Time-Variation and Structural Change in the Forward Discount: Implications for the Forward Rate Unbiasedness Hypothesis

by

Georgios Sakoulis and Eric Zivot*
Department of Economics, University of Washington

January, 2000
This Draft:February, 2001

Abstract

It is a well accepted empirical result that forward exchange rate unbiasedness is rejected in tests using the "differences regression" of the change in the logarithm of the spot exchange rate on the forward discount. The result is referred to in the International Finance literature as the forward discount puzzle. Competing explanations of the negative bias of the forward discount coefficient include the possibilities of a time-varying risk premium or the existence of "peso problems." We offer an alternative explanation for this anomaly. One of the stylized facts about the forward discount is that it is highly persistent. We model the forward discount as an AR(1) process and argue that its persistence is exaggerated due to the presence of structural breaks. We document the temporal variation in persistence, using a time-varying parameter specification for the AR(1) model, with Markov-switching disturbances. We also show, using a stochastic multiple break model, suggested recently by Bai and Perron (1998), that for the G-7 countries, with the exception of Japan, the forward discount persistence is substantially less, if one allows for multiple structural breaks in the mean of the process. These breaks could be identified as monetary shocks to the central bank's reaction function, as discussed in Eichenbaum and Evans (1995). Using Monte Carlo simulations we show that if we do not account for structural breaks which are present in the forward discount process, the forward discount coefficient in the "differences regression" is severely biased downward, away from its true value of 1.

* The authors would like to thank Charles Engel and the participants of the macroeconomics seminar at the New York Federal Reserve Bank for helpful comments and suggestions and Jushan Bai for generously providing the GAUSS code to estimate the multiple break models. The usual disclaimer applies.

1. Introduction

A recurring theme in the international finance literature is the investigation of forward market efficiency. Starting with Bilson (1981) and Fama (1984), the regression that most people have looked at when they test the forward rate unbiasedness hypothesis (FRUH) is the "differences regression":

$$\Delta s_{t+k} = \alpha + \beta(f_{t,k} - s_t) + \varepsilon_{t+k} \qquad (1)$$

where, s_t is the log of the spot exchange rate, $f_{t,k}$ is the log of the k-period forward exchange rate at time t, $f_{t,k}$ - s_t is the forward discount, which, under covered interest parity, equals the interest rate differential between two countries, ε_{t+k} is the regression error and Δ is the difference operator. FRUH stipulates that under the joint hypothesis of risk neutrality and rational expectations, the current forward rate is an unbiased predictor of the future spot rate; that is, under FRUH, $\alpha = 0$, $\beta = 1$, and $E_t(\varepsilon_{t+1}) = 0$, so that a domestic investor who invests in a foreign market cannot gain "excess returns" from foreign currency between times t and $t+k$. Nevertheless, the typical finding in the literature is that FRUH is compellingly rejected; not only the forward rate is not an unbiased predictor of the future spot rate, but typical estimates of β in (1) are significantly negative[1]. This anomalous empirical finding is so well documented that it is referred to as the forward discount anomaly.

A large number of researchers focused on the puzzling estimates of β from (1) and tried to explain what could be causing them to deviate from the theoretical value of 1. Two of the competing explanations that have been put forth are discussed in Engel (1996) and Lewis (1995). The first of these explanations was pioneered by Fama (1984). Fama suggests that the anomaly is due to an omitted variables problem. He shows that if risk neutrality fails, then negative estimates of β are consistent with a time-varying exchange rate risk premium rp_t, which is correlated with the forward discount so that equation (1) is mispecified. In the case of a negative estimate of β, the covariance of the risk premium with the expected change in the spot exchange rate must be negative and the variance of the risk premium must be greater than the variance of the expected change. Nevertheless,

[1] Froot (1990) reports an average value for β of –0.88 over 75 published studies.

empirical models of the risk premium thus far, have been unable to adequately address the anomaly. Engel (1996) concludes:

"...First, empirical tests routinely reject the null hypothesis that the forward rate is a conditionally unbiased predictor of future spot rates. Second, models of the risk premium have been unsuccessful at explaining the magnitude of this failure of unbiasedness..."

The second explanation is based upon the idea of systematic forecast errors being made by the foreign exchange market participants. Frankel and Froot (1987) and Froot and Frankel (1989) show, using various measures of expectations based on survey data, that excess returns are mainly due to systematic forecast errors and not risk premia. At any given time, some of the market participants do not use all available information efficiently, or in other words, they form expectations in an irrational manner. Their behavior generates additional risk in asset prices that has a two-fold effect: First, irrational agents earn higher expected returns because they bear higher risk. Secondly, rational agents, being more risk-averse, are not necessarily able to drive the first group out of the market by aggressively betting against them. In terms of equation (1), such behavior could bias the estimate of β, if the forecast error is negatively correlated with the forward discount.

Lewis (1989) and Lewis and Evans (1995) suggest a different reason why the forecast error could be negatively correlated with the forward discount. They attribute systematic errors in the presence of "learning" or "peso" problems. Briefly, the economy undergoes infrequent regime changes, due to shocks hitting the real, as well as the nominal side of the economy. In the case of "peso" problems, economic agents revise their future expectations in a rational manner, while trying to incorporate in their information set the probability of being in a different regime next period. If the anticipated regime is not realized within the sample examined, serial correlation in the forecast errors could be introduced in small samples. Although these models can partially explain the puzzle, Lewis (1995) admits that a substantial amount of variability in excess returns remains unexplained.

Alternatively, some authors have investigated statistical reasons for the anomaly focussing on the time series properties of exchange rates and the forward discount in the differences regression (1). It is well established that nominal exchange rates behave like $I(1)$ processes so that Δs_{t+k} is $I(0)$. However, one of the stylized facts about the forward discount is that it is highly persistent. The high persistence of the forward discount means that the differences regression is potentially "unbalanced"; that is, the amount of persistence in the dependent variable is much less than the amount in the regressor. It is well known that in unbalanced regressions the coefficient on the highly persistent regressor is potentially downward biased[2].

At one extreme, Crowder (1994) and Lewis and Evans (1995) have gone as far as to conclude that the forward discount, appears to have a unit root component. This would make (1) are regression of an $I(0)$ variable, Δs_{t+k}, on an $I(1)$ variable, $f_t - s_t$, and so the least squares estimate of β converges in probability to zero. A unit root in the forward discount, however, is unappealing for the following reason. Consider the following decomposition, in the presence of a time-varying risk premium $rp_{t+1} = f_t - E_t[s_{t+1}]$, originally due to Fama (1984):

$$f_t = s_{t+1} + rp_{t+1} + \eta_{t+1} \tag{2}$$

This equation relates today's forward rate, f_t, to next period's spot rate, s_{t+1}, a risk premium, rp_{t+1}, and a rational expectations forecast error term, η_{t+1}. We can rewrite the future spot rate as:

$$s_{t+1} = s_t + \Delta s_{t+1} \tag{3}$$

If we substitute equation (3) into (2) and rearrange, we get

$$f_t - s_t = \Delta s_{t+1} + rp_{t+1} + \eta_{t+1} \tag{4}$$

Equation (4) shows that the forward discount consists of three components: the change in the spot exchange rate; the risk premium; and a rational forecast error. Since both the change in the spot exchange rate and the forecast error are $I(0)$, the supposed unit root

[2] See Kim and Nelson (1993?) – Journal of Finance, Stambaugh (19xx?) – predictive regressions paper.

component of the forward discount is identified as the risk premium. A unit root risk in the risk premia would be very hard to rationalize since standard models of time-varying risk premia imply them to be $I(0)$ since they depend on the time series properties of other I(0) variables, such as the growth rate in consumption.[3]

More recently, models with long memory or fractional integration in the forward discount have been put forth in an attempt to address the possible connection between the forward discount bias and its persistence. These models include Baillie and Bollerslev (1994, 2000), as well as Maynard and Phillips (1998) among others. Baillie and Bollerslev (2000) model the forward discount as a mean-reverting fractionally integrated, $I(d)$, process , where d is the order of fractional differencing, such that the autocorrelation function decays very slowly. They show, using Monte Carlo simulations that in this case $\hat{\beta}$ in (1) will converge to its true value of unity, very slowly. Maynard and Phillips (1998) develop an asymptotic theory to provide theoretical justification for these results. Together, these results in these papers suggest that the forward discount anomaly is just a statistical artifact. It takes place exactly because the autocorrelations in the forward discount are very persistent and the sample size fairly small. Even if the forward discount is a biased predictor of the future spot rate, it is not possible to statistically reach a definite conclusion, given the typical size of exchange rate samples.

The main criticism against using models of fractional integration is whether fractionally integrated processes occur in the actual economy. Granger (1999) argues that such processes are at very low spectral frequencies where information accumulates very slowly. As a result long time series are required to provide estimates of d, the order of fractional integration, which are significantly different from 0 or 1. Typical macroeconomic series are not long enough to provide us with such evidence. Granger advocates the use of non-linear models as plausible alternatives to fractional integration. For example, using simulated data, as well as daily absolute returns for the S&P 500 index, Granger shows that the stochastic break model developed by Bai (1997) can

Mention what is required for the bias to be highly negative.

[3] This is also the argument made in Evans and Lewis (1995). They cite Grossman and Shiller (1981), Backus, Gregory, and Zin (1989) as examples of studies of time-varying risk premia. For a complete discussion of theoretical models of foreign exchange risk premia see Engel (1995).

produce many of the "long memory" properties of the data. In the present context, Choi and Zivot (2001) provide evidence that the long-memory properties of the forward discount can be largely explained by multiple breaks in the mean of the forward discount. Moreover, Diebold and Inoue (1999) analytically show that stochastic regime-switching is observationally equivalent to long memory, even asymptotically, thus offering additional evidence of the empirical relevance of such models.

Our starting point, as in Baillie and Bollerslev (1998), is an investigation into the time series properties of the forward discount, f_t - s_t. We start with the prior that the forward discount does not have a unit root or long-memory and that its observed persistence is due to structural changes that take place in the economy during the time period of our sample. Starting with Perron (1989), it is well documented in the econometrics literature that structural breaks could induce $I(1)$ as well as $I(d)$ like behavior in observed time series. We hypothesize that the forward discount is subject to structural breaks, and perhaps there is more than one instance of structural breaks in the data. Such breaks could be arising from changes in monetary policy objectives of the central banks of different countries, discrete change in policy where new initiatives take form such as the Plaza Agreement, as well as exogenous shocks to the decision rule of the monetary authority. For example, Eichenbaum and Evans (1995) consider three different measurements of the latter type of shocks. Applying VAR techniques, they find that contractionary shocks to U.S. monetary policy result in persistent increases in U.S. interest rates and persistent decreases in the spread between foreign and U.S. interest rates. Cushman and Zha (1997), Kim and Rubini (1995), Clarida and Gertler (1997) reach similar conclusions when applying different monetary policy shock measures to the foreign policy maker's decision rule[4]. Eichenbaum and Evans attribute the source of these policy shocks to political factors, factors pertaining to the views of the members of the FOMC, as well as technical factors such as measurement error in the data available to the FOMC.

To illustrate the evidence for structural change in the forward discount, we utilize an AR(1) model with a time–varying autoregressive parameter and Markov-switching

[4] These results, as well as the general issues concerning monetary policy shocks, are discussed in great detail in Christiano, Eichenbaum, and Evans (1998) "Monetary Policy Shocks: What Have We Learned and to What End?"

variance. Using data from G-7 countries, we find that the forward discount appears to be highly persistent at the very beginning of the sample and again starting at the late 1980s, and not before. The timing of the most noticeable changes in the time-varying coefficient suggests the presence of multiple structural breaks. While this model captures the temporal instability of the forward discount it does not explain its source. Based on the idea that changes in the mean of a process can induce both parameter instability and persistence in the AR coefficient, we estimate a stochastic multiple break model, using the methodology developed recently by Bai and Perron (1998). Similar approaches have been used, for instance, by Wang and Zivot (2000) in a Bayesian framework, as well as by Garcia and Perron (1995) in a Markov-switching framework. Interestingly enough, we find that once we account for structural breaks in the mean of the forward discount, it is not as persistent, even after the late 1980s[e1].

