Interest Rate Caps "Smile" Too! But Can the LIBOR Market Models Capture It?

ROBERT JARROW[a], HAITAO LI[b], AND FENG ZHAO[c]

December, 2004

[a] Jarrow is from Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853 (raj15@cornell.edu). [b] Li is from Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853 (hl70@cornell.edu). [c] Zhao is from School of Business, Fordham University, New York, NY 10023 (fz17@cornell.edu). We thank Warren Bailey, Jefferson Duarte, Richard Green (the editor), Marti Subrahmanyam, especially an anonymous referee, and seminar participants at the 2003 European Finance Association Meeting, 2004 Econometric Society Winter Meeting, 2004 Western Finance Association Meeting, 2004 Bank of Canada's fixed income workshop and Rutgers Business School for helpful comments. We are responsible for any remaining errors.

Interest Rate Caps "Smile" Too! But Can the LIBOR Market Models Capture It?

ABSTRACT

Using three years of interest rate caps price data, we provide one of the first comprehensive documentations of volatility smiles in the caps market. Using a multifactor term structure model with stochastic volatility and jumps, we develop a closed-form solution for cap prices and test the performance of our new models in capturing the volatility smile. We show that although a three-factor stochastic volatility model can price at-the-money caps well, significant negative jumps in interest rates are needed to capture the smile. The volatility smile contains information that is not available using only at-the-money caps, and this information is important for understanding term structure models.

JEL Classification: C4, C5, G1

The extensive literature on multifactor dynamic term structure models (hereafter DTSMs) of the last decade has mainly focused on explaining bond yields and swap rates.[1] Pricing and hedging over-the-counter interest rate derivatives, such as caps and swaptions, has attracted attention only in recent years. World wide, caps and swaptions are among the most widely traded interest rate derivatives. According to the Bank for International Settlements, in recent years, their combined notional value has been more than 10 trillion dollars, which is many times bigger than that of exchange-traded options. As a result, accurate and efficient pricing and hedging of caps and swaptions has enormous practical importance. Cap and swaption prices may also contain additional information on term structure dynamics not contained in bond yields or swap rates. Therefore, Dai and Singleton (2003) argue that there is an "enormous potential for new insights from using (interest rate) derivatives data in model estimations."[2]

The current literature on interest rate derivatives has here-to-fore primarily focused on two issues.[3] The first issue is the so-called "unspanned stochastic volatility" (hereafter USV) puzzle. Although caps and swaptions are derivatives written on LIBOR and swap rates, Collin-Dufresne and Goldstein (2002) and Heidari and Wu (2003) show that there appear to be risk factors that drive cap and swaption prices not spanned by the factors explaining LIBOR or swap rates. While Fan, Gupta, and Ritchken (2003) argue that swaptions might be spanned by bonds, Li and Zhao (2004) show that multifactor DTSMs have serious difficulties in hedging caps and cap straddles. The second issue is the relative pricing between caps and swaptions. A number of recent papers, including Hull and White (2000), Longstaff, Santa-Clara and Schwartz (2001) (hereafter LSS), and Jagannathan, Kaplin and Sun (2003) show that there is a significant and systematic mispricing between caps and swaptions using various multi-factor term structure models. As pointed out by Dai and Singleton (2003), these two issues are closely related and the "ultimate resolution of this 'swaptions/caps puzzle' may require time-varying correlations and possibly factors affecting the volatility of yields that do not affect bond prices."

The evidence of USV shows that, contrary to a fundamental assumption of most existing DTSMs, interest rate derivatives are not redundant securities and therefore they contain unique information about term structure dynamics that is not available in bond yields and swap rates. USV also suggests that existing DTSMs need to be substantially extended to explicitly incorporate USV for pricing interest rate derivatives. As shown by Collin-Dufresne and Goldstein (2002), however, it is rather difficult to introduce USV in traditional DTSMs: highly restrictive assumptions need to be imposed on model parameters to guarantee that certain factors that affect

[1]Dai and Singleton (2003) and Piazzesi (2003) provide excellent surveys of the literature.

