Working Paper Series CENTER FOR FINANCIAL ECONOMETRICS

STOCHASTIC SKEW IN CURRENCY OPTIONS

Peter Carr
Liuren Wu

Stochastic Skew in Currency Options*

PETER CARR†

Bloomberg L.P. and Courant Institute

LIUREN WU‡

Zicklin School of Business, Baruch College

First draft: July 8, 2003

This version: May 13, 2004

Filename: newcurrency04.tex

*We thank Nasir Afaf, Emanuel Derman, Bruno Dupire, Brian Healy, Dilip Madan, John Ryan, Harvey Stein, Arun Verma, and participants at Baruch College, UC Riverside, and Columbia University for comments. Any remaining errors are ours. We welcome comments, including references to related papers we have inadvertently overlooked.

†499 Park Avenue, New York, NY 10022; Tel: (212) 893-5056; Fax: (917) 369-5629; Email: pcarr4@bloomberg.com; Homepage: http://www.math.nyu.edu/research/carrp/.

‡One Bernard Baruch Way, Box B10-225, New York, NY 10010-5585; Tel: (646) 312-3509; Fax: (646) 312-3451; Email: Liuren_Wu@baruch.cuny.edu; Homepage: http://faculty.baruch.cuny.edu/lwu/.

Stochastic Skew in Currency Options

ABSTRACT

We document the behavior of over-the-counter currency option prices across moneyness, maturity, and calendar time on two of the most actively traded currency pairs over the past eight years. We find that the risk-neutral distribution of currency returns is relatively symmetric on average. However, on any given date, the conditional currency return distribution can show strong asymmetry. This asymmetry varies greatly over time and often switch directions. We design and estimate a class of models that capture these unique features of the currency options prices and perform much better than traditional jump-diffusion stochastic volatility models.

Stochastic Skew in Currency Options

Options markets have enjoyed tremendous growth during the past decade. In conjunction with this growth, researchers have developed numerous new option pricing models to account for the various pricing biases in the classic Black and Scholes (1973) model. Most recently, a series of papers synthesize and test the performance of a number of different models for pricing equity index options, e.g., Bakshi, Cao, and Chen (1997, 2000a,b), Bates (2000), Andersen, Benzoni, and Lund (2002), Pan (2002), Eraker (2003), and Huang and Wu (2004). However, studies on currency option pricing have been relatively sparse.

At first glance, this relative paucity of study is surprising since foreign exchange is the largest of the global financial markets. Currently, daily trading volume in the currency markets stands at over 1.5 trillion U.S. dollars. It is widely appreciated that the dynamic behavior of foreign exchange rates has important economic repercussions. It is also widely appreciated that currency option prices reveal important information about the conditional risk-neutral distribution of the underlying currency returns over different horizons.

The most likely reason for the relative scarcity of research on currency options is the absence of a publicly available database for currency option prices. Currency options trade on the Philadelphia Options Exchange (PHLX), but volume in this market has thinned during the past five years as trading activity has shifted to the over-the-counter (OTC) market. The OTC currency options market is very liquid and deep. The bid-ask spreads for major currency options are narrower than those on equity index options, and trading volume is measured in trillions of U.S. dollars per year. Hence, the over-the-counter currency options market constitutes an economically important market for academic research.

We obtain a data set of OTC option quotes on two of the most actively traded currency pairs during the past eight-year span from January 1996 to January 2004. The two currency pairs are the U.S. dollar price of Japanese yen (JPYUSD) and the U.S. dollar price of the British pound (GBPUSD). For each option at each date, we have a cross-section of 40 option quotes from a matrix of five strikes and eight maturities.

Using this data set, we analyze the behavior of option implied volatility along the dimensions of moneyness, maturity, and calendar time. As an industry standard, the foreign exchange market measures the moneyness of an option in terms of the option's delta according to the Black-Scholes formula. Moving across moneyness at a fixed maturity, we find that the time series average of the implied volatility is fairly symmetric about at the money, with the average out-of-the-money implied volatility higher than the average at-the-money implied volatility. This well-known smile pattern for the implied volatility across moneyness suggests that the risk-neutral conditional distribution of currency returns is fat-tailed, but on average symmetric. For each currency pair, the average implied volatility smile persists as the option maturity increases from one week to one and half years. The persistence of the smile over long maturities indicates that the average conditional currency return distribution remains highly fat-tailed even at long conditioning horizons.

When we investigate the dynamic behavior of the implied volatility surface over calendar time, we find that the relative curvature of the implied volatility smile is stable over both calendar time and the two currency pairs. In contrast, the *slope* of the implied volatility in moneyness varies greatly over calendar time and across the two currency pairs. Although implied volatility smiles are symmetric on average, they can be highly asymmetric on any given date. As a result, the risk-neutral skewness of the return distribution can be quite large in absolute terms on any given date.

Existing currency option pricing models, such as the jump-diffusion stochastic volatility model of Bates (1996b), readily accommodate the average shape of the implied volatility surface. In the Bates model, the Merton (1976) jump component captures the short-term curvature of the implied volatility smile, whereas the Heston (1993) stochastic volatility component generates smiles at longer maturities. It is a tribute to the ingenuity of the option pricing modelers that they can capture the average shape of the implied volatility surface while operating under the constraints of no arbitrage.

Although these models do represent an impressive application of option pricing technology, they cannot generate the strong time-variation in the risk-neutral skewness of the currency return distribution. The purpose of this paper is to design and test a new class of models that can capture this unique feature of the OTC currency options market.

If we start from the jump-diffusion stochastic volatility model of Bates (1996b), it would be tempting to attempt to capture stochastic skewness by randomizing the mean jump size parameter and/or the correlation parameter between the currency return and the stochastic volatility process. In the Bates model, these two parameters govern the risk-neutral skewness at short and long maturities, respectively. However, randomizing either parameter is not amenable to analytic solution techniques that greatly aid econometric estimation. In this paper, we attack the problem from a different perspective. We apply the very general framework of time-changed Lévy processes developed in Carr and Wu (2004). However, the subclass of models that we extract from this framework to price currency options are far from standard in the option pricing literature.

In our models, innovations in currency returns are driven by two Lévy processes. The two independent Lévy processes generate positive and negative jumps, respectively. We further apply separate random time changes to these two Lévy components. As a result, the total volatility and the relative contributions from positive and negative jumps can both vary stochastically over time. These random variations are controlled by two activity rate processes, which are specified in terms of traditional stochastic volatility processes. The variation in the relative proportion of positive and negative jumps generates variation in the risk-neutral skewness of the currency return distribution. Within this class of models, we propose various jump specifications that exhibit finite and infinite activities, respectively.

To econometrically estimate the models using our OTC currency options data, we cast the estimation problem into a state-space form. We define the state propagation equations based on the two activity rate processes that control the positive and negative jump Lévy components. We build the measurement equations based on the option prices at different levels of moneyness and maturity. We first extract the unobservable activity rate state variables using a relatively new filtering technique, the unscented Kalman Filter. We then estimate the model parameters using the quasi maximum likelihood method. The methodology estimates the activity rate dynamics under both the risk-neutral measure and the objective measure.

Our new models have about the same number of free parameters as the jump-diffusion stochastic volatility model pioneered by Bates (1996b). However, our models generate much better performance

in terms of both root mean squared pricing errors and log likelihood values, both in sample and out of sample. The stochastic volatility component in the Bates model can capture the time variation in overall volatility, but it cannot capture the variation in the relative proportion of positive and negative jumps. As a result, the Bates model, or any other existing one-factor stochastic volatility model, fails to capture a large proportion of the variation in the currency options data. In contrast, the two activity rates in our new models generate not only stochastic volatility, but also the stochastic skew that we have observed in the currency options.

In other related works, Bates (1996a) investigates the distributional properties of the currency returns implied from currency futures options. Campa and Chang (1995, 1998) and Campa, Chang, and Reider (1998) study the empirical properties of the OTC currency options. Bollen (1998) and Bollen, Gray, and Whaley (2000) propose regime-switching models for currency option pricing. Nevertheless, Bollen and Raisel (2003) find that the jump-diffusion stochastic volatility model of Bates (1996b) outperforms regime-switching and GARCH-type models in matching the observed behaviors of OTC currency options. Therefore, we regard the Bates model as the state of the art for currency option pricing and as our benchmark for model comparison.

The paper is organized as follows. Section I systematically documents the empirical properties of OTC currency options. Section II designs a class of models that capture the unique properties of the currency options. Section III proposes an estimation strategy that estimates both the risk-neutral and time-series dynamics of the activity rates simultaneously. Section IV reports the estimation results of the new models and compares their performance to the Bates (1996b) model. Section V concludes.

I. The Over-the-Counter Currency Options Data

Trades and quotes in OTC currency options differ from those on exchange-listed options in several important aspects. First, the OTC quotes are not directly on option prices, but rather on the Black-Scholes implied volatility. Given the quote on the implied volatility, the invoice price is computed based on the Black-Scholes model, with mutually agreed-upon inputs on the underlying spot currency price and interest rates. Second, when a transaction takes place, it involves not only the exchange of the

option position, but also the corresponding delta hedge in the underlying currency. Third, the implied volatilities are not quoted on a fixed strike price, but rather on a fixed Black-Scholes delta. This delta quote directly determines the amount of the underlying currency that change hands in the transaction. Given the delta, the strike price of the option is computed using the Black-Scholes formula and the implied volatility quote.

This unique market design greatly facilitates the liquidity and depth of the OTC currency options market. In an exchange-listed options market, only options are involved in each transaction and the market makers provide direct quotes on the option prices. This practice places severe burdens on market makers due to the derivative nature of the options market. Whenever the underlying currency moves, the options market maker needs to adjust the quotes on hundreds of options written on this currency. If the market-making technology does not allow the option quotes to be updated in a timely fashion, the market maker will have to protect him- or herself by posting wider bid-ask spreads. Furthermore, when a customer acts on private information regarding the directional move of the underlying currency, the correlated nature of all of the options on the same currency can force the market maker into large exposures. For example, if a customer believes that the British pound will strengthen against the dollar, the customer can in principle buy all the calls and sell all the puts on the pound against the dollar. Therefore, the market maker's risk exposure is greatly aggravated due to the highly correlated nature of all the options on the same asset. To protect him- or herself, the market maker has to further reduce quote sizes. These concerns have dried up liquidity in the exchange-traded currency options market.

The unique design of the OTC currency option market addresses these concerns and improves the liquidity and depth of the market. The exchange of the covered position, rather than a naked option position, significantly reduces the broker dealer's exposure to directional bets on the underlying currency. The quotation on the implied volatility rather than the option price itself further reduces the broker dealer's burden in constantly updating the option prices on every move in the underlying currency price. Although the covered position can still have a small dependence on the exchange rate, updates of the implied volatility are only necessary in practice when the broker dealer thinks that the second and higher central moments of the return distribution have changed. The quotation on

delta instead of on fixed strike prices further simplifies the transaction because the fixed delta directly determines the amount of the underlying currency that is involved in the option transaction. Finally, for large transactions, the over-the-counter market also has a mechanism that is similar to the "upstairs" market, where the broker dealer directly searches and matches buyers and sellers and hence secludes him- or herself from exposure to large inventory positions. As a result, the over-the-counter market can handle very large trades with small bid-ask spreads and little market impact, making it an ideal venue for institutional players to engage in large volumes of option trading.

A. The Black-Scholes Model and Notation

Since the market quotes for option value and moneyness are both defined in terms of the Black-Scholes formula, we first review the Black-Scholes model and fix the notation.

We use S_t to denote the time-t price of a foreign currency. A consequence of the Black-Scholes model is that under the risk-neutral measure $\mathbb{Q}$, the dynamics of S_t are governed by the following stochastic differential equation:

$$dS_t/S_t = (r_d - r_f)dt + \sigma dW_t, \tag{1}$$

where r_d and r_f denote the assumed constant instantaneous riskfree rate in the domestic and foreign currency, respectively. The term W_t is a standard Brownian motion, and σ is a constant denoting the instantaneous volatility of the currency return. Under this model, the risk-neutral distribution of the currency return $\ln(S_t/S_0)$ is normally distributed. Originally, Black and Scholes proposed this model for pricing stock options and corporate liabilities. Garman and Kohlhagen (1983) first applied this model to currency option pricing.

We use $c_t(K,\tau)$ and $p_t(K,\tau)$ to denote the time-t value of a European call option and a European put option, respectively. The arguments of the functions indicate that the currency options have a strike

price K and time to maturity $\tau = T - t$. We use $F_t = S_t e^{(r_d - r_f)\tau}$ to denote the forward price of the currency at the corresponding maturity. The Black-Scholes formulas for the option values are

$$c_t(K,\tau) = e^{-r_f\tau} S_t N(d_+) - e^{-r_d\tau} K N(d_-), \tag{2}$$

$$p_t(K,\tau) = -e^{-r_f\tau} S_t N(-d_+) + e^{-r_d\tau} K N(-d_-), \tag{3}$$

with

$$d_\pm = \frac{\ln(F_t/K)}{\sigma\sqrt{\tau}} \pm \frac{1}{2}\sigma\sqrt{\tau}. \tag{4}$$

Delta is defined as the partial derivative of the option value with respect to the underlying spot price. Under the Black-Scholes model, the delta of the call and put options are given by

$$\delta(c) = e^{-r_f\tau} N(d_+), \quad \delta(p) = e^{-r_f\tau} N(-d_+). \tag{5}$$

The delta for a put option is negative, but the convention is to quote the absolute magnitude and indicate that it is on a put or a call option. In the OTC currency options market, moneyness is conventionally quoted in terms of this Black-Scholes delta rather than the strike price. The Black-Scholes implied volatility refers to the parameter σ that a broker dealer must input into the Black-Scholes formulae in equations (2) and (3) so that option values match the market prices.

If the central conclusion of the Black-Scholes model in equation (1) were correct, we would only need one σ input for all the options on each currency. In practice, however, the market is well aware of the deficiencies of the Black-Scholes model. To compensate for these deficiencies, the market uses a different volatility input at each moneyness, maturity, and calendar time. We denote the Black-Scholes implied volatility at a certain delta (δ), time-to-maturity (τ), and calendar time (t) as $IV_t(\delta,\tau)$. We use IV instead of the parameter σ to notationally distinguish between the market quote and the model assumption. The fact that the market uses the Black-Scholes model to present option quotations does not mean that the market agrees with the assumptions or conclusions of the Black-Scholes model. Instead, the market is merely using the model as a monotonically linear transformation tool to convert

option prices into a more stable measure. Furthermore, the market also uses the Black-Scholes model to achieve approximately delta-neutral transactions.

Given the implied volatility quote $IV_t(\delta,\tau)$ at a certain delta and maturity, we can infer the strike price of the option contract,

$$K = F_t \exp\left[\mp IV_t(\delta,\tau)\sqrt{\tau}N^{-1}\left(\pm e^{r_f\tau}\delta\right) + \frac{1}{2}IV_t(\delta,\tau)^2\tau\right] \tag{6}$$

Each delta corresponds to two strike prices, one for the call option contract and the other for the put option contract.

B. Data Description

We have obtained OTC currency options quotes from several broker dealers and data vendors. These data sets cover different sample periods, sampling frequency, and currency pairs. We use the common samples of these different data sets to cross-validate the quality of the data. In this paper, we present the stylized evidence and estimate our models using two currency pairs from one data source because the samples on these two currency pairs span the longest time period from January 24, 1996 to January 28, 2004. The data are available in daily frequency, but to avoid weekday effects in model estimation, we sample the data weekly, on every Wednesday of each week. When market closes on a Wednesday, we use the quotes from the previous market open date. For each series, we have 419 weekly observations.

The two currency pairs are the U.S. dollar of Japanese yen (JPYUSD) and the U.S. dollar of British pound (GBPUSD). Options on each pair have eight maturities: one week, one month, two months, three months, six months, nine months, 12 months, and 18 months. Quotes on longer maturities from two to five years are also available, but careful inspection shows that these long-maturity quotes are merely extrapolations of the shorter-maturity quotes and do not contain much extra information.

At each maturity, the quotes are available at five strikes in the form of (1) delta-neutral straddle implied volatility (ATMV), (2) ten-delta risk reversal (RR10), (3) ten-delta strangle margin (SM10), (4) 25-delta risk reversal (RR25), and (5) 25-delta strangle margin (SM25).

A straddle combines a call option with a put option at the same strike. For the straddle to be delta-neutral, we need

$$\delta(c) + \delta(p) = 0. \tag{7}$$

From the definitions of deltas in equation (5), we have

$$N(d_+) - N(-d_+) = 0, \tag{8}$$

or $N(d_+) = 0.5$ and hence $d_+ = 0$. The strike price is very close to the spot or forward price of the currency for the delta-neutral straddle. Hence, we refer to this quote as the at-the-money implied volatility (ATMV) quote.

