epsilon_t \tag{A1}$$

where $\mathbf{Y}_t = \mathbf{H}\mathbf{W}_t$. Depending upon the order of the ARIMA process, the vector $\mathbf{W}_t$ includes current and lagged values for $\mathbf{Y}_t$ and lagged errors ϵ_t. $\mathbf{B}_0$ is a vector of constants, $\mathbf{B}$ a matrix of autogressive parameters, and $\mathbf{C}$ a matrix of constants. The matrix $\mathbf{H}$ selects the elements in $\mathbf{Y}_t$ from $\mathbf{W}_t$. In the case of the random walk specification examined in the text, $\mathbf{W}_t = \mathbf{Y}_t$, $\mathbf{B}_0 = \mathbf{A}$, and $\mathbf{B} = \mathbf{C} = \mathbf{I}_2$.

As an alternative example, suppose that the fundamentals in each regime followed an ARIMA(1, 1, 1):

$$\Delta y_t(1) \;=\; b_1 \Delta y_{t-1}(1) + \epsilon(1) + c_1 \epsilon_{t-1}(1)$$

$$\Delta y_t(0) = b_0 \Delta y_{t-1}(0) + \epsilon_t(0) + c_0 \epsilon_{t-1}(0)$$

The companion form for this model is

$$\mathbf{W}_t \equiv \begin{bmatrix} y_t(1) \\ y_t(0) \\ y_{t-1}(1) \\ y_{t-1}(0) \\ \epsilon_t(1) \\ \epsilon_t(0) \end{bmatrix}$$

$$= \begin{bmatrix} b_1+1 & 0 & -b_1 & 0 & c_1 & 0 \\ 0 & b_0+1 & 0 & -b_0 & 0 & c_0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1}(1) \\ y_{t-1}(0) \\ y_{t-2}(1) \\ y_{t-2}(0) \\ \epsilon_{t-1}(1) \\ \epsilon_{t-1}(0) \end{bmatrix}$$

$$+ \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_t(1) \\ \epsilon_t(0) \end{bmatrix} \tag{A2}$$

where $\mathbf{Y}_t = [\mathbf{I}_2, 0]\mathbf{W}_t$.

To solve for the equilibrium vector of possible exchange rates, we posit the solution,

$$\mathbf{S}_t = \mathbf{\Pi}_0 + \mathbf{\Pi}_1 \mathbf{W}_t \tag{A3}$$

where $\mathbf{\Pi}_0$ is a 2×1 vector and $\mathbf{\Pi}_1$ is a matrix of undetermined coefficients. Using this expression to substitute for $\mathbf{S}_t$ and $E_t\mathbf{S}_{t+1}$ in Equation (2) in the text, we obtain

$$\Pi_0 + \mathbf{\Pi}_1 \mathbf{W}_t = \phi \mathbf{\Lambda}[\mathbf{\Pi}_0 + \mathbf{\Pi}_1 E_t \mathbf{W}_{t+1}] + \mathbf{Y}_t.$$

Substituting for $E_t\mathbf{W}_{t+1}$ with Equation (A1), this expression can be rewritten as

$$\mathbf{\Pi}_0 + \mathbf{\Pi}_1 \mathbf{W}_t = \phi \mathbf{\Lambda}[\mathbf{\Pi}_0 + \mathbf{\Pi}_1(\mathbf{B}_0 + \mathbf{B}_1 \mathbf{W}_t)] + \mathbf{H}\mathbf{W}_t. \tag{A4}$$

Since Equation (A4) holds for all values of $\mathbf{W}_t$, $\mathbf{\Pi}_0$ and $\mathbf{\Pi}_1$ must satisfy

$$\mathbf{\Pi}_0 = \phi \mathbf{\Lambda}[\mathbf{\Pi}_0 + \mathbf{\Pi}_1 \mathbf{B}_0] \text{ and } \mathbf{\Pi}_1 = \phi \mathbf{\Lambda} \mathbf{\Pi}_1 \mathbf{B}_1 + \mathbf{H}.$$

Explicit solutions for the unknown elements of Π_0 and Π_1 are easily found from these equations:

$$\mathrm{Vec}(\mathbf{\Pi}_1) = (\mathbf{I} - \mathbf{B}_1' \otimes \phi\mathbf{\Lambda})^{-1}\,\mathrm{Vec}(\mathbf{H})$$

$$\mathbf{\Pi}_0 = (\mathbf{B}_0' \otimes [(\mathbf{I} - \phi\mathbf{\Lambda})^{-1}\phi\mathbf{\Lambda}])(\mathbf{I} - \mathbf{B}_1' \otimes \phi\mathbf{\Lambda})^{-1}\,\mathrm{Vec}(\mathbf{H}) \tag{A5}$$

[To derive these expressions we used the fact that for any conformable matrices, $\mathbf{M}$, $\mathbf{N}$, and $\mathbf{O}$, $\mathrm{Vec}(\mathbf{MNO}) = (\mathbf{O}' \otimes \mathbf{M})\,\mathrm{Vec}(\mathbf{N})$.]