[2]Jagannathan, Kaplin and Sun (2003) also show that it is important to use both the underlying Libor and swap rates, and prices of caps and swaptions for estimating DTSMs.

[3]For a review of the current empirical literature on interest rate derivatives, see Section 5 of Dai and Singleton (2003).

derivative prices do not affect bond prices. In contrast, it is much easier to introduce USV in the Heath, Jarrow, and Morton (1992) (hereafter HJM) class of models.[4] Any HJM model in which the forward rate curve has a stochastic volatility exhibits USV. Therefore, in addition to the commonly known advantages of HJM models (such as perfectly fitting the initial yield curve), they have the additional advantage of easily accommodating USV.[5]

Recently, several HJM models with USV have been developed and applied to price caps and swaptions. Collin-Dufresne and Goldstein (2003) develop a random field model with stochastic volatility and correlation in forward rates. Applying the transform analysis of Duffie, Pan, and Singleton (2000), they obtain closed-form formulae for a wide variety of interest rate derivatives. However, they do not calibrate their models to market prices of caps and swaptions. Han (2002) extends the model of LSS (2001) by introducing stochastic volatility and correlation in forward rates. Han (2002) shows that stochastic volatility and correlation are important for reconciling the mispricing between caps and swaptions.

Our paper makes both theoretical and empirical contributions to the fast growing literature on interest rate derivatives. Theoretically, we develop a multifactor HJM model with stochastic volatility and jumps in LIBOR forward rates. We allow LIBOR rates to follow the affine jump-diffusions (hereafter AJDs) of Duffie, Pan and Singleton (2000) and obtain closed-form solutions for cap prices. Given that a small number of factors can explain most of the variations of bond yields, we consider low dimensional model specifications based on the first few (up to three) principal components of historical forward rates. While similar to Han (2002) in this respect, our models have several advantages. The first advantage is that while Han's formulae, based on the approximation technique of Hull and White (1987), work well only for ATM options, our formulae, based on the affine technique, work well for all options. The second advantage is that we also explicitly incorporate jumps in LIBOR rates, making it possible to differentiate the importance of stochastic volatility versus jumps for pricing interest rate derivatives.

Our empirical investigation also substantially extends the existing literature by studying the relative pricing of caps with different strikes. Using a new dataset consisting of three years of cap prices, we provide one of the first comprehensive documentations of volatility smiles in the caps market. To our knowledge, we also conduct the first empirical analysis of term structure models with USV and jumps in capturing the smile. Caps and swaptions are traded over-the-counter and the common data sources, such as DataStream, only supply ATM option prices. As a result, the majority of the existing literature uses only at-the-money (ATM) caps and swaptions and there

[4]We refer to models that take the yield curve as given, such as the LIBOR models of Brace, Gatarek and Musiela (1997), and Miltersen, Sandmann and Sondermann (1997), the random field models Goldstein (2000), and the string models of Santa-Clara and Sornette (2001), broadly as HJM models.

[5]Of course, the trade-off here is that in HJM models the yield curve becomes an input to, not a prediction of, the model.

are almost no studies documenting the relative pricing of caps with different strike prices. In contrast, the attempt to capture the volatility smile in equity option markets is voluminous and it has been the driving force behind the development of the equity option pricing literature for the past quarter of a century (see Bakshi, Cao, and Chen and references therein).[6] Analogously, studying caps and swaptions with different strike prices could provide new insights about existing term structure models that are not available from using only ATM options.

Our analysis shows that a low dimensional LIBOR rate model with three principal components, stochastic volatility for each component, and strong negative jumps are necessary to capture the volatility smile in the cap market reasonably well. The three yield factors capture the variations of the levels of LIBOR rates, while the stochastic volatility factors are essential to capture the time varying volatilities of LIBOR rates. Even though a three-factor stochastic volatility model can price ATM caps reasonably well, it fails to capture the volatility smile in the cap market. Significant negative jumps in LIBOR rates are needed to do this. These results highlight the statement that additional information is contained in the volatility smile - the importance of negative jumps is revealed only through the pricing of caps across moneyness.

