Pricing foreign currency options with stochastic volatility* 1

Angelo Melino, Stuart Turnbull

Journal of Econometrics

Journal of Econometrics 45 (1990) 239–265. North-Holland

PRICING FOREIGN CURRENCY OPTIONS WITH STOCHASTIC VOLATILITY*

Angelo MELINO and Stuart M. TURNBULL

University of Toronto, Toronto M5S 1A1, Canada

This paper investigates the consequences of stochastic volatility for pricing spot foreign currency options. A diffusion model for exchange rates with stochastic volatility is proposed and estimated. The parameter estimates are then used to price foreign currency options and the predictions are compared to observed market prices. We find that allowing volatility to be stochastic results in a much better fit to the empirical distribution of the Canada–U.S. exchange rate, and that this improvement in fit results in more accurate predictions of observed option prices.

1. Introduction

Recent attempts to use option models to price foreign currency options have been relatively disappointing. Although observed foreign currency option prices satisfy the boundary conditions that are implied by all option pricing models [Bodurtha and Courtadon (1986)], attempts to specify models that lead to accurate point forecasts of option prices have not been very successful.

Bodurtha and Courtadon (1987) assume a lognormal probability distribution for exchange rates and predict option prices using implied volatilities. The implied volatilities are chosen so as to maximize the correspondence between their model's predictions and the option prices observed on the previous market day. They find that this procedure produces fairly small biases in the average predicted price for most categories of options, although on average they consistently overestimate call and put option prices. Bodurtha and Courtadon also report that their model produces pricing errors that have a large dispersion. In particular, the average ratio of the absolute forecast error to the actual price is about 13 percent for both the puts and the calls.[1] Melino and Turnbull (1987) question the assumption of lognormal-

*Financial support for this paper was provided by the Social Sciences and Humanities Research Council of Canada under Grant 410-87-0850. This paper has benefited from comments by Ken Singleton and from workshop presentations at the National Bureau of Economic Research, Berkeley, and Stanford. Ken Vetzal provided excellent research assistance. We are responsible for any remaining errors.

[1]See tables 4b, 5b, and 4a, respectively, in Bodurtha and Courtadon (1987).

ity. They consider a class of probability distributions for the exchange rate process which includes the lognormal as a limiting case, and use historical estimates of volatility to price the options. None of their models perform particularly well. The relatively poor performance of their models arises because the volatility estimates from actual exchange rate data are significantly smaller than those implied by observed option prices.

There are many potential explanations for the poor predictive performance of the foreign currency option models. A leading candidate is misspecification in the assumed distribution of exchange rates. Boothe and Glassman (1987) attempt to determine the appropriate probability distribution with which to describe changes in exchange rates over the period 1973–1984. They report that the distributions of daily and weekly exchange rate changes display kurtosis far in excess of the normal distribution, but find that the normal distribution provides a good approximation to the distribution of quarterly changes in exchange rates. This tendency of the distribution of exchange rate changes to resemble more closely the normal as we take longer differencing intervals suggests that the standard central limit theorems apply to daily exchange rate changes. This casts doubt on models (such as the stable Paretian) which account for the thick tails of the distribution of daily exchange rate changes but have infinite variance. In a similar period, Melino and Turnbull (1987) find evidence that the volatility parameter is highly unstable over time. Such a finding is important because, as noted by Press (1968), McFarland et al. (1982), and Engle (1982), finite but heteroskedastic variances can generate the fat tail feature that seems to characterize exchange rate distributions, but heteroskedastic variances also are consistent with the limiting normal distribution reported by Boothe and Glassman (1987).

The presence of stochastic volatility has important implications for option pricing. The effects of stochastic volatility upon stock option prices have recently been examined by Hull and White (1987), Johnson and Shanno (1987), Scott (1987), and Wiggins (1987). These papers demonstrate that European predicted option prices tend to be less than Black–Scholes option prices for at the money options. If the option is deep in the money, predicted option prices are usually greater than Black–Scholes. For deep out of the money options, the results are sensitive to the parameters of the stochastic process describing changes in volatility and the correlation between changes in volatility and stock price. Cooper et al. (1986) examine the pricing of European foreign currency options. They estimate the parameters of the stochastic process describing changes in volatility and then compare predicted prices to Black–Scholes prices. Their results are similar in nature to those found in the studies on stock options.

In this paper, we focus on obtaining a closer correspondence to the empirical distribution of exchange rates and on the subsequent consequences for option pricing. The first objective of this study is to estimate the parameters of the stochastic process describing changes in exchange rates and volatility, assuming that volatility is stochastic. The fundamental difficulty in the estimation process is that volatility cannot be directly observed. Cooper et al. (1986), Scott (1987), and Wiggins (1987) use a method of moments approach to estimate the parameters. They found that the parameter estimates were sensitive to the moments which they fitted, but they were unable to pool the available information or to test if the different parameter estimates which they obtained were due simply to sampling error. We generalize their approach along the lines of Hansen (1982), so that we can address these issues.

The second objective of this paper is to examine whether the consideration of stochastic volatility translates into important differences in the implied option prices. The estimated parameters for the volatility process and an estimate of the volatility at each point in time are used to price the corresponding currency options. Predictions then are compared to observed option prices, using transactions data from the Philadelphia Exchange (PHLX). Due to very high computing costs, only one foreign currency is considered: the Canadian dollar.

The paper is organized as follows: Section 2 reviews briefly the pricing of foreign currency options when volatility is stochastic. The data are examined in section 3. The econometric strategy is described and the estimates of the volatility process are given in section 4. Section 5 presents the results of the option pricing model and compares the predicted to the observed option prices. Summary and conclusions are given in section 6.

2. The pricing of foreign currency options

We assume the following:

A1. No transactions costs, no differential taxes, no borrowing or lending restrictions, and trading takes place continuously.

A2. The term structure of interest rates in both the domestic and foreign country are flat and nonstochastic.

A3. The underlying state variables are the spot exchange rate (S) and the level of volatility (v).

A4. The transition probabilities for the state variables are summarized by

$$\mathrm{d}S = (a + bS)\,\mathrm{d}t + vS^{\beta/2}\,\mathrm{d}Z_S \tag{1}$$

and

$$\mathrm{d}v = \mu(v)\,\mathrm{d}t + \sigma(v)\,\mathrm{d}Z_v, \tag{2}$$

where Z_S and Z_v are standard Wiener processes whose increments have instantaneous correlation ρ; a, b, and β are parameters; $\mu(v)$ is the instantaneous mean and $\sigma^2(v)$ is the instantaneous variance of the volatility process. It is assumed that $0 \leq \beta \leq 2$, and that zero is an absorbing barrier for S.

Assumptions A1 and A2 are standard.[2] Economic theory provides little guidance on the choice of state variables or on their equations of motion. Many authors assume that the exchange rate itself is a sufficient statistic for its transition probabilities[3] and posit equation (1) with $\beta = 2$ and constant volatility. This model was investigated by Melino and Turnbull (1987), where it was found to be inadequate. The model (i) failed to adequately capture several empirically important features of the distribution of daily exchange rate changes, and (ii) using historical estimates of volatility, it led to predicted option prices that were systematically lower on average than observed transactions prices. Including volatility as a state variable with transition probabilities given by (2) may help account for these inadequacies.

Details about the historical behavior of the Canada–U.S. exchange rate are provided in section 3. At this stage, it is useful to point out three important stylized facts which any descriptive model of exchange rate movements must capture: (a) daily exchange rate changes are virtually unpredictable, (b) squared daily exchange rate changes have a nontrivial dynamic structure that suggests autoregressive conditional heteroskedasticity, and (c) the distribution of daily exchange rate changes has relatively fat tails. The lack of predictability of daily exchange rates can be fit in our specification by drift parameters, a and b, in eq. (1) that are essentially zero.[4] The remaining two features must be accounted for by the diffusion term.

With volatility (v) constant, the process given by (1) does generate exchange rate changes that display conditional heteroskedasticity, that is, heteroskedasticity that depends upon the level of lagged exchange rates, unless $\beta = 0$. Because the exchange rate displays a nontrivial dynamic structure, this conditional heteroskedasticity shows up (*inter alia*) as nonzero

[2] A more complete treatment would also relax A2, include the domestic and foreign interest rates as state variables, and allow for a possible reward for interest rate risk. While conceptually straightforward, allowing for stochastic volatility and stochastic foreign and domestic interest rates presents an imposing computational challenge. We leave this for future research.