The more general solution for $\mathbf{S}_t$ in Equations (A3) and (A5) shares features with the solution presented in the text. As there, the dynamics of exchange rates within a regime mirror the dynamics of fundamentals, in this case given by the vector $\mathbf{W}_t$. Furthermore, insofar as the elements in the rows of $\mathbf{B}_0$ and $\mathbf{B}_1$ differ from each other, our solution shows that the rows of $\mathbf{\Pi}_0$ and $\mathbf{\Pi}_1$ will also differ. This means that changes in the regime induced by switches in the fundamentals process will generally be accompanied by jumps in the exchange rate.

To this point, we have only considered examples where a switch in the fundamentals process is accompanied by a jump in y_t. To emphasize that the exchange rate will jump when the regime switches even when there is no jump in y_t, suppose fundamentals followed the process

$$y_t = \mu(z_t) + y_{t-1} + \epsilon_t. \tag{A6}$$

Here only the drift in the random walk varies when z_t changes from 1 to 0. There is no accompanying jump in y_t. Engel and Hamilton (1990) propose a similar process for the nominal exchange rate.

To solve for the exchange rate, we first rewrite Equation (A6) in vector form

$$\mathbf{Y}_t \equiv \begin{bmatrix} y_t(1,1) \\ y_t(0,1) \\ y_t(1,0) \\ y_t(0,0) \end{bmatrix} \tag{A6a}$$

$$= \begin{bmatrix} \mu(1) \\ \mu(0) \\ \mu(1) \\ \mu(0) \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{t-1}(1,1) \\ y_{t-1}(0,1) \\ y_{t-1}(1,0) \\ y_{t-2}(0,0) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \epsilon_t,$$

where $y_t(i, j)$ represents the value for y_t when $z_t = i$ and $z_{t-1} = j$. Next, we rewrite Equation (2) to conform with Equation (A7). For this purpose, we redefine S_t as $[s_t(1,1)s_t(0,1)s_t(1,0)s_t(0,0)]'$ and the matrix of transition probabilities

$$\mathbf{\Lambda} \equiv \begin{bmatrix} \lambda_1 & 1-\lambda_1 & 0 & 0 \\ 0 & 0 & 1-\lambda_0 & \lambda_0 \\ \lambda_1 & 1-\lambda_1 & 0 & 0 \\ 0 & 0 & 1-\lambda_0 & \lambda_0 \end{bmatrix}.$$

We can solve for the equilibrium $\mathbf{S}_t$ with Equation (A7) and the revised version of Equation (2) following the steps described above. The solutions for $\mathbf{\Pi}_0$ and $\mathbf{\Pi}_1$ are again given by Equation (A5) with $\mathbf{H} = \mathbf{I}_4$. Simplifying these expressions and substituting the results into Equation (A3) gives

$$\begin{aligned} \mathbf{s}_t(1) &= \{[\lambda_1(1-\phi) + \phi(1-\lambda_0)]\mu(1) + (1-\lambda_1)\mu(0)\}\phi/\psi(1-\phi) \\ &\quad + [1/(1-\phi)]y_t(1) \end{aligned}$$

and

$$\begin{aligned} \mathbf{s}_t(0) &= \{[\lambda_0(1-\phi) + \phi(1-\lambda_1)]\mu(0) + (1-\lambda_0)\mu(1)\}\phi/\psi(1-\phi) \\ &\quad + [1/(1-\phi)]y_t(0), \end{aligned}$$

where $\psi = (1-\phi)[1+\phi-\phi(\lambda_1+\lambda_0)]$, $s_t(1) = s_t(1,1) = s_t(1,0)$ and $s_t(0) = s_t(0,1) = s_t(0,0)$. Here we see that the constants in both equations differ from each other so long as $\mu(1) \neq \mu(0)$. There will therefore be a jump in the exchange rate when the trend in the fundamentals processes switches.

A.2 Estimation methods for the cointegrating regressions

In this appendix, we describe Stock and Watson's (1989) method to adjust for simultaneous equation bias in the cointegrating regression in Equation (14). For a more detailed and thorough discussion, see Stock and Watson (1989).