The rest of this paper is organized as follows. In Section I, we introduce the data and document the volatility smile in cap markets. In Section II, we introduce our new market models with stochastic volatility and jumps, and the statistical methods for parameter estimation and model comparison. Section III reports the empirical findings and Section IV concludes.

I. A Volatility Smile in the Interest Rate Cap Markets

In this section, using three years of cap price data, we provide one of the first comprehensive documentations of volatility smiles in the cap market. The data are obtained from SwapPX and includes daily information on LIBOR forward rates (up to ten years), and prices of caps with different strikes and maturities from August 1, 2000 to September 23, 2003.[7] The data were collected every day when the market was open between 3:30 and 4:00 pm. To reduce noises in the data and computational burdens, we use weekly data (every Tuesday) in our empirical analysis.[8] After excluding missing data, in total we have 164 weeks in our sample.

[6] For reviews of the equity option literature, see Duffie (2002) and Campbell, Lo and MacKinlay (1997).

[7] Jointly developed by GovPX and Garban-ICAP, SwapPX is the first widely distributed service delivering 24-hour real-time rates, data and analytics for the world-wide interest rate swaps market. GovPX was established in early 1990s by the major U.S. fixed-income dealers as a response to regulators' demands to increase the transparency of the fixed-income markets. It aggregates quotes from most of the largest fixed-income dealers in the world. Garban-ICAP is the world's leading swap broker specializing in trades between dealers and between dealers and large customers. According to Harris (2003), "Its securities, derivatives, and money brokerage businesses have daily transaction volumes in excess of 200 billion dollars".

[8] If Tuesday is not available, we first use Wednesday followed by Monday.

Interest rate caps are portfolios of call options on LIBOR rates. Specifically, a cap gives its holder a series of European call options, called caplets, on LIBOR forward rates. Each caplet has the same strike price as the others, but with different expiration dates. Suppose $L(t,T)$ is the 3-month LIBOR forward rate at $t \leq T$, for the interval from T to $T+\frac{1}{4}$. A caplet for the period $\left[T, T+\frac{1}{4}\right]$ struck at K pays $\frac{1}{4}\left(L(T,T)-K\right)^{+}$ at $T+\frac{1}{4}$.[9] Note that while the cash flow of this caplet is received at time $T+\frac{1}{4}$, the LIBOR rate is determined at time T. Hence, there is no uncertainty about the caplet's cash flow after the LIBOR rate is set at time T. In summary, a cap is just a portfolio of caplets whose maturities are three months apart. For example, a five-year cap on three-month LIBOR struck at six percent represents a portfolio of 19 separately exercisable caplets with quarterly maturities ranging from 6 months to 5 years, where each caplet has a strike price of 6%.

As the caps in our data are written on three-month LIBOR rates, our model and analysis focus on modeling the LIBOR forward rate curve. The dataset provides three-month LIBOR spot and forward rates at 9 different maturities (3 and 6 month, 1, 2, 3, 4, 5, 7, and 10 year). As shown in Figure 1, the forward rate curve is relatively flat at the beginning of the sample period and it declines over time, with the short end declining more than the long end. As a result, the forward rate curve becomes upward sloping in the later part of the sample.

The existing literature on interest rate derivatives has mainly focused on ATM contracts. One advantage of our data is that we observe prices of caps over a wide range of strikes and maturities.[10] For example, every day for each maturity, there are ten different strike prices, which are 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 8.0, 9.0, and 10.0 percent between August 1, 2000 and October 17, 2001, and 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, and 7.0 percent between November 2, 2001 and September 23, 2003.[11] Throughout the whole sample period, caps have fifteen different maturities, which are 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 6.0, 7.0, 8.0, 9.0, and 10.0 years. This cross-sectional information on cap prices allows us to study the performance of existing term structure models in pricing and hedging caps for different maturity and moneyness.