[3] For example Grabbe (1983), Shastri and Tandon (1986), and Bodurtha and Courtadon (1987).

[4] Although economic theory has a good deal to say about what causes exchange rates to move, we have little interest in considering specifications involving more than a linear drift for the exchange rate process because with high frequency data (basically daily) of the sort we plan to use (i) the drift is notoriously difficult to estimate and (ii) the diffusion parameter estimates are fairly robust to the precise specification of the drift (because the centered and uncentered second moments converge). In addition, the drift parameters do not appear explicitly in the pricing equations for the currency options.

auto-correlations of squared exchange rate changes. In addition, the conditional heteroskedasticity leads to fatter tails and higher kurtosis. Therefore, nonstochastic volatility is consistent with the important qualitative features of the exchange rate distribution described above. Nonetheless, it is not satisfactory because it cannot account for the magnitudes that appear in actual data. The values of the autocorrelations of squared daily changes and excess kurtosis that can be generated by (1) depend critically upon both the sample size and the level of volatility. Two limiting arguments can be used to clarify this dependence. For a fixed level of volatility, these values become very large as the sample size increases.[5] However, for a fixed sample size, these values will be virtually zero, if v is small enough. Some simulations that we have performed using estimates of v based on historical data of the Canada–U.S. exchange rate indicate that values of the autocorrelations of squared daily changes and excess kurtosis that can be generated by (1) are virtually zero, even for sample sizes that correspond to 15 years of daily data. Related simulations show that allowing v to be stochastic considerably increases the size of the autocorrelations and excess kurtosis that can be generated.

There are many specifications that one might consider, *a priori*, for the v process. Our empirical work will assume that volatility changes can be described by

$$\mathrm{d}\ln v = (\alpha + \delta \ln v)\,\mathrm{d}t + \gamma\,\mathrm{d}Z_v, \tag{2'}$$

so that $\mu(v) = v(\alpha + \gamma^2/2 + \delta \ln v)$ and $\sigma(v) = \gamma v$. This specification of the volatility process is admittedly ad hoc.[6] It was chosen because it respects the nonnegativity of volatility and because it is tractable. Similar specifications have been suggested by Cooper et al. (1986), Scott (1987), and Wiggins (1987).

Melino and Turnbull (1987) found that the option prices predicted by A1, A2, and eq. (1) with constant volatility (estimated from the historical data) systematically underpredicted observed transactions prices. In their sample, Canadian dollar call options were underpredicted by just over 25%, on average. For the other currencies they examined, predicted call and put option prices ranged from 4% to 22% lower, on average, than corresponding

[5]For example, consider the case where $\beta = 2$ and we take the initial exchange rate, S_0, as fixed. $\mathrm{E}(\Delta S_t^4)/\mathrm{E}^2(\Delta S_t^2)$ is unbounded for large t, and, using values of v of the order suggested by the historical estimates, the $\mathrm{cov}(\Delta S_t^2, \Delta S_{t-1}^2)/\mathrm{var}(\Delta S_t^2)$ goes to a value slightly in excess of 1/3.

[6]A more general specification of the volatility process would allow for a possible dependence of volatility on days of the week, or perhaps on whether markets were open. We chose not to consider this more general specification because it would introduce formidable difficulties into the econometric estimation and the solution of the predicted option prices.

observed option prices. Treating volatility (v) as stochastic and generated by (2′) may account for this systematic underprediction.

Given A1 to A4, the price of options written on spot foreign currency can be developed using familiar methods.[7] The option price (C) satisfies the following partial differential equation:

$$\tfrac{1}{2}\left(v^2 S^\beta C_{SS} + 2\rho\sigma v S^{\beta/2} C_{Sv} + \sigma^2 C_{vv}\right)$$

$$+ (r_D - r_F) S C_S + (\mu(v) - \lambda\sigma(v)) C_v - r_D C + C_t = 0, \tag{3}$$

where r_D and r_F denote the domestic and foreign risk-free rates, respectively, and λ is a risk premium that arises because volatility is not a traded asset.

Note that the drift in the exchange rate does not appear in eq. (3). The derivation of the option price does assume, however, knowledge of all the parameters of the volatility process (including its drift), as well as the level of volatility and the risk premium at each point in time. Section 4 describes a method to estimate the parameters and level of the volatility process.

Without an explicit description of endowments, preferences, and technology, it is difficult to say very much about the risk premium. To rule out arbitrage we must require $\lambda = 0$ if $v\sigma(v) = 0$. With specification (2′), however, this constraint is never binding. Many authors simply assume that λ is identically zero. This follows from the assumption that innovations to exchange rate volatility are uncorrelated with innovations to aggregate consumption. It seems worthwhile to consider more general models for λ that include $\lambda = 0$ as a special case. For computational reasons, it is useful to restrict attention to functional forms for the risk premium that are smooth functions and involve at most the exchange rate, the level of volatility and time. As a first approximation we shall simply assume that λ is a constant.[8] This can be justified if preferences are logarithmic. We will treat the risk premium, λ, as a free parameter in our empirical work and try to infer its value from observed option prices.

When v is constant, (3) simplifies considerably and reduces to the expression derived in Melino and Turnbull (1987). Explicit solutions for the option price satisfying (3) are not available, except for certain special cases. However, eq. (3) can be solved (subject to the appropriate boundary conditions) by numerical methods.

[7]See Cox, Ingersoll, and Ross (1985).

[8]It would be interesting to investigate alternative models for λ, especially in light of the sensitivity of predicted option prices to λ reported in section 5. We leave this for future research.

Table 1a
Canada–U.S. exchange rate properties, descriptive statistics.

Variable	Mean	Standard deviation	Skewness	Excess kurtosis	Minimum value	Maximum value
		75/01/02–86/12/10 (3011 observations)				
S	85.116	8.94	0.37	−0.75	69.498	103.863
ΔS	−0.009	0.197	−0.31	4.98	−1.886	1.073
$(\Delta S)^2$	0.039	0.103	16.30	474.66	0.000	3.556
$\lvert\Delta S\rvert$	0.143	0.135	2.51	14.45	0.000	1.886
		75/01/02–80/12/31 (1513 observations)				
S	92.096	6.51	0.26	−1.48	82.590	103.863
ΔS	−0.011	0.207	−0.56	6.27	−1.886	1.073
$(\Delta S)^2$	0.043	0.124	17.41	447.57	0.000	3.556
$\lvert\Delta S\rvert$	0.152	0.140	2.89	20.44	0.000	1.886
		81/01/02–86/12/10 (1498 observations)				
S	78.067	4.37	−0.23	−1.46	69.498	84.983
ΔS	−0.007	0.187	0.52	2.83	−0.993	0.920
$(\Delta S)^2$	0.035	0.077	5.59	42.60	0.000	0.851
$\lvert\Delta S\rvert$	0.134	0.130	2.02	5.97	0.000	0.923
		83/02/28–84/02/10 (241 observations)				
S	80.979	0.47	−0.74	−0.27	79.853	81.806
ΔS	−0.005	0.101	−0.14	0.30	−0.275	0.329
$(\Delta S)^2$	0.010	0.016	2.87	10.59	0.000	0.109
$\lvert\Delta S\rvert$	0.079	0.064	1.12	1.15	0.000	0.329
		84/02/13–85/01/24 (240 observations)				
S	76.796	1.31	1.06	0.34	74.923	80.321
ΔS	−0.020	0.150	0.01	2.12	−0.588	0.636
$(\Delta S)^2$	0.023	0.046	4.81	31.10	0.000	0.405
$\lvert\Delta S\rvert$	0.112	0.102	1.81	4.68	0.000	0.636

3. Data description

Our empirical work requires the following data: (i) the spot Canada–U.S. exchange rate, (ii) the price of exchange rate options written on the Canadian dollar, and (iii) the term structure of interest rates for Canada and the U.S.

Daily data on the spot Canada–U.S. exchange rate for the period January 2, 1975 to December 10, 1986 were obtained from the Bank of Canada. These data refer to rates prevailing at mid-day on the interbank market in Canada. We used these data to identify the stochastic process for the exchange rate. Note that we *did not* use these data directly to price the exchange rate options (see below).