The ten-delta risk reversal (RR10) quote measures the difference in implied volatility between a ten-delta call option and a ten-delta put option,

$$RR10 = IV(10c) - IV(10p), \tag{9}$$

where we use $10p$ and $10c$ to denote a ten-delta put and call, respectively. Hence, the risk reversal is a measure of asymmetry, or slope, of the implied volatility smile across moneyness.

The ten-delta strangle margin (SM10) measures the difference between the average implied volatility of the two ten-delta options and the delta-neutral straddle implied volatility,

$$SM10 = \left(IV(10c) + IV(10p)\right)/2 - ATMV. \tag{10}$$

Hence, a strangle margin measures the average curvature of the implied volatility smile. The market also refers to a strangle margin as a butterfly spread. The 25-delta risk-reversal and strangle margins are analogously defined.

From the five quotes, we obtain the implied volatilities at the five deltas as

$$\begin{aligned} IV(0s) &= ATMV; \\ IV(25c) &= SM25 + ATMV + RR25/2; \\ IV(25p) &= SM25 + ATMV - RR25/2; \\ IV(10c) &= SM10 + ATMV + RR10/2; \\ IV(10p) &= SM10 + ATMV - RR10/2, \end{aligned} \tag{11}$$

where we use $(0s)$ to denote the delta of the straddle at $d_+ = 0$.

Altogether, we have 16,760 implied volatility quotes for each of the two currency pairs, spanning 419 weeks with a cross-section of 40 option quotes per date (five strikes multiplied by eight maturities). Figure 1 plots the time series of the 40 implied volatility series for each currency pair. Historically, implied volatilities on JPYUSD have varied in a wide range from 5.89 percent to 45.34 percent. The large spike in late 1998 corresponds to the hedge fund crisis, when most hedge funds had gone short on Yen before the crisis and were then forced to use options to cover their positions during the crisis. Implied volatilities on GBPUSD vary in a much narrower range between 3.5 percent and 15.95 percent.

The data set also contains the underlying spot currency price. To convert the implied volatility quotes into option prices, we also need information on domestic and foreign interest rates. We construct our interest rate series using LIBOR and swap rates from the three countries. We download the LIBOR an swap rates data from Bloomberg. The LIBOR rates are simply compounded, with maturities from one week to 12 months. We directly convert them into continuously compounded interest rates. For the interest rates at 18 months, we bootstrap them from the LIBOR and swap rates.

C. Stylized Features of Currency Option Implied Volatilities

Using the currency option implied volatility quotes, we document a series of important features of the data that a reasonable currency option pricing model should accommodate.

C.1. Relatively Symmetric Mean Implied Volatility Smile

When we plot the time series average of the implied volatility against the delta at each maturity, we observe a relatively symmetric average implied volatility smile across all maturities and the two currency pairs. Figure 2 plots the average implied volatility smile across moneyness at selected maturities for the two currency pairs: one month (solid lines), three months (dashed lines) and one year (dash-dotted lines). In the graphs, we denote the x-axis in terms of approximate put option delta. In particular, we approximately denote the ten-delta call as a 90-delta put in the graph, and denote the delta-neutral straddle at 50 delta.

The constant return volatility assumption of the Black-Scholes model implies a normal risk-neutral distribution for currency returns. The smile shape of the implied volatility across moneyness has long been regarded as evidence for return non-normality under the risk-neutral measure. The curvature of the smile reflects fat-tails or positive excess kurtosis in the risk-neutral return distribution. The asymmetry of the smile reflects asymmetry or skewness in the currency return distribution. The relatively symmetric mean implied volatility smiles show that on average, the risk-neutral distribution of the currency return is fat-tailed, but not highly asymmetric.

C.2. The Mean Implied Volatility Smile Persists with Increasing Maturity

Suppose that we model currency returns as being generated by a discrete time process with independent and identically distributed (iid) non-gaussian increments with finite return variance. By design, the short-term return distribution is non-normal and could potentially be consistent with the short-term implied volatility smiles. However, this non-normality disappears rapidly as we consider a longer time horizon for the return. By virtue of the classic central limit theorem, the return skewness declines like the reciprocal of the square root of the time horizon, and the kurtosis declines like the reciprocal of the time horizon. Mapping this declining non-normality to the implied volatility smile at different maturities, we would expect the smile to flatten out rapidly at longer option maturities.

The maturity pattern of the mean implied volatility smiles in Figure 2 indicates otherwise for currency options. The smiles remain highly curved as the option maturity increases from one week to one year. This maturity pattern indicates that the conditional risk-neutral distribution for the currency return remains highly non-normal as the conditioning horizon increases. Thus, an iid return distribution with finite return variance cannot generate this maturity pattern of the implied volatility smile. The continuous-time equivalent of the iid return distribution is to model the currency return as following a Lévy process. To slow down the convergence of return distribution to normality with increasing maturity, researchers, e.g., Bates (1996b), have proposed incorporating a persistent stochastic volatility process.

C.3. Strangle Margin is Stable, But Risk Reversal Varies Greatly Over Time

The market quotes on risk reversals and strangle margins provide direct and intuitive measures of the asymmetry and curvature of the implied volatility smile, respectively. In Figure 3, we plot the time series of the ten-delta risk reversal (solid lines) and strangle margin (dashed lines), both normalized as percentages of the corresponding at-the-money implied volatility level. The multiple lines for both the risk reversals and the strangle margins represent the different option maturities, which we do not distinguish in the plot. To reduce clustering, we only plot three maturities (one, three, and 12 months).

We observe that the ten-delta strangle margins (dashed lines) are consistently at about ten percent of the at-the-money implied volatility level during the eight-year span at all three option maturities and for both currency pairs. Therefore, the curvature of the smile is relatively stable over option maturity, calendar time, and for different currency pairs. This feature of the data shows that excess kurtosis in the currency return distribution is a robust and persistent feature of the OTC currency options market.

In stark contrast to the stability of the strangle margins, the risk reversals (solid lines) vary greatly over calendar time. The dispersion of the risk reversals across different option maturities is also larger. For JYPUSD, the ten-delta risk reversals have moved from -30 percent to 60 percent of the at-the-money implied volatility level. In contrast, the ten-delta strangle margins have only moved within a 20 percentage range. For GBPUSD, the swing of the ten-delta risk reversal is smaller from between

-20 percent to 20 percent, but the movement of the ten-delta strangle margin is even smaller within a narrow band of 10 percent, except at the very early years.

Table I reports the mean, standard deviation, and the weekly autocorrelation of risk reversals, strangle margins, and at-the-money straddle implied volatilities. We again normalize the risk reversals and strangle margins as percentages of the at-the-money implied volatility.

For JPYUSD, the sample averages of the risk-reversals are positive, implying that the out-of-money call options are more expensive than the corresponding out-of-money put options during the sample period. The average strangle margins are around 12 percent at ten delta and three to four percent at 25 delta. For GBPUSD, the average implied volatility smile is much more symmetric as the average risk-reversals are close to zero. The average strangle margins are only slightly smaller than the corresponding averages for JPYUSD. The average strangle margins for GBPUSD are around nine percent at ten delta and less than three percent at 25 delta.

For both currencies, the standard deviations of the risk reversals are much larger than the standard deviations of the same-delta strangle margins. For JPYUSD, the standard deviations are around 15 percent for ten-delta risk reversals and are just about three to four percent for ten-delta strangle margins. The standard deviations of 25-delta risk reversals are about eight percent, but that for the 25-delta strangle margins are about one percent or less. The same pattern holds for GBPUSD. The standard deviations for the risk reversals are about three times larger than that for the corresponding strangle margins. The at-the-money implied volatilities have standard deviations around three for JPYUSD and less than two for GBPUSD.

These numbers are consistent with our observations from Figure 3. The major variation in the currency option implied volatilities comes from the risk reversal, that is, the difference in volatility between calls and puts of the same delta. The variations in the curvature of the volatility smile are much smaller.

Mapping the implied volatility pattern to the risk-neutral distribution of the currency return, we conclude that the skewness of the risk-neutral currency return distribution varies greatly over time. The kurtosis of the return distribution varies much less.

All implied volatility series exhibit strong serial correlation. The weekly autocorrelation ranges from 0.69 to 0.98. Furthermore, we do not observe a significant difference in autocorrelation between the volatility portfolios (risk reversals and strangles) and the single volatility series (ATMV), especially at long maturities. These serial dependence reflects the time-series dynamics of the return volatility.

C.4. Changes in Risk Reversals are Positively Correlated with Currency Returns

Table II reports the cross-correlation between currency returns and the weekly changes in risk reversals, strangle margins, and at-the-money implied volatilities. Again, risk reversals and strangles are measured in percentages of the at-the-money implied volatility. We find that risk reversals exhibit very strong positive correlations with currency returns. This strong correlation is present at all maturities and for both currency pairs, at both ten and 25 deltas. This positive correlation implies that whenever a foreign currency appreciates and hence generates a positive return, the risk reversal also increases and hence the risk-neutral return distribution is more likely to be positively skewed.

We also measure the cross-correlations at different leads and lags. Figure 4 plots the sample estimates of the cross-correlations between the currency return and changes in the one-month ten-delta risk reversals at different leads and lags. For both currencies, the cross-correlation between return and the risk reversal is mainly contemporaneous. We do not identify any significant cross-correlations with leads and lags. This pattern also holds for other maturities.

In contrast to the strong and positive correlation with the risk reversals, the currency return has very little correlation with the changes in strangle margins. Furthermore, we obtain positive correlation estimates between the currency return and changes in the at-the-money implied volatility for JPYUSD, but the estimates for GBPUSD are essentially zero. Hence, the only persistent and universal correlation pattern is between the currency returns and the risk reversals.

Using different currency pairs, sample periods, and different data sources, we have cross-validated the above-documented evidence on currency options. The above findings are all robust to sample variations and data sources. The most striking, and the most talked-about feature among currency options traders, is the strong time variation of the risk reversals, and the lack of models that can capture this feature.

II. Modeling Currency Returns For Option Pricing

In this section, we propose a class of models that can capture not only the average behavior of currency option implied volatilities across moneyness and maturity, but also the dynamic properties of at-the-money implied volatilities and risk reversals.

We use $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ to denote a complete stochastic basis defined on a risk-neutral probability measure $\mathbb{Q}$. We assume constant interest rates mainly for notational clarity. We let r_d and r_f denote the continuously-compounded domestic and foreign riskfree rates, respectively.

For option pricing, we first specify the currency return process $s_t = \ln(S_t/S_0)$ under the risk-neutral measure $\mathbb{Q}$. The historical or traditional approach to option pricing has been to derive unique risk-neutral dynamics as a consequence of no arbitrage, continuous trading opportunities, and a specification of the statistical process that leads to market completeness. It is increasingly being recognized that realistic statistical processes and trading opportunities render markets incomplete. As a result, there are multiple risk-neutral processes, consistent with a given realistic statistical process for the underlying asset price and market setting. Since different risk-neutral processes lead to different option prices, a more pragmatic approach for obtaining unique option prices begins by specifying a parametric family of risk-neutral processes for the underlying currency. Then, the option pries are used to identify the parameters and thereby select a unique risk-neutral process.

After specifying the family of risk-neutral processes governing currency returns $s_t = \ln(S_t/S_0)$, we derive the generalized Fourier transform of the currency return. We use this transform to price options based on the fast Fourier inversion method of Carr and Madan (1999). When we perform dynamic

estimation, we also specify the dynamics under the physical measure $\mathbb{P}$, which we assume is absolutely continuous with respect to $\mathbb{Q}$.

We derive option pricing models by specifying asset returns as following time-changed Lévy processes. Carr and Wu (2004) show that most stochastic processes used in traditional option pricing models can be cast as special cases of time-changed Lévy processes. Huang and Wu (2004) apply this framework successfully to pricing equity index options.

We assume that the log currency return obeys the following time-changed Lévy process under the risk-neutral measure $\mathbb{Q}$,

$$s_t \equiv \ln S_t/S_0 = (r_d - r_f)t + \left(L^R_{T^R_t} - \xi^R T^R_t\right) + \left(L^L_{T^L_t} - \xi^L T^L_t\right), \tag{12}$$

where L^R and L^L denote two Lévy processes that exhibit right (positive) and left (negative) skewness, respectively. The terms ξ^R and ξ^L denote concavity adjustments of the two Lévy processes, needed so that the exponential of each process is a martingale. Each Lévy process can have a continuous martingale component, and both must have a jump component to generate the required skewness. We further apply separate stochastic time changes T^R_t and T^L_t to the two Lévy components so that the relative proportion of the two components can vary over time.

In principle, the generic specification in equation (12) can capture all the salient features of currency options. First, by setting the unconditional weights of the two Lévy components equal to each other, we can obtain a relatively symmetric unconditional distribution with fat tails for the currency return under the risk-neutral measure. This unconditional property captures the relatively symmetric feature of the sample averages of the implied volatility smile.

Second, by applying separate time changes to the two components, aggregate return volatility can vary over time so that the model can generate stochastic volatility.

Third, the relative weight of the two Lévy components can also vary over time due to the separate time changes. When the weight of the right-skewed Lévy component L^R is higher than the weight of the left-skewed Lévy component L^L, the model generates a right-skewed conditional return distribution

and hence positive risk reversals. When the opposite is the case, the model generates left-skewed conditional return distributions and negative risk reversals. Thus, we can generate variations and even sign changes on the risk reversals via the separate time changes.

Finally, the model captures the instantaneous correlation between the return and the risk reversal through the correlations between the Lévy components and the time change. To stress the ability of our family of models described by equation (12) in capturing stochastic skews of the currency return distribution, we christen this family as **stochastic skew models** (SSM).

For each model considered in this paper, we first derive its generalized Fourier transform and then price European options using a fast Fourier transform method. The generalized Fourier transform of the currency return is defined as

$$\phi_s(u) \equiv E\left[e^{ius_t}\right], \quad u \in \mathcal{D} \subset \mathbb{C}, \tag{13}$$

where $\mathcal{D}$ is a subset of the complex domain $\mathbb{C}$ on which the expectation in equation (13) is finite. When u takes only real values, $\phi_s(u)$ denotes the characteristic function of the currency return. See Titchmarsh (1986) for details on the extension of u to the complex plane.

In what follows, we propose parsimonious specifications for the two Lévy components and the stochastic time change.

A. The Lévy Components

We consider a one-dimensional Lévy process X_t that is adapted to $\mathcal{F}_t$. The sample paths of X are right-continuous with left limits, and $X_u - X_t$ is independent of $\mathcal{F}_t$ and distributed as X_{u-t} for $0 \leq t < u$. By the Lévy-Khintchine Theorem, the characteristic function of X_t has the form,

$$\phi_x(u) \equiv E\left[e^{iuX_t}\right] = e^{-t\psi_x(u)}, \quad t \geq 0, \tag{14}$$

where the characteristic exponent $\psi_x(u), u \in \mathbb{R}$, is given by (Bertoin (1996)),

$$\psi_x(u) = -iu\mu + \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}^0} \left(1 - e^{iux} + iux1_{|x|<1}\right) \nu(x)dx. \tag{15}$$

The triplet $\left(\mu, \sigma^2, \nu\right)$ defines the Lévy process X and is referred to as the *Lévy characteristics*. The first member of the triplet, μ, describes the constant drift of the process. The second member σ^2 describes the constant variance rate of the diffusion component of the Lévy process. The third member $\nu(x)$ describes the jump structure and determines the arrival rate of jumps of size x. The term $\nu(x)dx$ is referred to as the Lévy measure, with $\nu(x)$ being the Lévy density. To value options, we extend the characteristic function parameter u to the complex plane, $u \in \mathcal{D} \subseteq \mathbb{C}$.

In equation (15), $1_{|x|<1}$ is an indicator function that equals one when $|x| < 1$ and zero otherwise. This truncation is meant to guarantee that the integral is well defined around the singular point of zero (Bertoin (1996)). There are other commonly used truncation functions for the same purpose. In principle, we can use any truncation functions, $h : \mathbb{R} \to \mathbb{R}$, which are bounded, with compact support, and satisfy $h(x) = x$ in a neighborhood of zero (Jacod and Shiryaev (1987)).

For our model design, we make the following generic decomposition on the two Lévy components in equation (12),

$$L_t^R = J_t^R + \sigma W_t^R, \quad L_t^L = J_t^L + \sigma W_t^L,$$

where (W_t^R, W_t^L) denote two independent, standard Brownian motions and (J_t^R, J_t^L) denote two pure jump Lévy components with positive and negative skewness in distribution, respectively.