For notational simplicity, note that Equation (14) may be written as

$$s_{t+k} = a_1 f_t + u_{1t} \tag{A7}$$

$$\Delta f_t = u_{2t} \tag{A8}$$

where u_{1t} and u_{2t} are stationary, and the constant term is omitted for simplicity. We are interested in testing $a_1 = 1$. Since the forward rate f_t contains a stochastic trend component that is the cumulation of a component of u_{1t}, $\text{Cov}(f_t, u_{1t}) \neq 0$. Therefore, the estimate of a_1 will be biased in any finite sample. Inferences based on the estimates of a_1 must take account of this bias. Note that this bias arises even though the estimate of a_1 is consistent. We use two versions of the Stock and Watson procedure to adjust for the bias.

A.2.1 Single equation estimation and testing. To test whether $a_1 = 1$ equation by equation, we rewrite Equation (A7) as

$$s_{t+k} = a_0 + a_1 f_t + \beta(L)\Delta f_t + u_{1t} \tag{A9}$$

where $\beta(L)$ is a polynomial in the lag operator, L, that is, $\beta(L) = (\beta_n L^n + \beta_{n-1}L^{n-1} + \cdots + \beta_1 L + 1 + \beta_{-1}L^{-1} + \cdots + \beta_{-n+1}L^{-n+1} + \beta_{-n}L^{-n})$. The idea to rewriting Equation (A7) in this form is to include as many leads and lags of Δf_t on the righthand side of the equation to make u_{1t} independent of f_t. This implies that the asymptotic distribution of the OLS estimator of a_1 is normal. Intuitively, including the leads and lags of Δf_t on the right-hand side "soaks up" the simultaneous equation bias. Note that since u_{1t} will be serially correlated in general, the Wald test of $a_1 = 1$ from Equation (A9) should also use the Newey and West (1987) estimator for the covariance matrix. The results reported in Table 2 are based on estimates of a_1 from Equation (A9) that include three leads and lags of Δf_t. The upper p-values are calculated from Wald statistics that allow for conditional heteroskedasticity in u_{1t}. The lower the p-values are calculated from the Wald statistics that allow for conditional heteroskedasticity and MA(2) serial correlation in u_{1t}. The results reported in the table are not sensitive to these choices.

A.2.2 Joint equation estimation and testing For the joint equation estimation, we rewrite equations Equations (A7) and (A8) in their stacked equation form:

$$\underline{\mathbf{s}}_{t+k} = \underline{\mathbf{a}}_1 \underline{\mathbf{f}}_t + \underline{\mathbf{u}}_{1t} \tag{A7a}$$

$$\Delta\underline{\mathbf{f}}_t = \underline{\mathbf{u}}_{2t} \tag{A8a}$$

where $\underline{\mathbf{s}}_t$ and $\underline{\mathbf{f}}_t$ are 3×1 vectors of spot and forward rates, $\underline{\mathbf{a}}_1$ is a 3×3 matrix, and $\underline{\mathbf{u}}_{it}$ are 3×1 vectors of errors. Writing the errors in terms of the primary innovations, we have the representation

$$\underline{\mathbf{u}}_{1t} = \mathbf{C}_{12}(\mathbf{L})\underline{\epsilon}_{2t} + \mathbf{C}_{11}(\mathbf{L})\underline{\epsilon}_{1t}$$
$$\underline{\mathbf{u}}_{2t} = \mathbf{C}_{22}(\mathbf{L})\underline{\epsilon}_{2t},$$

where the vectors of innovations $\underline{\epsilon}_{1t}, \underline{\epsilon}_{2t}$ are independent, $\mathbf{C}_{11}(\mathbf{L})$ and $\mathbf{C}_{22}(\mathbf{L})$ are one sided and $\mathbf{C}_{12}(\mathbf{L})$ is a two-sided polynomial matrix in the lag operator. For more details, see Stock and Watson (1989).

Using this representation, we can write Equations (A7a) and (A8a) as

$$\underline{\mathbf{s}}_{t+k} = \mathbf{a}_1 \underline{\mathbf{f}}_t + \mathbf{d}(\mathbf{L})\Delta \underline{f}_t + \underline{e}_t \tag{A10}$$

$$\Delta\underline{\mathbf{f}}_t = \mathbf{C}_{22}(\mathbf{L})\underline{\epsilon}_{2t} \tag{A11}$$

where $\mathbf{d}(\mathbf{L}) \equiv \mathbf{C}_{12}(\mathbf{L})\mathbf{C}_{22}(\mathbf{L})^{-1}$ and $\underline{\mathbf{e}}_t \equiv \mathbf{C}_{11}(\mathbf{L})\underline{\epsilon}_{1t}$.