Ideally, we would like to study caplet prices, because they provide clear predictions of model performance across maturity. Unfortunately, we only observe cap prices. To simplify the empirical analysis, we consider the difference between the prices of caps with the same strike and adjacent

[9] It can be shown that a caplet behaves like a put option on a zero-coupon bond.

[10] To our knowledge, the only other studies that consider caps with different strikes are Gupta and Subrahmanyam (2003) and Deuskar, Gupta and Subrahmanyam (2003). The data used in the former is obtained from Tullett and Tokoyo Liberty, and it covers a shorter time period (March 1 to December 31, 1998), it has a narrower spectrum of strikes and maturities (four choices for each), and it has a maximum maturity that is only five years. The data used in the latter paper covers prices of Euro caps and floors from January 1999 to May 2001. To our knowledge, our dataset is the most comprehensive available for caps written on dollar Libor rates.

[11] The strike prices are lowered to 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0 and 5.5 percent between October 18 and November 1, 2001.

maturities, which we refer to as *difference caps*. Thus, our analysis deals with only the sum of the few caplets between two neighboring maturities with the same strike. For example, for the rest of the paper, 1.5 year *difference caps* with a specific strike represent the sum of the 1.25 and 1.5 year caplet with the same strike.

Due to daily changes in LIBOR rates, *difference caps* have a different set of moneyness (defined as the ratio between the strike price and the average LIBOR forward rates underlying the few caplets that form the difference cap) on each day. Therefore, throughout our analysis, we focus on the prices of *difference caps* at given fixed moneyness. That is, each day we interpolate difference cap prices with respect to the strike price to obtain prices at fixed moneyness. Specifically, we use local cubic polynomials to preserve the shape of the original curves and to attain smoothing over the grid points. We refrain from extrapolation and interpolation over grid points without nearby observations, and we eliminate all observations that violate various arbitrage restrictions.[12]

Figure 2.a plots the average Black implied volatilities of *difference caps* across moneyness and maturity, while Figure 2.b plots the average implied volatilities of ATM *difference caps*, over the whole sample period. Consistent with the existing literature, the implied volatilities of *difference caps* with a moneyness between 0.8 to 1.2 have a humped shape with a peak at around two year maturity. However, the implied volatilities of all other *difference caps* decline with maturity. There is also a pronounced volatility skew for *difference caps* with all maturities, with the skew being stronger for short-term *difference caps*. The pattern is similar to that of equity options: ITM *difference caps* have higher implied volatilities than OTM *difference caps*. The implied volatilities of the very short-term *difference caps* are more like a symmetric smile than a skew. Figure 3.a, b, and c plot the time series of Black implied volatilities for 2, 5, and 8 year *difference caps* across moneyness, respectively, while Figure 3.d plots the time series of ATM implied volatilities of the three contracts. It is clear that the implied volatilities are time varying and have increased dramatically (especially for 2 year *difference caps*) over our sample period. As a result of changing interest rates and strike prices, there are more ITM caps in the later part of our sample.

II. Market Models with Stochastic Volatility and Jumps: Theory and Estimation

In this section, we develop a multifactor HJM model with stochastic volatility and jumps in LIBOR forward rates to capture volatility smiles in the cap market. We estimate model parameters using the implied-state generalized method of moments (IS-GMM) of Pan (2002) and compare model performance using a statistic developed by Diebold and Mariano (1995) in the time series forecast literature.

A. Market Models with Stochastic Volatility and Jumps

[12] We eliminate observations with zero prices, and violate monotonicity and convexity with respect to strikes.

The volatility smile observed in the cap market suggests that the lognormal assumption of the standard LIBOR market models of Brace, Gatarek and Musiela (1997), and Miltersen, Sandmann and Sondermann (1997) is violated. Given the overwhelming evidence of stochastic volatility and jumps in interest rates,[13] we develop a multifactor HJM model of LIBOR rates with stochastic volatility and jumps to capture the smile. Instead of modeling the unobservable instantaneous spot rate or forward rate, we focus on the LIBOR forward rates which are observable and widely used in the market.