To maintain parsimony, we assume relative symmetry for the unconditional return distribution. We set the instantaneous volatility (σ) of the two diffusion components to be the same. We also set the two pure jump Lévy components J_t^R and J_t^L to be mirror images of each other. From equation (15), we have the characteristic exponent of the two diffusion components as

$$\psi^R(u) = \psi^L(u) = \frac{1}{2}u^2\sigma^2. \tag{16}$$

The concavity adjustment for the diffusion component is

$$\xi^R = \xi^L = -\psi(-i) = \frac{1}{2}\sigma^2.$$

For the pure jump components, we propose a simple yet flexible Lévy density,

$$\nu^R(x) = \begin{cases} \lambda e^{-\frac{|x|}{v_j}}|x|^{-\alpha-1}, & x > 0, \\ 0, & x < 0. \end{cases}, \quad \nu^L(x) = \begin{cases} 0, & x > 0, \\ \lambda e^{-\frac{|x|}{v_j}}|x|^{-\alpha-1}, & x < 0. \end{cases} \tag{17}$$

so that the right-skewed jump component only allows positive jumps and the left-skewed jump component only allows negative jumps. For both jumps, we use the same parameters $(\lambda, v_j) \in \mathbb{R}^+$ and $\alpha \leq 2$ for parsimony. This specification has its origin in the CGMY model of Carr, Geman, Madan, and Yor (2002). We label it as CG jump. The Lévy density of the CG specification follows an exponentially dampened power law. Depending on the magnitude of the power coefficient α, the sample paths of jump process can exhibit finite activity ($\alpha < 0$), infinite activity with finite variation ($0 \leq \alpha < 1$), or infinite variation ($1 \leq \alpha \leq 2$). We need $\alpha \leq 2$ to maintain finite quadratic variation. Therefore, this parsimonious specification can capture a wide range of jump behaviors. We can thus let the data determine the exact jump behavior for currency prices.

Given the Lévy density specifications in equation (17), we can derive the characteristic exponents for the two jump components by applying the integral in equation (15). When $\alpha \neq 0$ and $\alpha \neq 1$, we have (Wu (2004)),

$$\psi^R(u) = \lambda\Gamma(-\alpha)\left[\left(\frac{1}{v_j}\right)^\alpha - \left(\frac{1}{v_j} + iu\right)^\alpha\right] + iuC_+, \tag{18}$$

$$\psi^L(u) = \lambda\Gamma(-\alpha)\left[\left(\frac{1}{v_j}\right)^\alpha - \left(\frac{1}{v_j} + iu\right)^\alpha\right] + iuC_-, \tag{19}$$

where C_+ and C_- are immaterial drift terms due to the truncation that will eventually be cancelled out with the corresponding terms in the concavity adjustments. The concavity adjustment terms are

$$\xi^R = -\lambda\Gamma(-\alpha)\left[\left(\frac{1}{v_j}\right)^\alpha - \left(\frac{1}{v_j} - 1\right)^\alpha\right] - C_+, \quad \xi^L = -\lambda\Gamma(-\alpha)\left[\left(\frac{1}{v_j}\right)^\alpha - \left(\frac{1}{v_j} + 1\right)^\alpha\right] - C_-.$$

Within this general jump specification, we consider three special cases, each one representing a different jump type.

A.1. KJ: Finite Activity Jumps

For a finite-activity jump process, the number of jumps within any finite time interval is finite. The CG specification generates finite-activity jumps when $\alpha < 0$. Here, we consider the special example of $\alpha = -1$. The Lévy density becomes,

$$\nu^R(x) = \begin{cases} \lambda e^{-\frac{|x|}{v_j}}, & x > 0, \\ 0, & x < 0. \end{cases}, \quad \nu^L(x) = \begin{cases} 0, & x > 0, \\ \lambda e^{-\frac{|x|}{v_j}}, & x < 0. \end{cases} \tag{20}$$

This jump specification exhibits finite activity because the integral of the Lévy density is finite,

$$\int_0^\infty \lambda e^{-\frac{x}{v_j}} dx = \lambda v_j. \tag{21}$$

The quantity (λv_j) is often referred to as the jump intensity or mean arrival rate. Conditional on one jump occurring, the jump size for each component has a one-sided exponential distribution.

The characteristic exponents of the two jump components follow equations (18) and (19) with $\alpha = -1$. We can also rewrite them as

$$\psi^R(u) = \lambda \int_0^\infty \left(1 - e^{iux}\right) e^{-\frac{x}{v_j}} dx = -\lambda v_j \frac{iuv_j}{1 - iuv_j}, \tag{22}$$

$$\psi^L(u) = \lambda \int_{-\infty}^0 \left(1 - e^{iux}\right) e^{\frac{x}{v_j}} dx = \lambda v_j \frac{iuv_j}{1 + iuv_j}. \tag{23}$$

For finite-activity jumps, the integrals are well-behaved around zero. Hence, we do not need the truncation term $(iux1_{|x|<1})$ in (15).

Combining the positive and negative jumps, we obtain the characteristic exponent of a symmetric compound Poisson double-exponential model of Kou (2002),

$$\psi^R(u)+\psi^L(u)=\lambda\nu_j\frac{2u^2\nu_j^2}{1+u^2\nu_j^2}.$$

We label this finite-activity jump specification as KJ. The concavity adjustment terms are,

$$\xi^R=\frac{\lambda\nu_j^2}{1-\nu_j},\quad \xi^L=-\frac{\lambda\nu_j^2}{1+\nu_j}.$$

In the estimation, we reparameterize $\lambda=\lambda\nu_j^2$ for numerical stability.

A.2. VG: Infinite Activity with Finite Variation Jumps

An infinite-activity jump process generates an infinite number of jumps within any finite interval. The CG specification generates infinite activity jumps when $\alpha\geq 0$. Here, we consider the special case of $\alpha=0$. The Lévy density becomes,

$$\nu^R(x)=\begin{cases}\lambda e^{-\frac{|x|}{\nu_j}}|x|^{-1}, & x>0,\\ 0, & x<0.\end{cases},\quad \nu^L(x)=\begin{cases}0, & x>0,\\ \lambda e^{-\frac{|x|}{\nu_j}}|x|^{-1}, & x<0.\end{cases} \tag{24}$$

By having the power term $|x|^{-1}$, the arrival rate of small jumps increases dramatically so that as $|x|\rightarrow 0$, the Lévy density approaches infinity. Under this specification, the integral of the Lévy density in equation (21) is no longer finite. Thus, the sample paths of the process exhibit infinite activity. Nevertheless, the following integral remains finite

$$\int_{\mathbb{R}^0}x1_{|x|<1}\nu(dx)<\infty. \tag{25}$$

Hence, the specification has finite variation.

With $\alpha = 0$, the characteristic exponents take different forms from equations (18) and (19). They are

$$\psi^R(u) = \lambda \ln(1 - iuv_j), \quad \psi^L(u) = \lambda \ln(1 + iuv_j). \tag{26}$$

Combining the two components, we obtain the characteristic exponent of a symmetric variance-gamma model (Madan, Carr, and Chang (1998) and Madan and Seneta (1990)),

$$\psi^R(u) + \psi^L(u) = \lambda \ln\left(1 + u^2 v_j^2\right).$$

We label this jump specification as VG. The concavity adjustment terms are

$$\xi^R = -\lambda \ln(1 - v_j), \quad \xi^L = -\lambda \ln(1 + v_j).$$

Recently, Madan and Daal (2004) empirically show that the VG model performs better than the Merton (1976) jump-diffusion model in capturing both the time series dynamics of currency returns and the behavior of currency options.

A.3. CJ: Infinite Variation Jumps

The sample paths of the VG jumps exhibit infinite activity, but nevertheless finite variation. When $\alpha \geq 1$, the integral in equation (25) also becomes infinite and the sample paths of the jumps will exhibit infinite variation. We consider the special case of $\alpha = 1$. The Lévy densities are,

$$\nu^R(x) = \begin{cases} \lambda e^{-\frac{|x|}{v_j}} |x|^{-2}, & x > 0, \\ 0, & x < 0. \end{cases}, \quad \nu^L(x) = \begin{cases} 0, & x > 0, \\ \lambda e^{-\frac{|x|}{v_j}} |x|^{-2}, & x < 0. \end{cases} \tag{27}$$

The characteristic exponents of this jump specification also take unique forms. For the right-skewed jump component, we have

$$\begin{aligned}\psi^R(u) &= \lambda\int_0^\infty \left(1-e^{iux}+iux1_{|x|<1}\right)e^{-\frac{|x|}{v_j}}|x|^{-2}dx \\ &= -\lambda(1/v_j-iu)\ln(1-iuv_j)-iu\lambda(1+\mathcal{E}_1(\beta)).\end{aligned} \tag{28}$$

We need to incorporate a truncation term ($iux1_{|x|<1}$) into the integral to maintain finiteness for the infinite-variation jump specification. The term $\mathcal{E}_1(\beta)$ denotes the standard exponential integral function,

$$\mathcal{E}_1(\beta)=\int_\beta^\infty e^{-x}x^{-1}dx. \tag{29}$$

The characteristic exponent of the left-skewed jump component can be similarly derived as,

$$\begin{aligned}\psi^L(u) &= \lambda\int_{-\infty}^0 \left(1-e^{iux}+iux1_{|x|<1}\right)e^{-\frac{|x|}{v_j}}|x|^{-2}dx \\ &= -\lambda(1/v_j+iu)\ln(1+iuv_j)+iu\lambda(1+\mathcal{E}_1(\beta)).\end{aligned} \tag{30}$$

Combining the two components, we obtain the characteristic exponent of a symmetric infinite-variation model,

$$\psi^R(u)+\psi^L(u)=-\lambda(1/v_j+iu)\ln(1+iuv_j)-\lambda(1/v_j-iu)\ln(1-iuv_j).$$

If we drop the exponential term in the Lévy density, we obtain the Lévy density for a Cauchy process. Thus, we label this jump specification as CJ. The concavity adjustment terms are

$$\xi^R=\lambda(1/v_j-1)\ln(1-v_j)+\lambda(1+\mathcal{E}_1(\beta)),\quad \xi^L=\lambda(1/v_j+1)\ln(1+v_j)-\lambda(1+\mathcal{E}_1(\beta)).$$

The characteristic exponents for the concavity-adjusted Lévy components simplify to,

$$\psi^R(u) = -\lambda(1/v_j - iu)\ln(1-iuv_j) + iu\lambda(1/v_j - 1)\ln(1-v_j), \tag{31}$$

$$\psi^L(u) = -\lambda(1/v_j + iu)\ln(1+iuv_j) + iu\lambda(1/v_j + 1)\ln(1+v_j). \tag{32}$$

Here, we observe that the drift term $iu\lambda(1+\mathcal{E}_1(\beta))$ drop out of the characteristic exponents for the concavity-adjusted Lévy components. Hence, they are immaterial for our estimation.

All together, we consider four jump specifications: CG, KJ, VG, and CJ, with the last three as special cases of the encompassing CG specification. By comparing their relative performance in pricing currency options, we can infer the jump behavior of currency prices.

B. Activity Rates

We assume a differentiable and therefore continuous time change and let

$$v_t^R \equiv \frac{\partial T_t^R}{\partial t}, \quad v_t^L \equiv \frac{\partial T_t^R}{\partial t},$$

denote the instantaneous activity rates of the two Lévy components. We model the two activity rates as following the square-root process of Heston (1993),

$$\begin{aligned} dv_t^R &= \kappa\left(1-v_t^R\right)dt + \sigma_v\sqrt{v_t^R}dZ_t^R, \\ dv_t^L &= \kappa\left(1-v_t^L\right)dt + \sigma_v\sqrt{v_t^L}dZ_t^L. \end{aligned} \tag{33}$$

For identification reasons, we normalize the long-run mean of both processes to one. For parsimony and symmetry, we set the mean-reversion parameter κ and volatility of volatility parameter σ_v to be the same for both activity rate processes.

We allow the two Brownian motions (W_t^R, W_t^L) in the return process and the two Brownian motions (Z_t^R, Z_t^L) in the activity rates to be correlated as follows,

$$\rho^R dt = E^{\mathbb{Q}}\left[dW_t^R dZ_t^R\right], \quad \rho^L dt = E^{\mathbb{Q}}\left[dW_t^L dZ_t^L\right].$$

The four Brownian motions are assumed to be independent otherwise.

The activity rate specification is the same as in Bates (1996b) except that we have two activity rates that govern two Lévy components of different skewness whereas Bates uses one stochastic variance rate process to govern the overall volatility level. Eraker, Johannes, and Polson (2003) propose to incorporate a jump component in the variance rate dynamics when modeling the time-series dynamics of index returns, but both Eraker (2003) and Broadie, Chernov, and Johannes (2004) show that the option-pricing impacts of jumps in the activity rate processes are minimal, even if they are present in the time series dyanmics. Hence, we choose the more parsimonious but equally effective pure-diffusion specification in (33).

C. The Generalized Fourier Transform of the Currency Return

For time-changed Lévy processes, Carr and Wu (2004) show that the problem of deriving the generalized Fourier transform can be converted into an equivalent problem of deriving the Laplace transform of the time change under a new, complex-valued measure:

$$\begin{aligned}\phi_s(u) &= e^{iu(r_d-r_f)t}E^{\mathbb{Q}}\left[e^{iu\left(L^R_{T_t^R}-\xi^R T_t^R\right)+iu\left(L^L_{T_t^L}-\xi^L T_t^L\right)}\right] \\ &= e^{iu(r_d-r_f)t}E^{\mathbb{M}}\left[e^{-\psi^\top T_t}\right] \equiv e^{iu(r_d-r_f)t}\mathcal{L}_T^{\mathbb{M}}(\psi), \end{aligned} \tag{34}$$

where $\psi \equiv \left[\psi^R, \psi^L\right]^\top$ denotes the vector of the characteristic exponents of the concavity-adjusted right- and left-skewed Lévy components, respectively, and $\mathcal{L}_T^{\mathbb{M}}(\psi)$ represents the Laplace transform of the

stochastic time T_t under a new measure $\mathbb{M}$. The measure $\mathbb{M}$ is absolutely continuous with respect to the risk-neutral measure $\mathbb{Q}$ and is defined by a complex-valued exponential martingale,

$$\left.\frac{d\mathbb{M}}{d\mathbb{Q}}\right|_t \equiv \exp\left[iu\left(L^R_{T^R_t} - \xi^R T^R_t\right) + iu\left(L^L_{T^L_t} - \xi^L T^L_t\right) + \psi^R T^R_t + \psi^L T^L_t\right]. \tag{35}$$

Equation (34) reduces the problem of obtaining a generalized Fourier transform of a time-changed Lévy process into a simpler problem of deriving the Laplace transform of the stochastic clock. The solution to this Laplace transform depends on the specification of the instantaneous activity rate $v(t)$ and the characteristic exponents.

Since the Laplace transform of the time change in equation (34) is defined under the complex measure $\mathbb{M}$, we need to obtain the activity rate process under $\mathbb{M}$. By Girsanov's Theorem, under measure $\mathbb{M}$, the diffusion coefficient of $v(t)$ remains the same as $\sigma_v\sqrt{v^j_t}, j = R,L$. The drift terms adjust as follows:

$$\texttt{drift}(v^j_t)^{\mathbb{M}} = \kappa(1 - v^j_t) + iu\sigma\sigma_v\rho^j v^j_t, \quad j = R,L. \tag{36}$$

The instantaneous drift and variance of the two activity rate processes are affine under both the probability measure $\mathbb{Q}$ and the new complex valued measure $\mathbb{M}$. Under affine activity rates, the Laplace transform of T_t is exponential-affine in the current level of the activity rates, $[v^R_0, v^L_0]$:

$$\mathcal{L}^{\mathbb{M}}_T(\psi) = \exp\left(-b^R(t)v^R_0 - c^R(t) - b^L(t)v^L_0 - c^L(t)\right), \tag{37}$$

where

$$\begin{aligned} b^j(t) &= \frac{2\psi^j\left(1-e^{-\eta^j t}\right)}{2\eta^j - (\eta^j - \kappa^j)\left(1-e^{-\eta^j t}\right)}, \\ c^j(t) &= \frac{\kappa}{\sigma_v^2}\left[2\ln\left(1 - \frac{\eta^j - \kappa^j}{2\eta^j}\left(1 - e^{-\eta^j t}\right)\right) + (\eta^j - \kappa^j)t\right], \end{aligned} \tag{38}$$

and

$$\eta^j = \sqrt{(\kappa^j)^2 + 2\sigma_v^2\psi^j}, \quad \kappa^j = \kappa - iu\rho^j\sigma\sigma_v, \quad j = R,L.$$

The characteristic exponents, $\psi^j, j = R,L$, depend on the specification of the Lévy components. We summarize them in Table III.

D. Traditional Jump-Diffusion Stochastic Volatility Models

The jump-diffusion stochastic volatility model of Bates (1996b) represents the state of the art in the currency option pricing literature. This model combines the jump-diffusion specification of Merton (1976) with the stochastic volatility specification of Heston (1993). We label this model as MJDSV, where MJD denotes the Merton jump-diffusion specification and SV denotes its stochastic volatility feature.