To estimate Equation (A10), we stack the three individual equations:

$$\begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 & 0 & 0 \\ 0 & \mathbf{x}_2 & 0 \\ 0 & 0 & \mathbf{x}_3 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix}$$

$$\mathbf{Y} = \mathbf{X}\delta + \mathbf{e}$$

where $\mathbf{s}_i = (s_{i1+k-n}, \ldots, s_{iT-n})'$, $\mathbf{e}_i = (e_{i1}, \ldots, e_{iT-k-n})'$, and the tth row of x_i is $(f_{it}, \Delta f_{it-n}, \ldots, \Delta f_{it+n})$. Then, defining $\boldsymbol{\Omega}$ as a consistent estimate of the long-run covariance matrix of $\underline{\mathbf{e}}_t$, the estimator of δ is given by

$$\hat{\delta} = (\mathbf{X}'(\boldsymbol{\Omega} \otimes \mathbf{I}_T)\mathbf{X})^{-1}\mathbf{X}'(\Omega \otimes \mathbf{I}_T)\mathbf{Y}. \tag{A12}$$

The asymptotic variance of this estimator is given by

$$\text{Var}(\hat{\delta}) = (\mathbf{M}'\mathbf{X})^{-1}\mathbf{M}'\mathbf{V}\mathbf{M}(\mathbf{X}'\mathbf{M})^{-1}, \tag{A13}$$

where $\mathbf{M} \equiv (\boldsymbol{\Omega}^{-1} \otimes \mathbf{I}_T)\mathbf{X}$ and $\mathbf{V} \equiv \mathbf{E}(\underline{e}\ \underline{e}')$. Thus, $\mathbf{M}'\mathbf{V}\mathbf{M}$ can be estimated with the Newey and West (1987) estimator allowing for both serial correlation and heteroskedasticity. Notice that our estimator is essentially an instrumental variable estimator using $\mathbf{M}$ as instruments. The properties of this estimator are discussed in Hansen and Phillips (1990).

The p-values for joint tests for $a_1 = 1$ reported in Table 2 are calculated from the Wald test of $a_1 = 1$ in all three equations. These tests use the elements of $\hat{\delta}$, and so account for the finite sample bias in the OLS estimates of a_1. The estimates are calculated allowing for three leads and lags of Δf_t in each of the equations. The upper p-values are calculated from a Wald statistic that allows for conditional heteroskedasticity in the construction of $\text{Var}(\hat{\delta})$. The lower p-values are calculated from a Wald statistic that allows for conditional heteroskedasticity and MA(2) serial correlation in the construction of $\text{Var}(\hat{\delta})$. Again, the results reported in the table are not sensitive to these choices.

A.2.3 Description of the Monte Carlo experiments in Table 3

The Monte Carlo p-values reported in columns (3) and (4) of Table 3 we calculated as follows:

(i) We estimated the Stock-Watson version of Equation (14) [see Equation (A9)] with monthly data saving the residuals and the coefficient estimates. As discussed in Stock and Watson (1989), the residuals from this equation are independent of the entire sequence of the right-hand side variable and so can be treated as strictly exogenous. In cases where $k = 3$, the residuals contain overlapping forecast errors and so it is necessary to remove the induced serial correlation before sampling. For this purpose we estimated AR(2) models for the residuals and used the estimated errors in the sampling procedure described in (ii) below.

(ii) In the experiments that assumed conditional homoskedasticity, we then drew randomly from the distribution of residuals in the data. In the experiments that allowed for conditional heteroskedasticity we first estimated a third-order ARCH process with the residuals. We then scaled the residuals by the predictions of the ARCH model and drew randomly from the scaled distribution. Finally, we rescaled the distribution of residuals using predictions from the ARCH process.

(iii) We took the coefficient estimates for a_0 (and the Stock-Watson coefficients), set $a_1 = 1$, and generated a time series for s_{t+k} equal to the number of observations in the sample. In cases where $k = 3$, we used the estimates of our AR(2) model to generate a new set of serially correlated residuals (due to the forecast overlap) from which we then generated a new time series for s_{t+k}.

(iv) Using the generated time series for s_{t+k} on the left-hand side, we estimated the cointegrating regressions in Equation (A9) and saved the Wald test statistic for H_0: $a_1 = 1$. These statistics use estimates of the covariance matrix that ignore the presence of heteroskedasticity and serial correlation in the residuals.

(v) We then repeated steps (ii) through (iv) 1000 times to form empirical distributions for the Wald statistics under the null hypothesis of $a_1 = 1$.

