Economics Bulletin

Volume 31, Issue 2

Structural change in the forward discount: a Bayesian analysis of forward rate unbiasedness hypothesis

Yi-Chi Chen
National Cheng Kung University

Wei-Choun Yu
Winona State University

Abstract

Using Bayesian methods, we reexamine the empirical evidence from Sakoulis et al. (2010) regarding structural breaks in the forward discount for G-7 countries. Our Bayesian framework allows the number and pattern of structural changes in level and variance to be endogenously determined. We find different locations of breakpoints for each currency; mostly, fewer breaks are present. We find little evidence of moving toward stationarity in the forward discount after accounting for structural change. Our findings suggest that the existence of structural change is not a viable justification for the forward discount anomaly.

Citation: Yi-Chi Chen and Wei-Choun Yu, (2011) "Structural change in the forward discount: a Bayesian analysis of forward rate unbiasedness hypothesis", *Economics Bulletin*, Vol. 31 no.2 pp. 1807-1826.
Submitted: Nov 11 2010. **Published:** June 19, 2011.

1. Introduction

With the assumptions of market efficiency, rational expectation and risk neutrality, the forward rate unbiasedness hypothesis states that the forward rate should be an unbiased forecaster for the future spot rate. The empirical testing for the validity of the hypothesis usually involves the differences regression:

$$\xi_{t+l} - \xi_t = \alpha + \beta\left(f_{t,l} - \xi_t \right) + \varepsilon_{t+l} \quad (1)$$

where ξ_t and $f_{t,l}$ denote the log of the spot and forward exchange rates at time t and l is the length of forward contract. $f_{t,l} - \xi_t$ is the forward discount, which equals the interest rate differential based on the covered interest parity. Under the null hypothesis of unbiasedness such that $\alpha = 0$, $\beta = 1$, and $E_t(\varepsilon_{t+l}) = 0$, it is expected that the estimate of β should equal to unity. However, the empirical consensus is that the estimate of β is significantly negative and the unbiasedness hypothesis is strongly rejected. The negative estimate is known as the forward discount anomaly.[1]

The forward discount anomaly has puzzled researchers for a long time. The literature provides four possible explanations: (1) the time-varying risk premium, (2) the peso problem, (3) the irrational expectation and the speculative bubble, and (4) the international market friction and inefficiency. However, recent studies suggest that the anomaly is exaggerated because of improper treatments of the forward discount regression. For example, Baillie and Bollerslev (2000) suggest that long memory process of forward discount could explain the anomaly. Sakoulis *et al*. (2010), hereafter SZC, argue that the structural change process, instead of long memory process, accounts for the overstated forward discount anomaly. SZC use a stochastic multiple break model, proposed by Bai and Perron (1998, 2003), to test the forward rate unbiasedness hypothesis and find that the forward discounts for the G-7 countries tend to be less persistent when they allow for structural change in the mean of the process.

However, SZC only investigate structural breaks in mean, taking no account for possible structural changes in volatility. Considering both potential structural breaks in mean and volatility, we use a Bayesian approach with the Gibbs-sampling algorithm to reexamine the empirical findings of SZC. The Bayesian methodology is different from the classical methodology used in SZC and has several advantages. First, the Bayesian inference allows for non-nested model comparison and selection that determines the optimal number and form of structural changes. Second, the Bayesian approach simplifies complicated estimations and inference procedures in multiple structural change models and allows for finite-sample

[1] See Hodrick (1987), Engle (1996), Baillie and Bollerslev (2000) and Sarno (2005) for a comprehensive survey of the literature.

inferences. Finally, the Bayesian methodology incorporates model and parameter uncertainty.

The remainder of the paper is organized as follows. Section 2 displays the model. Section 3 reviews the Bayesian methodology. Section 4 provides empirical findings. Section 5 concludes.

2. Model

Following Zivot (2000) and SZC, we model the series of the forward discount, y_t, as an AR(1) process implied by a Vector Error Correction Model (VECM):

$$y_t = a_t + \phi y_{t-1} + s_t u_t \tag{2}$$
$$u_t \mid \Omega_t \sim iid\ N(0,1) \quad \text{for } t = 1, 2, \ldots, T$$

where Ω_t denotes all available information up to time t. We assume that the mean/level, a_t, and volatility, s_t, parameters are subject to $m < T$ structural changes. The corresponding break dates are denoted by k_1, k_2,..., k_m such that $1 < k_1 < k_2 < ... < k_m \leq T$ giving m+1 possible regimes in T observations. For each regime i (i=1,…,m+1) the parameters a_t and s_t are given by the values α_i and σ_i for $k_{i-1} \leq t < k_i$ with $k_0 = 1$ and $k_{m+1} = T+1$. The AR parameters ϕ are assumed to be identical across regimes.

We consider two models. The first is a more general model, called **Design I**, which allows for unrestricted structural changes in level and volatility such that $a_t = \alpha_i$ and $s_t = \sigma_i$ for i=1,…,m+1. The second model, **Design II,** only allows for structural changes in level, holding the volatility constant across regimes, so that (2) becomes:

$$y_t = a_t + \phi y_{t-1} + \sigma u_t \quad t = 1, 2, \ldots, T \tag{3}$$

Letting I_{E_i} denote an indicator/state variable for the event $E_i = \{k_{i-1} \leq t < k_i\}$, Equations (2) and (3) can be expressed as the linear regression:

$$y_t = x_t' \mathbf{B} + s_t u_t \tag{4}$$

where

$$\mathbf{x}_t = (I_{E_1}, \ldots, I_{E_{m+1}}, y_{t-1})',$$
$$\mathbf{B} = (\alpha_1, \ldots, \alpha_{m+1}, \phi)'.$$

The $(3m+3) \times 1$ vector of unknown parameters is denoted $\boldsymbol{\theta} = \left(\mathbf{B}', \boldsymbol{\sigma}', \mathbf{k}'\right)'$. Given the normality assumption and the observed data $\mathbf{Y} = \left(y_1, \ldots, y_T\right)$, the likelihood function of (4) is:

$$L(\theta \mid \mathbf{Y}) \propto \left(\prod_{t=1}^{T} s_t\right)^{-1} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t - x_t'\mathbf{B}\right)^2}{s_t^2}\right\} \tag{5}$$

3. Bayesian Inference

In this section, we illustrate the Bayesian framework with the Gibbs-sampling algorithm developed by Wang and Zivot (2000) and Chen and Zivot (2010).