To compare the MJDSV model to our SSM specification, we cast the MJDSV model into the time-changed Lévy process framework and write the log return process under measure $\mathbb{Q}$ as

$$s_t = (r_d - r_f)t + (J_t(\lambda) - \xi t) + \left(\sigma W_{T_t} - \frac{1}{2}\sigma^2 T_t\right), \tag{39}$$

where $J_t(\lambda)$ denotes a compound Poisson pure jump process with a Poisson arrival rate λ. Conditional on one jump occurring, the jump size in log returns is normally distributed with mean μ_j and variance v_j. The term W_t denotes a standard Brownian motion, and T_t denotes the stochastic clock with activity rate given by $v_t = \partial T_t / \partial t$. The activity rate follows a square-root process:

$$dv_t = \kappa(1 - v_t)\,dt + \sigma_v \sqrt{v_t}\, dZ_t,$$

with $\rho dt = E^{\mathbb{Q}}[dW_t dZ_t]$. Equation (39) makes it obvious that the MJDSV model generates stochastic volatility purely from the diffusion component while keeping the jump arrival rate constant over time. Furthermore, if we set $\lambda = 0$ and delete the jump component, the Bates (1996b) model degenerates into the pure-diffusion stochastic volatility model of Heston (1993). We also estimate this restricted version and denote it as HSTSV.

Equation (39) also makes it obvious that both HSTSV and MJDSV can generate stochastic volatility via the stochastic time change of the diffusion component, but neither can generate stochastic skew. Under HSTSV, the average skew is determined by the correlation parameter ρ between the diffusion in the currency return and the diffusion in the activity rate. With a fixed correlation parameter, the model

cannot generate dramatically varying skews. Under MJDSV, the mean jump size μ_j also helps in generating an average skew shape. However, since it is also a fixed parameter, the MJDSV model cannot generate large variations in the conditional skew, either. Besides, the fact that the sample averages of the implied volatility smiles are relatively symmetric dictates that to capture the average smile shape, both parameters should be set around zero.

Under the MJDSV model, the generalized Fourier transform of the currency return is given by

$$\phi_s(u) = e^{iu(r_d - r_f)t - t\psi^J - b(t)v_0 - c(t)}, \tag{40}$$

where the characteristic exponent of the concavity-adjusted jump component is

$$\psi^J = \lambda\left[iu\left(e^{\mu_j + \frac{1}{2}\sigma_j^2} - 1\right) - \left(e^{iu\mu_j - \frac{1}{2}u^2\sigma_j^2} - 1\right)\right], \tag{41}$$

and the coefficients $b(t)$ and $c(t)$ for the diffusion component are the same as in (38) with $\psi^j = \psi^D = \frac{1}{2}\sigma^2(iu + u^2)$. Equation (40) also applies to the HSTSV model with $\psi^J = 0$ as the Heston (1993) model does not have a jump component.

We estimate the MJDSV model and its restricted version HSTSV. We compare their performance with our stochastic skew models. The number of free parameters for the MJDSV model is about the same as the number of free parameters in our SSM specifications with KJ, VG, and CJ jump structures.

III. Quasi-Maximum Likelihood With Unscented Kalman Filter

To estimate the dynamic models to the time series data of implied volatilities, we cast the models into a state-space form and estimate the models using the quasi-maximum likelihood method.

To capture the time-series dynamics, we need to specify the currency return and activity rate dynamics under the objective measure $\mathbb{P}$. Since the return process under measure $\mathbb{P}$ has limited relevance for option pricing, we focus on the activity rate processes and leave the market price of return risk

unspecified. We assume that the market price of risk on the activity rates (volatility) is proportional to the square root of the activity rates:

$$\gamma(v_t^j) = \gamma\sqrt{v_t^j}, \quad j = L, R. \tag{42}$$

For symmetry, we use the same parameter γ for both activity rates. Then, the $\mathbb{P}$-processes governing the activity rates become

$$\begin{aligned} dv_t^R &= \kappa^P(\theta^P - v_t^R)dt + \sigma_v\sqrt{v_t^R}dZ_t^R, \\ dv_t^L &= \kappa^P(\theta^P - v_t^L)dt + \sigma_v\sqrt{v_t^L}dZ_t^L, \end{aligned} \tag{43}$$

with

$$\kappa^P = \kappa - \sigma_v\gamma, \quad \theta^P = \frac{\kappa}{\kappa - \sigma_v\gamma}. \tag{44}$$

Thus, the activity rates also follow square root processes under the objective measure $\mathbb{P}$. We make analogous assumptions on the volatility risk premium on the Heston (1993) model and Bates (1996b) model.

In the state-space form, we regard the two activity rates as the unobservable states and specify the state propagation equation using an Euler approximation of equation (43):

$$v_t = (1-\varphi)\theta^P + \varphi v_{t-1} + \sigma_v\sqrt{v_{t-1}\Delta t}\varepsilon_t, \tag{45}$$

where $\varphi = \exp(-\kappa^P\Delta t)$ denotes the autocorrelation coefficient with Δt being the length of the discrete time interval, and ε denotes iid bivariate standard normal innovation. With weekly sampling frequency, we set $\Delta t = 7/365$. The term $v_t = \left[v_t^R, v_t^L\right]$ denotes the bivariate vector of activity rates for our SSM models and a scalar for the Heston (1993) model and the Bates (1996b) model. For notational clarity, we normalize the discrete time interval to one.

Under this specification, the conditional covariance matrix of the state vector is a diagonal matrix with state-dependent diagonal elements:

$$Q_t = diag\langle\sigma_v^2 v_{t-1}\Delta t\rangle,$$

where $diag\langle\cdot\rangle$ denotes a diagonal matrix with the diagonal elements given by the vectors inside the bracket.

We construct the measurement equations based on the observed out-of-money option prices, assuming additive, normally-distributed measurement errors:

$$y_t = O(v_t;\Theta) + e_t, \quad cov(e_t) = \mathcal{R}, \tag{46}$$

where y_t denotes the observed option prices at time t and $O(v_t;\Theta)$ denotes the model-implied value as a function of the parameter set Θ and the state vector v_t. The term e_t denotes the pricing errors, with covariance matrix $\mathcal{R}$. We convert the implied volatility quotes into out-of-money option prices and scale all option prices by their Black-Scholes vega. With this scaling, we assume that the pricing errors are iid normal with zero mean and variance matrix, $\mathcal{R} = \sigma_r I$, with σ_r being a scalar and I being an identity matrix of the relevant dimension. The dimension of the measurement equation is 40, capturing the 40 options quotes on each date.

How to define the pricing error and how to weigh the pricing error are important yet delicate issues. Since the risk-neutral distributional properties of the currency return show up most vividly on the implied volatility surface across moneyness and maturity, it would be ideal to define the pricing error as the difference between the Black-Scholes implied volatility quote and its model-implied fair value. Nevertheless, our algorithm generates option prices from the return characteristic function. Converting the option prices into Black-Scholes implied volatility involves an additional minimization routine that can be time-consuming when embedded in the global optimization procedure. By dividing the out-of-the-money option prices by its vega, we are essentially converting the option price into the implied volatility space via linear approximation. Under the Black-Scholes model, the vega is given by

$$\frac{\partial C}{\partial IV} = Se^{-r_f\tau}\sqrt{\tau}N'(d_+). \tag{47}$$

The scaling of $\sqrt{\tau}$ makes the option prices relatively on the same level across maturities. The scaling by the normal probability density adjusts for the fact that out-of-money options are less expensive than at-

the-money options. For the estimation, we first convert the implied volatility quotes into out-of-money option prices in percentages of the underlying spot. Then, we ignore the interest rate effect and apply time-homogeneous weighting on options prices at fixed delta (δ) and time-to-maturity (τ),

$$w(\delta,\tau) = \frac{1}{100\sqrt{\tau}N'(N^{-1}(\delta))}. \tag{48}$$

We use $\overline{v}_t, \overline{P}_t, \overline{y}_t, \overline{A}_t$ to denote the time-$(t-1)$ ex ante forecasts of time-t values of the state vector, the covariance of the state vector, the measurement series, and the covariance of the measurement series, respectively. We use $\widehat{v}_t$ and $\widehat{P}_t$ to denote the ex post update, or filtering, on the state vector and its covariance at the time t based on observations (y_t) at time t. In the case of linear measurement equations,

$$y_t = Hv_t + e_t, \tag{49}$$

the Kalman (1960) filter provides the most efficient updates. The ex ante predictions are,

$$\begin{aligned} \overline{v}_t &= (1-\varphi)\,\theta^P + \varphi\widehat{v}_{t-1}; \\ \overline{P}_t &= \varphi\widehat{P}_{t-1}\varphi^\top + Q_{t-1}; \\ \overline{y}_t &= H\overline{v}_t; \\ \overline{A}_t &= H\overline{P}_t H^\top + \mathcal{R}, \end{aligned} \tag{50}$$

and the ex post filtering updates are,

$$\begin{aligned} \widehat{v}_{t+1} &= \overline{v}_{t+1} + K_{t+1}\left(y_{t+1} - \overline{y}_{t+1}\right), \\ \widehat{P}_{t+1} &= \overline{P}_{t+1} - K_{t+1}\overline{A}_{t+1}K_{t+1}^\top, \end{aligned} \tag{51}$$

where K_{t+1} is the Kalman gain, given by,

$$K_{t+1} = S_{t+1}\left(\overline{A}_{t+1}\right)^{-1}, \quad S_{t+1} = \overline{P}_{t+1}H.$$

Thus, we can obtain a time series of the ex ante forecasts and ex post updates on both the mean and covariance of the state vectors and the observed series, via the iterative procedure defined by (50) and (51).

In our application, the measurement equation in (46) is nonlinear. Traditionally, nonlinearity is often handled by the Extended Kalman Filter (EKF), which approximates the nonlinear measurement equation with a linear expansion, evaluated at the predicted states,

$$y_t \approx H(\overline{v}_t;\Theta)\, v_t + e_t,$$

where

$$H(\overline{v}_t;\Theta) = \left.\frac{\partial O(\overline{v}_t;\Theta)}{\partial v_t}\right|_{v_t=\overline{v}_t}. \tag{52}$$

The prediction and the updates follow equations (50) and (51). The extended Kalman filter uses only one point (the conditional mean) from the prior filtering density for the prediction and filtering updates.

In this paper, we use a relatively new filtering technique called the unscented Kalman filter (UKF). The UKF uses a set of (sigma) points to match not only the mean and variance, but also the higher moments of the state distribution. If we let k denote the number of states (one in the Bates/Heston model and two in our SSM models) and let $\zeta > 0$ denote a control parameter, we generate a set of $2k+1$ sigma vectors χ_i according to the following equations,

$$\begin{aligned}
\chi_{t,0} &= \widehat{v}_t, \\
\chi_{t,i} &= \widehat{v}_t \pm \sqrt{(k+\zeta)(\widehat{P}_t + Q_t)_j}, \quad j = 1,\cdots,k; \quad i = 1,\cdots,2k,
\end{aligned}$$

with the corresponding weights w_i given by,

$$w_0 = \delta/(k+\zeta), \quad w_i = 1/[2(k+\zeta)], \quad j = 1,\cdots,2k.$$

We can regard these sigma vectors as forming a discrete distribution with w_i being the corresponding probabilities. Then, we can verify that the mean, covariance, skewness, and kurtosis of this distribution

are $\widehat{v}_t$, $\widehat{P}_t + Q_t$, 0, and $k+\zeta$ respectively. Thus, we can use the control parameter ζ to accommodate conditional non-normalities in the state propagation equation.

Given the sigma points, the prediction steps are given by:

$$\begin{aligned}
\overline{\chi}_{t,i} &= (1-\varphi)\,\theta^P + \varphi\chi_{t,i};\\
\overline{v}_{t+1} &= \sum_{i=0}^{2k} w_i(\overline{\chi}_{t,i});\\
\overline{P}_{t+1} &= \sum_{i=0}^{2k} w_i(\overline{\chi}_{t,i}-\overline{v}_{t+1})(\overline{\chi}_{t,i}-\overline{v}_{t+1})^\top;\\
\overline{y}_{t+1} &= \sum_{i=0}^{2k} w_i O\left(\overline{\chi}_{t,i};\Theta\right);\\
\overline{A}_{t+1} &= \sum_{i=0}^{2k} w_i\left[O\left(\overline{\chi}_{t,i};\Theta\right)-\overline{y}_{t+1}\right]\left[O\left(\overline{\chi}_{t,i};\Theta\right)-\overline{y}_{t+1}\right]^\top + \mathcal{R},
\end{aligned} \tag{53}$$

and the filtering updates are given by

$$\begin{aligned}
\widehat{v}_{t+1} &= \overline{v}_{t+1} + K_{t+1}\left(y_{t+1}-\overline{y}_{t+1}\right);\\
\widehat{P}_{t+1} &= \overline{P}_{t+1} - K_{t+1}\overline{A}_{t+1}K_{t+1}^\top,
\end{aligned} \tag{54}$$

with

$$K_{t+1} = S_{t+1}\left(\overline{A}_{t+1}\right)^{-1};\quad S_{t+1} = \sum_{i=0}^{2k} w_i\left[\overline{\chi}_{t,i}-\overline{v}_{t+1}\right]\left[O\left(\overline{\chi}_{t,i};\Theta\right)-\overline{y}_{t+1}\right]^\top.$$

We refer to Julier and Uhlmann (1997) for general treatments of the UKF.

To estimate the model parameters, we define the log-likelihood for each day's observation assuming that the forecasting errors are normally distributed:

$$l_{t+1}(\Theta) = -\frac{1}{2}\log\left|\overline{A}_t\right| - \frac{1}{2}\left(\left(y_{t+1}-\overline{y}_{t+1}\right)^\top\left(\overline{A}_{t+1}\right)^{-1}\left(y_{t+1}-\overline{y}_{t+1}\right)\right). \tag{55}$$

We choose model parameters to maximize the log likelihood of the data series, which is a summation of the daily log likelihood values,

$$\Theta \equiv \arg\max_{\Theta} \mathcal{L}(\Theta, \{y_t\}_{t=1}^{N}), \quad \texttt{with} \quad \mathcal{L}(\Theta, \{y_t\}_{t=1}^{N}) = \sum_{t=0}^{N-1} l_{t+1}(\Theta), \tag{56}$$

where $N = 419$ denotes the number of weeks in our sample.

For each currency, we estimate six models, which include the Heston (1993) model (HSTSV), the Bates (1996b) model (MJDSV), and four SSM models. The four SSM models differ in their respective jump specifications. We label them as KJSSM, VGSSM, CJSSM, and CGSSM, with KJ, VG, CJ, and CG denoting the four different jump structures.

The Bates (1996b) model has nine free parameters $\Theta_B = \left[\sigma_r, \sigma^2, \lambda, \mu_j, v_j, \kappa, \sigma_v, \rho, \kappa^P\right]$. The Heston (1993) constitutes a restricted version with $\lambda = v_j = \mu_j = 0$. Our SSM models with KJ, VG, or CJ jumps also have nine parameters, $\Theta_S = \left[\sigma_r, \sigma^2, \lambda, v_j, \kappa, \sigma_v, \rho^R, \rho^L, \kappa^P\right]$. The SSM model with CG jump specification (CGSSM) has one extra free parameter α that controls the type of the jump process. This extra parameter α is fixed at -1, 0, and 1 for KJ, VG, and CJ, respectively. Furthermore, the four SSM models have two state variables (v_t^R, v_t^L) that generate both stochastic volatility and stochastic skew in the currency return distribution. The Bates model and the Heston model have only one state variable v_t that controls the instantaneous variance of the diffusion component.

IV. Results and Discussion

In this section, we discuss the estimation results and address the following questions: Which model best captures the time series and cross-sectional behaviors of currency option implied volatilities? How the estimated activity rate dynamics relate to the observed time variation in implied volatilities and risk reversals?

A. In-Sample Model Performance Comparison

We compare the in-sample model performance along two dimensions. First, we investigate how our new SSM models perform against traditional jump-diffusion stochastic volatility models, e.g., Bates (1996b) (MJDSV). Second, within our new SSM model framework, we investigate which jump structure delivers the best performance in capturing the currency option price behavior.

Table IV reports the parameter estimates and standard errors (in parentheses) for the six models on the two currency pairs. In the last two rows of the table, we also report the root mean squared pricing error and the maximized log likelihood value for each model and each currency. The results in Table IV are obtained based on the whole sample of eight years of data. Hence, performance comparisons are in sample.

When we compare the performance of our SSM models to the traditional jump-diffusion stochastic volatility model of Bates (1996b) (MJDSV), we find that our SSM models markedly outperform the MJDSV model in terms of both the in-sample log likelihood values and the root mean squared pricing errors. For the currency pair JPYUSD, the log likelihood value for MJDSV is lower than values for the four SSM models by 2,605, 2,619, 2,637, and 2,685, respectively. The root mean squared error is 0.984 for MJDSV and is 0.822 or lower for the four SSM models.