(vi) The p-values reported in columns (3) and (4) are calculated by comparing the empirical distributions of the Wald statistics against Wald statistics for H_0: $a_1 = 1$ from the actual data, ignoring the presence of heteroskedasticity and serial correlation in the residuals.

References

Backus, D., A. Gregory, and S. Zin, 1989, "Risk Premiums in the Term Structure: Evidence from Artificial Economies," *Journal of Monetary Economics*, 24, 371–399.

Baillie, R. T., and T. Bollerslev, 1989, "Common Stochastic Trends in a System of Exchange Rates," *Journal of Finance*, 44, 167–181.

Baily, W., and L. Kropywiansky, 1994, "Jitters, Jump Fears, and the Peso Problem in Foreign Exchange," working paper, Massachusetts Institute of Technology.

Bates, D. S., 1994, "Dollar Jump Fears 1984–1992: Distributional Abnormalities Implicit in Foreign Exchange Futures Options," working paper, Wharton School, University of Pennsylvania; forthcoming in *Journal of International Money and Finance*.

Bekaert, G., and R. J. Hodrick, 1992, "Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets," *Journal of Finance*, 47, 467–509.

Bekaert, G., and R. J. Hodrick, 1993, "On Biases in the Measurement of Foreign Exchange Risk Premiums," *Journal of International Money and Finance*, 12, 115–138.

Campbell, J. Y., and P. Perron, 1991, "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots," in O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual*, MIT Press, Cambridge, Mass.

Diebold, F. X., and M. Nerlove, 1990, "Unit Roots in Economic Time Series: A Selective Survey," *Advances in Econometrics*, 8, 3–69.

Engel, C., and J. D. Hamilton, 1990, "Long Swings in the Dollar: Are They in the Data and Do the Markets Know It?" *American Economic Review*, 80, 689–713.

Fama, E., 1984, "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 14, 319–338.

Froot, K., and J. A. Frankel, 1989, "Forward Discount Bias: Is It An Exchange Rate Risk Premium?" *The Quarterly Journal of Economics*, 104, 139–161.

Grossman, S., and R. Shiller, 1981, "The Determinants of the Variability of Stock Market Prices," *American Economic Review*, 71, 222–227.

Hakkio, C. S., and M. Rush, 1989, "Market Efficiency and Cointegration: An Application to the Sterling and Deutsche Mark Exchange Markets," *Journal of International Money and Finance*, 8, 75–88.

Hamilton, J., 1989, "A New Approach to the Econometric Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.

Hansen, B., and P. C. B. Phillips, 1990, "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies*, 57, 99–126.

Hansen, L. P., and R. Jagannathan, 1991, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.

Hodrick, R. J., 1987, *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*, Harwood Academic Publishers, Chur, Switzerland.

Hodrick, R. J., and S. Srivastava, 1986, "The Covariation of Risk Premiums and Expected Future Spot Exchange Rates," *Journal of International Money and Finance*, 5, S5–S21.

Johansen, S., 1988, "Statistical Analysis of Cointegrating Factors," *Journal of Economic Dynamics and Control*, 12, 231–254.

Kaminsky, G. L., and K. K. Lewis, 1993, "Does Foreign Exchange Intervention Signal Future Monetary Policy?" Working Paper 4298, National Bureau of Economic Research, Cambridge, Mass.

Lewis, K. K., 1992, "Peso Problems," in J. Eatwell, M. Milgate, and P. Newman (eds.), *The New Palgrave Dictionary of Money and Finance*, Macmillan Press, London.

McCallum, B. T., 1992, "A Reconsideration of the Uncovered Interest Parity Relationship," Working Paper 4113, National Bureau of Economic Research, Cambridge, Mass.

Mark, N., 1995, "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability and Overshooting," *American Economic Review*, 85, 201–218.

Meese, R. A., and K. Rogoff, 1983, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics*, 21, 3–24.

Meese, R. A., and K. J. Singleton, 1982, "On Unit Roots and the Empirical Modeling of Exchange Rates," *Journal of Finance*, 37, 1029–1054.

Newey, W., and K. West, 1987, "A Simple Positive Semi-Definite Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–706.

Rogoff, K. S., 1980, "Essays on Expectations and Exchange Rate Volatility," Ph.D. dissertation, Massachusetts Institute of Technology.

Stambaugh, R. F., 1988, "The Information in the Forward Rates: Implications for Models of the Term Structure," *Journal of Financial Economics*, 21, 41–69.

Stock, J., 1987, "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica*, 55, 1035–1056.

Stock, James, and Mark Watson, 1989, "A Simple MLE of Cointegrating Vectors in Higher Order Integrated Systems," Working Paper 83, National Bureau of Economic Research, Cambridge, Mass.