3.1. Prior specification

We assume that the vectors **k**, **B** and $\boldsymbol{\sigma}^2$ are mutually independent and that the elements of $\boldsymbol{\sigma}^2$ are independent. For the specification of the prior beliefs about unknown parameters, we use proper priors for **k**, **B** and $\boldsymbol{\sigma}^2$. The break points, **k,** are assumed to follow a discrete uniform distribution over all ordered subsequences of (2,3,...,*T*) of length of *m*. This is a diffuse prior which does not impose any information about the location of the break dates. With regard to the remaining parameters, we employ natural conjugate priors. The prior distribution of **B** in equation (4) is given by a multivariate normal (MVN) distribution, $\mathbf{B} \sim MVN(\mathbf{B}_0, \boldsymbol{\Sigma}_B)$, where $\mathbf{B_0}$ and $\boldsymbol{\Sigma}_\mathbf{B}$ are the prior mean and prior covariance matrix of **B**, respectively. The prior for $\boldsymbol{\sigma}^2$ specifies that each element follows an independent inverted Gamma (IG) distribution. That is, for each regime *i* (*i* = 1,..., *m*+1), $\sigma_i^2 \sim IG(v_0, \delta_0)$. To represent a diffuse prior, we set $\mathbf{B}_0 = 0$, $v_0 = 1.001$, $\delta_0 = .001$, and $\boldsymbol{\Sigma}_\mathbf{B}$ equal to a diagonal matrix with each diagonal element equal to 1,000.

3.2. Gibbs-sampling algorithm

The posterior distributions of the parameters are derived using the Gibbs sampler (Geman and Geman 1984; Gelfand and Smith 1990; Gelfand *et al*. 1990; Casella and George 1992; Gelman *et al*. 1995; Chib and Greenberg 1996). The basic principle of the Gibbs sampler is to approximate the joint and marginal posterior distributions by sampling from conditional distributions. Given the full conditionals $f(\theta_i \mid \theta_{-i}, \mathbf{Y})$, where θ_{-i} denotes the vector of $\boldsymbol{\theta}$ excluding the element θ_i, the Gibbs-sampling algorithm allows us to draw samples of $\boldsymbol{\theta}$ iteratively from the full conditional densities. After sufficient iteration, the draws of these random variables will converge to the target posterior distribution $f(\boldsymbol{\theta} \mid \mathbf{Y})$, and the marginal distribution of θ_i can be approximated by the empirical distribution of the draws.

Before proceeding with the Gibbs sampler, we first describe the full conditionals of the unknown parameters. For a given break date, k_i, the sample space only depends on the neighboring break points k_{i-1} and k_{i+1}. Accordingly, the posterior conditional density of k_i is of the form:

$$f(k_i \mid \theta_{-k_i}, \mathbf{Y}) \propto f(k_i \mid k_{i-1}, k_{i+1}, \mathbf{B}, \boldsymbol{\sigma}, \mathbf{Y}) \tag{6}$$

where *i* = 1,..., *m*. The breakpoint k_i can be drawn from a multinomial distribution with a sample size parameter equal to the number of dates between k_{i-1} and k_{i+1} and probability parameter proportional to the likelihood function. For the posterior conditional distribution of

$\mathbf{B}$, the normal prior for $\mathbf{B}$ combined with the normal likelihood of (5) yields a MVN conditional posterior:

$$\mathbf{B} \mid \theta_{-\mathbf{B}}, \mathbf{Y} \sim MVN\left(\tilde{\mathbf{B}}, \Sigma_{\tilde{\mathbf{B}}}\right) \tag{7}$$

where $\tilde{\mathbf{B}} = \Sigma_{\tilde{\mathbf{B}}}\left(\Sigma_{\mathbf{B}}^{-1}\mathbf{B}_0 + \mathbf{X}'\mathbf{S}^{-2}\mathbf{Y}\right)$ and $\Sigma_{\tilde{\mathbf{B}}} = \left(\Sigma_{\mathbf{B}}^{-1} + \mathbf{X}'\mathbf{S}^{-2}\mathbf{X}\right)^{-1}$. Here, $\mathbf{S}$ is a diagonal matrix with $(s_1,...,s_T)$ along the diagonal. Finally, with the natural conjugate IG prior for σ_i^2 and the normal likelihood (5), the posterior conditional for σ_i^2 also follows an IG distribution:

$$\sigma_i^2 \mid \theta_{-\sigma_i^2}, \mathbf{Y} \sim IG(\nu_i, \delta_i) \tag{8}$$

where $\nu_i = \nu_0 + n_i/2$, n_i represents the number of observations in regime i, $\delta_i = \delta_0 + \frac{1}{2}\left(\mathbf{Y}^i - \mathbf{X}^i\mathbf{B}\right)'\left(\mathbf{Y}^i - \mathbf{X}^i\mathbf{B}\right)$, $\mathbf{Y}^i$ is the vector of y_t values and $\mathbf{X}^i$ is the matrix of x_t values in regime i.