Finally, we investigate, using Monte Carlo simulations, the implications of our finding of structural change in the forward discount for the forward discount puzzle. Assuming that the true generating process of the forward discount is given by our stochastic break model, we construct spot and forward rates based on our estimated parameters and test for unbiasedness using equation (1). We show that even when the true β coefficient in (1) is equal to 1, the least squares estimate of β is significantly biased downward. Therefore, although our modeling strategy of the forward discount is different than that of Baillie and Bollerslev, we arrive to a similar conclusion. The forward discount anomaly is not as bad as we think and it is, at least partly, due to the statistical properties of the data[e2].

The plan of the paper is as follows: In Section 2 we present some stylized facts of exchange rate data. In Section 3 we present the two alternative models design to capture structural change in the forward discount. In section 4 we discuss the empirical results. In section 5 we develop the Monte Carlo simulations based on the estimated parameters. We conclude in section 6.

2. Exchange Rate data

Let s_t denote the log of the spot exchange rate in month t and f_t denote the log of the forward exchange rate in the same month. We consider monthly data for which the maturity date of the forward contract is the same as the sampling interval, in order to avoid modeling complications arising from overlapping data, so $k = 1$ in (1). Our exchange rates are spot and forward rates which are obtained from Datastream. The data are end of month, average of bid and ask rates, for the German Mark, French Franc, Italian Lira, Canadian Dollar, British Pound, and Japanese Yen. Exchange rates are expressed as the home country price of the foreign currency, where the foreign currency is the US dollar. They span the period 1976:01-1999:01 except in the case of Japan where they span the period 1978:07-1999:01. All logs have been multiplied by 100, so that the final series of the forward discount and changes in the exchange rates, are expressed in percentage differences. Figure 1 plots the forward discount, f_t - s_t, for all the currencies. Notice that the forward discount is much more volatile at the beginning of the sample and especially so between 1980-83. Table 1 gives some summary statistics of the data. Spot and forward rates behave very similarly and exhibit random walk type behavior. The forward premiums are all highly autocorrelated. The variances of Δs_{t+1} and Δf_{t+1} are roughly ten times larger than the variance of the forward discount. Finally, with the exception of France and Italy, for all currencies, Δs_{t+1} and Δf_{t+1} are negatively correlated with f_t - s_t.

3. The Models

Godbout and von Norden (1995), Mark et al (1998), Mark and Wu (1998), and Zivot (2000) among others, show that the stylized facts of the monthly exchange rate data reported in the previous section can be captured by a simple cointegrated VAR(1) model for $y_t = (f_t, s_t)'$:

$$\Delta y_t = \mu + \Pi y_{t-1} + \varepsilon_t \tag{5}$$

where $\varepsilon_t \sim iid\ (0, \Sigma)$ and Σ has elements σ_{ij} $(i, j = f, s)$. Under the assumption that spot and forward rates are $I(1)$ and cointegrated, Π has rank 1 and there exist 2×1 vectors β and γ such that $\Pi = \beta\gamma'$. Using the normalization $\gamma = (1, -\gamma_s)'$, (5) becomes a vector error correction model (VECM) with equations:

$$\Delta f_t = \mu_f + \beta_f (f_{t-1} - \gamma_s s_{t-1}) + \varepsilon_{ft}, \tag{6a}$$

$$\Delta s_t = \mu_s + \beta_s(f_{t-1} - \gamma_s s_{t-1}) + \varepsilon_{st,} \tag{6b}$$

Since spot and forward rates usually do not exhibit a systematic tendency to drift up or down it may be more appropriate to restrict the intercepts in (6) to the error correction term, so that $\mu_f = -\beta_f\mu_c$ and $\mu_s = -\beta_s\mu_c$. Under this restriction s_t and f_t are $I(1)$ without drift and the cointegrating residual, $f_t - \gamma_s s_t$, is allowed to have a non-zero mean μ_c.

With the intercepts in (6) restricted to the error correction term, the VECM can be solved to give a simple AR(1) model for the co-integrating residual $\gamma' y_t - \mu_c = f_t - \gamma_s s_t - \mu_c$.

$$f_t - \gamma_s s_t - \mu_c = \phi(f_{t-1} - \gamma_s s_{t-1} - \mu_c) + \eta_t \tag{7}$$

where $\phi = 1 + \gamma'\beta = 1 + (\beta_f - \gamma_s\beta_s)$ and $\eta_t = \gamma'\varepsilon_t = \varepsilon_{ft} - \gamma_s\varepsilon_{st}$. Notice that according to (7) if $\gamma_s = 1$, then the forward discount is $I(0)$ and follows an AR(1) process and the VECM (6) becomes[e3]

$$\Delta f_t = \beta_f(f_{t-1} - s_{t-1} - \mu_c) + \varepsilon_{ft,} \tag{8a}$$

$$\Delta s_t = \beta_s(f_{t-1} - s_{t-1} - \mu_c) + \varepsilon_{st,} \tag{8b}$$

(8b) is exactly equation (1) which is used to test the FRUH, where $\beta = \beta_s$.

Using similar data as that used in this paper, Zivot (2000) estimates γ_s using Stock and Watson's (1993) dynamic OLS (DOLS) and dynamic GLS (DGLS) lead-lag estimator, and Johansen's (1995) reduced rank MLE. The hypothesis that $\gamma_s = 1$ cannot be rejected using the appropriate asymptotic *t*-tests. Zivot also uses various tests of the null of no co-integration between spot and forward rates, imposing the cointegrating vector (1,-1)′, and finds mixed evidence that $f_t - s_t$ is I(0[e4]).

We choose to model the forward discount as the AR(1) process that is implied from the VECM (8**.** Since our purpose is to model and hopefully capture structural change effects, we look at two different kinds of models. The first model is a hybrid of a time-varying parameter AR(1) model and a Markov-switching model. Specifically, we allow the autoregressive coefficient to be time varying and the error variance to be Markov-switching. The idea here is that structural change is better captured in a continuous framework for some of the parameters of the model, while discrete changes are more appropriate for others. Plots of the estimated time-varying coefficients provide us with information on how structural change takes place continuously over time. We model the variance as a Markov-switching process in order to capture the stylized fact of high and low volatility regimes in the forward discount process.

The second model is one which allows multiple stochastic structural breaks in some of the parameters. We use this model to capture breaks in the mean of the level of the forward discount that could potentially have two effects. First, in a regular regression where breaks are not accounted for, they could bias the autoregressive coefficient upward. Second, a structural break in the mean has a more natural interpretation as the direct effect of an economic shock to the level of a process that could also explain the temporal variation in the time-varying parameter model.

3.1 The Time-varying Coefficient with Markov-switching Variance Model

One way to model time variation in a regression coefficient is to treat it as an unobserved component which evolves according to a transition equation. We start with an AR(1) process for the forward discount as in (7), but assume that the autoregressive coefficient ϕ is time-varying

$$f_t - s_t - \mu_c = \phi_t (f_{t-1} - s_{t-1} - \mu_c) + \eta_t \tag{9}$$

where μ_c is the mean of the process, which for the time being is assumed to be constant over time and ϕ_t is the time-varying coefficient. We assume that ϕ_t follows a random walk process

$$\phi_t = \phi_{t-1} + v_t \tag{10}$$

where v_t is an *iid* $(0, \sigma_v^2)$ process, independent of η_t.

Engle and Watson (1987) suggest that for most economic series a unit root specification for the evolution of the unobserved component is appropriate. Garbade (1977) shows using Monte Carlo simulations that a random walk specification is a parsimonious way of modeling the transition equation of regression coefficients as long as the true parameters follows a persistent AR(1) processes.

To compute the high and low volatility states of the forward discount, we specify a two-state Markov-switching representation for η_t:

$$\eta_t \sim iid\ N(0, \sigma^2_{\eta,S_t})$$
$$\sigma^2_{\eta,S_t} = \sigma^2_{\eta,0}(1-S_t) + \sigma^2_{\eta,1} S_t$$
$$\sigma^2_{\eta,1} > \sigma^2_{\eta,0}$$

where the binary state variable S_t describing the high and low volatility states follows a first order Markov process with transition probabilities given by:

$$\Pr[S_t = 1 \mid S_{t-1} = 1] = p \qquad \text{and} \qquad \Pr[S_t = 0 \mid S_{t-1} = 0] = q$$

The estimates of the hyper-parameters can be obtained via maximum likelihood estimation based on the prediction error decomposition of the log likelihood, as described in Kim and Nelson (1999).[5]

3.2 Partial Structural Break Model

Although the time-varying parameter model appears to be adequate in capturing the essential time-series properties of the forward discount, it only provides us with information on how the forward discount behavior has changed over time. More specifically, it tells us how the persistence has varied over time. We are interested in finding out why the forward discount behavior has changed. We hypothesize that structural breaks in the mean are mainly responsible for inflating the estimated persistence. If our prior has some merit, accounting for such structural breaks should take away what we hope to be considerable upward bias from the autoregressive coefficient. Since it does not seem that restricting the number of different regimes is appropriate, we turn to the class of multiple break models considered by Bai and Perron (1998), BP hereafter.

BP consider multiple structural changes in a linear regression model, which is estimated by minimizing the sum of squared residuals. They consider models of both pure structural change, where all the regression coefficients are subject to change, and partial structural change models, where only some of the coefficients are subject to

[5] For details about the filter and parameter estimation, see Appendix 1.A.

change. Their models allow heterogeneity in the regression errors but they do not provide methods for parametrically estimating this heterogeneity. We use the partial structural change model, since we want to address potential the upward bias to the autoregressive coefficient of the forward discount. This model given by:

$$f_t - s_t = c_j + \phi(f_{t-1} - s_{t-1}) + u_t, \quad t = T_{j-1}+1,...,T_j \tag{11}$$

for $j = 1,...,m+1$, $T_0 = 0$ and $T_{m+1} = T$. The process is subject to m breaks (m+1 regimes), c_j is the constant of the regression[6], subject to structural change, ϕ is the autoregressive coefficient of the lagged forward discount, which is not subject to structural change and is estimated using the entire sample. $(T_1,...,T_m)$ are the unknown break points.

Using BP's technique we are able to estimate the regression coefficients along with the break points, given T observations of the forward discount. Briefly the method of estimation is as follows[7]. In the case of a pure structural break model, i.e., both c and ϕ change, for each possible m-partition $(T_{1,...,}T_{\mathrm{m}})$ the least squares estimators of c and ϕ are obtained by minimizing the sum of square residuals. Then the estimated break points are the ones for which

$$\left(\hat{T}_1,\ldots,\hat{T}_m\right) = \arg\min_{T_1,\ldots,T_m} S_T(T_1,\ldots,T_m) \tag{12}$$

where $S_T(T_1,...,T_{\mathrm{m}})$ denotes the sum of squared residuals. Since the minimization takes place over all possible partitions, the break-point estimators are global minimizers. BP use a very efficient algorithm for estimating the break points which is based on dynamic programming techniques. In the partial structural break model case, we can estimate the c_js over the sub-samples defined by the break points, but the estimate of ϕ depends on the optimal partition $(T_1,...,T_m)$. BP modify a recursive procedure discussed in Sargan (1964) that makes the estimation possible.[8] Briefly, they first minimize the sum of square residuals with respect to the vector of the changing parameters, keeping ϕ fixed and then minimize with respect to both the vector of changing parameters and ϕ. For appropriate

---

[6] The implied mean in each regime is simply $\mu_j = \dfrac{c_j}{1-\phi}$.