Throughout our analysis, we restrict the cap maturity T to a finite set of dates $0 = T_0 < T_1 < ... < T_K < T_{K+1}$, and assume that the intervals $T_{k+1} - T_k$ are equally spaced by δ, a quarter of a year. Let $L_k(t) = L(t, T_k)$ be the LIBOR forward rate for the actual period $[T_k, T_{k+1}]$, and similarly let $D_k(t) = D(t, T_k)$ be the price of a zero-coupon bond maturing on T_k. Thus, we have

$$L(t, T_k) = \frac{1}{\delta}\left(\frac{D(t, T_k)}{D(t, T_{k+1})} - 1\right), \qquad \text{for } k = 1, 2, ...K. \tag{1}$$

For LIBOR-based instruments, such as caps, floors and swaptions, it is convenient to consider pricing under the forward measure. Thus, we will focus on the dynamics of the LIBOR forward rates $L_k(t)$ under the forward measure $\mathbb{Q}^{k+1}$, which is essential for pricing caplets maturing at T_{k+1}. Under this measure, the discounted price of any security using $D_{k+1}(t)$ as the numeraire is a martingale. Therefore, the time t price of a caplet maturing at T_{k+1} with a strike price of X is

$$Caplet(t, T_{k+1}, X) = \delta D_{k+1}(t) E_t^{\mathbb{Q}^{k+1}}\left[(L_k(T_k) - X)^+\right], \tag{2}$$

where $E_t^{\mathbb{Q}^{k+1}}$ is taken with respect to $\mathbb{Q}^{k+1}$ given the information set at t. The key to valuation is modeling the evolution of $L_k(t)$ under $\mathbb{Q}^{k+1}$ realistically and yet parsimoniously to yield closed-form pricing formula. To achieve this goal, we rely on the flexible AJDs of Duffie, Pan, and Singleton (2000) to model the evolution of LIBOR rates.

We assume that under the physical measure $\mathbb{P}$, the dynamics of LIBOR rates are given by the following system of SDEs, for $t \in [0, T_k)$ and $k = 1, ..., K$,

$$\frac{dL_k(t)}{L_k(t)} = \alpha_k(t)\,dt + \sigma_k(t)\,dZ_k(t) + dJ_k(t), \tag{3}$$

where $\alpha_k(t)$ is an unspecified drift term, $Z_k(t)$ is the k-th element of a Kdimensional correlated Brownian motion with a covariance matrix $\Psi(t)$, and $J_k(t)$ is the k-th element of a Kdimensional independent pure jump process assumed independent of $Z_k(t)$ for all k. To introduce stochastic volatility and correlation, we could allow the volatility of each LIBOR rate $\sigma_k(t)$ and each individual element of $\Psi(t)$ to follow a stochastic process. But, such a model is unnecessarily complicated

[13] Andersen and Lund (1997) and Brenner, Harjes and Kroner (1996) show that stochastic volatility or GARCH significantly improve the performance of pure diffusion models for spot interest rates. Das (2003), Johannes (2004) and Piazzesi (2004) show that jumps are important for capturing interest rate dynamics.

and difficult to implement. Instead, we consider a low dimensional model based on the first few principal components of historical LIBOR forward rates. We assume that the entire LIBOR forward curve is driven by a small number of factors $N < K$ ($N \leq 3$ in our empirical analysis). By focusing on the first N principal components of historical LIBOR rates, we can reduce the dimension of the model from K to N.