Given the full conditionals (6)-(8), the Gibbs-sampling algorithm can be iterated J times to obtain a vector sample of size J such that $\boldsymbol{\theta}^{(j)} = \left(\mathbf{k}^{(j)}, \mathbf{B}^{(j)}, \boldsymbol{\sigma}^{(j)}\right)$, $j = 1,..., J$.[2]

3.3. Posterior estimation

In order to generate the simulated draws from the Gibbs sampler, we use the method of the MCMC algorithm suggested by Geyer (1992). Specifically, given $N = n_0 + n_1$ iterations in the Markov chain, we only keep n_1 simulated samples for further inference by discarding the first n_0 sample as a burn-in. However, the output of the Gibbs sampler is a dependent sequence of parameter values forming a Markov chain. As a result, the series is serially correlated but stationary and ergodic. Then given $(\theta_i^{(1)}, \theta_i^{(2)}, \ldots, \theta_i^{(n_1)})$ post-convergent sample draws, the sample mean of these values can be used to estimate the posterior mean:

$$\bar{\theta}_i \equiv \frac{1}{n_1}\sum_{j=1}^{n_1}\theta_i^{(j)} \tag{9}$$

In addition, the Newey-West covariance matrix estimator that is consistent in the presence of both heteroskedasticity and autocorrelation:

$$\Gamma_0 + 2\sum_{j=1}^{q}\left(1 - \frac{j}{q+1}\right)\Gamma_j \tag{10}$$

where Γ_j is the jth-order sample autocovariance of θ_i from n_1 simulated draws and q is an

[2] Details of the Gibbs sampler for the structural break models are described in Wang and Zivot (2000). The C and Gauss codes for implementing Gibbs sampler were kindly provided by Jiahui Wang and Eric Zivot.

integer of the truncation lag such that $q = 4(n_1 / 100)^{1/4}$, can be used to estimate the variance of the posterior mean.

3.4. Model selection

The Bayesian framework provides a natural way of determining the number and form of structural breaks as a model selection problem.[3] We use Schwarz's Bayes information criterion (BIC) to select the best structural change model for the aggregate output series. The BIC for a model with m breaks is defined as:

$$\mathrm{BIC}(m) = 2 \times \ln L(\hat{\boldsymbol{\theta}} \,|\, \mathbf{Y}) - \lambda \ln(T) \tag{11}$$

where the likelihood function of $L(\cdot|\cdot)$ is equation (5) evaluated at the posterior mean of $\boldsymbol{\theta}$ based on the output of the Gibbs sampler, λ denotes the number of estimated parameters in model with m structural breaks, and T denotes the effective number of observations. By the definition of (11), the model with the highest posterior probability has the largest BIC value.

3.5. Comparison with classical approach

In contrast to our Bayesian approach, Bai and Perron (1998, 2003) consider the classical estimation of multiple structural changes in a linear model like (4) with a fixed number of breaks m, separated by a minimum number of observations, by global minimization of the sum of squared residuals. They do not estimate regime specific error variances but they can allow for general forms of serial correlation and conditional heteroskedasticity in the error terms. To determine the number of breaks, they consider a test of the null hypothesis of no break versus the alternative hypothesis of some unknown number of breaks between 1 and some upper bound M. Their tests, called double maximum tests, are based on the maximum of the (possibly weighted) individual tests for the null of no break versus m breaks (m = 1,..., M). The double maximum tests are particularly useful to determine whether some structural change is present since a sequential testing procedure can be unreliable for particular forms of multiple changes (see Bai and Perron, 2006). They also consider the use of model selection criteria to determine the number of breaks.

The asymptotic theory used in Bai and Perron (1998) assumes non-trending data and needs to be modified for trending data. While Bai and Perron (1998) consider the case of constant volatility, Bai (2000) advances the theory to allow for the breaks in the variance of the error term. For inference on the break dates, the asymptotic theory assumes that the magnitudes of the structural changes in the parameters shrink as the sample size increases and

[3] Wang and Zivot (2000) used several model selection criteria to determine the number and type of structural changes. Specifically, they used marginal likelihoods, posterior odds ratios and Schwarz's Bayes information criterion (BIC) to select the model with the most appropriate pattern of structural breaks that best describes the data-generating process of the series. Based on a set of Monte Carlo experiments they found that model selection based on the BIC performed the best.

can give perverse results for large parameter shifts. In contrast, inference using the Bayesian methodology is the same for trending and non-trending data and for any magnitude of structural changes on the parameters. For model selection, both the Bayesian and classical methodologies treat the number of breaks as unknown.

The classical least-squares and Bayesian estimation with a uniform prior on the break dates lead to similar results for the location of structural changepoints and point estimates for the regression coefficients for Design II. The least-squares estimation of the coefficients conditions on the least-squares estimates of the break dates. This approach is asymptotically justified as the break fractions converge faster than the regression coefficients. In particular, with a short time series, the inferences about the location of the break dates are subject to considerable uncertainty. Bayesian inference accounts for this uncertainty by integrating over, rather than conditioning on the break dates. In this case, Bayesian standard errors for coefficients are likely to be larger than the least-squares ones because the Bayesian inferences explicitly account for parameter uncertainty.

4. Empirical Findings

We use the log spot and 1-month log forward exchange rates compiled by SZC.[4] The exchange rates are end-of-month national currency units per US dollar quoted by the arithmetic average of the bid and ask rates for six G-7 countries (Canada, France, Germany, Italy, and U.K.) ranging from 1976:01 to 1998:12 and for Japan from 1978:07. The log values of the forward discount have been multiplied by 100 and therefore approximate to percentage differences. Figure 1 plots the forward discount, $f_{t,l}-\xi_t$, for all the currencies. It is clear to see that the forward discount is much more volatile during the period of the 1970s and 1980s compared to the 1990s. In particular, there appear to be regime shifts in the forward discount across the sample. Table 1 gives the summary statistics of the data. The noted observations of the forward discount, such as skewness, leptokurtosis, and significant Bera-Jarque statistics, are present in the data.

Table 2 reports posterior estimates without accounting for structural change. Most of the exchange rates are significantly persistent with the AR term above 0.8, while the French franc (0.699) and Italian lira (0.798) are relatively less persistent. Tables 3 and 4 present the results which consider structural breaks in the forward discount. In order to determine the number and pattern of structural breaks, we estimate the models of Design I and II with *m* breaks (*m* = 0,1,…,5) and then choose the model that maximizes the BIC. Inferences are based on 2,000 draws of Gibbs sampler after dropping the first 500 simulations as the burn-in period. Taking the Canadian dollar as an example, the Design I model with 4 breaks has the highest BIC value at 617, while the BIC for Design II with 4 breaks is 517, which is substantially lower

[4] The authors thank Eric Zivot for generously providing the data.

than those with any number of breaks in Design I. Table 3 shows that Design I is superior to Design II in the model fitting. By and large, our model selects fewer break points than those in SZC. For example, while SZC suggests 5 breaks for the German mark, our approach indicates 2 breaks.