[7] Bai and Perron (2000) provide a very detailed discussion of the estimation algorithm.

initial values of ϕ, convergence to the global minimum is attained, in most cases after only one iteration.

BP show that the break fractions $\hat{k}_i = \hat{T}_i / T$ converge to their true value k_i^0 at a rate T, making the estimated break fractions super-consistent. Hence, we can estimate the rest of the parameters, which converge to their true values at rate $T^{1/2}$, taking the break dates as known. BP's procedure allows for the estimation of the parameters and the confidence intervals under very general conditions regarding the structure of the data and the errors across segments. In particular, their method is robust to heterogeneous variances of the residuals, which is the case we are interested in.

4. Empirical Results

Before we proceed discussing the results of the models presented in the previous section it is useful to present the standard OLS results for both the differences regression and the forward discount without accounting for structural change. Table 2a presents OLS estimates of the differences regression (1) as well as *t*-statistics for the hypothesis that $\beta = 1$. OLS estimates of an AR(1) specification of the forward discount is presented in table 2b. Notice that the β estimates for the French Franc and the Italian Lira are positive and are not statistically different than 1 at the 95% significance level, although the point estimates are 0.352 and 0.518 respectively. In the case of Germany, β is different than 1 at the 95% but not the 99% significance level. For Canada, UK, and Japan, the point estimates as well as the *t*-tests confirm the usual finding of the forward discount being a biased predictor of the change in the future spot rate. In all cases, R^2 is very small, ranging from 0.001 in the case of the French Franc, to 0.034 for the Japanese Yen. Also notice in table 2b that for France and Italy, the forward discount appears to be less persistent than in the other countries.

4.1 Time-Varying Parameter with Markov-Switching Variance Model

Figure 2 displays the results of the model applied to the six currencies. The filtered inferences use information up to time *t*, and smoothed inferences use information

---

[8] The complete details of the estimation technique can be found in Bai and Perron (1998) "Computation

from the entire sample, although all inferences are conditional on the hyper-parameters of the model, which are estimated using the entire sample. With the exception of Japan, ϕ_t exhibits substantial time variation. For the countries where time-variation is present, the forward discount is not highly persistent throughout. Typically, the forward discount starts out quite persistent at the beginning of the sample only to decline during the first part of the 1980s, as low as 0.60 in the case of Germany for example. It becomes very persistent again starting roughly at 1988. In all the cases considered, after that year ϕ_t abruptly rises toward or even above unity and continues to exhibit unit-root-like behavior until 1993. A possible interpretation for this behavior could be given along the lines of Siklos and Granger (1996): There exist processes which are cointegrated most of the time but not all the time. Perhaps forward and spot exchange rates fall into this category. While this is an issue that deserves further investigation, we continue to assume throughout the rest of the paper that spot and forward rates are and remain cointegrated with a cointegrating vector of $(1,-1)'$.

Table 3a reports the maximum likelihood estimates of the hyper-parameters. Again, with the exception of Japan, the estimates of the variance for the time-varying coefficient are all significant and of the same order of magnitude as the estimates of the variance for the forward discount process in the low variance regime. The likelihood ratio test statistics for the null hypothesis of no time variation are presented in table 3b. For Germany, France, Italy, and Canada, the null hypothesis of parameter stability can be rejected at the 1% level, while for the UK the same hypothesis can be rejected at the 5% leve. However, parameter stability cannot be rejected in the case of Japan, even at the 10% level. Table 3b also presents likelihood ratio tests for the null hypothesis of constant variance. It should be noted here that since the transition probabilities are not identified under the null hypothesis, standard assumptions of asymptotic distribution theory do not hold and the likelihood ratio test does not have a χ^2 distribution. Hansen (1992) suggests a computationally intensive method to determine the asymptotic distributions of the relevant statistics. Instead, following Kim, Morley and Nelson (1999) we use a likelihood ratio test using the critical values of Garcia (1995). Garcia derives asymptotic distributions for a simple two-state Markov-switching model. The null hypothesis,

and Analysis of Multiple Stuctural Change Models."

$H_0 : \sigma^2_{\eta,0} = \sigma^2_{\eta,1}$, is one of no Markov-switching. We compare these estimates to Garcia's most conservative critical values for a two-state Markov-switching mean and variance model. He reports a critical value of 17.52 for a 1% significance level test. The likelihood statistics for all the countries in our sample are much higher than this critical value.

Figure 3 displays the filtered and smoothed probabilities of a low variance state. The results are very similar across the different countries. It appears that the most volatile state was the period during the latter part of the 1970s and the beginning of the 1980s. Changes in monetary policy, abandonment of interest rates as an instrument and attention to the monetary base, as well as the 1981-82 recession in the United States seem to be the driving force. Other high volatility periods appear mostly as spikes during 1986, right after the Plaza Accord Agreement, and again in September of 1992, when the ERM collapsed. While we do not explicitly model possible volatility feedback effects to the mean of the forward discount, figure 3 raises the possibility that the forward discount is potentially subject to events that could lead to structural change.

Table 4 presents some diagnostic tests for the model. We test for serial correlation in both standardized forecast errors and the squares of the standardized forecast errors. There is evidence of serial correlation for Germany and Canada for the standardized forecast errors, as well as evidence of serial correlation for Germany, France and less so UK in the square of the standardized forecast errors. This evidence suggests that our two-state Markov-switching variance model has not captured completely the heteroskedasticity pattern of the forward discount for these countries.[9] However, the time-varying parameter model seems to adequately capture the time series properties of the forward discount process.

The estimates of the autoregressive coefficient provide evidence of parameter instability, while the two distinct and persistent regimes of the variance suggest that changes in policy could substantially affect the volatility of the forward discount. High variance regimes are significant and are consistent with periods where policy changes are

[9] We also tried a three-state Markov-switching variance specification for the countries in our sample. Although, the diagnostic tests where somewhat improved, the qualitative inferences regarding the autoregressive coefficient of the forward discount did not change. The results from the three-state specifications are available upon request.

in effect. The time-varying parameter model describes how the forward discount behavior has changed over time. The next question that we are interested in addressing is why has this behavior changed over the period of our sample. Eichenbaum and Evans (1995) show that monetary shocks have a direct effect on the mean of the interest rate differential. Thus in the next section we present such evidence using the Bai and Perron methodology regarding partial multiple structural break models.

4.2 Partial Structural Break Model

Tables 5 and 6 present the results for the partial structural break models based on (11). The determination of the existence of structural change and the selection of the number of breaks depends on the values of various test statistics for structural change when break dates are estimated and deserves some discussion. Let $\sup F_T(l)$ denote the F-statistic for testing the null of no breaks ($c_j = c$ for all j) against the alternative of l breaks ($c_1 \neq c_2 \neq \ldots \neq c_l$) where the break dates are selected according to (12). Define the double maximum statistic $UD\max = \max_{1 \leq l \leq L} \sup F_T(l)$, where L is an upper bound on the number of possible breaks. BP also consider a version of this statistic, denoted WDmax, that applies weights to $\sup F_T(l)$ such that the marginal p-values are equal across values of l. These statistics test the null hypothesis of no breaks against the alternative of an unspecified number of breaks subject to a specified upper bound on the number of breaks. Next, let $\sup F_T(l+1|l)$ denote the F-statistic for testing the null of l breaks against the alternative of l+1 breaks. For this test the first l breaks are estimated and taken as given. The statistic $\sup F_T(l+1|l)$ is then the maximum of the F-statistics for testing no further structural change in the intercept against the alternative of one additional change in the intercept when the break date is varied over all possible dates. All of these test statistics have non-standard asymptotic distributions and BP provide the relevant critical values.

BP (1998, 2000) suggest the following strategy for selecting the number of breaks based on the above statistics. We first look at the UDmax or WDmax tests to see if at least one break is present. If the null of no breaks is rejected, then the number of breaks can be determined by looking at the sequential $supF_T(l+1|l)$ statistics. We select the number of breaks for which the $supF_T(l+1|l)$ statistic is significant at least at the 5% level.

Table 5 presents the values of all the tests used to determine the number of breaks for each country.[10] In the case of Germany, the *UD*max, and *WD*max tests point to the presence of multiple breaks. The $supF_T(l+1|l)$ tests suggests the use of a model with five structural breaks since the $supF_T(5|4)$ test is significant at the 1% level. In the case of France, the *UD*max and *WD*max tests reject the null hypothesis of no breaks versus the alternative of an unknown number of breaks. The $supF_T(l+1|l)$ suggest that a model of four breaks should be chosen over a model with three breaks. Notice that the $supF_T(2|1)$ does not reject the null hypothesis of one break versus two. At the same time, the $supF_T(l)$ does not reject the null of no breaks versus two but does reject the null when the number of breaks is one, three, or four. Therefore, we estimate a model with four breaks for France. For Japan neither the *UD*max nor any of the *WD*max tests point to a number of breaks which is significantly different than zero. Hence, we do not estimate a model of multiple structural breaks for Japan. Both the time-varying parameter model and the stochastic break model single out Japan as the case where parameter instability, or structural change is not statistically significant within our sample period.

The results of the estimated break model (11) for the countries except Japan are presented in table 6. Notice that all the point estimates of the ϕ coefficients across the countries have dropped significantly compared to the estimates in table 2b, where we do not account for structural breaks. For instance the autoregressive coefficient of the forward discount was 0.939 for Germany and 0.907 for the UK when structural breaks where not taken into account. The corresponding estimates are 0.666 and 0.728 respectively, when we allow for such breaks in the process.

In table 6 we also report the estimates of the break dates with their respective 95% confidence intervals. Most of the break dates have been estimated quite accurately given that the estimates of the confidence intervals span the period of about two years[11]. Also, with the exception of the first and second break dates for Germany, the confidence intervals do not overlap, suggesting that the number of breaks has also been estimated accurately. Finally, most of the break dates estimated for each country are very similar to the break dates for the rest of the countries, further attesting to the robustness of our

[10] Critical values for these tests can be found in Bai and Perron (1998)

[11] Because there is a lagged dependent variable in the break model (11), the adjustment after the break is gradual and depends on the value of the autoregressive coefficient.

results. For all the countries, almost half of the breaks take place during the beginning of the sample, coinciding with the first period of high volatility captured by the time-varying coefficient model. This period is consistent with the change in the US central bank's policy objectives, as well as the subsequent recession of 1981-82. The break dates that correspond to the US 1981-82 recession are identified for France, Italy, Canada, and the UK as 1981:03, 1981:04, 1980:10, and 1981:07 respectively. Given our definition of the exchange rate and covered interest parity, we can write:

$$f_t - s_t = i_t^* - i_t^{US} \tag{13}$$

where i_t^{US} and i_t^* represent the nominal interest rates in the US and the foreign country, respectively. In all the cases, the implied unconditional means of the interest rate differentials have changed significantly, for the duration of the regimes immediately following the break dates.[12]

The final question we ask is the following: What are the implications of our model, namely that the forward discount is not as persistent when structural breaks are taken into account, for the forward discount puzzle? The estimate of β in equation (1) is found consistently to be biased away from its theoretical value 1. Could our partial structural change model explain some of this bias? In the next section we show using Monte Carlo simulations that this is the case indeed. Although we impose FRUH, the least squares point estimates of the coefficient turn out to be significantly biased downward.