Table 1a provides descriptive statistics for the exchange rate over various sample periods. The data are expressed as the number of U.S. cents required to purchase one Canadian dollar. The change in the Canada–U.S. exchange rate tends to display a small negative mean over our sample. The average daily change, however, is swamped by daily volatility as proxied by either the average absolute change or the daily standard deviation. These proxies for volatility are fairly stable if we look at very long samples, but the sharp differences in the estimates obtained from the two shorter samples are too large to be rationalized by sampling error. Changes in the daily Canada–U.S. exchange rate display virtually no skewness, but consistently display positive excess kurtosis.[9]

The dynamic properties of exchange rates are examined in table 1b. The autocorrelations for the level of the Canada–U.S. exchange rate strongly indicate a random walk component. The change in daily exchange rates is virtually white noise. Because of the size of the sample, one can reject the hypothesis that the population autocorrelations of exchange rate changes are exactly zero. However, the estimated autocorrelations are very small which implies there is very little linear structure in the mean of exchange rate changes. By contrast, volatility, as proxied by either the square or the absolute value of daily exchange rate changes, displays a very interesting dynamic structure. The autocorrelations are well below one and they initially damp toward zero. However, the higher autocorrelations stay positive and seem to cluster around a small positive constant for very many more than the twelve autocorrelations reported here. Also, the positive constant around which the higher autocorrelations cluster declined when we looked at shorter subsamples. As pointed out by Schwert (1987), this behavior of the higher autocorrelations indicates a random walk component in both of these proxies for volatility. Note that finding a random walk component in the squared or the absolute daily change does not mean that volatility as defined in eq. (1) has a random walk component. According to our specification, the autocorrelations of these proxies for volatility will depend upon the dynamic structure of volatility *and* the level of exchange rates (unless $\beta = 0$). Therefore, the random walk component in these proxies for volatility may simply reflect the random walk component in the level of exchange rates. The bottom half of table 1b looks at $\rho(|\Delta S(t)|, \Delta S(t-j))$, the cross-correlations between the absolute daily change and the daily change itself. The correlations between the current absolute daily change and future levels of the daily change are small and display no consistent pattern which implies that the size of absolute changes has no predictive content for subsequent changes in the level of the exchange rate. However, the contemporaneous correlation and the correlation with lagged levels of the daily change are consistently negative and often large. This implies that large declines in the exchange rate are associated with large absolute movements on the current and subsequent

[9]Excess kurtosis refers to the value of the kurtosis coefficient in excess of the normal. Cramér (1946, §15.8) calls this the 'coefficient of excess'.

Table 1b

Canada–U.S. exchange rate, descriptive statistics, 75/01/02–86/12/10.

	Autocorrelations												
	1	2	3	4	5	6	7	8	9	10	11	12	
S	0.999	0.998	0.997	0.996	0.994	0.993	0.992	0.991	0.990	0.988	0.987	0.986	
ΔS	0.095	0.006	0.003	0.009	0.064	0.021	0.044	0.002	0.038	−0.007	−0.011	−0.002	
$(\Delta S)^2$	0.271	0.149	0.114	0.123	0.076	0.093	0.049	0.058	0.057	0.062	0.055	0.044	
$\|\Delta S\|$	0.232	0.190	0.162	0.164	0.161	0.150	0.122	0.123	0.103	0.092	0.097	0.099	
	Crosscorrelations $\rho(\|\Delta S(t)\|, \Delta S(t-j))$												
	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0
	1	2	3	4	5	6	7	8	9	10	11	12	
	0.051	0.042	0.033	−0.009	0.033	0.024	0.034	0.014	0.011	0.053	−0.001	−0.044	−0.104
	−0.081	−0.067	−0.077	−0.061	−0.048	−0.033	−0.060	−0.001	−0.035	−0.040	−0.036	−0.020	

days. Increases in the exchange rate, therefore, tend to precede periods of lower volatility.

To price foreign currency options it is necessary to obtain simultaneously sampled data on option prices, spot currency prices, and domestic and foreign interest rates for default-free claims matching the maturities of the option contracts. Our source for simultaneous option prices and currency prices is the transaction surveillance report compiled by the Philadelphia Exchange and described by Bodurtha (1984). For each option trade, the following data are recorded (*inter alia*): the date of the trade, the time of the trade; maturity, exercise price, option price; prevailing bid and ask spot quotes at the time of the trade, and value of the last spot quote reported by Telerate from the interbank market. Our sample of option trades runs from February 28, 1983 to January 24, 1985.

For short-term interest rates, daily closing quotes for 90 and 180 day Canadian and U.S. Treasury bills are used. This data were supplied by the Bank of Canada. For Canada, these data were supplemented by using quotes on 30-, 60-, 90-, and 180-day Treasury bills, collected on a weekly basis and provided by a leading Canadian securities firm.

4. Empirical results

4.1. Econometric methods

We assume that the model described by eqs. (1) and (2′) holds in calendar (as opposed to market) time. We have no direct measurements on volatility, so inferences about the parameters in (1) and (2′) must be based on a set of unevenly spaced observations $\{S(t_1), \ldots, S(t_n)\}$ on the exchange rate. The uneven spacing reflects weekends, trading holidays, and missing observations.

It does not appear possible to obtain closed form expressions for the exact conditional densities generated by eqs. (1) and (2′), except for certain special cases (such as v constant). Our strategy will be to approximate the continuous time stochastic process for exchange rates and volatility by a discrete time stochastic process. We will then proceed as if the data are generated from the discrete time model to recover parameter estimates.

Define $h_i = t_i - t_{i-1}$ and $h = \min\{h_i\}$. The continuous time process described by (1) and (2′) will be approximated by

$$S(t_i) = ah_i + (1 + bh_i)S(t_{i-1}) + v(t_{i-1})S(t_{i-1})^{\beta/2}h_i^{1/2}e(t_i), \qquad (4)$$

$$\ln v(t_i) = \alpha h + (1 + \delta h)\ln v(t_i - h) + \gamma h^{1/2}u(t_i), \qquad (5)$$

where

$$\begin{bmatrix} e(t_i) \\ u(t_i) \end{bmatrix} \sim \text{i.i.d. } \mathrm{N}_2\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right).$$

There are an infinity of discrete time models which differ only by terms of order $o(h)$ that converge to (1) and (2) as $\max\{h_i\} \to 0$.[10] *A priori*, it is not clear which of the discrete time models will serve as the best approximation. Eqs. (4) and (5) display a certain asymmetry because the unobserved volatility process has even spacing, but the exchange rate process has irregular spacing that matches the available data. This asymmetry turns out to be unimportant in our application but serves to simplify our calculations. If $\delta < 0$ and we assume that appropriate initial conditions are satisfied, then the even spacing in the discrete time approximation to the unobserved volatility process leads to the result that volatility will be stationary and, in particular, that

$$\ln v_t \sim \mathrm{N}(\mu_v, \sigma_v^2), \tag{6}$$

where

$$\mu_v = -\alpha/\delta \quad \text{and} \quad \sigma_v^2 = h\gamma^2 \big/ \left[1 - (1 + \delta h)^2\right].$$

The quality of the discrete time approximation depends (among other things) upon how tightly sampled are the data. In our application, the mean value of h_i is approximately 1.45 days and h_i never exceeds 5 days. With v constant, it is possible to obtain the exact conditional density for exchange rates and test the quality of the discrete time approximation directly. We did so in previous work [Melino and Turnbull (1987)] and found that the Gaussian quasi-likelihood implied by (4) provided an excellent (and extremely convenient) approximation to the true process.