Table 4 displays the Bayesian estimations of Design I based on the preferred model suggested in Table 3. The upper panel of the table presents a 95% highest posterior density (HPD) region of structural change with the posterior mode for the break years in bold. The lower panel shows the posterior means of the estimated parameters followed by the standard deviations associated with the estimates and the 2.5% and 97.5% posterior quantiles of the parameters. Note that the AR(1) terms for all of the G-7 countries become significantly larger, ranging between 0.84 to 0.96, than those in Table 2. That said, the presence of structural breaks makes the forward discount more persistent, opposing the less persistent evidence as observed by SZC. For example, the AR coefficient for the Italian lira is 0.798 in Table 2 but is 0.929 in Table 4. Our evidence rebukes the prior findings in SZC that allowing structural breaks in the forward discount can explain away the high prescience.

Table 5 compares the timing of breaks obtained by our Bayesian method and SZC. In the SZC study, the Sterling crisis of 1976 and the establishment of the European Monetary System (EMS) of 1979 were the major events that caused the breaks for most of the G-7 countries. In our results, most G-7 countries have been affected by the spillovers of the US economic downturn, including Canada (1983:08), France (1981:05 and 1983:05), Italy (1983:05) and Japan (1982:10). The burst of the EMS crisis has a significant impact on the European countries, such as France (1991:12), Germany (1993:07), Italy (1993:04) and the UK (1993:02). Only Canada has the break in the late 1970s that may be justified by the weakness of the Canadian dollar due to the political uncertainty, imminent inflationary pressure and the current account deficit. The steady decline of the Canadian dollar beginning in 1992 reflects the expansionary monetary policy and large current account deficits (Powell, 2005). Other breaks are consistent with SZC.

In our study, we do not require breaks to be separated by at least five years in the search for potential break dates, and we allow for the possibility that an outlier observation can be detected. For example, we find outlier observations for Canada in the second half of 1992. Past studies have documented volatility changes in exchange rate series and have shown that changes in exchange rate volatility can be confused with changes in level. As a result, using Design II could produce misleading inferences. This is most evident in the two-break models for the German mark, the Italian lira and the British pound. Design I detects breaks in both level and variance after the establishment of the EMS (1987:03, 1983:05 and 1985:04, respectively). In contrast, Design II translates the variance breaks into larger mean changes after 1977 as in SZC. In this case, Design I implies similar means across regimes, whereas Design II shows distinct forward discount dynamics in which a higher implied mean of the exchange rate was followed by a big decline in the forward discount.

Finally, Figures 2(a)-(f) plot the marginal posterior distributions of the break dates with the forward discount series superimposed. For instance, in Figure 2(d) for the Italia lira, the two structural breaks most likely occurred in 1983:05 and in 1993:04 with the highest posterior probability being around 0.58 and 0.22, respectively. The first break date is very precisely estimated with only k_1=1983:03 and k_1=1983:11 contained in a 95% HPD region.

5. Conclusions

Using Bayesian methods, we search for the most appropriate structural break specification to model the changes in the processes of G-7 countries for 20 years of monthly spot and forward exchange rate data. We find that the US recession and the EMS crisis have played a crucial role in explaining the breaks in the forward discount process. We find evidence that the forward discounts have experienced breaks in variance. Our posterior estimates indicate that the forward discount remains highly persistent even after accounting for structural change. Our findings cast doubt on the proposition by SZC that the existence of structural change is a viable justification for the forward discount anomaly.

References

Bai, J. and P. Perron. (1998) "Estimating and Testing Linear Models with Multiple Structural Changes" *Econometrica* **66**, 47-78.

Bai, J. (2000) "Vector Autoregressive Models with Structural Changes in Regression Coefficients and in Variance-Covariance Matrices" *Annals of Economics and Finance* **1**, 303-39.

Bai, J. and P. Perron (2003) "Computation and Analysis of Multiple Structural Change Models" *Journal of Applied Econometrics* **18**, 1-22.

Bai, J. and P. Perron (2006) "Multiple Structural Change Models: A Simulation Analysis" in *Econometric Theory and Practice* by D. Corbea, S. Durlauf and B.E. Hansen, Eds., Cambridge University Press: New York, 212-34.

Baillie, R.T. and T. Bollerslev (2000) "The Forward Premium Anomaly is not as Bad as You Think" *Journal of International Money and Finance* **19**, 471-88.

Casella, G. and E. I. George (1992) "Explaining the Gibbs Sampler" *American Statistician* **46**, 167-74.

Chen, Y.-C. and E. Zivot (2010) "Postwar Slowdowns and Long-Run Growth: A Bayesian Approach of Structural-Break Models" *Empirical Economics*, **39**, 897-921.

Chib, S. and E. Greenberg (1996) "Markov Chain Monte Carlo Simulation Methods in Econometrics" *Econometric Theory* **12**, 409-31.

Engel, C. (1996) "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence" *Journal of Empirical Finance* **3**, 123-92.

Gelfand, A.E. and A.F.M. Smith (1990) "Sampling-Based Approaches to Calculating Marginal Densities" *Journal of the American Statistical Association* **85**, 398-409.

Gelfand, A.E., S.E. Hills, A. Racine-Poon and A.F.M. Smith (1990) "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling" *Journal of the American Statistical Association* **85**, 972-85.

Gelman, A., J.S. Carlin, H.S. Stern and D.B. Rubin (1995) *Bayesian Data Analysis*, Chapman and Hall: London.

Geman, S. and D. Geman (1984) "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721-41.

Geyer, C.J. (1992) "Practical Markov Chain Monte Carlo" *Statistical Science* **7**, 473-83.