5. Monte-Carlo Simulations

In this section we use Monte Carlo simulations in order to assess the implications of the presence of structural breaks in the forward discount process for FRUH. We estimate the "differences regression" (1) and the AR(1) specification for the forward discount with

[12] The estimates of the unconditional mean for France, Italy, Canada and the UK, before these breaks are –1.207, 0.450, -0.027, and –0.433 respectively. The means implied after the structural break date are 0.725, 1.015, 0.169, 0.007.

and without structural breaks in the mean of the forward discount. We report the performance of the tests for whether $\beta = 1$ versus the alternatives of $\beta \neq 1$ and $\beta < 1$ in (1). We also report the performance of the unit root test for the autoregressive coefficient of the forward discount. All experiments are based on 5000 replications.

5.1 Design of the Experiments

We employ an alternative yet equivalent representation of the cointegrating system (9) as our data generating process. This representation is due to Phillips' (1991) and is called a triangular representation. For our purposes, the general form of the triangular representation for y_t is

$$f_t = \mu_c + s_t + u_{ft}, \tag{12a}$$

$$s_t = s_{t-1} + u_{st} \tag{12b}$$

where the vector of errors $u_t = (u_{ft}, u_{st})' = (f_t - s_t - \mu_c, \Delta s_t)'$ has the VAR(1) representation $u_t = Cu_{t-1} + e_t$ where

$$C = \begin{pmatrix} \phi & 0 \\ \beta_S & 0 \end{pmatrix}, \; V = \begin{pmatrix} \sigma_{\eta\eta} & \sigma_{\eta s} \\ \sigma_{s\eta} & \sigma_{ss} \end{pmatrix}$$

Equation (12a) models the structural co-integrating relationship and (12b) is a reduced form relationship describing the stochastic trend in the spot rate. The VAR(1) representation for u_t implies

$$u_{ft} = \phi\, u_{f,t-1} + \eta_t, \tag{12c}$$

$$u_{st} = \beta_s u_{ft-1} + \varepsilon_{st} \tag{12d}$$

Equation (12c) models the disequilibrium error (which equals the forward premium) as an AR(1) process and (12d) allows the lagged error to affect the change in the spot rate. Letting $e_t = (\eta_t, \varepsilon_{st})'$. Note ϕ is the autoregressive coefficient of the forward discount and β is the forward discount coefficient from equation (1). In our simulations we set $\beta_s = 1$ so that UIP holds. In our monthly exchange rate data $\sigma_{\eta s} \approx 0$. We calibrate forward and spot exchange rates using the parameter estimates from the partial structural break model of the forward discount reported in table 6. We also calibrate spot and forward rates under the assumption of no break in the mean of the forward discount using the estimates reported in table 2b. We estimate the differences regression (1), as well as the AR(1) model of the forward discount using ordinary least squares. We also report the rejection rates for testing that $\phi = 1$ in the forward discount and $\beta_s = 1$ in the differences regression. Since $\phi = 1$ is a unit root test its rejection rate really measures the power of the augmented Dickey–Fuller test against the stationary alternative. Finally, in all the experiments, we set the sample size, $T = 250$ to reflect the number of observations in our actual data sample.

5.2 Monte Carlo Results

Tables 7 through 11 summarize the results of the Monte Carlo simulations for each country. Notice that in the case of no structural breaks, both ϕ and β_s are estimated correctly. The adf test has very high power and the size of the t-test is a correct 5%. The point estimates of β_s range from 0.997 in the case of Italy, to 1.045 in the case of Germany. The point estimates for ϕ are also extremely close to their true values.

When the data are generated under the assumption of structural breaks the results of the Monte Carlo are quite different. Structural breaks, which are unaccounted for, seem to produce two different yet interrelated results. First, the autoregressive coefficient of the forward discount is estimated to be very high and the power of the adf test is seriously reduced, with the exception of Italy and Canada. In both of these cases though, the point estimates of ϕ are quite higher than their true values. For Germany, France and the UK, the power of the adf test is 55.7%, 62.3%, and 25.2 % respectively.

Secondly, the point estimate of β_s is seriously biased downward away from its true value of 1. The point estimates range from 0.162 for Germany, which is the most severe case of bias to 0.526 in the case of Canada. Moreover, the size of the *t*-test for the null hypothesis of $\beta_s = 1$ is distorted, forcing one to reject the null hypothesis more often than she should. On average, at the 5% level, the two-sided *t*-test rejects the true null hypothesis about 20% of the time, while the one-sided *t*-test's rejection rate is even worse at about 30%.

The results of the Monte Carlo simulations seem to justify our prior that the forward premium is not as persistent as it appears to be. The presence of structural breaks is responsible for generating I(1)-like behavior in its process, that contributes a considerable degree of downward bias to the point estimate of β in equation (1). The median estimate of β in our experiments is not negative, as usually is when actual exchange rate data is estimated by least squares. Nevertheless, one would still mistakenly reject the null hypothesis of forward rate unbiasdeness, if one did not account for the presence of structural breaks in the forward discount.

6. Conclusion

We employ two different models of the forward discount under the prior that structural breaks in its process could explain away the highly stylized fact of its high persistence. The first model is a time varying parameter model with Markov-switching variance that help us document the pattern of the persistence. We overwhemingly reject the null hypothesis of no parameter instability for all G-7 countries with the exception of Japan. The time varying parameter model is able to capture structural change that takes place in a continuing fashion. The timing of the changes suggested the possibility of structural breaks in the mean of the process. Thus, we proceed to use a stochastic partial break model developed by Bai and Perron that explicitly allows for the incorporation and estimation of structural breaks in the mean of a process. The stochastic break model can be viewed as a plausible alternative to the fractionally integrated model of the forward discount used by Baillie and Bollerslev (1998) and Maynard and Phillips (1998). We find that breaks in the mean are present, and their timing coincides, at least for the case of the 1981-82 US recession with the types of monetary shocks reported by Eichenbaum and Evans (1995). A contractionary shock to the US monetary policy increases the US interest rates and, given our definition of the nominal exchange rate, also persistently decreases the level of the forward discount which under covered interest parity is equal to the interest rate differential of the two countries. Once these breaks are estimated, the forward discount's persistence is considerably lower than previously thought.

This finding has potentially important implications for what is known in the International Finance literature as the "Forward Discount Anomaly." In the absence of a time-varying risk premium, we simulate spot and forward exchange rates under the assumption of FRUH with and without incorporating breaks to the mean of the forward discount. We find that when breaks are not accounted for, the least squares' coefficient of the forward discount in the "differences regression" is severely biased downward, away from its theoretical value of 1. Furthermore, usual one- and two-sided t-tests suffer from significant size distortion, forcing one to reject the null hypothesis of FRUH too often. The forward discount puzzle is, to a considerable degree, a statistical artifact arising from breaks in the mean of the forward discount. Since the median Monte Carlo estimates of the forward discount coefficient in the "differences regression" are not negative, as is usually reported when actual data is used, competing explanations of the bias may still be valid and worth examining in light of our results. In particular, if a time-varying foreign exchange rate risk premium does exist, what is its contribution to the bias, once structural change has been accounted for? We hope to address this and related issues in the future.

APPENDIX A

Time-Varying Coefficient with Markov-Switching Variance Model[13]

Letting $p_{ij} = \Pr[S_t = j \mid S_{t-1} = i]$ with i=0,1 and j=0,1 , the Kalman filter for the model described by equations (10) and (11) is given by:

$$\beta_{t|t-1}^{(i,j)} = \beta_{t-1|t-1}^{i} \tag{A.1}$$

$$P_{t|t-1}^{(i,j)} = P_{t-1|t-1}^{i} + \sigma_v^2 \tag{A.2}$$

$$\eta_{t|t-1}^{(i,j)} = y_t - \beta_{t|t-1}^{(i,j)} x_t \tag{A.3}$$

$$f_{t|t-1}^{(i,j)} = x_t P_{t|t-1}^{(i,j)} x_t^{'} + \sigma_{\varepsilon j}^2 \tag{A.4}$$

$$\beta_{t|t}^{(i,j)} = \beta_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} x_t^{'} f_{t|t-1}^{(i,j)^{-1}} \eta_{t|t-1}^{(i,j)} \tag{A.5}$$

$$P_{t|t}^{(i,j)} = (I - P_{t|t-1}^{(i,j)} x_t^{'} f_{t|t-1}^{(i,j)^{-1}} x_t) P_{t|t-1}^{(i,j)} \tag{A.6}$$

where $\beta_{t|t-1} \equiv E[\beta_t \mid \Psi_{t-1}]$, is the expectation of β_t conditional on information up to time t-1; $P_{t|t-1}$ is the variance of $\beta_{t|t-1}$; η_{t-1} is the forecast error and $f_{t|t-1}$ is the variance of the forecast error; $y_t \equiv f_t - s_t$, and $x_t \equiv f_{t-1} - s_{t-1}$. Equations (A.1)-(A.4) are the prediction equations of the Kalman Filter, while equations (A.5)-(A.6) are the updating equations.

We also need to use Hamilton's (1989) filter which is given in the following three steps:

Step 1: Given $\Pr[S_{t-1} = i \mid \Psi_{t-1}]$, calculate

$$\Pr[S_t = j, S_{t-1} = i \mid \Psi_{t-1}] = \Pr[S_{t=j} \mid S_{t-1} = i] \Pr[S_{t-1} = i \mid \Psi_{t-1}] \tag{A.7}$$

where $\Pr[S_{t=j} \mid S_{t-1} = i\,]$ is the transition probability

Step 2: Calculate the joint density of y_t, S_t, S_{t-1} and collapse across all possible states to find the marginal density of y_t:

[13] This discussion follows Kim and Nelson (1999) "State-Space Models with Regime Switching"

$$f(y_t, S_{t=j}, S_{t-1} = i \mid \Psi_{t-1}) = f(y_t \mid S_t = j, S_{t-1} = i, \Psi_{t-1}) \Pr[S_t = j, S_{t-1} = i \mid \Psi_{t-1}] \tag{A.8}$$

Then the marginal density of y_t is given by:

$$f(y_t \mid \Psi_{t-1}) = \sum_{j=0}^{1}\sum_{i=0}^{1} f(y_t, S_{t=j}, S_{t-1} = i \mid \Psi_{t-1}) =$$

$$\sum_{j=0}^{1}\sum_{i=0}^{1} f(y_t \mid S_t = j, S_{t-1} = i, \Psi_{t-1}) \Pr[S_t = j, S_{t-1} = i \mid \Psi_{t-1}] \tag{A.9}$$

where $f(y_t \mid S_t = j, S_{t-1} = i, \Psi_{t-1}) =$

$$(2\pi)^{\frac{T}{2}} \mid f_{t|t-1}^{(i,j)} \mid^{-\frac{1}{2}} \exp\{-\frac{1}{2}\eta_{t|t-1}^{(i,j)}\,' f_{t|t-1}^{(i,j)\ -1}\eta_{t|t-1}^{(i,j)}\} \tag{A.10}$$

Step 3: Update the joint probability of S_t and S_{t-1} given y_t and collapse across all possible values of S_{t-1}:

$$\Pr[S_t = j, S_{t-1} = i \mid \Psi_{t-1}] = \frac{f(y_t, S_t = j, S_{t-1} = i \mid \Psi_{t-1})}{f(y_t \mid \Psi_{t-1})} \tag{A.11}$$

$$\Pr[S_t = j \mid \Psi_t] = \sum_{i=0}^{1} \Pr[S_t = j, S_{t-1} = i \mid \Psi_t] \tag{A.12}$$

Finally, as in Kim (1994) to complete the Kalman filter we collapse $\beta_{t|t}^{(i,j)}$ and $P_{t|t}^{(i,j)}$ across all possible values for S_{t-1}:

$$\beta_{t|t}^{j} = \frac{\sum_{i=0}^{1} \Pr[S_t = j, S_{t-1} = i \mid \Psi_t]\beta_{t|t}^{(i,j)}}{\Pr[S_t = j \mid \Psi_t]} \tag{A.13}$$

$$P_{t|t}^{j} = \frac{\sum_{i=0}^{1} \Pr[S_t = j, S_{t-1} = i \mid \Psi_t]\{P_{t|t}^{(i,j)} + (\beta_{t|t}^{j} - \beta_{t|t}^{(i,j)})(\beta_{t|t}^{j} - \beta_{t|t}^{(i,j)})'\}}{\Pr[S_t = j \mid \Psi_t]} \tag{A.14}$$

In order to get get the maximum likelihood estimates of the parameters as well as the filtered inferences for $\beta_{t|t}$ and $P_{t|t}$, we iterate through equations (A.1.1)-(A.1.14) for given initial values for $\beta_{0|0}^{i}$, $P_{0|0}^{i}$ and $\Pr[S_0 = i]$. The initial values for the probability are given by

$$\Pr[S_0 = 0] = \frac{1-p}{2-p-q} \text{ and } \Pr[S_0 = 1] = \frac{1-q}{2-q-p} \quad (A.15)$$

Since $\beta_{t|}$ has no unconditional expectation under the random walk specification, we have no choice but to make an arbitrary guess as to its initial value and then assign a very large variance to our guess, i.e., $\beta^i_{0|0} = 0$ and $P^i_{0|0} \approx \infty$. We then use the first τ observations to determine $\beta^i_{\tau|\tau}$ and $P^i_{\tau|\tau}$ and use these values as the initial values for the maximum likelihood estimation.