For the currency pair GBPUSD, the log likelihood values for the four SSM models are also higher than the value for the MJDSV model, with the difference ranging from 1,537 to 1,561. The root mean squared pricing error is 0.421 for MJDSV and is 0.378 or lower for the four SSM models.

From MJDSV to its restricted version HSTSV, we observe a further reduction in likelihood values and an further increase in root mean squared pricing errors. The likelihood difference is 409 for JPYUSD and 604 for GBPUSD. The root mean squared error difference is 0.03 for JPYUSD and 0.011 for GBPUSD. These differences show that the jump component in MJDSV does improve the model performance over the pure-diffusion stochastic volatility model of Heston (1993) (HSTSV).

Within our SSM framework, we estimate four models with different jump specifications. In contrast to the large difference in log likelihood values between the SSM models and the MJDSV model,

the likelihood value differences among the four SSM models are much smaller. For JPYUSD, we detect a marginal increase in the likelihood value as we move from the KJ jump structure to VG and then to CJ. These three jump specifications differ by a power term in the Lévy density. The performance ranking corresponds to an increase in the power coefficient α and an increase in jump frequency. When we estimate the CGSSM model where α is a free parameter, the estimate for α is 1.556, higher than all the three restricted versions. This encompassing model also generates a higher likelihood value, potentially indicating that a high-frequency jump specification is favored for modeling currency options. Nevertheless, when we compare the root mean squared pricing errors for the four SSM models, we can hardly distinguish the differences among the four jump types.

For GBPUSD, the estimate of α under the CGSSM model is 1.180, smaller than the estimate for JPYUSD but still higher than all three restricted versions. Nevertheless, for this currency pair, the performance differences of the four SSM models are negligible in terms of both the log likelihood values and the room mean squared pricing errors. Therefore, we conclude that our currency options data cannot effectively distinguish between different jump types. There is only weak evidence that favors a high-frequency jump specification with infinite variation for the JPYUSD currency pair.

Our results on the nature of the jump specification for currency options are not as strong as those in Carr and Wu (2003) and Huang and Wu (2004) for equity index options. Both studies find that infinite-activity jump specifications significantly outperform finite-activity jump specifications for pricing S&P 500 index options. Madan and Daal (2004) also find evidence that the infinite-activity VG model performs better than the finite-activity Merton (1976) jump in pricing currency options. Those studies use exchange-traded options that include very deep out-of-the-money contracts. The over-the-counter currency options data that we use in this paper have only five strikes for each maturity, all located within approximately the tenth and 90th percentile of the return distribution. Hence, the currency options data that we use do not provide much information on the tail (beyond the tenth percentile) of the currency return distribution. However, it is exactly in the tails of the currency return distribution where the alternative jump specifications display their differences.

To test the statistical significance of the performance difference between different models, we adopt the likelihood ratio statistic constructed by Vuong (1989) for non-nested models. Formally, we let $LR(\Theta_i, \Theta_j)$ denote the log likelihood ratio between models i and j,

$$LR(\Theta_i, \Theta_j) \equiv \mathcal{L}_i(\Theta_i) - \mathcal{L}_j(\Theta_j). \tag{57}$$

Vuong constructs a test statistic based on this log likelihood ratio,

$$\mathcal{M} = \sqrt{N} LR(\Theta_i, \Theta_j)/\widehat{s}, \tag{58}$$

where N denotes the number of days in the time series and $\widehat{s}^2$ denotes the variance estimate of the daily log likelihood ratio $(l_i - l_j)$. Vuong proves that $\mathcal{M}$ is asymptotically normally distributed $N(0,1)$ under the null hypothesis that the two models are equivalent in terms of likelihood:

$$H_0 : E\left[l_i - l_j\right] = 0. \tag{59}$$

Based on the daily log likelihood estimates, we compute the sample mean and standard deviation of the likelihood ratio between each pair of models and then construct the test statistic in equation (58). In estimating $\widehat{s}$, we adjust serial dependence in the daily log likelihood ratios according to Newey and West (1987) with the lags optimally chosen following Andrews (1991) under an AR(1) specification.

Table V reports the pairwise log likelihood ratio test statistics. For each currency pair, we report the statistics in a (6×6) matrix, with the (i, j)th element being the statistic on $(l_i - l_j)$. The diagonal terms are zero by definition. For both currency pairs, all the off-diagonal elements in the first column are positive and strongly significant, indicating that HSTSV is the worst performing of all six estimated models. The last four elements in the second column are also strongly positive and significant, indicating that the performance of MJDSV is significantly worse than the four SSM models.

However, as we move to the (4×4) block in the right bottom corner, none of the elements are significant for either currency pair. This block compares the performance among the four SSM models. Therefore, we conclude that within the SSM modeling framework, our currency options data cannot

effectively distinguish the different jump specifications. Nevertheless, our SSM models significantly outperform the MJDSV model and the HSTSV model.

B. Out-of-Sample Performance Comparison

To study the out-of-sample performance, we re-estimate the six models using the first six years of data from January 24, 1996 to December 26, 2001, 310 weekly observations for each series. Then, we use these estimated model parameters to compare the model performance both in sample during the first six years and out of sample during the last two years from January 2, 2002 to January 28, 2004 (109 weekly observations for each series). If our estimation generates stable model parameters and the currency option price behaviors have not dramatically changed during the last two years, we would expect that the out-of-sample performance for each model is similar to its in-sample performance. We also investigate whether the superior performance of our SSM models over the traditional specifications such as HSTSV and MJDSV extends to out-of-sample comparison .

Table VI reports the model parameter estimates and standard errors using the subsample of six years of data. Both the estimates and standard errors are close to what we have obtained from the full sample in Table IV, indicating that the currency option price behaviors have not experienced dramatic changes over the past two years. The one exception is the estimates on the CGSSM model on the currency GBPUSD. The α estimate has changed from 1.180 in the full sample estimation to -1.162 in the subsample estimation. But the new estimate shows has a very large standard error, indicating potential identification problems for this encompassing specification.

Table VII compares the in-sample and out-of-sample performance of the six models based on the subsample estimation. We report the root mean squared pricing error (rmse), the mean daily log likelihood value ($\mathcal{L}/N$), and the pairwise likelihood ratio test statistics defined in equation (58). To facilitate comparison between in- and out-of-sample performance, we normalize the likelihood value ($\mathcal{L}$) by the number of weeks (N) for each sample period and report the mean daily log likelihood estimate ($\mathcal{L}/N$). The in-sample comparison is based on the first 310 weeks of data. The out-of-sample comparison is based on the last 109 weeks of data.

For each currency pair and each model, we first compare the in-sample and out-of-sample performance in terms of the root mean squared pricing error and the mean daily log likelihood value. We find that the in-sample and out-of-sample estimates are very close to one another. For JPYUSD, most models generate slightly larger out-of-sample pricing errors and smaller out-of-sample likelihood values than their in-sample counterpart. The one exception is the MJDSV model, which actually generates smaller out-of-sample error and larger out-of-sample likelihood value. For the currency pair GBPUSD, all models actually generate smaller out-of-sample pricing errors and larger out-of-sample likelihood values. Therefore, we do not observe much obvious degeneration or variation in out-of-sample performance. These results confirm our inference from the parameter estimates that the currency options behaviors during the past two years are not dramatically different from their earlier behaviors. The model parameter estimates from the first six years of data can be readily applied to the recent two years of data with no obvious degeneration in performance.

We now compare the performance of different models both in sample and out of sample. The root mean square and the log likelihood values show that the four SSM models perform much better than the MJDSV and HSTSV models, both in sample and out of sample. The likelihood ratio test statistics $\mathcal{M}$ tell the same story. For both in-sample and out-of-sample tests, the off-diagonal terms in the first column of the $\mathcal{M}$ matrix are all strongly positive for both currencies, indicating that all other models significantly outperform the Heston (1993) model. The last four elements of the second column are also strongly positive, indicating that our four SSM models all significantly outperform the MJDSV model of Bates (1996b).

Among the four SSM models, the in-sample $\mathcal{M}$ statistics show that the four models are not statistically different from one another for both currencies. When we look at the out-of-sample statistics for JPYUSD, we find that the CG jump structure significantly outperforms all three restricted jump specifications (KJ, VG, and CJ). Among the three restricted jump specifications, CJ significantly outperforms KJ and VG; VG significantly outperforms KJ, thus generating the following statistically significant performance ranking in descending order: CG, CJ, VG, and KJ. The qualitative conclusion is the same

as from the in-sample comparison, but statistically stronger: High frequency jumps perform better in capturing the option price behavior on JPYUSD.

For GBPUSD, the out-of-sample performance ranking among the four jump specifications under SSM goes the opposite direction, albeit with less statistically significance. In particular, although the encompassing CG jump specification generates slightly better in-sample performance, its out-of-sample performance is significantly worse than KJ and VG. Thus, options on GBPUSD seem to ask for a more parsimonious and less frequent jump specification.

Reviewing the options behavior on JPYUSD and GBPUSD, we find that historically, JPYUSD options have generated much larger skews (risk reversals) than options on GBPUSD. Thus, we conclude that high-frequency jump specification is needed for capturing large non-normalities, but a finite-activity jump specification suffices for capturing moderate non-normalities in the return distribution.

C. Pricing Biases

Another way to investigate the robustness and performance of different models is to check for remaining structures in the pricing errors of these models. Since we have documented the evidence mainly in the implied volatility space, here we convert the model-implied option prices into Black-Scholes implied volatilities. We define the pricing error in the volatility space as the difference between the observed implied volatility quote and the corresponding values computed from the model.

The meaning pricing error of a good model should be close to zero and show no obvious structures along both the moneyness and the maturity dimensions. Figure 5 plots the mean pricing error in volatility percentage points along the moneyness dimension at selected maturities of one month (solid lines), three months (dashed lines), and 12 months (dash-dotted lines). Since the in-sample and out-of-sample performances are similar for all models, from now on we only report results based on the full-sample model estimation. To further reduce graphics clustering, we henceforth focus on two models, one from our four SSM specifications and one from the two traditional specifications. The four SSM models generate similar performance, we choose KJSSM as the representative. Of the two traditional models,

the Bates model (MJDSV) performs better than the pure-diffusion Heston model (HSTSV). We choose the better performing MJDSV and compare its behaviors to KJSSM.

Under the MJDSV model, the mean pricing errors show obvious remaining structures for JPYUSD along both the moneyness and maturity dimensions. At short maturities, the mean pricing errors exhibit a smile shape along the moneyness dimension, implying that the MJDSV model cannot fully account for the implied volatility smile at short maturities. At longer maturities, the mean pricing errors show an inverse smile shape along the moneyness dimension, implying that the MJDSV model generates an overly curved implied volatility smile at these maturities.

In contrast, under our SSM model, the mean pricing errors are very close to zero and do not show any obvious remaining structures. For both currencies, the mean pricing errors under the SSM model are all well within half a percentage point, the average bid-ask spread for the implied volatility quotes.

While the mean pricing error plots can reveal the remaining structures and deficiencies of a model, a plot of the mean absolute pricing errors illustrates the average performance of the model in fitting the observed implied volatility quotes. Figure 6 plots the mean absolute pricing error in implied volatility under both MJDSV and KJSSM. Under both models, the mean absolute pricing errors are smaller for GBPUSD than for JPYUSD.

Under the MJDSV model, the mean absolute pricing errors are larger on out-of-money options than on at-the-money options, indicating that the MJDSV model cannot fully account for the observed implied volatility smile. The mean absolute pricing errors are also larger at very short and long maturities than at moderate maturities, indicating that the model cannot fully account for the term structure of the implied volatilities.

The mean absolute pricing errors under the SSM model are smaller than those under the MJDSV model across all moneyness, maturities, and the two currency pairs, showing the universal better performance of the SSM model over the MJDSV model. Furthermore, under the SSM model, the mean absolute pricing error are flat across moneyness under all maturities, indicating that the model can capture the volatility smile very well. Along the maturity dimension, the mean absolute pricing errors are

smaller at moderate maturities than at very short and very long maturities, indicating that the model has some remaining tensions along the term structure dimension.

D. The Activity Rate Dynamics

Under the SSM models, the risk-neutral dynamics of the two activity rates are mainly controlled by two parameters: κ and σ_v. The parameter κ controls the speed of mean-reversion for the activity rate processes. The parameter σ_v controls the instantaneous volatility of the process. Furthermore, the activity rate processes interact with the currency return innovation through the instantaneous correlation parameters ρ^R and ρ^L. Under the physical measure, the time-series dynamics of the activity rates differ from the risk-neutral dynamics in terms of the mean-reverting speeds κ^P and the long-run mean θ^P. The difference between κ and κ^P captures the market price of volatility risk. When the market price of risk coefficient γ is positive, the time-series dynamics of the activity rates are more persistent and also have a larger long-run mean than the risk-neutral dynamics. The opposite is true when the coefficient is negative.

Table IV reports the full-sample parameter estimates. For JPYUSD, the estimate for the risk-neutral mean-reversion speed κ varies from 0.387 to 0.465 as we change the jump specification. The mean-reversion speeds under the time-series measure κ^P are larger and range from 0.502 to 0.586. The difference between the two sets of parameters imply that the market price of activity rate is negative.

For GBPUSD, the κ estimates for the SSM models are larger and between 1.18 and 1.211. The corresponding time-series estimates are between 1.158 and 3.296, implying negative market price of risk except under CJSSM. Nevertheless, we caution the interpretation of the market price of risk as we observe that the standard errors for the time-series estimates κ^P are much larger than that for the risk-neutral counterparts κ. The observation holds for both currencies.

The estimates for the instantaneous volatility coefficient of the activity rates σ_v are also quite stable across different jump specifications under the SSM framework. The estimates are between 1.566 and 1.675 for JPYUSD and between 1.429 and 1.505 for GBPUSD.

The estimates for the instantaneous correlation are significantly positive between the positively skewed Lévy component and its activity rate, and are strongly negative between the negatively-skewed Lévy component and its activity rate. As a result, the model-implied innovation in the risk reversal is positively correlated with the currency return, consistent with the observations from the data (Table II).

Under the HSTSV and MJDSV models, a scalar activity rate process controls the overall stochastic volatility. The estimates for the persistence parameters κ and κ^P and for the instantaneous volatility parameter σ_v are similar to those obtained under the SSM models. However, the instantaneous correlation ρ estimates are close to zero under both currencies, consistent with our observation that the currency returns and changes in volatilities do not have strong cross-correlations.

The unscented Kalman Filter provides a fast way to update the activity rates to achieve an approximate fit to the implied volatility surface. The top two panels in Figure 7 plot the filtered activity rates for the MJDSV model. In the bottom two panels, we plot the filtered activity rates of both the right-skewed (solid lines) and left-skewed (dashed lines) return components under the KJSSM model.

Under both models, the overall time variation of the activity rates match the ups and downs in the time series of the implied volatilities in Figure 1. Hence, both models can capture the stochastic volatility feature of the currency options. For example, the implied volatilities on JPYUSD show a large spike between 1998 and 1999, reflecting the market stress during the Russian bond crisis and the ensuing hedge fund crisis. The single activity rate process under MJDSV shows a similar spike, and the two activity rates under the SSM model are both high during this period.

Furthermore, under the SSM model, the relative variation of the two activity also matches the time variation in the risk reversals plotted in Figure 3. When the risk reversal is positive, the activity rate for the right-skewed return component (v_t^R, solid lines) is higher than the activity rate for the left-skewed component (v_t^L, dashed lines), and vice versa. To see this feature more clearly, we plot in Figure 8 the percentage differences between the two activity rates, defined as $100 \times (v_t^R - v_t^L)/(v_t^R + v_t^L)$. We also re-plot the ten-delta risk reversal quotes for comparison. We see that the movement of the percentage differences in the activity rates matches the movements of the risk reversals very well.

E. Theory and Evidence on the Stochastic Skew

The key feature that we have observed from the OTC currency options market is the strong time variation in the risk reversal, and hence the stochastic skew. Using the filtered time series on the activity rates, we compute the model-implied option prices and Black-Scholes implied volatilities. From the implied volatilities, we re-construct the model-implied risk reversals and compare them with the market observations.

Figure 9 compares the time series of the observed risk reversals to the model-implied values. For clarity, we only plot one time series for each currency pair: the ten-delta risk reversal at three-month maturity in percentages of the at-the-money implied volatility of the same maturity. The dash-dotted lines denote data quotes, the solid lines are the values computed from the estimated models.

The MJDSV model can generate the overall stochastic volatilities observed in the data, but the top two panels in Figure 9 show that this model fails miserably in capturing the observed strong variation in risk reversals. Compared to the strong variations in the data (dashed lines), the model-implied values vary very little.

In contrast, the bottom two panels in Figure 9 show that our SSM models can generate risk reversals that match the data very closely. The matches are close to perfection except under extreme realizations. Therefore, our SSM modeling framework contributes to the literature by capturing the strong and unique feature of the OTC currency options market.