Hodrick, R.J. (1987). *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*, Harwood: London.

Powell, J. (2005) *A History of the Canadian Dollar*, Bank of Canada: Ottawa.

Sakoulis, G., E. Zivot and K. Choi (2010) "Structural Change in the Forward Discount: Implications for the Forward Rate Unbiasedness Hypothesis" *Journal of Empirical Finance* **17**, 957-66.

Sarno, L. (2005) “Viewpoint: Towards a Solution to the Puzzles in Exchange Rate Economics: Where do We Stand?” *Canadian Journal of Economics* **38**, 673-708.

Wang, J. and E. Zivot (2000) “A Bayesian Time Series Model of Multiple Structural Changes in Level, Trend, and Variance” *Journal of Business and Economic Statistics* **18**, 374-86.

Ziovt, E. (2000) “Cointegration and Forward and Spot Exchange Rate Regressions” *Journal of International Money and Finance* **19**, 785-812.

Table 1. Summary statistics for the forward discount

	German Mark	French Franc	Italian Lira	Canadian Dollar	British Pound	Japanese Yen
Mean	-0.163	0.176	0.501	0.114	0.216	-0.296
Median	-0.192	0.138	0.413	0.109	0.187	-0.294
Std. Dev.	0.279	0.331	0.433	0.163	0.260	0.259
Skewness	0.589	1.203	2.052	-0.208	0.324	-0.388
Kurtosis	3.630	7.120	9.027	3.277	4.654	3.527
J-B	20.52*	261.76*	611.51*	2.88	36.28*	9.07*

Note: the sample period for the monthly forward discount runs from 1976:01 to 1998:12, except for Japanese yen from 1978:07. J-B is the Jarque–Bera test for normality. * indicates significance at the 5% level.

Table 2. Posterior estimates of the AR(1) specification of the forward discount

	Canada	France	Germany	Italy	Japan	U.K.
α	.017	.052	-.010	.099	-.021	.019
	(.006)	(.016)	(.007)	(.024)	(.009)	(.009)
	[.004, .030]	[.019, .084]	[-.023, .003]	[.051, .146]	[-.039, -.002]	[.002, .036]
σ	.088	.238	.096	.264	.096	.109
	(.004)	(.010)	(.004)	(.011)	(.004)	(.005)
	[.081, .095]	[.219, .259]	[.088, .104]	[.242, .287]	[.088, .105]	[.100, .119]
φ	.841	.699	.940	.798	.929	.908
	(.030)	(.041)	(.019)	(.035)	(.022)	(.024)
	[.780, .905]	[.617, .784]	[.900, .980]	[.728, .870]	[.884, .975]	[.860, .958]
LLK	277.01	2.03	251.94	-26.07	225.96	216.83

Note: the numbers in parentheses denote standard errors; those in square brackets are the 95% highest posterior density (HPD) regions.

Table 3. Choice of the number of breaks for the forward discount by BIC

	Canada		France		Germany		Italy		Japa
Design	I	II	I	II	I	II	I	II	I
m=0	537.170		-12.793		487.035		-68.995		435.
m=1	598.776	526.421	197.714	-21.806	647.932	476.715	224.826	-50.132	**582.**
m=2	602.452	516.599	273.169	-6.908	**660.003**	493.777	**261.547**	-58.242	580.
m=3	603.708	525.579	321.051	-3.332	657.814	465.326	256.076	-45.265	570.
m=4	**617.187**	517.377	303.912	0.015	656.111	474.784	248.662	-67.909	566.
m=5	614.351	516.136	**324.085**	39.847	642.452	491.239	254.548	-55.205	555.
Number of breaks	4		5		2		2		1
	(3)		(4)		(5)		(4)		(0)

Note: *m* denotes the number of breaks in the model. The Schwarz's BIC is calculated by 2*LLK- λ*log(T) where LLK is the margi
posterior mean of the parameter, λ is the number of parameters with *m* structural breaks and *T* is the number of observations. Thus,
BIC value. The maximum BIC value is highlighted in bold. The numbers in the parentheses denote the choice of breaks used in Sa

Table 4. Final structural break model results for the forward discount

	Canada	France	Germany	Italy	Japan	U.K.
Trend Breaks:						
k_1	[76:08, **78:02**, 80:05]	[80:01, **81:05**, 81:09]	[87:02, **87:03**, 89:03]	[83:03, **83:05**, 83:11]	[82:02, **82:10**, 83:07]	[84:09, **85:04**, 86:01]
k_2	[82:02, **83:08**, 84:03]	[83:03, **83:05**, 83:09]	[92:09, **93:07**, 96:09]	[92:09, **93:04**, 94:01]		[92:09, **93:02**, 94:05]
k_3	[92:06, **92:10**, 92:10]	[86:06, **86:06**, 87:10]				
k_4	[92:11, **92:11**, 93:02	[90:07, **91:12**, 92:10]				
k_5		[95:12, **96:05**, 96:09]				
Coefficients:						
α_1	.022 (.028) [-.010, .064]	.015 (.028) [-.033, .066]	-.013 (.014) [-.038, .011]	.055 (.053) [-.050, .153]	-.023 (.027) [-.080, .032]	.012 (.019) [-.026, .048]
σ_1	.060 (.033) [.037, .098]	.183 (.021) [.151, .221]	.132 (.009) [.116, .149]	.462 (.035) [.398, .537]	.185 (.020) [.153, .226]	.171 (.035) [.143, .210]
α_2	.008 (.024) [-.036, .055]	.105 (.163) [-.181, .386]	.009 (.010) [-.003, .021]	.033 (.015) [.005, .060]	-.011 (.006) [-.023, .000]	.031 (.023) [.007, .057]
σ_2	.153 (.047) [.117, .210]	.689 (.125) [.521, .933]	.044 (.005) [.037, .053]	.105 (.009) [.092, .120]	.052 (.003) [.047, .058]	.061 (.019) [.049, .077]
α_3	.023 (.008) [.009, .038]	.031 (.032) [-.020, .079]	-.010 (.005) [-.017, -.001]	.007 (.009) [-.008, .024]		.008 (.005) [-.001, .017]
σ_3	.056 (.012) [.041, .066]	.136 (.019) [.108, .173]	.023 (.006) [.017, .029]	.044 (.004) [.037, .053]		.028 (.007) [.020, .034]
α_4	.350 (.172) [.122, .485]	.022 (.010) [.005, .040]				
σ_4	.058 (.152) [.017, .215]	.054 (.006) [.045, .066]				
α_5	-.003 (.005) [-.013, .007]	.040 (.046) [-.016, .095]				
σ_5	.043 (.004) [.037, .051]	.134 (.017) [.109, .165]				
α_6		-.029 (.017) [-.048, -.011]				
σ_6		.020 (.004) [.015, .026]				
φ	.877 (.027) [.824, .928]	.842 (.090) [.745, .934]	.961 (.021) [.930. .990]	.929 (.026) [.883, .974]	.952 (.021) [.913, .991]	.910 (.044) [.855, .963]