Define $\theta = (a, b, \alpha, \delta, \gamma, \rho; \beta)$ and $w_i(\theta)$ by

$$w_i(\theta) = \frac{S(t_i) - ah_i - (1 + bh_i)S(t_{i-1})}{\left[h_i S^{\beta}(t_{i-1})\right]^{1/2}}. \tag{7}$$

The w_i represent the one-observation-ahead forecast errors normalized by a

[10] Nelson (1989) provides a set of sufficient conditions for a sequence of discrete time models to converge to a well-behaved continuous time limit.

term to reflect the contribution of observables to the conditional heteroskedasticity of these forecast errors. Given the discrete time approximation (4), w_i can also be written as

$$w_i(\theta) = v(t_{i-1})e(t_i), \tag{8}$$

so that w_i is the product of two stationary series. In general, the expectation of functions of w_i will depend upon θ. The method of moments estimates the value of θ by matching the computed sample average of these functions of w_i to their expected values. There are infinitely many functions of the w_i that may be considered. In this paper, we consider the following:

$$w_i^m(\theta), \qquad m = 1,2,3,\ldots, \tag{9a}$$
$$|w_i^m(\theta)|, \qquad m = 1,2,3,\ldots, \tag{9b}$$
$$w_i(\theta)w_{i-j}(\theta), \qquad j = 1,2,3,\ldots, \tag{9c}$$
$$|w_i(\theta)w_{i-j}(\theta)|, \qquad j = 1,2,3,\ldots, \tag{9d}$$
$$w_i^2(\theta)w_{i-j}^2(\theta), \qquad j = 1,2,3,\ldots, \tag{9e}$$
$$|w_i(\theta)|w_{i-j}(\theta), \qquad j = 0, \pm 1, \pm 2, \pm 3,\ldots. \tag{9f}$$

Expressions for the unconditional expectation of these functions are provided in an appendix that is available upon request. The list of moments was chosen with an eye to three criteria: familiarity, identification, and efficiency. The moments in (9a), (9c), and (9e) have been estimated and discussed by various authors, so we began with them. After some Monte Carlo simulations, we realized that including the absolute moments in (9b) and (9d) provided sharp increases in efficiency. These simulations also indicated that the sampling variability of the sample averages of the functions in (9a)–(9b) increase dramatically with m, so we only consider moments up to $m = 4$ in our empirical work. Some of these functions provide information about only a subset of the parameters. For example, the odd powers of w_i all have expectation zero, at the true value of θ. These moments can be used to infer a, b, and β, but they are uninformative about the remaining components of θ. Looking also at the absolute moments allows us to infer the values of μ_v and σ_v^2. The functions in (9c)–(9f) summarize some of the dynamic properties of exchange rates. These properties allow us to distinguish the parameters α, δ, and γ (which appear in μ_v and σ_v^2), allow us to identify ρ, and provide additional information about the remaining parameters. It turns out that consideration of (9f) is essential to uniquely identify ρ. In the appendix, it is shown that

$$\mathrm{E}|w_i|w_{i-1} = (2/\pi)^{1/2}\rho\gamma\exp\{2\mu_v + 2\sigma_v^2\}.$$

The sign of this expression is determined entirely by ρ. If $\rho < 0$, for example, then unexpected decreases (increases) in the exchange rate S tend to precede a period with high (low) volatility. The opposite is true if $\rho > 0$.

One way to estimate θ is to choose (judiciously) as many functions from (9) as there are parameters to estimate, and then match the sample averages of these functions to their expected values. This is essentially the approach suggested by Cooper et al. (1986), Scott (1987), and Wiggins (1987). We follow Hansen (1982) who provides a more general framework which allows us to optimally pool the information in as many functions of the w_i as we wish to consider.

Let $f_i(\theta) \in R^p$ denote a vector whose components are functions of w_i drawn from (9) less their unconditional expectations. Define $g_n(\theta)$ by

$$g_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} f_i(\theta). \tag{10}$$

We wish to choose an estimator which makes $g_n(\theta)$ 'small'. More precisely, our estimator is

$$\hat{\theta}_n = \underset{\theta\in\Theta}{\arg\min}\, g_n'(\theta)\hat{W}_n g_n(\theta), \tag{11}$$

where Θ denotes the permissible parameter space and $\hat{W}_n$ is a (possibly random) positive definite weighting matrix. Following previous work [Marsh and Rosenfeld (1983), Melino and Turnbull (1987)] we will minimize the expression given in eq. (11) for a grid of values for β. Under suitable regularity conditions,[11] $\hat{\theta}_n$ will be consistent and asymptotically normal, i.e.,

$$n^{1/2}(\hat{\theta}_n - \theta) \sim \mathrm{AN}(0, V_n). \tag{12}$$

[11]There are some delicate issues that prevent an immediate mapping into Hansen (1982). If the drift parameters are truly zero (which is not a bad first approximation), then there is a unit root in the level of the exchange rate. Also, because 0 is an absorbing barrier, for some parameter configurations the exchange rate process will not be ergodic. Finally, it is technically very difficult to deal rigorously with the approximation error introduced by using the discrete time model. We refer the reader to Duffie and Singleton (1988) for a discussion of these issues. Although we realize that there are technical difficulties in invoking the usual large sample theory results, simulations indicate that they are a useful guide to the finite sample distribution of g_n. A unit root in the level of the exchange rate would clearly complicate the distribution theory for the drift parameters, but, because they are virtually independent, the asymptotic distribution theory for the remaining parameters should still be a good guide.

The covariance matrix of the asymptotic normal distribution, V_n, can be consistently estimated by

$$\hat{V}_n = \left(D_n' \hat{W}_n D_n \right)^{-1} D_n' \hat{W}_n \hat{\Sigma}_n {}' \hat{V}_n D_n \left(D_n' \hat{W}_n D_n \right)^{-1}, \tag{13}$$

where $D_n'(\theta)$ is the Jacobian matrix,[12]

$$D_n'(\theta) = \mathrm{d}g_n(\theta)/\mathrm{d}\theta, \tag{14}$$

evaluated at any consistent estimator of θ, and $\hat{\Sigma}_n$ is any consistent estimator of Σ_n given by

$$\Sigma_n = \mathrm{E}\left(\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} f_i(\theta) f_j(\theta)' \right). \tag{15}$$

Our empirical work uses the Newey–West (1987) estimator of $\hat{\Sigma}_n$.[13] The asymptotic covariance matrix is minimized by choosing $\hat{W}_n = \hat{\Sigma}_n^{-1}$ in which case the expression in (13) also simplifies considerably.

In practice, we fix a value of β and then obtain initial consistent estimators of the remaining components of θ by fitting exactly an initially chosen vector $f_i(\theta)$ with $p = 6$ components. Then we construct a vector $f_i(\theta)$ with a large number ($p = 47$) of components. The initial consistent estimates of θ are used to construct $\hat{\Sigma}_n$ and $\hat{W}_n = \hat{\Sigma}_n^{-1}$, and we again minimize (11) to obtain efficient estimators. Although there is no asymptotic advantage, to reduce the importance of our initial choice of $f_i(\theta)$, we iterate to approximately a fixed point before reporting our results.

[12]Strictly speaking, $\mathrm{d}g_n(\theta)/\mathrm{d}\theta$ may not exist in any given sample, because the derivatives involving absolute moments have contributions from the individual observations, such as $\mathrm{d}|w_i(\theta)|/\mathrm{d}a$, which are not defined at $w_i(\theta) = 0$. Formally, the derivatives exist almost everywhere, so if we encounter a value of $w_i(\theta) = 0$ during our calculations, we are entitled to set its contribution to $\mathrm{d}g_n(\theta)/\mathrm{d}a$ to anything we like. Because of floating point arithmetic, this situation did arise in our work, and we chose to set the contribution to zero.

[13]There was a good deal of serial correlation in the components of $f_i(\theta)$. To decide on the number of autocorrelations to include in the estimate of the covariance matrix, we used the initial consistent estimates of θ and undertook a Monte Carlo study. We generated a collection of samples with 3,000 observations. For each sample, we constructed the Newey–West estimator for different lag lengths (up to 300) and constructed $g_n(\theta)$ as in eq. (10). We then averaged the Newey–West estimators across the samples and compared them to the covariance matrix for $g_n(\theta)$. Based on this comparison, we decided to include 50 autocorrelations in all our subsequent calculations.

Table 2

Canada–U.S. exchange rate, parameter estimates, 75/01/02–86/12/10.[a]

Beta	a	b	α	δ	γ	ρ	$ng'Wg$
0.0	0.037	$-0.0^{3}48$	−0.252	−0.127	0.190	−0.081	41.53
	(0.092)	$(0.0^{2}11)$	(0.083)	(0.042)	(0.031)	(0.088)	
1.0	0.042	$-0.0^{3}54$	−0.384	−0.091	0.153	−0.110	39.36
	(0.093)	$(0.0^{2}11)$	(0.164)	(0.039)	(0.033)	(0.110)	
2.0	0.044	$-0.0^{3}57$	−0.745	−0.116	0.192	−0.188	41.33
	(0.087)	$(0.0^{2}10)$	(0.204)	(0.032)	(0.028)	(0.090)	

[a]The parameter estimates were obtained from fitting the following moments:

$$
\begin{array}{ll}
w_i^m(\theta), & m = 1,3, \\
|w_i^m(\theta)|, & m = 1,\ldots,4, \\
w_i(\theta)w_{i-j}(\theta), & j = 1,\ldots,10, \\
|w_i(\theta)w_{i-j}(\theta)|, & j = 1,\ldots,10, \\
w_i^2(\theta)w_{i-j}^2(\theta), & j = 1,\ldots,10, \\
|w_i(\theta)|w_{i-j}(\theta), & j = 0,\ldots,10,
\end{array}
$$

Standard errors are in parentheses.