V. Conclusion

In this paper, we document the statistical properties of currency option implied volatilities across the dimensions of moneyness, maturity, and calendar time. We find that the market prices of OTC currency options exhibit several unique behaviors that challenge standard models in the option pricing literature. Chief among the challenging behaviors is the observation that although the average implied

volatility smile is relatively symmetric, the risk reversals can take large values on any given date and that these values vary greatly over time, so much so that the sign of the risk reversal can also change.

Using the time-changed Lévy process framework, we design and estimate a subclass of models that capture this unique stochastic skew behavior of currency option prices. Our estimation results show that our SSM models strongly outperform traditional jump-diffusion stochastic volatility models, both in sample and out of sample.

REFERENCES

Andersen, Torben G., Luca Benzoni, and Jesper Lund, 2002, An empirical investigation of continuous-time equity return models, *Journal of Finance* 57, 1239–1284.

Andrews, Donald, 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* 59, 817–858.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, 2003–2049.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 2000a, Do call prices and the underlying stock always move in the same direction?, *Review of Financial Studies* 13, 549–584.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 2000b, Pricing and hedging long-term options, *Journal of Econometrics* 94, 277–318.

Bates, David, 1996a, Dollar jump fears, 1984-1992: Distributional abnormalities implicit in foreign currency futures options, *Journal of International Money and Finance* 15, 65–93.

Bates, David, 1996b, Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options, *Review of Financial Studies* 9, 69–107.

Bates, David, 2000, Post-'87 crash fears in the S&P 500 futures option market, *Journal of Econometrics* 94, 181–238.

Bertoin, Jean, 1996, *Lévy Processes*. (Cambridge University Press Cambridge).

Black, Fisher, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–654.

Bollen, Nick P., 1998, Valuing options in regime-switch models, *Journal of Derivatives* 6, 38–49.

Bollen, Nick P., Stephen F. Gray, and Robert E. Whaley, 2000, Regime-switching in foreign exchange rates: Evidence from currency option prices, *Journal of Econometrics* 94, 239–276.

Bollen, Nick P., and Emma Raisel, 2003, The performance of alternative valuation models in the OTC currency option market, *Journal of International Money and Finance* 22, 33–64.

Broadie, Mark, Mike Chernov, and Michael Johannes, 2004, Specification and risk premiums: The information in S&P 500 futures options, Working paper, Columbia University.

Campa, Jose Manuel, and P. H. Kevin Chang, 1995, Testing the expectations hypothesis on the term structure of volatilities in foreign exchange options, *Journal of Finance* 50, 529–547.

Campa, Jose Manuel, P. H. Kevin Chang, and R. L. Reider, 1998, Implied exchange rate distributions: Evidence from otc options market, *Journal of International Money and Finance* 17, 117–160.

Carr, Peter, Hélyette Geman, Dilip Madan, and Marc Yor, 2002, The fine structure of asset returns: An empirical investigation, *Journal of Business* 75, 305–332.

Carr, Peter, and Dilip Madan, 1999, Option valuation using the fast Fourier transform, *Journal of Computational Finance* 2, 61–73.

Carr, Peter, and Liuren Wu, 2003, Finite moment log stable process and option pricing, *Journal of Finance* 58, 753–777.

Carr, Peter, and Liuren Wu, 2004, Time-changed Lévy processes and option pricing, *Journal of Financial Economics* 71, 113–141.

Eraker, Bjørn, 2003, Do stock prices and volatility jump? Reconciling evidence from spot and option prices, *Journal of Finance* forthcoming.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in equity index volatility and returns, *Journal of Finance* 58, 1269–1300.

Garman, Mark, and Steven Kohlhagen, 1983, Foreign currency option values, *Journal of International Money and Finance* 2, 231–237.

Heston, Stephen, 1993, Closed-form solution for options with stochastic volatility, with application to bond and currency options, *Review of Financial Studies* 6, 327–343.

Huang, Jingzhi, and Liuren Wu, 2004, Specification analysis of option pricing models based on time-changed Lévy processes, *Journal of Finance* 59, 1405–1439.

Jacod, Jean, and Albert N. Shiryaev, 1987, *Limit Theorems for Stochastic Processes*. (Springer-Verlag Berlin).

Julier, Simon J., and Jeffrey K. Uhlmann, 1997, A new extension of the kalman filter to nonlinear systems, Working paper, University of Oxford Sydney, Australia.

Kalman, Rudolph Emil, 1960, A new approach to linear filtering and prediction problems, *Transactions of the ASME–Journal of Basic Engineering* 82, 35–45.

Kou, Steven G., 2002, A jump-diffusion model for option pricing, *Management Science* 48, 1086–1101.

Madan, Dilip, and Eugene Seneta, 1990, The variance gamma (V.G.) model for share market returns, *Journal of Business* 63, 511–524.

Madan, Dilip B., Peter P. Carr, and Eric C. Chang, 1998, The variance gamma process and option pricing, *European Finance Review* 2, 79–105.

Madan, Dilip P., and Elton Daal, 2004, An empirical investigation of the variance-gamma model for foreign currency options, Working paper, University of Maryland.

Merton, Robert C., 1976, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics* 3, 125–144.

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

Pan, Jun, 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.

Titchmarsh, Edward Charles, 1986, *Introduction to the Theory of Fourier Integrals*. (Chelsea Publication Co. New York) 3rd edn.

Vuong, Quang H., 1989, Likelihood ratio tests for model selection and non-nested hypotheses, *Econometrica* 57, 307–333.

Wu, Liuren, 2004, Dampened power law: Reconciling the tail behavior of finanical asset returns, *Journal of Business* forthcoming.

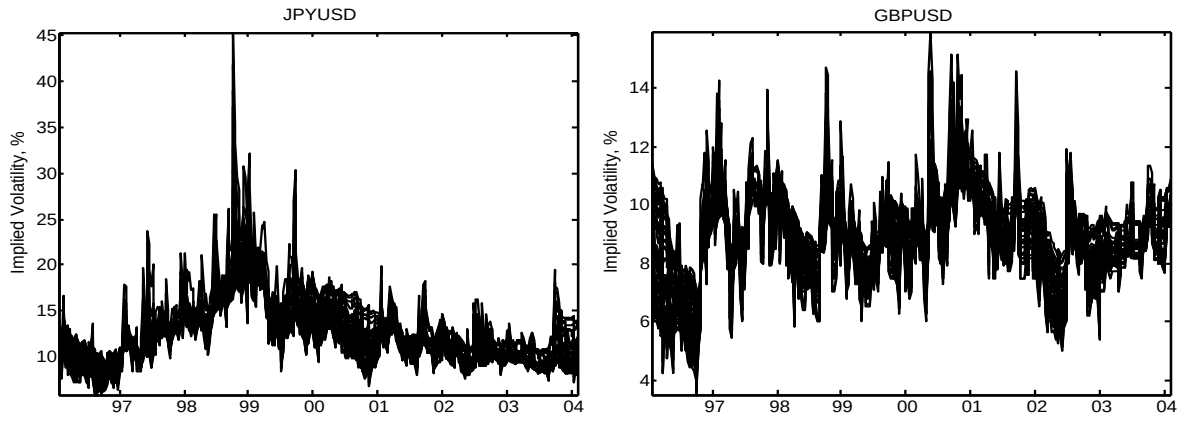

Figure 1. The time series of OTC currency option implied volatilities. Lines plot the time-series of 40 implied volatility quotes on the dollar price of yen (JPYUSD, left panel) and pound (GBPUSD, right panel). The 40 series are from eight maturities and five strike levels at each maturity. Data are weekly from January 24, 1996 to January 28, 2004, 419 observations for each series.

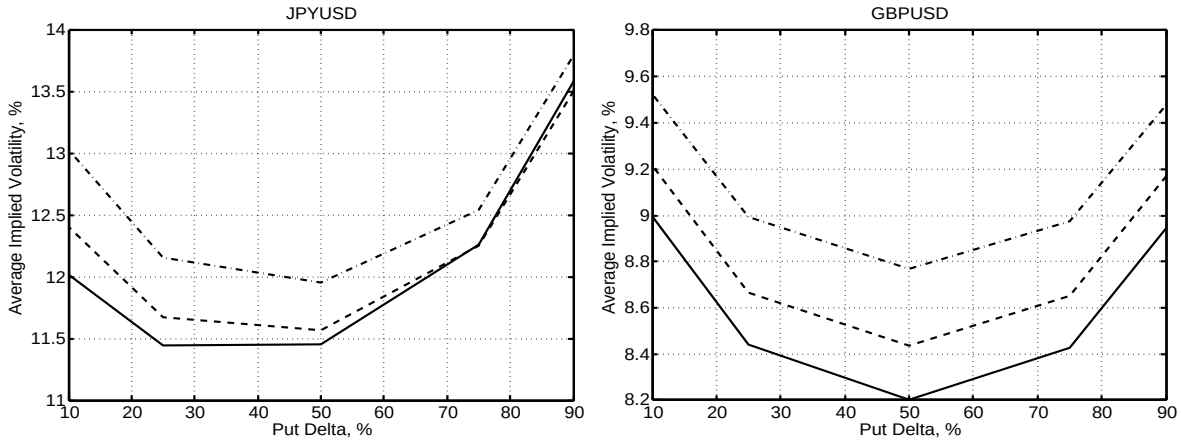

Figure 2. Mean implied volatility smiles on currency options. Lines plot the time-series average of the implied volatility against the delta of the currency options at three option maturities: one month (solid lines), three months (dashed lines), and one year (dash-dotted lines). The averages are on weekly data from January 24, 1996 to January 28, 2004, 419 observations for each series.

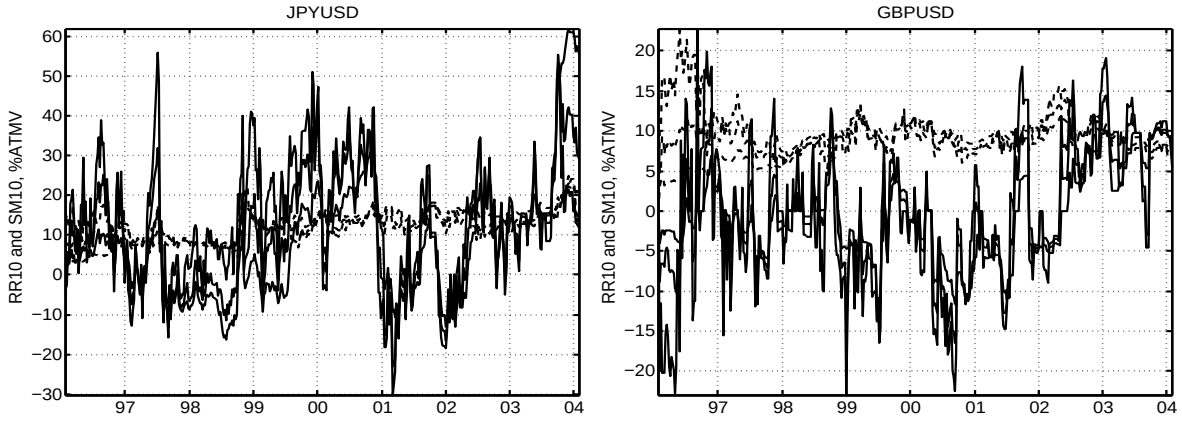

Figure 3. Risk reversals and strangle margins over calendar time. Solid lines are ten-delta risk reversals and dashed lines are ten-delta strangle margins, both in percentages of the at-the-money implied volatility. To reduce clustering, we plot the lines at three maturities (one, three, and 12 months).

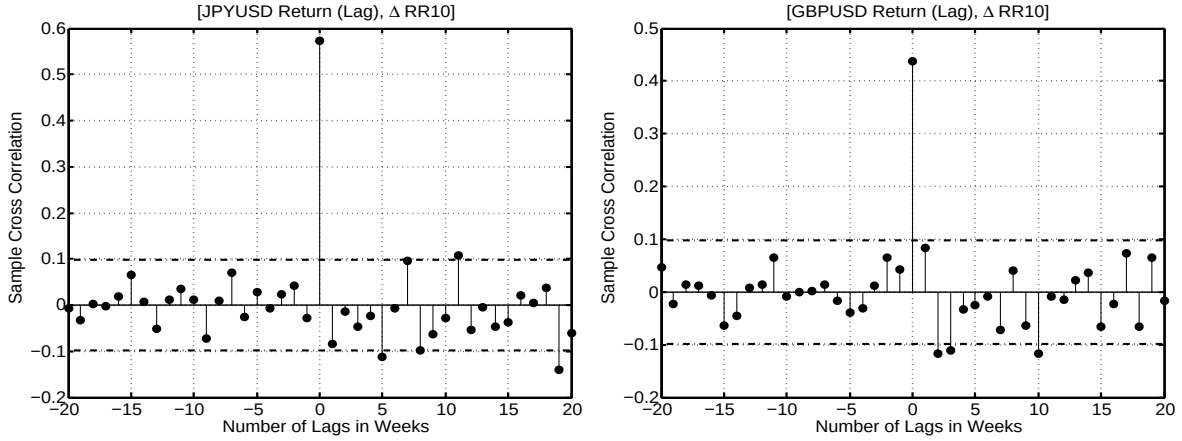

Figure 4. Cross-Correlogram between currency returns and changes in one-month ten-delta risk reversals. The stem bars represent the cross-correlation estimates between the currency returns and the weekly changes in one-month ten-delta risk reversals at different lags and leads. The two dashed lines in each panel denote the 95 confidence bands.

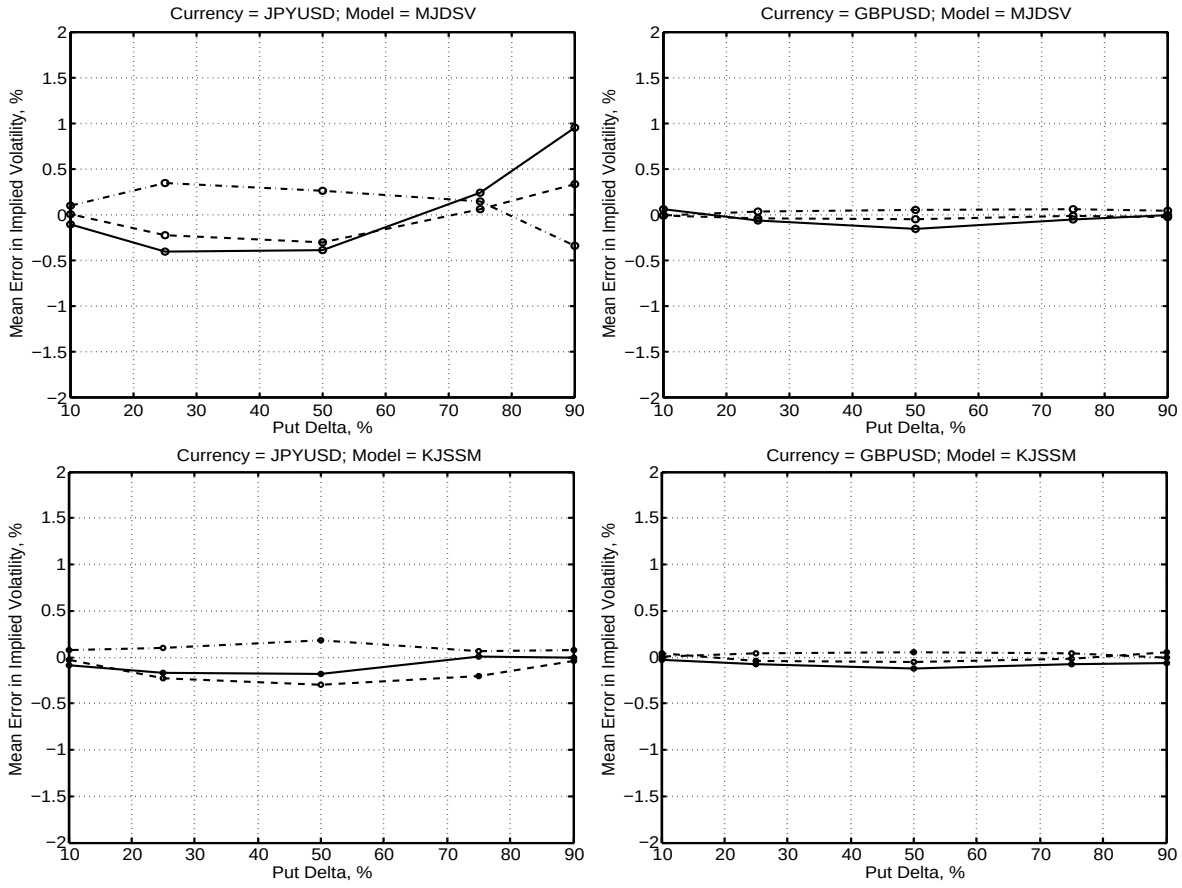

Figure 5. Mean pricing bias in implied volatility. We define the pricing error as the difference between the observed implied volatility quote and the corresponding value implied by the estimated models, both in percentages. We then compute the mean pricing error at each moneyness and maturity. The three lines represent three chosen maturities at one month (solid lines), three months (dashed lines), and 12 months (dash-dotted lines).