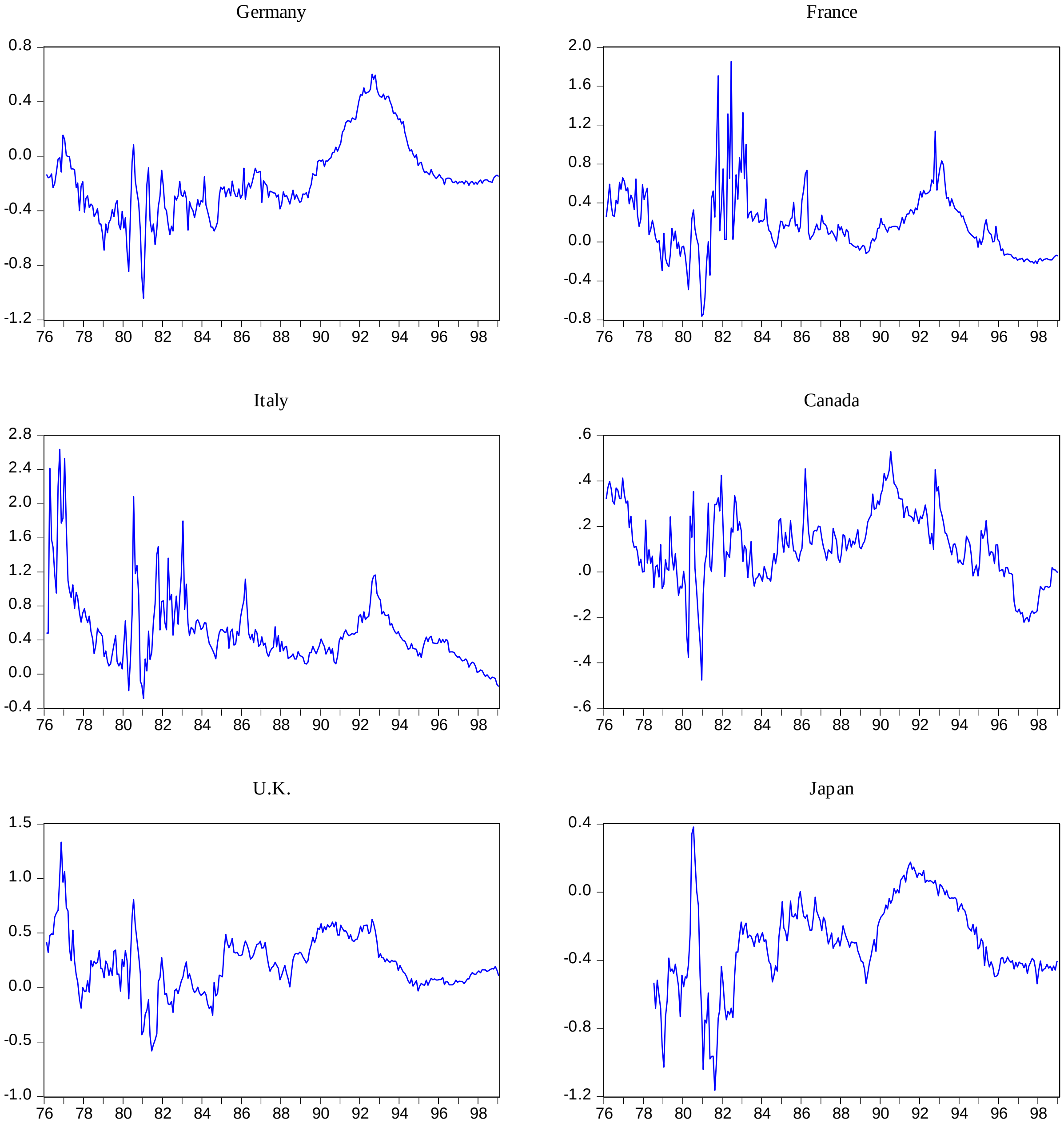
Note: the numbers in parentheses denote standard errors; those in square brackets are the 95% highest posterior density (HPD) regions with mode in bold for the break years for each country.

Table 5. Summary of break dates for the forward discount

Country	k_1	k_2	k_3	k_4	k_5
Canada	78:02	83:08	92:10	92:11	
	(77:01)	(80:10)	(95:10)		
France	81:05	83:05	96:06	91:12	96:5
	(78:01)	(81:03)	(83:01)	(95:09)	
Germany	87:03	93:07			
	(77:05)	(84:08)	(89:05)	(90:11)	(94:01)
Italy	83:05	93:04			
	(77:01)	(81:04)	(82:11)	(96:04)	
Japan	82:10				
	(—)				
U.K.	85:04	93:02			
	(77:01)	(80:06)	(81:07)	(84:08)	(92:08)

Note: Designs I and II refer to equations (1) and (2), respectively. *k* denotes the time of break over 1976:01~1998:12, or 1978:07~1998:12 for Japan. The dates given in the parentheses are the estimated break dates in Sakoulis *et al*. (2010) in which they do not find any significant break for the Japanese yen.