4.2. Parameter estimates

Parameter estimates for various values of β are reported in table 2. The drift parameters for the exchange rate, a and b, are imprecisely estimated, for all values of β. Although the estimates of b are negative, there is not much evidence of mean-reverting behavior in the level of exchange rates. The parameters of the volatility process are estimated fairly accurately. Allowing volatility to be stochastic appears to generate statistically significant improvements in the fit of the exchange rate distribution. The hypothesis of non-stochastic volatility, $\gamma = 0$, is overwhelmingly rejected for all the models considered. The estimates of δ are negative and significant and the point estimates indicate that shocks to volatility in the Canada–U.S. exchange rate are mostly short-lived, and there is a strong tendency to revert quickly to the mean level of volatility. The half-life of a shock to volatility implied by the parameter estimates is approximately one week. The correlation between innovations to volatility and the level of exchange rates appears to be negative, but the evidence is weak except for the case $\beta = 2$. The point estimates indicate that declines in the Canadian dollar tend to precede periods of high volatility.

Under the null hypothesis, the minimized value of the quadratic form in eq. (11) has a χ^2 distribution, with degrees of freedom equal to the difference between the number of moments fitted and the number of parameters estimated. All of the models reported in table 2 were estimated from 47 moments, so the χ^2 statistic is approximately equal to its expected value in

Table 3

Descriptive statistics from simulated samples (average and standard deviation across simulations).[a]

Variable	Mean	Standard deviation	Skewness	Excess kurtosis	Minimum value	Maximum value
S	85.999	6.29	0.63	−0.12	75.838	101.845
	2.861	1.24	0.56	1.02	2.904	1.547
ΔS	−0.007	0.20	−0.14	5.58	−1.472	1.325
	0.001	0.01	0.33	2.89	0.405	0.303
$(\Delta S)^2$	0.041	0.11	10.64	215.68	0.000	2.694
	0.003	0.02	5.46	267.35	0.000	1.521
$\|\Delta S\|$	0.142	0.14	2.62	13.05	0.000	1.595
	0.005	0.01	0.46	8.00	0.000	0.389

	Autocorrelations											
	1	2	3	4	5	6	7	8	9	10	11	12
S	0.998	0.996	0.994	0.993	0.991	0.989	0.987	0.985	0.983	0.982	0.980	0.978
	0.001	0.001	0.002	0.003	0.003	0.004	0.005	0.006	0.006	0.007	0.008	0.008
ΔS	−0.001	0.001	0.004	−0.002	0.001	0.000	0.002	0.000	0.001	−0.002	−0.003	0.001
	0.026	0.026	0.026	0.023	0.028	0.022	0.021	0.019	0.020	0.022	0.020	0.020
$(\Delta S)^2$	0.112	0.092	0.072	0.076	0.125	0.034	0.022	0.018	0.034	0.054	0.004	0.003
	0.087	0.058	0.056	0.076	0.179	0.065	0.027	0.024	0.029	0.046	0.021	0.024
$\|\Delta S\|$	0.160	0.131	0.105	0.101	0.152	0.055	0.041	0.033	0.051	0.082	0.012	0.011
	0.040	0.031	0.031	0.032	0.041	0.025	0.024	0.024	0.025	0.030	0.021	0.024

Crosscorrelations $\rho(\|\Delta S(t)\|, \Delta S(t-j))$												
−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0
1	2	3	4	5	6	7	8	9	10	11	12	
−0.001	0.003	−0.005	−0.001	0.000	0.001	0.001	−0.011	−0.003	−0.002	−0.005	−0.002	−0.065
0.023	0.020	0.025	0.020	0.023	0.021	0.019	0.028	0.026	0.027	0.024	0.028	0.057
−0.047	−0.036	−0.031	−0.027	−0.027	−0.017	−0.015	−0.012	−0.012	−0.008	−0.005	−0.003	
0.023	0.022	0.024	0.023	0.027	0.022	0.023	0.024	0.021	0.028	0.021	0.020	

[a]Number of observations in each sample = 3000. Number of samples = 100. The first row for each statistic is its mean value across the samples, and the second row is its standard deviation. The statistics were calculated from the parameter estimates corresponding to the $\beta = 2$ model.

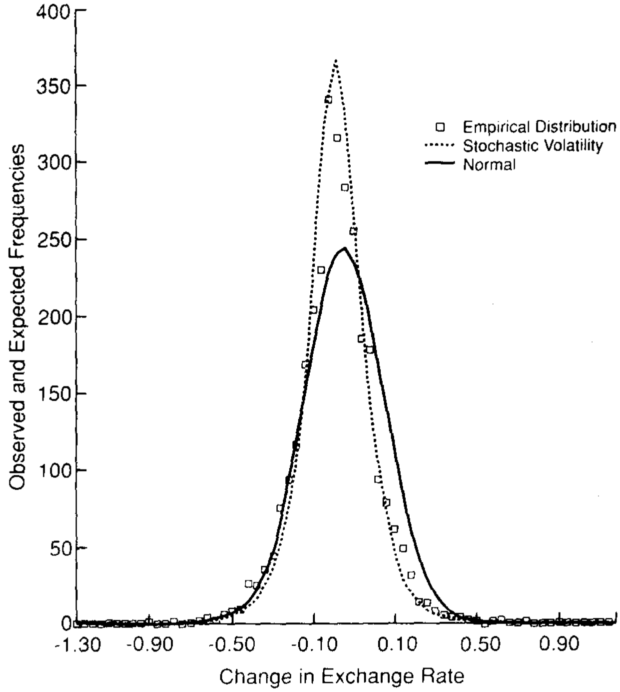

Fig. 1. A comparison of actual and predicted exchange rate changes.

all cases. There is very little variation in this goodness-of-fit statistic, so that the data appear to be indifferent to the choice of β.

The small χ^2 statistics in table 2 could be due to a good fit or to large sampling error. In either case, the statistic pertains to a model's ability to mimic the moments of w_t, while we are more directly concerned with the moments of exchange rate changes. In order to provide more information about the model's performance, exchange rate data were simulated according to the discrete time model given in eqs. (4) and (5), using the estimated parameter values. For each simulated sample of 3,000 observations, the same summary statistics as those examined in table 1 were calculated. Table 3 reports the mean and standard deviation of each statistic across the simulated samples. A good fit to the empirical exchange rate distribution requires the difference between the actual statistic and its mean across the simulated samples to be small, in some meaningful sense. A good statistical fit requires the difference to be small relative to the reported standard deviation. Because the results are virtually identical, only the results corresponding to the $\beta = 2$ case are reported.

The static properties of the exchange rate distribution (i.e., those described in table 1a) are matched very closely. Allowing for stochastic volatility generates the characteristic positive excess kurtosis of daily exchange rate changes and does an excellent job of matching the other moments as well. Fig. 1 compares the empirical distribution of exchange rate changes to that of our simulated exchange rates. The correspondence is striking, especially when compared to the normal approximation.

The fit between the actual and the simulated dynamic properties (i.e., those described in table 1b) is good. The simulated level of exchange rates display the characteristic random walk component. The simulated daily change in exchange rates generate essentially zero autocorrelations at all lags. However, the model is unable to fit the small but statistically significant autocorrelation displayed at lag one in the actual data. The model generates serial dependence in the squared daily change and the absolute daily change, but there are some noticeable differences between the actual and predicted autocorrelations. The model does not capture the large first-order autocorrelation in these proxies for volatility. The model does generate a tendency for the autocorrelations to converge to a small positive value as the order of the autocorrelations gets large, which indicates a unit root in these proxies for volatility, but it underestimates the limiting value of these autocorrelations by a large percentage. Since these autocorrelations are not accurately estimated, the discrepancy in fit is not statistically significant. Nonetheless, a specification which allows volatility to display both a transitory and a persistent component would be worth exploring in future research.