The filtered inferences about β_t and the confidence bands based on $P_{t|t}$ are given by:

$$\beta_{t|t} = \sum_{j=0}^{1} \Pr[S_t = j \mid \Psi_t]\beta^j_{t|t} \quad (A.16)$$

$$P_{t|t} = \sum_{j=0}^{1} \Pr[S_t = j \mid \Psi_t]\{P^j_{t|t} + (\beta_{t|t} - \beta^j_{t|t})(\beta^j_{t|t} - \beta^j_{t|t})'\} \quad (A.17)$$

The parameters of the model can be estimated by:

$$\max_{\theta} l(\theta) = \sum_{t=\tau+1}^{T} \ln f(y_t \mid \Psi_{t-1}) \quad (A.18)$$

Using Kim's (1994) smoothing algorithm, we can also obtain the smoothed probability $\Pr[S_t = 0 \mid \Psi_T]$. This is accomplished by iterating backward through the following equations (conditional on $S_t = j$ *and* $S_{t+1} = l$, where $j = 0,1$ and $l = 0,1$):

$$\Pr[S_{t+1} = l, S_{t+1} = j \mid \Psi_T] = \frac{\Pr[S_{t+1} = l \mid \Psi_T]\Pr[S_t = j \mid \Psi_t]\Pr[S_{t+1} = l \mid S_t = j]}{\Pr[S_{t+1} = l \mid \Psi_T]} \quad (A.19)$$

$$\Pr[S_{t+1} = j \mid \Psi_T] = \sum_{l=0}^{1} \Pr[S_{t+1} = l, S_{t+1} = j \mid \Psi_T] \quad (A.20)$$

Finally, we can obtain smoothed inferences about β_t conditional up to Information T, using the smoothed probabilities given by equations (A.19)-(A.20), and iterating backward the following two equations:

$$\beta^{(j,l)}_{t|T} = \beta^j_{t|t} + P^j_{t|t} P^{(j,l)^{-1}}_{t+1|t} (\beta^l_{t+1|T} - \beta^{(j,l)}_{t+1|T}) \quad (A.21)$$

$$P^{(j,l)}_{t|T} = P^j_{t|t} + (P^j_{t|t} P^{(j,l)^{-1}}_{t+1|t})(P^l_{t+1|T} - P^{(j,l)}_{t+1|t})(P^j_{t|t} P^{(j,l)^{-1}}_{t+1|t})' \quad (A.22)$$

References

Bai, Jushan (1997b): "Estimating multiple breaks one at a time," Econometric theory, 13, 315-352.

Bai, Jushan and Pierre Perron (1998): "Estimating and testing linear models with multiple structural changes," *Econometrica*, 66, 47-78.

____ (2000): "Computation and analysis of multiple structural change models," working paper, Department of Economics, Boston College.

Baillie, Richard T. and Tim Bollerslev (1994): "The long memory of the forward premium," *Journal of International Money and Finance*, 11, 208-219.

Baillie, Richard T. and Tim Bollerslev (2000): "The forward premium anomaly is not as bad as you think *Journal of International Money and Finance*.

Bilson, John F.0. (1981): "The 'speculative efficiency' hypothesis," *Journal of Business*, 54, 435-51.

Campbell, John Y. and Pierre Perron (1991): "Pitfalls and opportunities: What macroeconomists should know about unit roots and cointegration," NBER *Macroeconomics Annual*. MIT Press, Cambridge, MA.

Choi, K., and Zivot, E. (2001). "Long memory and structural breaks in the forward discount," manuscript in preparation, Department of Economics, University of Washington.

Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (): "Monetary policy shocks: What have we learned and to what end?" NBER, working paper 6400.

Clarida, Richard and Mark Gertler (1997): "How the Bundesbank conducts Monetary policy," in *Reducing inflation: motivation and strategy*, University of Chicago Press, Chicago.

Crowder, William J. (1994): "Foreign exchange market efficiency and common stochastic trends," *Journal of International Money and Finance*, 13, 551-564.

Cushman, David O. and Tao Zha (1997): "Identifying monetary policy in a small open economy under flexible exchange rates," *Journal of Monetary Economics*, 39, 4.

Diebold, Francis X. and Atsushi Inoue (1999): "Long memory and structural change," working paper, University of Pennsylvania.

Dominguez, Kathryn M. (1998): "Central bank intervention and exchange rate volatility," *Journal of International Money and Finance*, 17, 161-190.

Eichenbaum Martin, and Charles L. Evans (1995): "Some empirical evidence on the effects of shocks to monetary policy on exchange rates," *The Quarterly Journal of Economics*, Vo. 110, 4, 1975-1010.

Engel, Charles (1996): "The forward discount anomaly and the risk premium: A survey of recent evidence," *Journal of Empirical Finance*, 3, 123-192.

Engle, R.F and M. Watson (1987): Kalman filter: applications to forecasting and rational expectation models, *Advances in Econometrics*, Fifth World Congress, Vol.1,245-281.

Evans, Martin D.D. and Karen Lewis (1995): "Do long-term swings in the dollar affect estimates of the risk premia?" *Review of Financial Studies*, Vol. 8, No. 3, 709-742.

Fama, Eugene (1984a): "Forward and spot exchange rates," *Journal of Monetary Economics*, 14, 319-338.

Froot, Kenneth A. (1990): "Short rates and expected asset returns," working paper no. 3247, *National Bureau of Economic Research*, Cambridge, MA.

Froot, Kenneth A. and Jeffrey A. Frankel (1989): "Forward discount bias: is it an exchange risk premium?" *Quarterly Journal of Economics*, 104, 139-61.

Froot, Kenneth A. and Richard H. Thaler (1990): "Anomalies: Foreign exchange," *The Journal of Economic Perspectives*, Vol. 4, 3, 179-192.

Garbade, K. (1977): Two methods for examining the stability of regression coefficients, *Journal of the American Statistical Association*, 72, 54-63.

Garcia, Rene (1995): "Asymptotic null distribution of the likelihood ratio test in Markov switching models", manuscript, University of Montreal.

Garcia, Rene and Pierre Perron (1996): "An analysis of the real interest rate under regime shifts," *Review of Economics and Statistics*, 78, 111-125.

Godbout, Marie-Josee and Simon van Norden (1996): "Unit root tests and excess returns," working paper 96-10, Bank of Canada.

Granger, Clive W.J. (1999): "Aspects of research strategies for time series analysis," presentation to the conference on *New Developments in Time Series Economics*, New Haven, October 99.

Hai, Weike, Nelson Mark and Yangru Wu (1997): "Understanding spot and forward exchange rate regressions," *Journal of Apllied Econometrics*, Vol. 12, No. 6, 715-734.

Hamilton, James (1993): *Time Series Analysis*, Princeton University Press, Princeton, NJ.

Hansen, Bruce E. (1992): "The likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP," *Journal of Applied Econometrics*, 64, 307-333.

Horvath, Michael T.K and Mark W. Watson (1995): "Testing for cointegration when some of the cointegrating vectors are prespecified," *Econometric Theory*, 11, 984-1015.

Kim, Chang-Jin and Charles R. Nelson (1999): *State-space models with regime switching, classical and Gibbs-sampling approaches with applications*, MIT Press, Cambridge, Massachusetts.

Kim, Chang-Jin, James C. Morley, and Charles R. Nelson (1999): "Do changesin the market risk premiumexplain the empirical evidence of mean reversion in stock prices?" working paper, Department of Economics, University of Washington, Seattle, WA.

Kim, Souyong and Nouriel Roubini (1995): "Liquidity and exchange rates, a structural VAR approach," manuscript, New York University.

Lewis, Karen K. (1989): "Changing beliefs and systematic rational forecast errors with evidence from foreign exchange," *The American Economic Review*,Vol. 79, 4, 621-636.

____ (1995): Puzzles in international financial markets, *Handbook of International Economics Volume 3*, Elsevier Science B.V., Amsterdam, Netherlands

Mark, Nelson C. and Yangru Wu (1998): "Rethinking deviations from uncovered interest parity: The role of covariance risk and noise," *The Economic Journal*, 108, 1686-1706.

Marston, Richard C. (1995): *International financial integration: A study of interest differentials between the major industrial countries*, Cambridge University Press, New York, New York.

Maynard Alex and Peter C.B. Phillips (1998): "Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly,'" working paper, Yale University.

Perron, Pierre (1989): “The great crash, the oil price shock and the unit root hypothesis,” *Econometrica*, 57, 1361-1401.

Phillips, Peter C.B. (1991):Optimal inference in cointegrated systems,” *Econometrica*, 60, 119-143.