Fig. 1. Forward Discount for G7 Countries



Source: Sakoulis, Zivot and Choi (2010).

Fig. 2(a). Canada

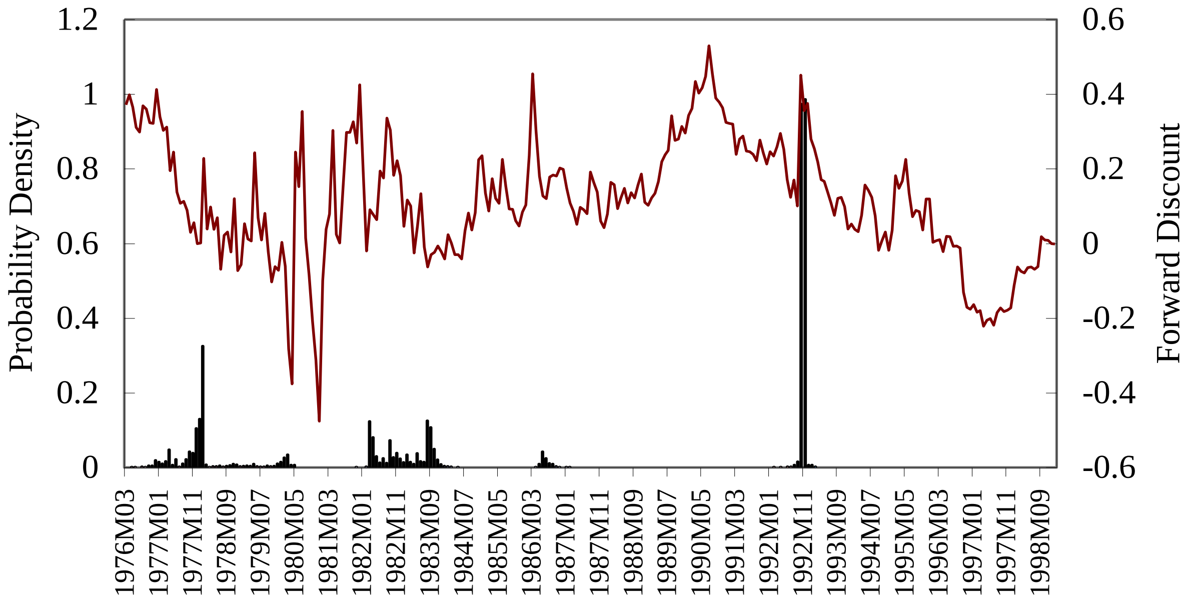


Fig. 2(b). France

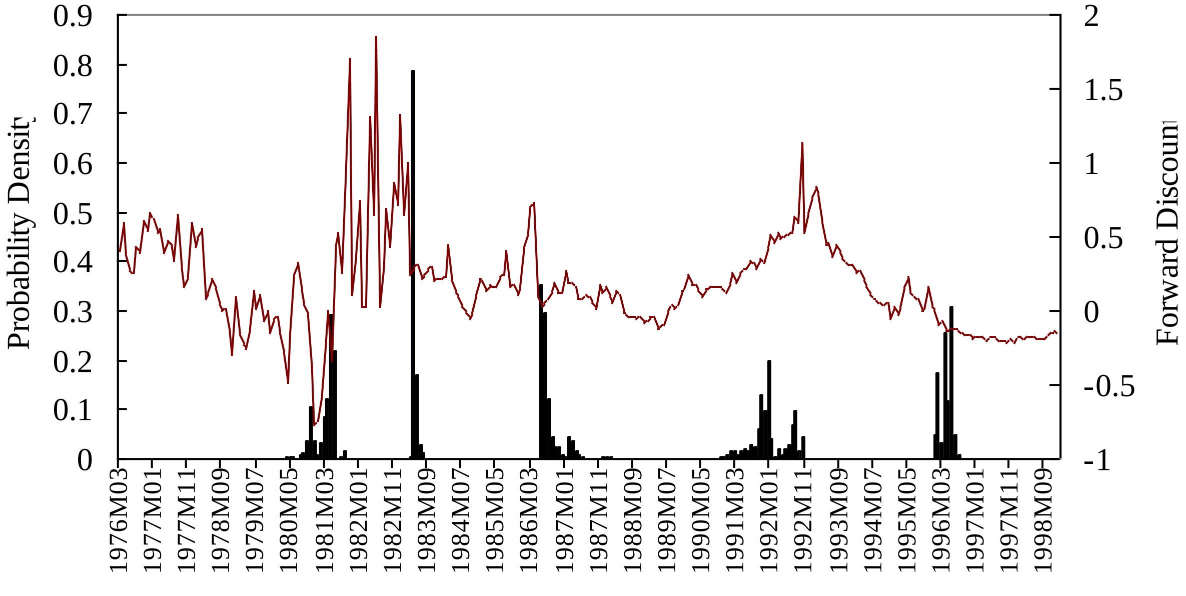


Fig. 2(c). Germany