5. Comparison of actual and predicted option prices

5.1. Obtaining daily estimates of volatility

In order to price the foreign currency options, it was necessary to obtain estimates of daily volatility. Our strategy was to apply the Extended Kalman Filter (EKF), described in Anderson and Moore (1979), to the discrete time model in section 4. Because the EKF is based on a Taylor series approximation, it is sensitive to nonlinear transformations of the model. After some experimentation, eq. (5) was taken as the state transition equation. Eq. (8) was transformed by taking the logarithm of the absolute value of both sides to obtain the measurement equation, and then the standard Kalman filter was applied. The daily estimates of volatility were obtained by exponentiating the filtered estimates of log-volatility produced by the EKF. Simulations based on our estimated parameter values indicate that this procedure produces essentially unbiased estimates of volatility that have a correlation of just over 0.5 with the true volatility process.[14]

5.2. Solving the partial differential equation

For European options, solutions to eq. (3) have been derived by Johnson and Shanno (1987), Hull and White (1987), and Cooper et al. (1986). Johnson and Shanno (1987) assume that there exists a marketable asset which is perfectly correlated with volatility. All of these studies use a risk-neutral valuation approach and use Monte Carlo techniques to price the options. To price American options, Monte Carlo techniques cannot be readily used, so other numerical methods must be employed. The alternating directions method algorithm of McKee and Mitchell (1970) was used to solve (3). The accuracy of this algorithm was tested by comparing numerical solutions to the Monte Carlo solutions published in Hull and White (1987).

5.3. Empirical results

Descriptive statistics for the actual and predicted option prices are reported in table 4. To reduce computing costs, we considered only the cases $\beta = 0$ and $\beta = 2$. The predicted American option values for the stochastic volatility model are based on the parameter estimates in table 2, and are evaluated at four different values of λ. We also report the European and American option values that are implied by eq. (1) with nonstochastic volatility. For purposes of comparison, these two predicted option values were also evaluated using our daily estimates of volatility.

As in Melino and Turnbull (1987), we find that the constant volatility American and European predicted option prices have standard deviations that are comparable to observed option prices, but that they tend to underestimate the market prices. The constant volatility models underestimate the average price of a put in our sample by about 15%, and the average price of a call by more than 25%.[15] From an investment perspective, pricing errors of this magnitude are enormous. The option prices predicted by the stochastic volatility model show a good deal of sensitivity to the value of λ. Both the mean and the standard deviation of the predicted prices increase as we decrease λ. A pretty close agreement to the observed sample mean and standard deviation of observed option prices is obtained for both values of β

[14]Adjustments that incorporated the conditional variance of the filtered estimates of log-volatility led to almost identical results.

[15]Note that this underpricing occurs during a sample subperiod where the *ex post* volatility is low compared to its sample average. See table 1.

Table 4

Descriptive statistics for actual and predicted option prices.

	$\beta = 0$			$\beta = 2$		
	Mean	S.D.	RMSE	Mean	S.D.	RMSE
	Call options (2465 observations)					
Stochastic volatility						
$\lambda = 0.1$	0.3318	0.3480	0.2137	0.2985	0.3467	0.2445
$\lambda = 0.0$	0.3870	0.3572	0.1772	0.3499	0.3552	0.2050
$\lambda = -0.1$	0.4550	0.3706	0.1547	0.4183	0.3697	0.1736
$\lambda = -0.2$	0.5360	0.3882	0.1732	0.5039	0.3909	0.1756
Nonstochastic volatility						
American	0.3531	0.3520	0.2037	0.3127	0.3464	0.2342
European	0.3349	0.3289	0.2259	0.2928	0.3198	0.2573
Observed	0.4797	0.3772		0.4797	0.3772	
	Put options (1887 observations)					
Stochastic volatility						
$\lambda = 0.1$	0.4968	0.4107	0.2167	0.4607	0.4047	0.2456
$\lambda = 0.0$	0.5544	0.4219	0.1836	0.5164	0.4152	0.2067
$\lambda = -0.1$	0.6234	0.4367	0.1682	0.5853	0.4230	0.1785
$\lambda = -0.2$	0.7000	0.4512	0.1912	0.6698	0.4505	0.1850
Nonstochastic volatility						
American	0.5260	0.4281	0.2081	0.4798	0.4210	0.2369
European	0.5275	0.4291	0.2118	0.4790	0.4213	0.2375
Observed	0.6305	0.4345		0.6305	0.4345	

and for both calls and puts by choosing $\lambda = -0.1$. For this value of λ, the root mean square error is in all cases at least 20% smaller than those of the constant volatility models.

To examine in more detail the differences between observed option prices (C) and predicted option prices ($\hat{C}$), two cross-sectional regressions were estimated:

Test 1:

$$C = \alpha_0 + \alpha_1 \hat{C} + \tilde{e},$$

Test 2:

$$C - \hat{C} = \gamma_0 + \gamma_1 (S - X)/X + \gamma_2 T + \gamma_3 r_D + \gamma_4 r_F + \tilde{e},$$

where S denotes the spot exchange rate, X the exercise price, T the maturity of the option (measured in days), r_D (r_F) the domestic (foreign) rate of interest, and $\tilde{e}$ are disturbance terms. In Test 1, a model provides an unbiased estimate of the actual premium if $\alpha_0 = 0$ and $\alpha_1 = 1$. Test 2 is designed to identify if the prediction errors are systematically related to the fundamental inputs used in pricing the options. The interpretation of γ_1 depends upon whether the option is a call or a put. For call options, a positive value of γ_1 means that the model under (over) predicts the price of an option that is in (out of) the money. If γ_2 is positive, then the longer the maturity of the option, the greater the degree of under pricing by the model. If γ_3 (γ_4) is positive, the degree of under pricing is an increasing function of the domestic (foreign) interest rate. If the pricing error is independent of the exercise price, maturity, domestic and foreign interest rates, then $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$.

Strictly speaking, under the null that a model is correctly specified, any discrepancy between observed and predicted option prices must be due to the sampling error in the parameters. The asymptotic normality of the parameter estimates and the mean value approximation can be used to construct asymptotically valid confidence regions for option prices. A formal test would simply check if observed option prices fell in these regions. The computational difficulties caused by the large number of observations on option prices in our sample and the use of American option pricing models which lack closed form solutions preclude our use of this large sample test. Neither Test 1 nor Test 2 should be viewed as formal tests of our option pricing models, but we believe that they provide useful descriptive summaries of each model's performance.

Test 1: A model provides an unbiased estimate of the actual premium if $\alpha_0 = 0$ and $\alpha_1 = 1$. This hypothesis is tested using the Wald statistic which is asymptotically distributed as $\chi^2(2)$. Table 5 shows that for all cases the null hypothesis can be overwhelmingly rejected. The slope coefficients tend to be close to one, but the intercepts are positive. In terms of fit and the size of the Wald test statistic, there is some preference for the $\beta = 0$ models, but the differences are not large. The American pricing model with nonstochastic volatility performs slightly better than the European pricing model. Both of these models are dominated by the stochastic volatility models with nonpositive λ. A choice of $\lambda = -0.1$ does reasonably well in all four cases.

Test 2: A number of general comments can be made about the Test 2 results reported in table 6. The results are once again comparable for the two values of β, but with a slight preference for $\beta = 0$. The values of γ_1 and γ_2 vary a great deal across the different models, but they are always smaller for the stochastic volatility models. For $\lambda = -0.1$, the time-to-maturity bias is

Table 5[a]