Following LSS (2001) and Han (2002), we assume that the instantaneous covariance matrix of changes in LIBOR rates share the same eigenvectors as the historical covariance matrix. Suppose that the historical covariance matrix can be approximated as $H = U\Lambda_0 U'$, where Λ_0 is a diagonal matrix whose diagonal elements are the first N largest eigenvalues in descending order, and the N columns of U are the corresponding eigenvectors.[14] Our assumption means that the instantaneous covariance matrix of changes in LIBOR rates with fixed time-to-maturity, Ω_t, share the same eigenvectors as H. That is

$$\Omega_t = U\Lambda_t U', \tag{4}$$

where Λ_t is a diagonal matrix whose i-th diagonal element, denoted by $V_i(t)$, can be interpreted as the instantaneous variance of the i-th common factor driving the yield curve evolution at t. We assume that $V(t)$ follows the square-root process that has been widely used in the literature for modeling stochastic volatility (see, e.g., Heston 1993):

$$dV_i(t) = \kappa_i(\bar{v}_i - V_i(t))\,dt + \xi_i\sqrt{V_i(t)}d\tilde{W}_i(t) \tag{5}$$

where $\tilde{W}_i(t)$ is the i-th element of an N-dimensional independent Brownian motion assumed independent of $Z_k(t)$ and $J_k(t)$ for all k.

While (4) and (5) specify the instantaneous covariance matrix of LIBOR rates with fixed time-to-maturity, in applications we need the instantaneous covariance matrix of LIBOR rates with fixed maturities Σ_t. At $t = 0$, Σ_t coincides with Ω_t; for $t > 0$, we obtain Σ_t from Ω_t through interpolation. Specifically, we assume that $U_{s,j}$ is piecewise constant,[15] i.e., for time to maturity $s \in (T_k, T_{k+1})$,

$$U_s^2 = \frac{1}{2}\left(U_k^2 + U_{k+1}^2\right). \tag{6}$$

We further assume that $U_{s,j}$ is constant for all caplets belonging to the same *difference cap*. For the family of the LIBOR rates with maturities $T = T_1, T_2, ...T_K$, we denote U_{T-t} the time-t matrix that consists of rows of U_{T_k-t}, and therefore we have the time-t covariance matrix of the LIBOR

[14] We acknowledge that with jumps in LIBOR rates, both the historical and instantaneous covariance matrix of LIBOR rates contain a component that is due to jumps. Our approach implicitly assumes that the first three principal components from the historical covariance matrix captures the variations in LIBOR rates due to continuous shocks and that the impact of jumps is only contained in the residuals.

[15] Our interpolation scheme is slightly different from that of Han (2002) for the convenience of deriving closed-form solution for cap prices.

rates with fixed maturities,

$$\Sigma_t = U_{T-t}\Lambda_t U'_{T-t}. \tag{7}$$

To stay within the family of AJDs, we assume that the random jump times arrive with a constant intensity λ_J, and conditional on the arrival of a jump, the jump size follows a normal distribution $N\left(\mu_J, \sigma_J^2\right)$. Intuitively, the conditional probability at time t of another jump within the next small time interval Δt is $\lambda_J \Delta t$ and, conditional on a jump event, the mean relative jump size is $\mu = \exp\left(\mu_J + \frac{1}{2}\sigma_J^2\right) - 1$.[16] We also assume that the shocks driving LIBOR rates, volatility, and jumps (both jump time and size) are mutually independent from each other.

Given the above assumptions, we have the following dynamics of LIBOR rates under the physical measure $\mathbb{P}$,

$$\frac{dL_k(t)}{L_k(t)} = \alpha_k(t)\,dt + \sum_{j=1}^{N} U_{T_k-t,j}\sqrt{V_j(t)}dW_j(t) + dJ_k(t),\, k = 1, 2, ..., K. \tag{8}$$