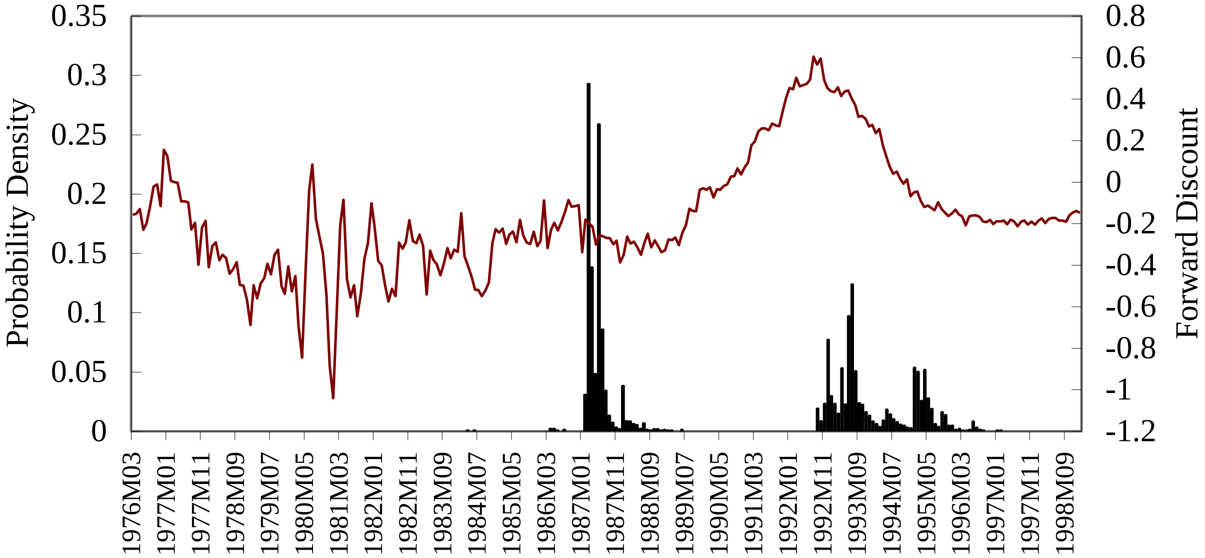
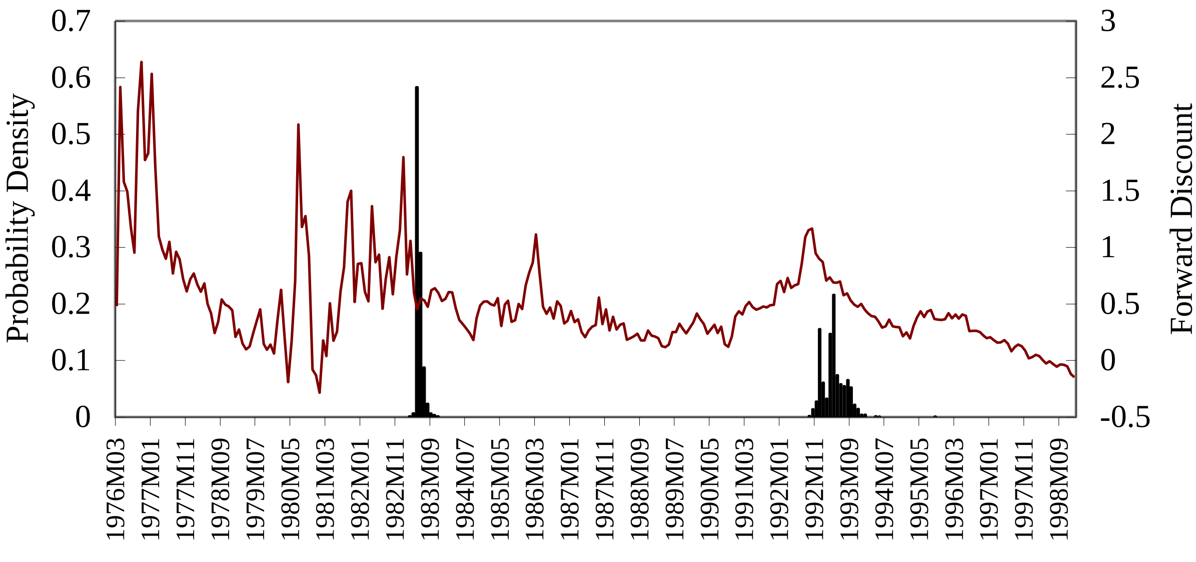


Fig. 2(d). Italy



Economics Bulletin, 2011, Vol. 31 no.2 pp. 1807-1826

Fig. 2(e). Japan

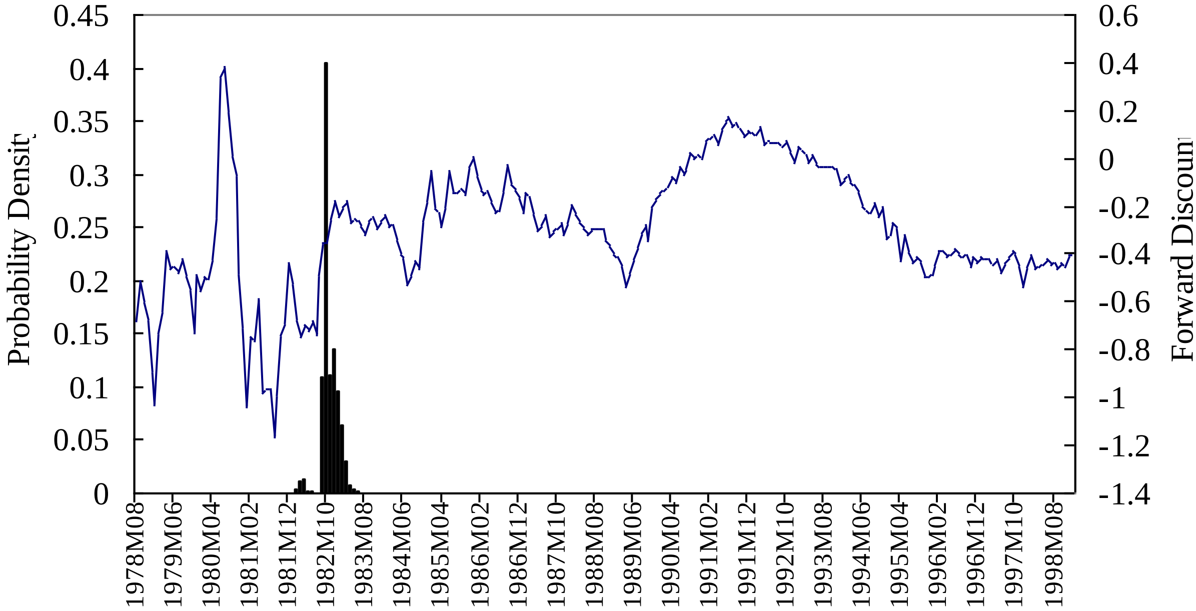


Fig. 2(f). UK