	Call options				Put options			
	Intercept	Slope	Wald statistic	R^2	Intercept	Slope	Wald statistic	R^2
				$\beta = 0$				
$\lambda = 0.1$	0.1515 (0.0107)	0.9893 (0.0290)	223.2	0.8329	0.1470 (0.0196)	0.9732 (0.0196)	59.4	0.8465
$\lambda = 0$	0.1050 (0.0099)	0.9683 (0.0322)	119.4	0.8407	0.1028 (0.0192)	0.9519 (0.0265)	34.4	0.8543
$\lambda = -0.1$	0.0551 (0.0098)	0.9331 (0.0356)	33.1	0.8405	0.0564 (0.0191)	0.9209 (0.0225)	19.5	0.8567
$\lambda = -0.2$	0.0064 (0.0056)	0.8830 (0.0147)	9.6	0.8259	0.0106 (0.0194)	0.8856 (0.0260)	19.5	0.8460
American	0.1367 (0.0115)	0.9716 (0.0270)	158.5	0.8219	0.1432 (0.0201)	0.9265 (0.0267)	53.3	0.8336
European	0.1385 (0.0113)	1.0187 (0.0376)	151.7	0.7891	0.1455 (0.0195)	0.9195 (0.0269)	60.5	0.8249
				$\beta = 2$				
$\lambda = 0.1$	0.1872 (0.0115)	0.9800 (0.0296)	308.8	0.8113	0.1790 (0.0198)	0.9801 (0.0196)	86.2	0.8337
$\lambda = 0$	0.1423 (0.0106)	0.9641 (0.0339)	196.1	0.8241	0.1341 (0.0193)	0.9613 (0.0206)	55.9	0.8439
$\lambda = -0.1$	0.0932 (0.0102)	0.9240 (0.0386)	85.7	0.8202	0.0863 (0.0188)	0.9299 (0.0227)	31.9	0.8469
$\lambda = -0.2$	0.0431 (0.0114)	0.8667 (0.0430)	17.0	0.8065	0.0380 (0.0188)	0.8846 (0.0259)	22.1	0.8415
American	0.1730 (0.0118)	0.9807 (0.0254)	248.9	0.8109	0.1804 (0.0199)	0.9382 (0.0265)	84.2	0.8267
European	0.1744 (0.0115)	1.0427 (0.0362)	238.4	0.7815	0.1814 (0.0199)	0.9376 (0.0265)	84.5	0.8266

[a]Figures in parentheses are standard errors based on Newey and West (1987).

virtually zero. There is no 'in-the-money' bias for calls, but the model underpredicts the value of 'out-of-the-money' put options. There is considerable evidence that the forecast errors are related to the level of interest rates, which suggests there could be a useful payoff to relaxing our assumption A2. The forecast errors are in all cases negatively related to the level of the domestic (U.S.) interest rate and positively related to the level of the foreign interest rate. A 100 basis point increase in the foreign risk-free rate is associated with a decrease in the forecast error of calls of about 0.07 U.S. cents. In addition to specific biases, we also find that the predictability of the forecast errors from the various models is rather large. The Test 2 regression achieves an R^2 of about 0.3 for calls and 0.2 for puts. The Wald statistic to

Table 6[a]

	γ_0	γ_1	γ_2	γ_3	γ_4	R^2	W_1	W_2
			$\beta = 0$, Call options					
$\lambda = 0.1$	−0.4566	0.9287	0.0010	−5.2475	9.1446	0.2991	55.7	251.4
	(0.1269)	(0.5561)	0.0002)	(1.7973)	(2.0150)			
$\lambda = 0$	−0.4853	0.2760	0.0007	−5.7089	9.6094	0.2807	40.3	134.9
	(0.1264)	(0.5571)	(0.0002)	(1.7877)	(2.0216)			
$\lambda = -0.1$	−0.5186	−0.4225	0.0002	−6.2028	10.1016	0.2959	29.2	49.7
	(0.1265)	(0.5812)	(0.0002)	(1.7795)	(2.0324)			
$\lambda = -0.2$	−0.5674	1.0714	−0.0003	−6.8263	10.7970	0.3571	28.9	32.1
	(0.1295)	(0.6362)	(0.0002)	(1.8060)	(2.0796)			
American	−0.1756	1.1423	0.0010	−6.7391	7.7476	0.1739	45.9	180.7
	(0.1231)	(0.5134)	(0.0002)	(2.0637)	(2.0724)			
European	−0.2179	3.2686	0.0013	−8.0473	9.3399	0.2417	96.0	258.1
	(0.1288)	(0.5455)	(0.0003)	(2.1241)	(2.1100)			
			$\beta = 0$, Put options					
$\lambda = 0.1$	−0.1931	3.5060	0.0012	−5.2626	6.3124	0.2005	80.1	289.8
	(0.1813)	(0.6969)	(0.0003)	(2.5070)	(3.1080)			
$\lambda = 0$	−0.2004	3.8326	0.0007	−5.5116	6.4074	0.1694	59.0	168.6
	(0.1855)	(0.6973)	(0.0003)	(2.5430)	(3.1701)			
$\lambda = -0.1$	−0.2122	4.1395	0.0002	−5.8370	6.5710	0.1731	44.0	77.9
	(0.1898)	(0.7138)	(0.0003)	(2.5791)	(3.2306)			
$\lambda = -0.2$	−0.2460	4.2044	−0.0002	−6.5607	7.1929	0.2125	36.8	37.3
	(0.1989)	(0.7714)	(0.0003)	(2.6665)	(3.3616)			
American	0.1177	3.9614	0.0011	−6.7370	4.5986	0.1602	64.7	212.0
	(0.1755)	(0.7100)	(0.0003)	(2.5425)	(2.9681)			
European	0.1010	3.9146	0.0011	−7.0800	5.0295	0.1544	61.6	196.5
	(0.1862)	(0.7063)	(0.0003)	(2.6904)	(3.1949)			
			$\beta = 2$, Call options					
$\lambda = 0.1$	−0.4639	1.3309	0.0012	−5.7177	9.7630	0.3217	66.9	334.1
	(0.1312)	(0.5923)	(0.0002)	(1.8876)	(2.0663)			
$\lambda = 0$	−0.4959	0.6523	0.0009	−6.2292	10.3249	0.3038	53.1	209.5
	(0.1299)	(0.5778)	(0.0002)	(1.8336)	(2.0327)			
$\lambda = -0.1$	−0.5465	−0.1461	0.0004	−7.1942	11.3875	0.3175	39.1	93.9
	(0.1315)	(0.5933)	(0.0002)	(1.9032)	(2.1005)			
$\lambda = -0.2$	−0.6066	−0.9989	−0.0002	−8.0873	12.4316	0.3840	35.3	36.9
	(0.1330)	(0.6348)	(0.0002)	(1.9191)	(2.1323)			
American	−0.1958	1.7281	0.0013	−6.7632	8.1199	0.2232	65.9	285.2
	(0.1225)	(0.5253)	(0.0002)	(2.0495)	(2.0404)			
European	−0.2345	4.0929	0.0016	−8.0696	9.6991	0.3081	136.8	409.7
	(0.1276)	(0.5564)	(0.0002)	(2.1108)	(2.0699)			

Table 6[a] (continued)

	γ_0	γ_1	γ_2	γ_3	γ_4	R^2	W_1	W_2
			$\beta = 2$, Put options					
$\lambda = 0.1$	−0.2206	3.1163	0.0014	−5.7774	7.1361	0.2300	88.9	373.8
	(0.1312)	(0.5923)	(0.0002)	(1.8876)	(2.0663)			
$\lambda = 0$	−0.2329	3.4802	0.0010	−6.0820	7.3387	0.1904	66.3	239.5
	(0.1855)	(0.6974)	(0.0003)	(2.5542)	(3.1846)			
$\lambda = -0.1$	−0.2554	3.8408	0.0005	−6.6020	7.7706	0.1820	47.6	121.6
	(0.1908)	(0.7072)	(0.0003)	(2.6172)	(3.2795)			
$\lambda = -0.2$	−0.2892	4.1863	−0.0002	−7.2382	8.3600	0.2312	40.5	51.3
	(0.1970)	(0.7396)	(0.0003)	(2.6743)	(3.3756)			
American	0.0779	3.7563	0.0014	−6.8193	5.2069	0.1888	78.7	311.9
	(0.1755)	(0.7100)	(0.0003)	(2.5425)	(2.9681)			
European	0.0799	3.7718	0.0014	−6.8034	5.1846	0.1883	77.7	311.1
	(0.1276)	(0.5564)	(0.0002)	(2.1108)	(2.0699)			

[a]Figures in parentheses are standard errors based on Newey and West (1987). W_1 is the Wald test statistic for $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$. W_2 is the Wald test statistic for $\gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$.

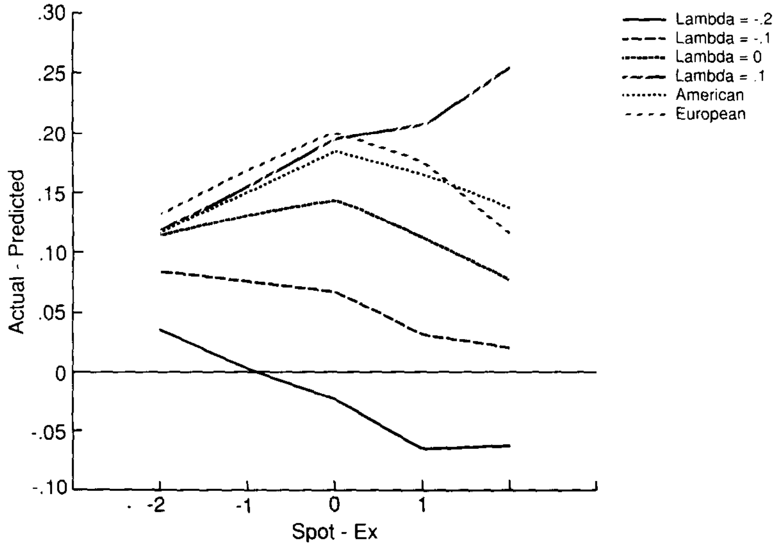

Fig. 2. Average pricing errors, call options ($\beta = 2$).