To price caps, we need the dynamics of LIBOR rates under the appropriate forward measure. The existence of stochastic volatility and jumps results in an incomplete market and hence the non-uniqueness of forward martingale measures. Our approach for eliminating this nonuniqueness is to specify the market prices of both the volatility and jump risks to change from the physical measure $\mathbb{P}$ to the forward measure $\mathbb{Q}^{k+1}$.[17] Following the existing literature, we model the volatility risk premium as $\eta_j^{k+1}\sqrt{V_j(t)}$, for $j = 1, ..., N$. For the jump risk premium, we assume that under the forward measure $\mathbb{Q}^{k+1}$, the jump process has the same distribution as that under P, except that the jump size follows a normal distribution with mean μ_J^{k+1} and variance σ_J^2. Thus, the mean relative jump size under $\mathbb{Q}^{k+1}$ is $\mu^{k+1} = \exp\left(\mu_J^{k+1} + \frac{1}{2}\sigma_J^2\right) - 1$. Our specification of the market prices of jump risks allows the mean relative jump size under $\mathbb{Q}^{k+1}$ to be different from that under $\mathbb{P}$, accommodating a premium for jump size uncertainty. This approach, which is also adopted by Pan (2002), artificially absorbs the risk premium associated with the timing of the jump by the jump size risk premium. In our empirical analysis, we make the simplifying assumption that the volatility and jump risk premiums are linear functions of time-to-maturity, i.e., $\eta_j^{k+1} = c_{jv}(T_k - 1)$ and $\mu_J^{k+1} = \mu_J + c_J(T_k - 1)$.[18]

[16] For simplicity, we assume that different forward rates follow the same jump process with constant jump intensity. It is not difficult to allow different jump processes for individual LIBOR rates and the jump intensity to depend on the state of the economy within the AJD framework.

[17] The market prices of interest rate risks are defined in such a way that the LIBOR rate is a martingale under the forward measure.

[18] In order to estimate the volatility and jump risk premiums, we need a joint analysis of the dynamics of LIBOR rates under both the physical and forward measure, as in Chernov and Ghysels (2000), Pan (2002), and Eraker (2004). In our empirical analysis, we only focus on the dynamics under the forward measure. Therefore, we can only identify the differences in the risk premiums between forward measures with different maturities. Our specifications of both risk premiums implicitly use the one year LIBOR rate as a reference point.

Given the above market prices of risks, we can write down the dynamics of $\log(L_k(t))$ under forward measure $\mathbb{Q}^{k+1}$,

$$d\log(L_k(t)) = -\left(\lambda_J \mu^{k+1} + \frac{1}{2}\sum_{j=1}^{N} U_{T_k-t,j}^2 V_j(t)\right)dt + \sum_{j=1}^{N} U_{T_k-t,j}\sqrt{V_j(t)}dW_j^{\mathbb{Q}^{k+1}}(t) + dJ_k^{\mathbb{Q}^{k+1}}(t).$$

For pricing purpose, the above process can be further simplified to the following one which has the same distribution,

$$d\log(L_k(t)) = -\left(\lambda_J \mu^{k+1} + \frac{1}{2}\sum_{j=1}^{N} U_{T_k-t,j}^2 V_j(t)\right)dt + \sqrt{\sum_{j=1}^{N} U_{T_k-t,j}^2 V_j(t)}dZ_k^{\mathbb{Q}^{k+1}}(t) + dJ_k^{\mathbb{Q}^{k+1}}(t), \tag{9}$$

where $Z_k^{\mathbb{Q}^{k+1}}(t)$ is a standard Brownian motion under $\mathbb{Q}^{k+1}$. Now the dynamics of $V_i(t)$ under $\mathbb{Q}^{k+1}$ becomes

$$dV_i(t) = \kappa_i^{k+1}\left(\bar{v}_i^{k+1} - V_i(t)\right)dt + \xi_i\sqrt{V_i(t)}d\tilde{W}_i^{\mathbb{Q}^{k+1}}(t) \tag{10}$$

where $\tilde{W}^{\mathbb{Q}^{k+1}}$is independent of $Z^{\mathbb{Q}^{k+1}}$, $\kappa_j^{k+1} = \kappa_j - \xi_j\eta_j^{k+1}$, and $\bar{v}_j^{k+1} = \frac{\kappa_j \bar{v}_j}{\kappa_j - \xi_j \eta_j^{k+1}}$, $j = 1, .., N$. The dynamics of $L_k(t)$ under the forward measure $\mathbb{Q}^{k+1}$ are completely captured by (9) and (10).