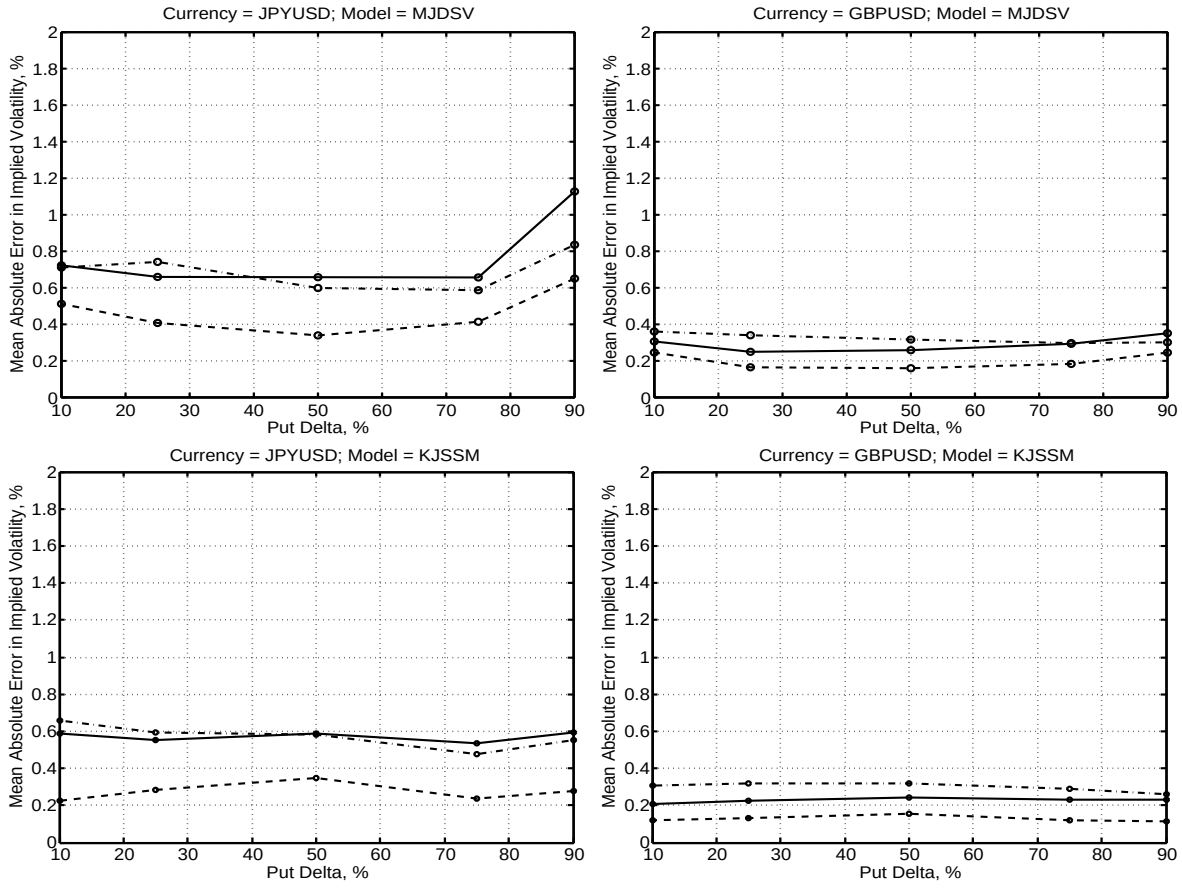

Figure 6. Mean absolute pricing error in implied volatility. We define the pricing error as the difference between the observed implied volatility quote and the corresponding value implied by the estimated models, both in volatility percentages. We compute the mean absolute value of the pricing errors at each moneyness and maturity. The three lines represent three chosen maturities at one month (solid lines), three months (dashed lines), and 12 months (dash-dotted lines).

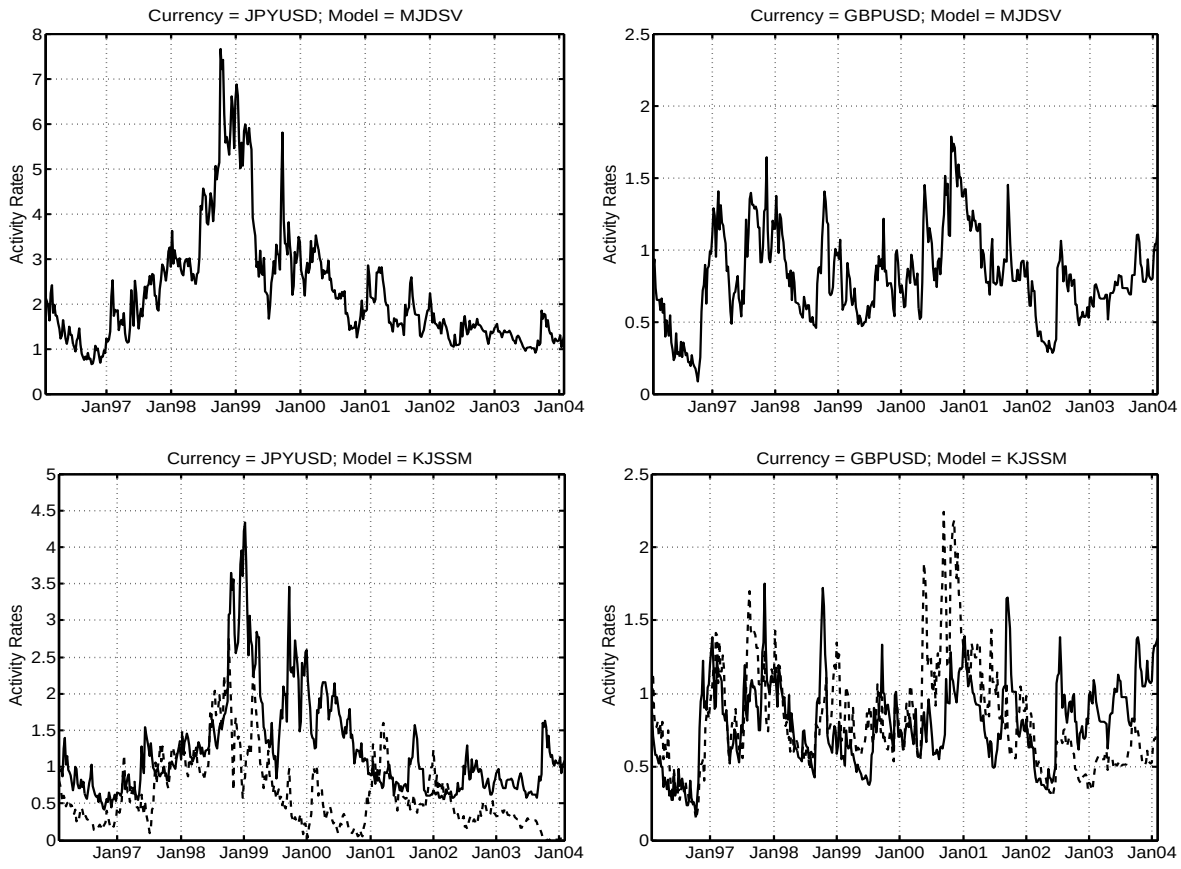

Figure 7. Filtered activity rates. The top two panels plot the single series of the activity rates from the MJDSV model. The bottom two panels plot the two activity rate series from the KJSSM model. The solid lines denote the activity rate for the right-skewed Lévy component and the dashed lines denote the activity rate for the left-skewed Lévy component under the SSM model. We extract the activity rates from the options data using unscented Kalman filter, based on the relevant estimated models using the whole sample of data.

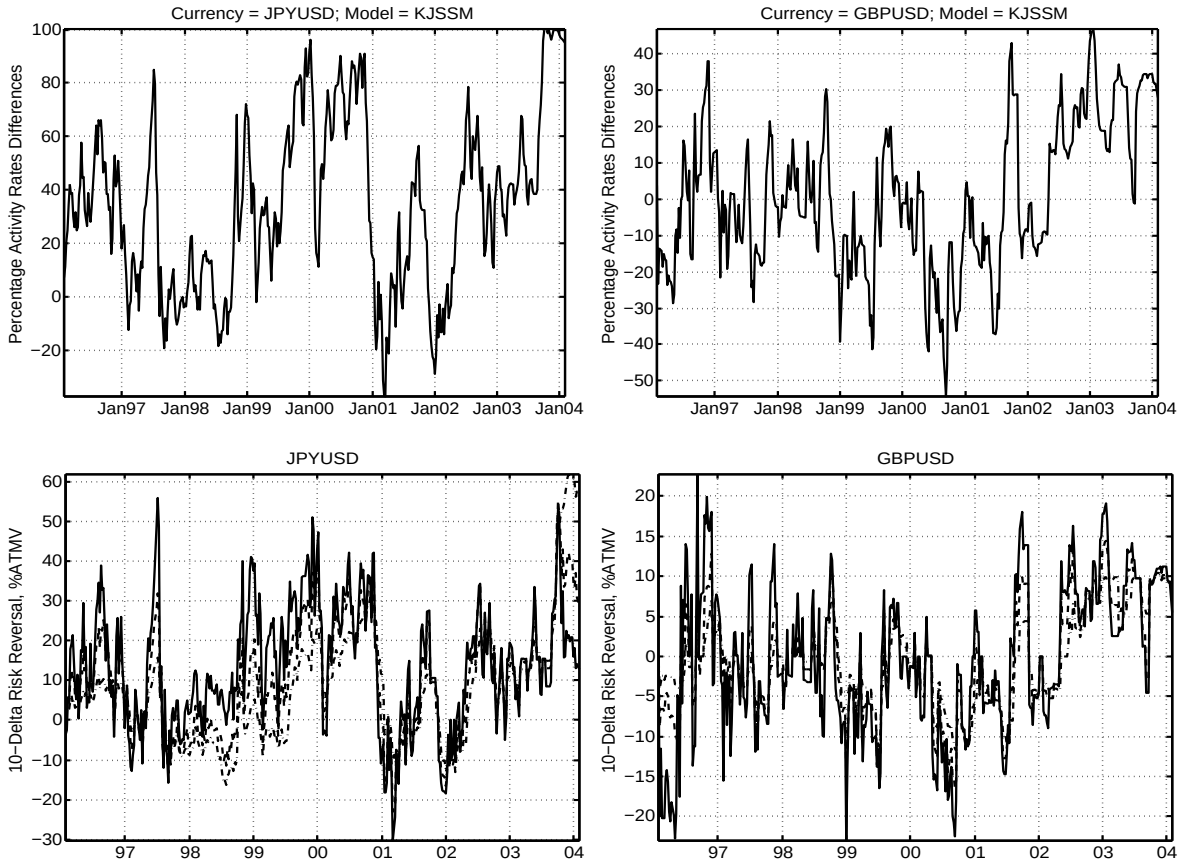

Figure 8. Activity rates differences and risk reversals. The top two panels plot the percentage difference in the two activity rates from the KJSSM model, defined as single series of the activity rates from the $100 \times (v_t^R - v_t^L)/(v_t^R + v_t^L)$. The bottom two panels plot the market quotes for ten-delta risk reversals for comparison. The three lines represent the three selected maturities at one, three, and 12 months.

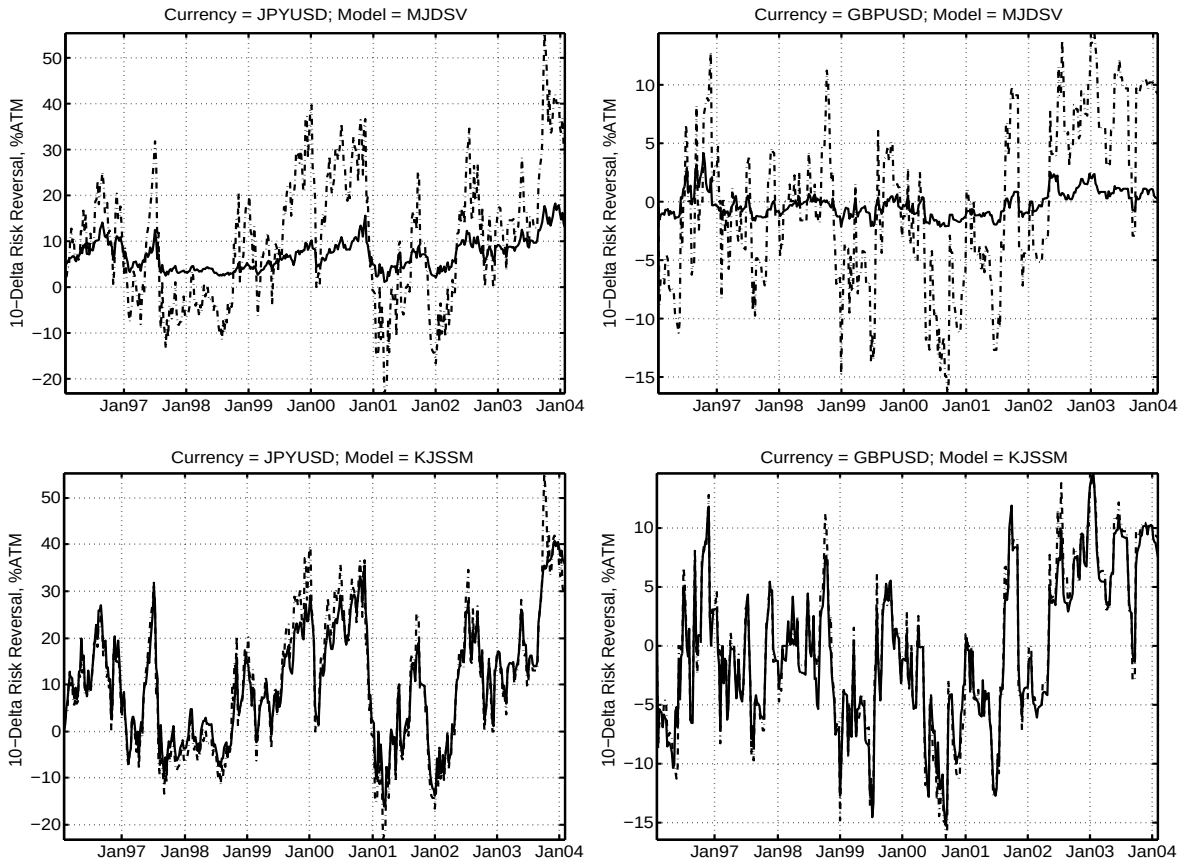

Figure 9. Theory and evidence on the stochastic skew. Dashed lines are the market quotes on three-month ten-delta risk reversals, in percentages of the at-the-money implied volatility of the same maturity. Solid lines are the values computed from the estimated models using the whole sample of data.

Table I
Summary Statistics of Currency Option Implied Volatilities

The three columns under each contract report the mean, standard deviation, and weekly autocorrelation of the contract on risk reversal (RR), strangle margin (SM), and at-the-money implied volatilities (ATMV). Risk reversals and strangles are in percentages of the at-the-money implied volatility. The numbers following RR and SM denote the delta of the contract. Data are weekly from January 24, 1996 to January 28, 2004, 419 observations for each series. The first column denotes the option maturities, with 'w' denoting weeks and 'm' denoting months.