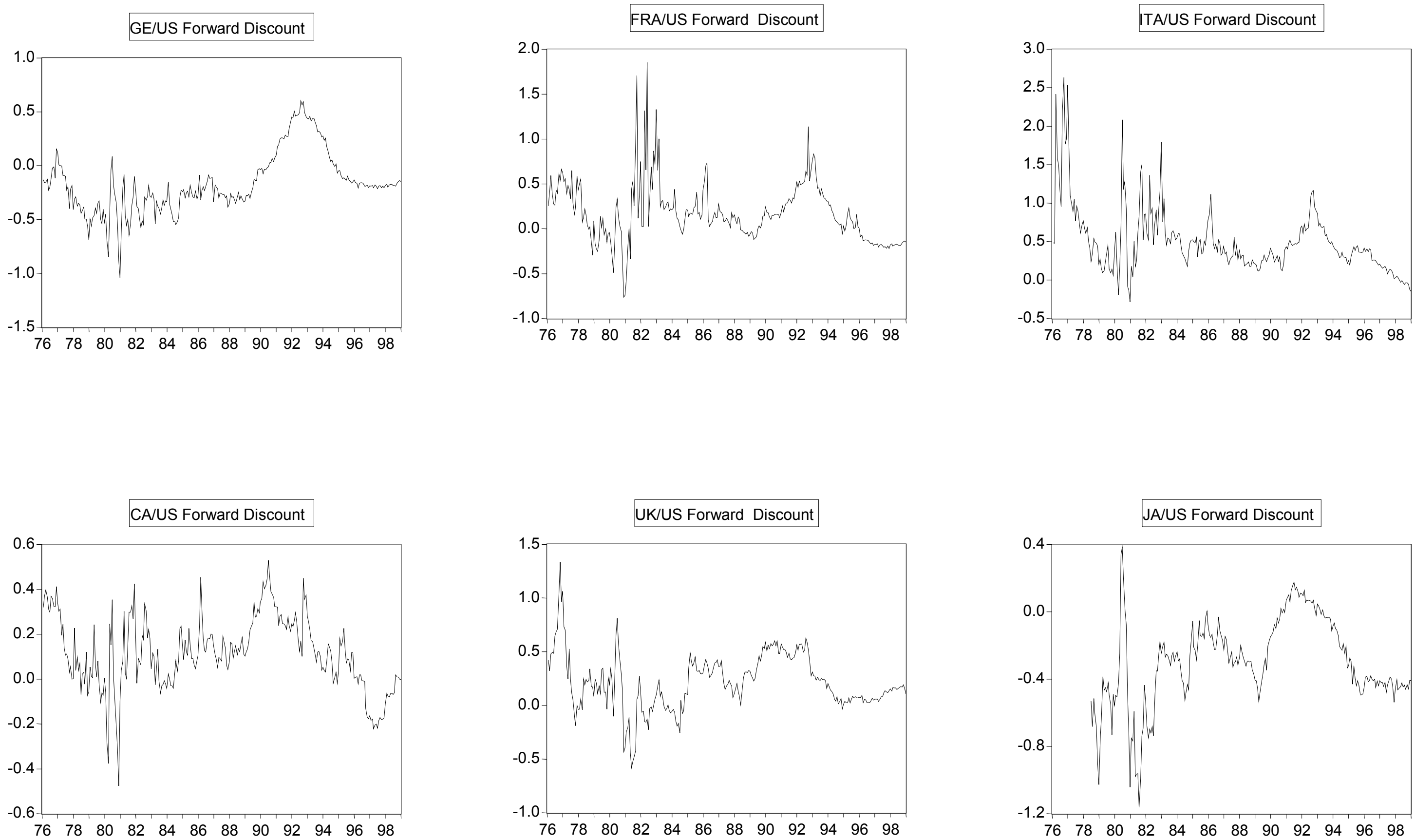
Sargan, J.D. (1964): “Wages and prices in the United Kingdom: a study in econometric methodology,” in P.E Hart, G. Mills and J.K. Whitaker (eds.), *Econometric Analysis for National Economic Planning*, London: Butterworths, 25-54.

Siklos, Pierre L. and Clive W. Granger (1996): “Temporary cointegration with an application to interest rate parity,” UCSD Discussion Paper No 96-11.

Wang, Jiahui and Eric Zivot (2000): “A Bayesian time series model of multiple structural changes in level, trend and variance,” *Journal of Business and Economic Statistics*.

Zivot, Eric (1999): “Cointegration and Forward and Spot Exchange Ratesforthcoming in the Journal of International Money and Finance.

Figure 1
Monthly Forward Discount



Source: Datastream

Table 1a: Summary Statistics For Exchange Rate Data

	German Mark			French Franc			Italian Lira		
	Δs_{t+1}	Δf_{t+1}	f_t-s_t	Δs_{t+1}	Δf_{t+1}	f_t-s_t	Δs_{t+1}	Δf_{t+1}	f_t-s_t
mean	-0.164	-0.157	-0.163	0.080	0.080	0.176	0.319	0.277	0.500
sd	3.325	3.326	0.279	3.235	3.197	0.331	3.241	3.176	0.432
ρ1	-0.012	-0.015	0.939	-0.011	-0.003	0.696	0.075	0.075	0.791
Correla	1.000	0.999	-0.057	1.000	0.996	0.036	1.000	0.996	0.07
tion		1.000	-0.062		1.000	0.005		1.000	0.042
Matrix			1.000			1.000			1.000

Table 1b: Summary Statistics For Exchange Rate Data

	Canadian Dollar			British Pound			Japanese Yen		
	Δs_{t+1}	Δf_{t+1}	f_t-s_t	Δs_{t+1}	Δf_{t+1}	f_t-s_t	Δs_{t+1}	Δf_{t+1}	f_t-s_t
mean	0.149	0.155	0.113	0.070	0.071	0.215	-0.360	-0.237	-0.296
sd	1.379	1.387	0.162	3.289	3.295	0.26	3.670	3.765	0.259
ρ1	-0.088	-0.092	0.839	0.084	0.086	0.907	-0.001	-0.004	0.928
Correla	1.000	0.997	-0.154	1.000	0.999	-0.124	1.000	0.999	-0.184
tion		1.000	-0.171		1.000	-0.131		1.000	-0.189
Matrix			1.000			1.000			1.000

Note: ρ_1 denotes the first order autocorrelation coefficient.

Table 2a: Estimates of the Differences Regression

$$\text{OLS: } \Delta s_{t+1} = \alpha + \beta(f_t - s_t) + \varepsilon_t$$

	German Mark	French Franc	Italian Lira	Canadian Dollar	British Pound	Japanese Yen
α	-0.269	0.0191	0.019	0.305	0.411	-1.031
	(0.240)	(0.239)	(0.293)	(0.099)	(0.239)	(0.316)
$t_{\alpha=0}$	-1.12083	0.0795	0.06485	3.08081	1.71967	-3.26266
	(0.868)	**(0.468)**	**(0.474)**	**(0.001)**	**(0.043)**	**(0.000)**
β	-0.686	0.352	0.518	-1.304	-1.568	-2.680
	(0.909)	(0.873)	(0.484)	(0.506)	(0.856)	(0.090)
$t_{\beta=1}$	-1.85	-0.74	-1.00	-4.55	-3.00	-40.89
	(0.032)	**(0.229)**	**(0.159)**	**(0.000)**	**(0.001)**	**(0.000)**
$\sigma^{1/2}_{ss}$	3.33	3.241	3.182	1.377	3.295	3.768
R^2	0.003	0.001	0.004	0.023	0.015	0.034

Note: White heteroskedasticity-consistent standard errors in parentheses. $t_{\beta=1}$ denotes the two-tail t-statistic for H_0: $\beta = 1$. p-values are in bold parentheses.

Table 2b: Estimates of the AR(1) specification of the Forward Discount

$$\text{OLS: } f_t - s_t = c + \phi(f_{t-1} - s_{t-1}) + v_t$$

	German Mark	French Franc	Italian Lira	Canadian Dollar	British Pound	Japanese Yen
c	-0.009	0.052	0.099	0.0169	0.018	-0.020
	(0.005)	(0.016)	(0.026)	(0.009)	(0.010)	(0.008)
ϕ	0.939	0.698	0.797	0.840	0.907	0.928
	(0.028)	(0.108)	(0.062)	(0.053)	(0.038)	(0.031)
$\sigma^{1/2}_{v}$	0.279	0.331	0.433	0.162	0.260	0.259
R^2	0.882	0.486	0.630	0.710	0.824	0.864

Note: White heteroskedasticity-consistent standard errors in parentheses

Table 3a
Maximum Likelihood Estimates of the Time-Varying Coefficient with Markov-Switching Variance Model of the Forward Discount

$$f_t - s_t - \mu_c = \phi_t(f_{t-1} - s_{t-1} - \mu_c) + \eta_t$$

$$\phi_t = \phi_{t-1} + v_t \text{ , } v_t \sim i.i.d.N(0, \sigma_v^2)$$

$$\eta_t \sim i.i.d.N(0, \sigma_{\eta t,t}^2)$$

$$\sigma_{\eta t}^2 = \sigma_{\eta 0}^2(1 - S_t) + \sigma_{\eta 1}^2 S_t$$

$$E[\eta_t v_t] = 0 \text{ , for all t}$$

$$\Pr[S_t = 1 | S_{t-1} = 1] = p \quad \text{and} \quad \Pr[S_t = 0 | S_{t-1} = 0] = q$$

	Germany	France	Italy	Canada	UK	Japan
μ	-0.309	0.121	0.366	0.081	0.084	-0.387
	(0.043)	(0.038)	(0.039)	(0.027)	(0.047)	(0.067)
σ_v	0.016	0.022	0.028	0.031	0.014	0.009
	(0.005)	(0.012)	(0.015)	(0.010)	(0.003)	(0.005)
$\sigma_{\eta 0}$	0.033	0.056	0.079	0.044	0.047	0.045
	(0.002)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
$\sigma_{\eta 1}$	0.127	0.417	0.428	0.145	0.167	0.169
	(0.008)	(0.038)	(0.042)	(0.013)	(0.013)	(0.017)
q	1.000	0.966	0.988	0.980	0.993	0.992
	(0.000)	(0.015)	(0.007)	(0.011)	(0.006)	(0.007)
p	0.991	0.915	0.944	0.946	0.973	0.957
	(0.008)	(0.015)	(0.031)	(0.031)	(0.018)	(0.030)
Loglik value	343.42	181.71	172.08	332.21	313.94	309.49

Note: Asymptotic standard errors are reported in parentheses. Samples are adjusted for lagged variables and staring values of the Kalman filter

Table 3b
Likelihood Ratio Tests for Parameter Constancy and Markov-Switching Specification

	Germany	France	Italy	Canada	UK	Japan
LRT for	25.866	7.278	11.824	8.275	3.755	1.357
H_o: $\sigma^2_{vt} = 0$	(0.000)	(0.006)	(0.000)	(0.004)	(0.0526)	(0.244)
LRT for	177.582	318.352	274.695	123.245	175.793	159.452
H_o: $\sigma^2_{\eta 0} = \sigma^2_{\eta 1}$	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: p-values in parentheses

Figure 2:
Filtered and Smoothed Probabilities of the Time-Varying Coefficient and Markov-Switching Variance Model of the Forward Discount
Filtered and Smoothed Inferences about ϕ_t with 95% Confidence Bands