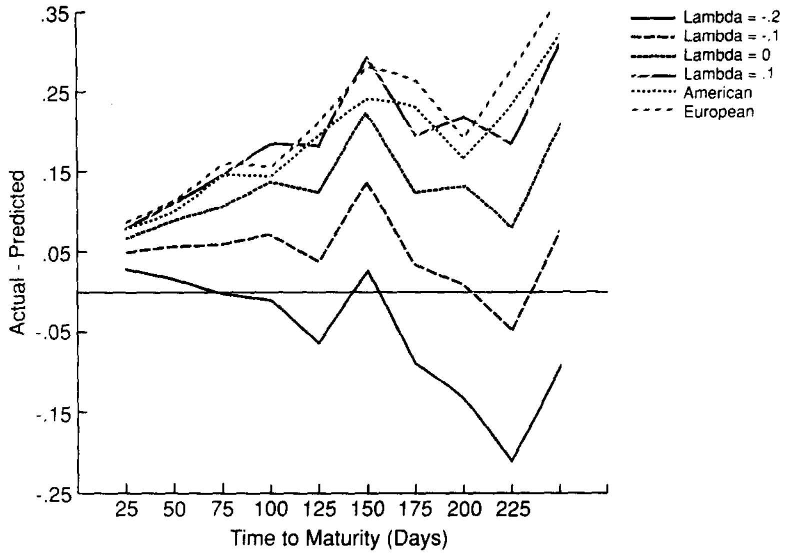

Fig. 3. Average pricing errors, call options ($\beta = 2$).

test the hypothesis $\gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$ is asymptotically distributed as $\chi^2(5)$. All the models are overwhelmingly rejected, although the stochastic volatility models with nonpositive λ fare considerably better than the American or European nonstochastic volatility models. A choice of $\lambda = -0.1$ or $\lambda = -0.2$ seems to be the best.

The regression tests indicate that the stochastic volatility models yield significantly better predictions than the usual American or European nonstochastic volatility models. A simple graph shows just how striking is the improvement. In order to examine the relationship between the mispricing of the option and the degree of being in- or out-of-the-money, we plotted $(S - X)$, the exchange rate minus the exercise price, against $(C - \hat{C})$, the observed option price minus its predicted value. To conserve space, we present only the graph made for calls. The graph for puts was qualitatively similar except that it shows more bias. Fig. 2 presents the graph corresponding to the $\beta = 2$ model. The $S - X$ axis was divided into a number of intervals with the size of the interval being chosen so as to equalize the number of observations. The graph connects the mean pricing error in each interval.

A similar procedure was used to examine the effects of maturity on the degree of mispricing. The results for calls are provided in fig. 3. They indicate considerably less mispricing by the stochastic volatility models. For the choice

$\lambda = -0.1$, there appears to be no systematic relationship between the pricing errors and maturity.

6. Conclusions

The standard model for pricing foreign currency options assumes a lognormal probability distribution for exchange rates and that volatility is constant. We have shown that such assumptions are inappropriate: a more plausible hypothesis is that volatility changes stochastically. Our specification of the volatility process provides a much closer fit to the empirical distribution of exchange rate changes than that produced by the lognormal model. It provides a good match to the dynamic properties of exchange rate changes, and an excellent correspondence to the static properties.

The option pricing equation contains a preference term, as volatility is a nontraded asset. It is not obvious *a priori* what constitutes a reasonable value for the price of volatility risk and many authors simply set it to zero. We treated this price as constant and examined predicted option prices over a grid of four values for this price. In all four cases, we found evidence of systematic prediction error. Nonetheless, for the two negative prices of risk which we examined, the predicted foreign currency option prices provided by the stochastic volatility model provided a striking improvement for all categories of option over the predictions of the standard model.

We conclude that the stochastic volatility model dominates the standard model. It does a superior job of matching the empirical exchange rate distribution and corrects many of the observed biases in predicted option prices. These results have important implications for the risk management of foreign currency option portfolios and the pricing of long-term foreign currency options.[16]

References

Anderson, B.D.O. and John B. Moore, 1979, Optimal filtering (Prentice-Hall, Englewood Cliffs, NJ).

Bodurtha, James N., 1984, PHLX currency option surveillance report and OEC exercise history: A data base, Mimeo.

Bodurtha, James N. and Georges R. Courtadon, 1986, Efficiency tests of foreign currency options markets, Journal of Finance 41, 151–162.

Bodurtha, James N. and Georges R. Courtadon, 1987, Tests of the American option pricing model on the foreign currency options market, Journal of Financial and Quantitative Analysis 22, 153–167.

Boothe, Paul and Debra Glassman, 1987, The statistical distribution of exchange rates, Journal of International Economics 22, 297–319.

Cooper, Ian, Costas Kaplanis, Anthony Newberger, and Stephen Schaefer, 1986, Option hedging, Mimeo. (London Business School, London).

[16]See Melino and Turnbull (1988).

Cox, J.C., J.E. Ingersoll, and S.A. Ross, 1985, An intertemporal general equilibrium model of asset prices, Econometrica 53, 363–384.

Cramér, Harald, 1946, Mathematical methods of statistics (Princeton University Press, Princeton, NJ).

Duffie, Darell and Kenneth J. Singleton, 1988, Simulated moments estimation of diffusion models of asset prices, Mimeo. (GSB, Stanford University, Stanford, CA).

Engle, Robert F., 1982, Autoregressive conditional heteroscedasticity with estimates of United Kingdom inflation, Econometrica 50, 987–1007.

Grabbe, J. Orlin, 1983, The pricing of call and put options on foreign exchange, Journal of International Money and Finance 2, 239–254.

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029–1054.

Hull, John and Alan White, 1987, The pricing of options on assets with stochastic volatilities, Journal of Finance 42, 281–300.

Johnson, Herb and David Shanno, 1987, Option pricing when the variance is changing, Journal of Financial and Quantitative Analysis 22, 143–151.

McFarland, James W., R. Richardson Pettit, and Sam K. Sung, 1982, The distribution of foreign exchange price changes: Trading day effects and risk measurement, Journal of Finance 37, 693–715.

McKee, S. and A.R. Mitchell, 1970, Alternating direction methods for parabolic equations in two space dimensions with a mixed derivative, The Computer Journal 13, 81–86.

Marsh, Terry A. and Eric R. Rosenfeld, 1983, Stochastic processes for interest rates and equilibrium bond prices, Journal of Finance 38, 635–646.

Marsh, Terry A. and Eric R. Rosenfeld, 1986, Non-trading, market making, and estimates of stock price volatility, Journal of Financial Economics 15, 359–372.

Melino, Angelo and Stuart M. Turnbull, 1987, The pricing of foreign currency options, Working paper 8720 (Department of Economics, University of Toronto, Toronto).

Melino, Angelo and Stuart M. Turnbull, 1988, The pricing of long term foreign currency options, Mimeo. (Department of Economics, University of Toronto, Toronto).

Nelson, Dan, 1989, ARCH models as diffusion approximations, this issue.

Newey, Whitney K. and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.

Press, S. James, 1968, A compound events model for security prices, Journal of Business 40, 317–335.

Schwert, G. William, 1987, Effects of model specification on tests for unit roots in macroeconomic data, Journal of Monetary Economics 20, 73–103.

Scott, Louis O., 1987, Option pricing when the variance changes randomly: Theory, estimation and an application, Journal of Financial and Quantitative Analysis 22, 419–438.

Shastri, Kuldeep and Kishore Tandon, 1986, Valuation of foreign currency options: Some empirical tests, Journal of Financial and Quantitative Analysis 21, 145–160.

Wiggins, James B., 1987, Option values under stochastic volatility: Theory and empirical estimates, Journal of Financial Economics 19, 351–372.