Mat	RR10			SM10			RR25			SM25			ATMV		
							JPYUSD								
1w	15.18	16.96	0.69	14.34	4.26	0.77	7.40	8.10	0.70	4.32	1.47	0.85	11.70	3.80	0.83
1m	13.32	15.21	0.85	12.15	3.40	0.89	6.90	8.04	0.87	3.60	0.88	0.87	11.45	3.10	0.92
2m	11.53	14.27	0.89	12.08	3.21	0.92	6.02	7.63	0.91	3.51	0.67	0.87	11.47	2.84	0.94
3m	10.16	14.14	0.92	12.20	3.29	0.94	5.34	7.60	0.93	3.47	0.64	0.89	11.57	2.70	0.96
6m	8.25	14.32	0.96	12.30	3.67	0.96	4.30	7.63	0.96	3.41	0.72	0.94	11.78	2.58	0.97
9m	7.77	14.66	0.97	12.42	4.11	0.98	4.01	7.74	0.97	3.39	0.82	0.96	11.87	2.55	0.98
12m	7.45	14.99	0.97	12.39	4.48	0.98	3.81	7.91	0.97	3.34	0.90	0.97	11.95	2.53	0.98
18m	7.95	14.42	0.97	12.03	4.95	0.98	4.00	7.61	0.97	3.17	1.00	0.97	12.00	2.49	0.98
							GBPUSD								
1w	-0.14	11.76	0.73	10.30	4.60	0.86	0.13	5.72	0.76	2.95	1.50	0.89	8.20	1.79	0.81
1m	-0.52	9.35	0.84	9.74	3.04	0.91	-0.11	4.68	0.84	2.95	0.86	0.88	8.20	1.47	0.90
2m	-0.33	7.48	0.88	9.22	1.83	0.87	-0.05	3.95	0.89	2.77	0.57	0.87	8.33	1.31	0.92
3m	-0.37	6.74	0.90	9.11	1.56	0.86	-0.10	3.55	0.91	2.72	0.47	0.84	8.43	1.20	0.93
6m	-0.44	5.92	0.94	8.80	1.72	0.92	-0.15	3.13	0.95	2.59	0.52	0.89	8.61	1.02	0.95
9m	-0.38	5.60	0.96	8.63	1.95	0.95	-0.14	2.98	0.96	2.55	0.56	0.92	8.69	0.95	0.95
12m	-0.36	5.45	0.96	8.46	2.11	0.96	-0.14	2.91	0.97	2.49	0.55	0.92	8.77	0.90	0.95
18m	-0.53	4.93	0.97	7.99	2.38	0.97	-0.24	2.63	0.97	2.26	0.61	0.94	8.88	0.89	0.95

Table II
Cross-correlation Between Currency Returns and Changes in Implied Volatilities

Entries report the contemporaneous correlation between log currency returns and changes in risk reversals (RR), strangle margins (SM), and at-the-money implied volatilities (ATMV). Risk reversals and strangles are in percentages of the at-the-money implied volatility level. The numbers following RR and SM denote the delta of the contract. The first column denotes the option maturities, with 'w' denoting weeks and 'm' denoting months. Data are weekly from January 24, 1996 to January 28, 2004, 419 observations for each series.

	JPYUSD					GBPUSD				
Mat	RR10	SM10	RR25	SM25	ATMV	RR10	SM10	RR25	SM25	ATMV
1w	0.46	-0.06	0.48	-0.14	0.41	0.38	-0.01	0.40	-0.02	-0.02
1m	0.57	-0.06	0.58	-0.14	0.44	0.44	0.01	0.45	0.01	-0.00
2m	0.58	-0.05	0.59	-0.10	0.40	0.46	-0.01	0.46	0.02	0.02
3m	0.59	-0.06	0.59	-0.08	0.35	0.47	0.03	0.47	0.03	0.00
6m	0.59	-0.04	0.59	-0.04	0.25	0.44	0.04	0.45	0.04	0.02
9m	0.56	-0.04	0.57	-0.02	0.21	0.42	0.03	0.43	0.03	0.04
12m	0.57	-0.03	0.58	0.00	0.18	0.39	0.05	0.40	0.05	0.04
18m	0.53	-0.05	0.55	-0.01	0.18	0.37	0.06	0.37	0.07	0.02

Table III
Characteristic Exponents of Different Lévy Components

All Lévy specifications have a diffusion component. The characteristic exponent for the diffusion component is $\psi^D = \frac{1}{2}\sigma^2\left(iu+u^2\right)$.

Model	Right-Skewed Component	Left-Skewed Component
KJ	$-iu\lambda\left[\frac{1}{1-iuv_j}-\frac{1}{1-v_j}\right]+\psi^D$	$iu\lambda\left[\frac{1}{1+iuv_j}-\frac{1}{1+v_j}\right]+\psi^D$
VG	$\lambda\ln\left(1-iuv_j\right)-iu\lambda\ln\left(1-v_j\right)+\psi^D$	$\lambda\ln\left(1+iuv_j\right)-iu\lambda\ln\left(1+v_j\right)+\psi^D$
CJ	$-\lambda\left(1/v_j-iu\right)\ln(1-iuv_j)$ $+iu\lambda\left(1/v_j-1\right)\ln(1-v_j)+\psi^D$	$-\lambda\left(1/v_j+iu\right)\ln(1+iuv_j)$ $+iu\lambda\left(1/v_j+1\right)\ln(1+v_j)+\psi^D$
CG	$\lambda\Gamma(-\alpha)\left[\left(\frac{1}{v_j}\right)^\alpha-\left(\frac{1}{v_j}-iu\right)^\alpha\right]$ $-iu\lambda\Gamma(-\alpha)\left[\left(\frac{1}{v_j}\right)^\alpha-\left(\frac{1}{v_j}-1\right)^\alpha\right]+\psi^D$	$\lambda\Gamma(-\alpha)\left[\left(\frac{1}{v_j}\right)^\alpha-\left(\frac{1}{v_j}+iu\right)^\alpha\right]$ $-iu\lambda\Gamma(-\alpha)\left[\left(\frac{1}{v_j}\right)^\alpha-\left(\frac{1}{v_j}+1\right)^\alpha\right]+\psi^D$

Table IV
Likelihood Estimates of Model Parameters

Entries report the quasi-maximum likelihood estimates of the model parameters, standard errors (in parentheses), root mean squared pricing errors (rmse), and log likelihood values ($\mathcal{L}$). For each currency pair, we estimate six models: the Heston (1993) model (HSTSV), the Bates (1996b) model (MJDSV), and our stochastic skew models (SSM) with four different jump specifications: KJ, VG, CJ, and CG. The estimation uses eight years of weekly option data from January 24, 1996 to January 28, 2004 (419 weekly observations for each series). The column under "Θ_B" denotes the parameter names for the Heston model and the Bates model. The column under "Θ_S" denotes the parameter names for our SSM models.

Currency		JPYUSD						GBPUSD					
Θ_B	Θ_S	HSTSV	MJDSV	KJSSM	VGSSM	CJSSM	CGSSM	HSTSV	MJDSV	KJSSM	VGSSM	CJSSM	CGSSM
σ^2	σ^2	0.020	0.006	0.006	0.005	0.004	0.003	0.010	0.008	0.003	0.003	0.002	0.002
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
λ	λ	—	0.016	0.059	1.708	0.035	0.004	—	0.422	0.079	6.869	0.080	0.032
		(—)	(0.001)	(0.003)	(0.151)	(0.002)	(0.001)	(—)	(0.044)	(0.005)	(0.700)	(0.005)	(0.015)
v_j	v_j	—	0.497	0.029	0.045	0.104	0.270	—	0.003	0.012	0.017	0.031	0.039
		(—)	(0.013)	(0.001)	(0.001)	(0.004)	(0.056)	(—)	(0.000)	(0.000)	(0.001)	(0.001)	(0.004)
κ	κ	0.559	0.569	0.387	0.394	0.421	0.465	1.532	1.044	1.205	1.206	1.211	1.180
		(0.006)	(0.011)	(0.005)	(0.006)	(0.007)	(0.010)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.008)
σ_v	σ_v	1.837	1.210	1.675	1.657	1.582	1.566	2.198	1.737	1.429	1.447	1.505	1.492
		(0.023)	(0.022)	(0.027)	(0.028)	(0.027)	(0.031)	(0.026)	(0.023)	(0.039)	(0.040)	(0.017)	(0.018)
ρ	ρ^R	0.076	0.123	0.395	0.393	0.400	0.424	-0.023	-0.061	0.848	0.848	0.849	0.836
		(0.005)	(0.065)	(0.017)	(0.018)	(0.022)	(0.056)	(0.003)	(0.017)	(0.040)	(0.043)	(0.017)	(0.016)
μ_j	ρ^L	—	-0.210	-0.739	-0.758	-0.851	-1.000	—	0.002	-1.000	-0.999	-1.000	-1.000
		(—)	(0.024)	(0.034)	(0.036)	(0.040)	(0.144)	(—)	(0.001)	(0.047)	(0.050)	(0.000)	(0.004)
κ^P	κ^P	0.745	0.258	0.522	0.502	0.544	0.586	1.276	0.800	2.062	2.092	1.158	3.296
		(0.396)	(0.114)	(0.289)	(0.288)	(0.251)	(0.261)	(0.345)	(0.236)	(0.213)	(0.213)	(0.006)	(0.223)
σ_r	σ_r	1.045	1.002	0.704	0.703	0.703	0.700	0.198	0.184	0.148	0.148	0.148	0.148
		(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
—	α	—	—	—	—	—	1.602	—	—	—	—	—	1.180
		—	—	—	—	—	(0.126)	—	—	—	—	—	(0.155)
rmse		1.014	0.984	0.822	0.822	0.822	0.820	0.445	0.424	0.376	0.376	0.376	0.378
$\mathcal{L}, \times 10^3$		-9.430	-9.021	-6.416	-6.402	-6.384	-6.336	4.356	4.960	6.501	6.502	6.497	6.521

Table V
In-Sample Likelihood Ratio Tests of Model Performance Differences

Entries report the pairwise likelihood ratio test statistics $\mathcal{M}$ constructed by Vuong (1989) on non-nested models. The statistic has an asymptotic standard normal distribution. We report the pairwise statistics in a (6×6) matrix, with the (i, j)th element denoting the statistic on model *i* versus model *j* such that a strongly positive estimate for this element indicates that model *i* significantly outperforms model *j*. The tests are in sample, based on the model estimations using the full sample of eight years of data for each currency.

$\mathcal{M}$	HSTSV	MJDSV	KJSSM	VGSSM	CJSSM	CGSSM
			JPYUSD			
HSTSV	0.00	-2.55	-4.92	-4.88	-4.75	-4.67
MJDSV	2.55	0.00	-5.39	-5.33	-5.22	-5.07
KJSSM	4.92	5.39	0.00	-1.11	-0.86	-1.20
VGSSM	4.88	5.33	1.11	0.00	-0.72	-1.21
CJSSM	4.75	5.22	0.86	0.72	0.00	-1.59
CGSSM	4.67	5.07	1.20	1.21	1.59	0.00
			GBPUSD			
HSTSV	0.00	-2.64	-4.70	-4.68	-4.63	-4.71
MJDSV	2.64	0.00	-3.85	-3.86	-3.89	-4.19
KJSSM	4.70	3.85	0.00	-0.04	0.34	-0.37
VGSSM	4.68	3.86	0.04	0.00	0.56	-0.39
CJSSM	4.63	3.89	-0.34	-0.56	0.00	-0.51
CGSSM	4.71	4.19	0.37	0.39	0.51	0.00

Table VI
Subsample Likelihood Estimates of Model Parameters

Entries report the quasi-maximum likelihood estimates of the model parameters and their standard errors (in parentheses). For each currency pair, we estimate six models: the Heston (1993) model (HSTSV), the Bates (1996b) model (MJDSV), and our stochastic skew models (SSM) with four different jump specifications: KJ, VG, CJ, and CG. The estimation uses the first six years of weekly option data from January 24, 1996 to December 26, 2001 (310 weekly observations for each series). The column under "Θ_B" denotes the parameter names for the Heston model and the Bates model. The column under "Θ_S" denotes the parameter names for our SSM models.

Currency		JPYUSD						GBPUSD					
Θ_B	Θ_S	HSTSV	MJDSV	KJSSM	VGSSM	CJSSM	CGSSM	HSTSV	MJDSV	KJSSM	VGSSM	CJSSM	CGSSM
σ^2	σ^2	0.022	0.011	0.006	0.006	0.005	0.002	0.010	0.009	0.003	0.003	0.002	0.003
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
λ	λ	—	0.016	0.074	2.486	0.053	0.004	—	2.027	0.087	7.829	0.091	1210
		(—)	(0.001)	(0.004)	(0.234)	(0.004)	(0.002)	(—)	(0.153)	(0.006)	(0.922)	(0.007)	(9439)
v_j	v_j	—	0.491	0.027	0.041	0.087	0.273	—	0.001	0.012	0.017	0.030	0.011
		(—)	(0.018)	(0.001)	(0.001)	(0.004)	(0.089)	(—)	(0.000)	(0.000)	(0.001)	(0.001)	(0.006)
κ	κ	0.810	0.846	0.660	0.665	0.686	0.739	1.449	1.015	1.177	1.178	1.183	1.173
		(0.006)	(0.013)	(0.006)	(0.007)	(0.008)	(0.012)	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)	(0.012)
σ_v	σ_v	1.943	1.171	1.945	1.922	1.881	1.777	2.091	2.041	1.428	1.452	1.523	1.518
		(0.025)	(0.024)	(0.031)	(0.031)	(0.032)	(0.037)	(0.030)	(0.028)	(0.047)	(0.048)	(0.023)	(0.053)
ρ	ρ^R	0.050	0.062	0.270	0.267	0.252	0.299	-0.056	-0.065	0.796	0.794	0.789	0.720
		(0.005)	(0.078)	(0.015)	(0.016)	(0.018)	(0.092)	(0.005)	(0.013)	(0.047)	(0.050)	(0.022)	(0.053)
μ_j	ρ^L	—	-0.212	-0.629	-0.642	-0.672	-1.000	—	-0.001	-1.000	-0.999	-1.000	-0.905
		(—)	(0.033)	(0.035)	(0.037)	(0.041)	(0.396)	(—)	(0.000)	(0.059)	(0.062)	(0.000)	(0.069)
κ^P	κ^P	1.090	0.636	0.924	0.879	0.822	0.813	1.308	2.529	2.022	2.060	1.166	2.192
		(0.390)	(0.155)	(0.392)	(0.385)	(0.364)	(0.331)	(0.451)	(0.238)	(0.263)	(0.260)	(0.270)	(0.263)
σ_r	σ_r	1.095	1.072	0.746	0.747	0.746	0.744	0.217	0.200	0.175	0.175	0.175	0.174
		(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
—	α	—	—	—	—	—	1.691	—	—	—	—	—	-1.162
		—	—	—	—	—	(0.175)	—	—	—	—	—	(15.37)

Table VII
In-Sample and Out-of-Sample Model Performance Comparison

Entries report the root mean squared pricing error (rmse), mean daily log likelihood value ($\mathcal{L}/N$), and the pairwise likelihood ratio test statistics $\mathcal{M}$ constructed by Vuong (1989) on non-nested models. The models are estimated using data from January 24, 1996 to December 26, 2001 (310 weekly observations for each series). The in-sample statistics are from the same period. The out-of-sample statistics are computed from the remaining two years of data from January 2, 2002 to January 28, 2004 (109 weekly observations for each series) based on model parameter estimates from the first subsample.

	HSTSV	MJDSV	KJSSM	VGSSM	CJSSM	CGSSM	HSTSV	MJDSV	KJSSM	VGSSM	CJSSM	CGSSM
	JPYUSD						GBPUSD					
	In-Sample Performance											
rmse	1.04	1.02	0.85	0.85	0.85	0.85	0.47	0.44	0.41	0.41	0.41	0.41
$\mathcal{L}/N$	-23.69	-23.03	-16.61	-16.60	-16.57	-16.47	8.36	10.06	12.27	12.27	12.26	12.28
$\mathcal{M}$												
HSTSV	0.00	-2.14	-4.44	-4.41	-4.33	-4.17	0.00	-3.34	-4.42	-4.39	-4.24	-4.33
MJDSV	2.14	0.00	-4.74	-4.70	-4.61	-4.42	3.34	0.00	-3.40	-3.39	-3.33	-3.33
KJSSM	4.44	4.74	0.00	-0.49	-0.51	-0.84	4.42	3.40	0.00	0.08	0.36	-0.42
VGSSM	4.41	4.70	0.49	0.00	-0.51	-0.89	4.39	3.39	-0.08	0.00	0.51	-0.42
CJSSM	4.33	4.61	0.51	0.51	0.00	-1.14	4.24	3.33	-0.36	-0.51	0.00	-0.55
CGSSM	4.17	4.42	0.84	0.89	1.14	0.00	4.33	3.33	0.42	0.42	0.55	0.00
	Out-of-Sample Performance											
rmse	1.06	1.00	0.90	0.90	0.89	0.89	0.39	0.37	0.27	0.27	0.27	0.27
$\mathcal{L}/N$	-24.01	-21.75	-18.47	-18.35	-18.23	-18.11	14.36	15.85	23.30	23.29	23.26	23.25
$\mathcal{M}$												
HSTSV	0.00	-6.01	-5.90	-6.01	-6.08	-6.12	0.00	-4.88	-7.06	-7.06	-7.05	-7.05
MJDSV	6.01	0.00	-3.11	-3.23	-3.32	-3.48	4.88	0.00	-5.98	-5.99	-5.99	-5.97
KJSSM	5.90	3.11	0.00	-7.76	-6.81	-5.27	7.06	5.98	0.00	0.64	1.47	4.51
VGSSM	6.01	3.23	7.76	0.00	-4.39	-3.67	7.06	5.99	-0.64	0.00	1.63	4.19
CJSSM	6.08	3.32	6.81	4.39	0.00	-3.11	7.05	5.99	-1.47	-1.63	0.00	0.23
CGSSM	6.12	3.48	5.27	3.67	3.11	0.00	7.05	5.97	-4.51	-4.19	-0.23	0.00