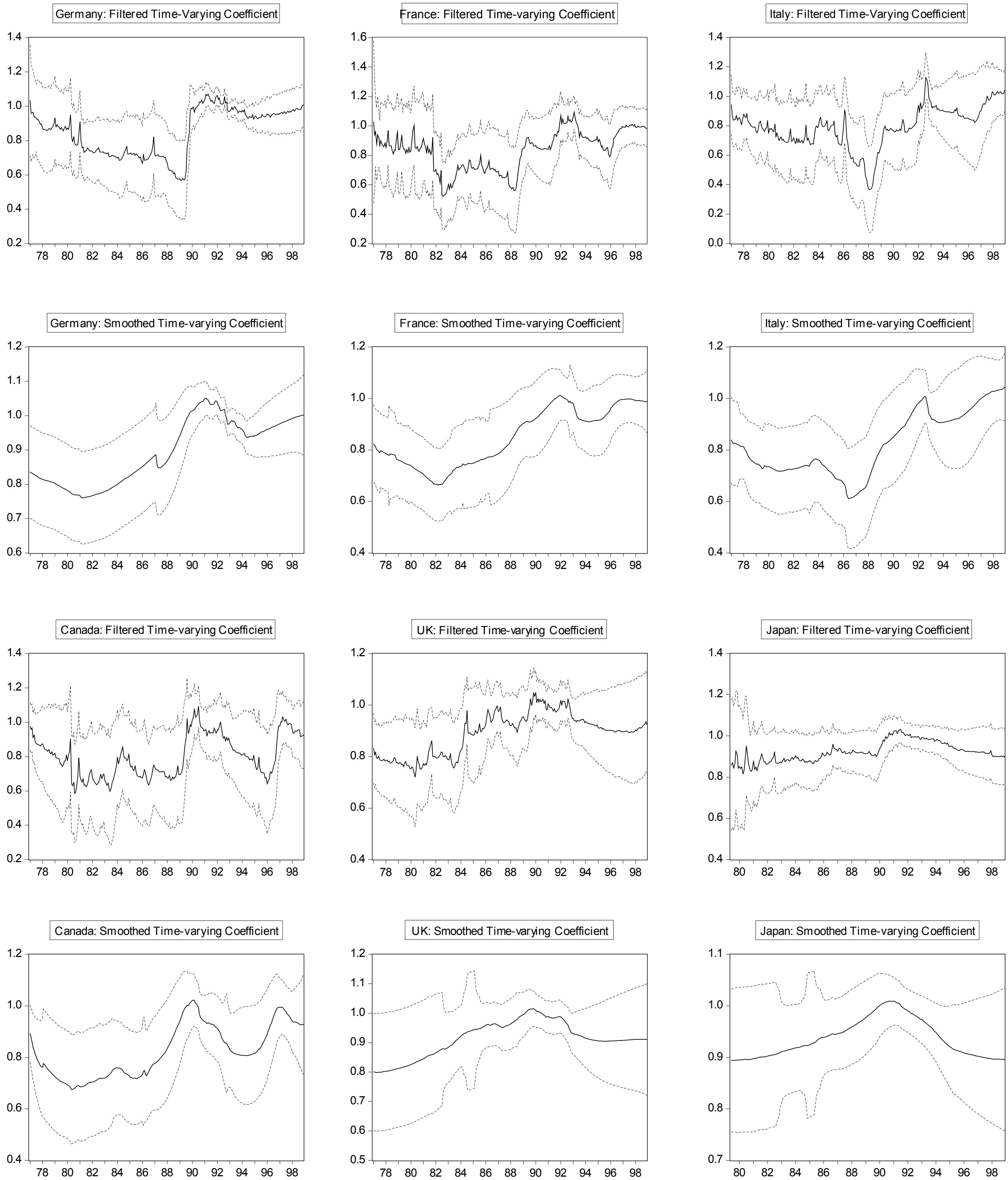


Figure 3
Time-Varying Coefficient and Markov-Switching Variance Model
Filtered and Smoothed Probabilities of the Low Variance State

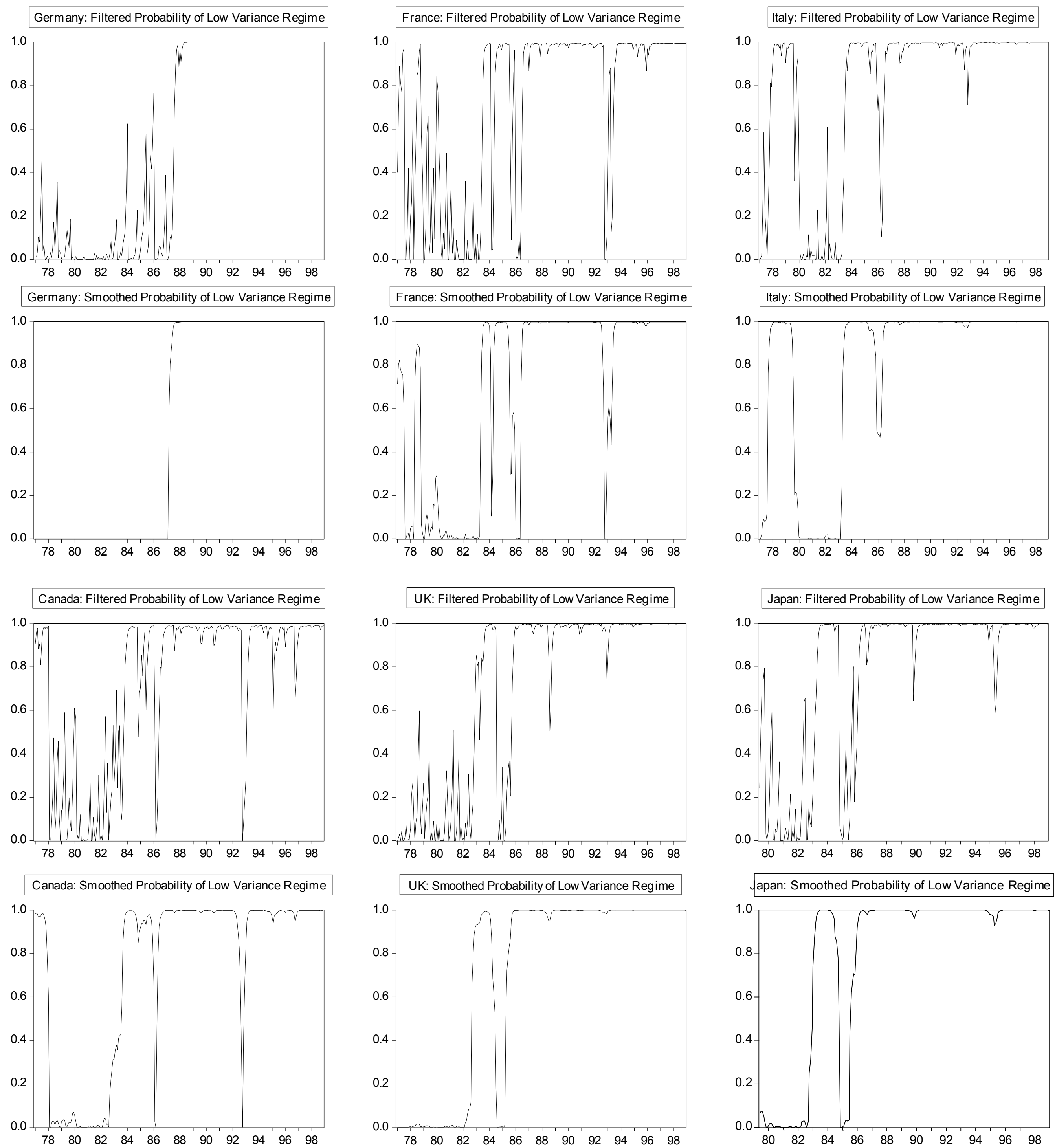


Table 4
Diagnostic Tests for the Time-Varying Parameter with Markov-Switching Variance Model

Q-stat of standardized residuals	GE/US Forward Discount	FR/US Forward Discount	IT/US Forward Discount	CA/US Forward Discount	UK/US Forward Discount	JA/US Forward Discount
Q(12)	31.315	26.966	20.670	38.654	9.525	9.223
	(0.002)	(0.008)	(0.055)	(0.000)	(0.657)	(0.684)
Q(24)	55.023	40.119	37.668	54.481	24.824	20.627
	(0.000)	(0.021)	(0.037)	(0.000)	(0.415)	(0.661)
Q(36)	62.470	49.854	46.417	64.523	31.685	25.581
	(0.004)	(0.062)	(0.076)	(0.002)	(0.674)	(0.850)

Q-stat of square of standardized residuals	GE/US Forward Discount	FR/US Forward Discount	IT/US Forward Discount	CA/US Forward Discount	UK/US Forward Discount	JA/US Forward Discount
Q(12)	84.321	194.12	9.333	17.030	39.979	29.859
	(0.000)	(0.000)	(0.674)	(0.148)	(0.000)	(0.003)
Q(24)	90.540	199.68	28.453	25.977	46.945	38.781
	(0.000)	(0.000)	(0.241)	(0.354)	(0.003)	(0.029)
Q(36)	102.61	207.32	53.546	35.858	57.476	41.520
	(0.000)	(0.000)	(0.030)	(0.475)	(0.013)	(0.243)

Note: p-values in parentheses

Table 5
Structural Break Tests for the Multiple Structural Break Forward Discount Model

Wdmax(1%)	20.510[a]	32.930[a]	17.752[a]	7.934[a]	14.446	9.017
SupFT(2|1)	31.183[a]	8.579	27.040[a]	10.486[c]	5.204	13.005[b]
SupFT(3|2)	10.550[c]	38.887[a]	7.69	19.932[a]	9.22	13.005[b]
SupFT(4|3)	12.807[c]	18.462[a]	30.293[a]	10.486	14.550[b]	13.005[b]
SupFT(5|4)	29.109[a]	6.461	7.69	8.128	14.550[b]	11.691[c]
Decision: number of breaks	5	4	4	3	5	0

Note: [a, b, c] denote 1%, 5%, and 10% levels of significance respectively

Table 6
Multiple Structural Break Model

$$f_t - s_t = c_j + \phi(f_{t-1} - s_{t-1}) + u_t, \qquad t = T_{j-1}+1,\ldots,T_j$$

	Germany	France	Italy	Canada	UK	Japan
c_1	-0.020	0.2686	0.768	0.098	0.217	-
	(0.019)	(0.0476)	(0.218)	(0.018)	(0.059)	-
c_2	-0.139	-0.7053	0.199	-0.008	0.050	-
	(0.025)	(0.0321)	(0.056)	(0.020)	(0.026)	-
c_3	-0.082	0.4237	0.449	0.049	-0.118	-
	(0.016)	(0.1345)	(0.104)	(0.008)	(0.045)	-
c_4	0.004	0.1209	0.191	-0.027	0.002	-
	(0.008)	(0.0228)	(0.035)	(0.007)	(0.017)	-
c_5	0.128	-0.0954	0.027	-	0.106	-
	(0.021)	(0.0142)	(0.012)	-	(0.018)	-
c_6	-0.048	-	-	-	0.028	-
	(0.007)	-	-	-	(0.006)	-
ϕ	0.666	0.4159	0.558	0.711	0.728	-
	(0.052)	(0.0901)	(0.074)	(0.043)	(0.045)	-
T_1	77:05	78:01	77:01	77:01	77:01	-
95% C.I.	76:01-77:10	77:08-77:03	76:09-78:05	76:01-77:02	76:02-78:03	-
T_2	84:08	81:03	81:04	80:10	80:06	-
95% C.I.	83:09-90:04	79:09-81:04	79:02-82:09	79:07-86:05	79:07-81:04	-
T_3	89:05	83:01	82:11	95:10	81:07	-
95% C.I.	89:02-89:12	82:11-84:08	82:08-84:12	95:07-96:06	80:10-83:04	-
T_4	90:11	95:09	96:04	-	84:08	-
95% C.I.	90:08-90:12	95:08-95:10	96:02-96:11	-	84:03-85:09	-
T_5	94:01	-	-	-	92:08	-
95% C.I.	93:11-94:02	-	-	-	92:05-93:05	-

Note: Asymptotic standard errors are reported in parentheses. $T_1,\ldots,T_5$ are the estimated break dates

Figure 4
Break Dates Estimated by the Multiple Structural Break Forward Discount Model

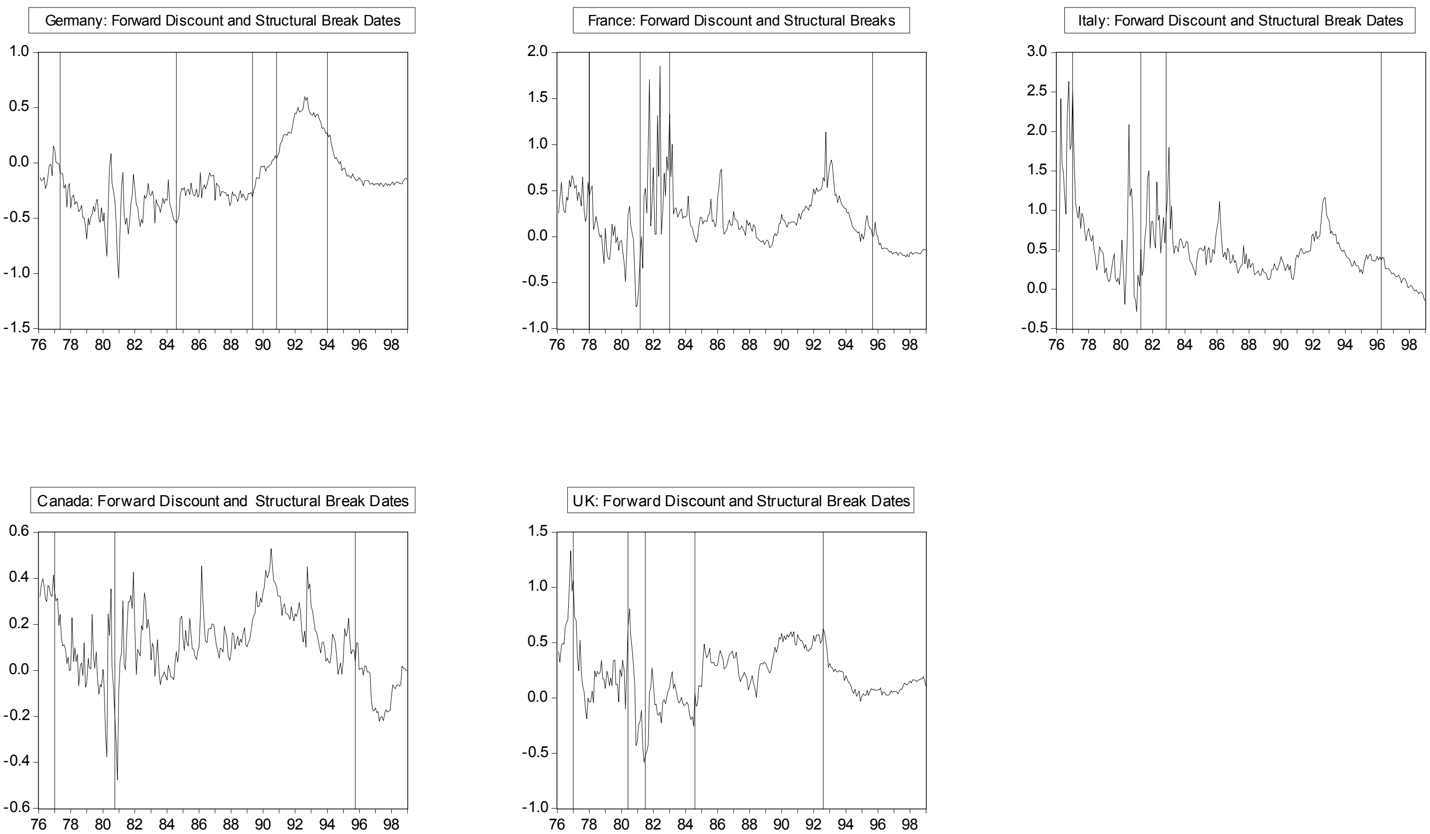


Table 7
Monte Carlo Estimates of the Forward Discount and "Differences" Regressions
Germany

$$f_t = \mu_c + s_t + u_{ft}, u_{ft} = \phi u_{ft-1} + \eta_t$$
$$s_t = s_{t-1} + u_{st}, u_{st} = \beta u_{ft-1} + \varepsilon_{st}$$
$$\begin{pmatrix} \eta_t \\ \varepsilon_{st} \end{pmatrix} = iidN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.088^2 & 0 \\ 0 & 3.325^2 \end{pmatrix}\right]$$

Estimated regression: $f_t - s_t = c + \phi(f_{t-1} - s_{t-1}) + e_t$

T = 250	Structural Breaks			No Structural Breaks		
	c	ϕ	ADF	c	ϕ	ADF
β=1	0.003	0.919	-2.92	-0.009	0.655	-6.251
ϕ =0.666			(0.557)[b]			(0.995)[b]

Estimated regression: $\Delta s_{t+1} = \alpha + \beta(f_t - s_t) + v_{t+1}$

T = 250	Structural Breaks			No Structural Breaks		
β=1	α	β	$t_{\beta=1}$	α	β	$t_{\beta=1}$
ϕ =0.666	-0.012	0.162	-1.198	0.027	1.045	-0.003
			(0.217)[a]			(0.055)[a]
			(0.323)[b]			(0.050)[b]

Note: $\mu_c = \frac{c}{1-\phi}$ is the unconditional mean of the forward discount. When we estimate the model with no structural breaks c = -0.009, the estimate from table 2b. When we estimate the model with structural breaks then c is set to the estimates given in table 6. We report the median estimates of the parameters based on 5,000 simulations. The values in parentheses indicate the empirical rejection frequency of nominal 5% two-sided tests[a] and one-sided tests[b].

Table 8
Monte Carlo Estimates of the Forward Discount and "Differences" Regressions
France

$$f_t = \mu_c + s_t + u_{ft}, u_{ft} = \phi u_{ft-1} + \eta_t$$
$$s_t = s_{t-1} + u_{st}, u_{st} = \beta u_{ft-1} + \varepsilon_{st}$$
$$\begin{pmatrix} \eta_t \\ \varepsilon_{st} \end{pmatrix} = iidN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.214^2 & 0 \\ 0 & 3.235^2 \end{pmatrix}\right]$$

Estimated regression: $f_t\text{-}s_t = c + \phi(f_{t-1}\text{-}s_{t-1}) + e_t$

T = 250	Structural Breaks			No Structural Breaks		
	c	ϕ	ADF	c	ϕ	ADF
β=1	0.004	0.874	-2.991	0.052	0.409	-8.048
ϕ =0.415			(0.623)[b]			(0.998)[b]

Estimated regression: $\Delta s_{t+1} = \alpha + \beta(f_t\text{-}s_t) + v_{t+1}$

T = 250	Structural Breaks			No Structural Breaks		
β=1	α	β	$t_{\beta=1}$	α	β	$t_{\beta=1}$
ϕ =0.415	-0.005	0.143	-2.545	-0.091	1.004	-0.032
			(0.719)[a]			(0.050)[a]
			(0.811)[b]			(0.052)[b]

Note: $\mu_c = \frac{c}{1-\phi}$ is the unconditional mean of the forward discount. When we estimate the model with no structural breaks c = 0.052, the estimate from table 2b. When we estimate the model with structural breaks then c is set to the estimates given in table 6. We report the median estimates of the parameters based on 5,000 simulations. The values in parentheses indicate the empirical rejection frequency of nominal 5% two-sided tests[a] and one-sided tests[b].

Table 9
Monte Carlo Estimates of the Forward Discount and "Differences" Regressions
Italy

$$f_t = \mu_c + s_t + u_{ft}, u_{ft} = \phi u_{ft-1} + \eta_t$$
$$s_t = s_{t-1} + u_{st}, u_{st} = \beta u_{ft-1} + \varepsilon_{st}$$
$$\begin{pmatrix} \eta_t \\ \varepsilon_{st} \end{pmatrix} = iidN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.242^2 & 0 \\ 0 & 3.241^2 \end{pmatrix}\right]$$

Estimated regression: f_t-s_t = $c + \phi(f_{t-1}$-$s_{t-1}) + e_t$

T = 250	Structural Breaks			No Structural Breaks		
	c	ϕ	ADF	c	ϕ	ADF
β=1	0.144	0.728	-4.191	0.101	0.547	-7.068
ϕ =0.558			(0.983)[b]			(0.999)[b]

Estimated regression: $\Delta s_{t+1} = \alpha + \beta(f_t\text{-}s_t) + v_{t+1}$

T = 250	Structural Breaks			No Structural Breaks		
β=1	α	β	$t_{\beta=1}$	α	β	$t_{\beta=1}$
ϕ =0.558	-0.248	0.472	-1.087	-0.219	0.997	-0.033
			(0.196)[a]			(0.053)[a]
			(0.288)[b]			(0.052)[b]

Note: $\mu_c = \frac{c}{1-\phi}$ is the unconditional mean of the forward discount. When we estimate the model with no structural breaks c = 0.099, the estimate from table 2b. When we estimate the model with structural breaks then c is set to the estimates given in table 6. We report the median estimates of the parameters based on 5,000 simulations. The values in parentheses indicate the empirical rejection frequency of nominal 5% two-sided tests[a] and one-sided tests[b].

Table 10
Monte Carlo Estimates of the Forward Discount and "Differences" Regressions
Canada

$$f_t = \mu_c + s_t + u_{ft}, u_{ft} = \phi u_{ft-1} + \eta_t$$
$$s_t = s_{t-1} + u_{st}, u_{st} = \beta u_{ft-1} + \varepsilon_{st}$$
$$\begin{pmatrix} \eta_t \\ \varepsilon_{st} \end{pmatrix} = iidN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.084^2 & 0 \\ 0 & 1.379^2 \end{pmatrix}\right]$$

Estimated regression: $f_t - s_t = c + \phi(f_{t-1} - s_{t-1}) + e_t$

T = 250	Structural Breaks			No Structural Breaks		
	c	ϕ	ADF	c	ϕ	ADF
β=1	-0.021	0.824	-3.928	0.017	0.700	-5.827
ϕ =0.711			(0.946)[b]			(0.995)[b]

Estimated regression: $\Delta s_{t+1} = \alpha + \beta(f_t - s_t) + v_{t+1}$

T = 250	Structural Breaks			No Structural Breaks		
β=1	α	β	$t_{\beta=1}$	α	β	$t_{\beta=1}$
ϕ =0.711	0.063	0.526	-0.888	-0.062	1.001	0.018
			(0.144)[a]			(0.045)[a]
			(0.228)[b]			(0.047)[b]

Note: $\mu_c = \frac{c}{1-\phi}$ is the unconditional mean of the forward discount. When we estimate the model with no structural breaks c = 0.0169, the estimate from table 2b. When we estimate the model with structural breaks then c is set to the estimates given in table 6. We report the median estimates of the parameters based on 5,000 simulations. The values in parentheses indicate the empirical rejection frequency of nominal 5% two-sided tests[a] and one-sided tests[b].

Table 11
Monte Carlo Estimates of the Forward Discount and Differences Regressions
UK

$$f_t = \mu_c + s_t + u_{ft}, u_{ft} = \phi u_{ft-1} + \eta_t$$
$$s_t = s_{t-1} + u_{st}, u_{st} = \beta u_{ft-1} + \varepsilon_{st}$$
$$\begin{pmatrix} \eta_t \\ \varepsilon_{st} \end{pmatrix} = iidN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.100^2 & 0 \\ 0 & 3.289^2 \end{pmatrix}\right]$$

Estimated regression: $f_t - s_t = c + \phi(f_{t-1} - s_{t-1}) + e_t$

T = 250	Structural Breaks			No Structural Breaks		
	c	ϕ	ADF	c	ϕ	ADF
β=1	0.017	0.897	-2.471	0.0187	0.717	-5.636
ϕ =0.728			(0.252)[b]			(0.996)[b]

Estimated regression: $\Delta s_{t+1} = \alpha + \beta(f_t - s_t) + v_{t+1}$

T = 250	Structural Breaks			No Structural Breaks		
β=1	α	β	$t_{\beta=1}$	α	β	$t_{\beta=1}$
ϕ =0.728	-0.031	0.214	-1.157	-0.064	1.019	-0.015
			(0.212)[a]			(0.048)[a]
			(0.307)[b]			(0.046)[b]

Note: $\mu_c = \frac{c}{1-\phi}$ is the unconditional mean of the forward discount. When we estimate the model with no structural breaks c = 0.018, the estimate from table 2b. When we estimate the model with structural breaks then c is set to the estimates given in table 6. We report the median estimates of the parameters based on 5,000 simulations. The values in parentheses indicate the empirical rejection frequency of nominal 5% two-sided tests[a] and one-sided tests[b].