Given that LIBOR rates follow AJDs under both the physical and forward measure, we can directly apply the transform analysis of Duffie, Pan and Singleton (2000) to derive closed-form formula for cap prices. Denote the state variables at t as $Y_t = (\log(L_k(t)), V_t)'$ and the time-t expectation of $e^{u \cdot Y_{T_k}}$ under the forward measure $\mathbb{Q}^{k+1}$ as $\psi(u, Y_t, t, T_k) \triangleq E_t^{\mathbb{Q}^{k+1}}\left[e^{u \cdot Y_{T_k}}\right]$. Let $u = (u_0, 0_{1\times N})'$, then the time-$t$ expectation of LIBOR rate at T_k equals,

$$\begin{aligned} E_t^{\mathbb{Q}^{k+1}}\{\exp[u_0 \log(L_k(T_k))]\} &= \psi(u_0, Y_t, t, T_k) \\ &= \exp\left[a(s) + u_0 \log(L_k(t)) + B(s)' V_t\right], \end{aligned}$$

where $s = T_k - t$ and closed-form solutions of $a(s)$ and $B(s)$ (an N-by-1 vector) are obtained by solving a system of Ricatti equations in the appendix.

Following Duffie, Pan and Singleton (2000), we define

$$G_{a,b}(y; Y_t, T_k, \mathbb{Q}^{k+1}) = E_t^{\mathbb{Q}^{k+1}}\left[e^{a \cdot \log(L_k(T_k))} 1_{\{b \cdot \log(L_k(T_k)) \le y\}}\right],$$

and its Fourier transform,

$$\begin{aligned} \mathcal{G}_{a,b}(v; Y_t, T_k, \mathbb{Q}^{k+1}) &= \int_R e^{ivy} dG_{a,b}(y) \\ &= E_t^{\mathbb{Q}^{k+1}}\left[e^{(a+ivb)\cdot \log(L_k(T_k))}\right] \\ &= \psi(a + ivb, Y_t, t, T_k). \end{aligned}$$

Levy's inversion formula gives

$$G_{a,b}(y;Y_t,T_k,\mathbb{Q}^{k+1}) = \frac{\psi\left(a+ivb,Y_t,t,T_k\right)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\operatorname{Im}\left[\psi\left(a+ivb,Y_t,t,T_k\right)e^{-ivy}\right]}{v}dv.$$

The time-0 price of a caplet that matures at T_{k+1} with a strike price of X equals

$$Caplet(0,T_{k+1},X) = \delta D_{k+1}\left(0\right)E_0^{\mathbb{Q}^{k+1}}\left[\left(L_k\left(T_k\right)-X\right)^+\right], \tag{11}$$

where the expectation is given by the inversion formula,

$$\begin{aligned} E_0^{\mathbb{Q}^{k+1}}\left[L_k(T_k)-X\right]^+ &= G_{1,-1}(-\ln X;Y_0,T_k,\mathbb{Q}^{k+1}) \\ &\quad -XG_{0,-1}(-\ln X;Y_0,T_k,\mathbb{Q}^{k+1}). \end{aligned}$$

The new models developed in this section nest some of the most important models in the literature, such as LSS (2001) (with constant volatility and no jumps) and Han (2002) (with stochastic volatility and no jumps). The closed-form formula for cap prices makes an empirical implementation of our model very convenient and provides some advantages over existing methods. For example, Han (2002) develops approximations of ATM cap and swaption prices using the techniques of Hull and White (1987). However, such an approach might not work well for away-from-the-money options. In contrast, our method would work well for all options, which is important for explaining the volatility smile.

In addition to introducing stochastic volatility and jumps, our multifactor HJM models also has advantages over the standard LIBOR market models of Brace, Gatarek and Musiela (1997), Miltersen, Sandmann and Sondermann (1997), and their extensions often applied to caps in practice.[19] While our models provide a unified multifactor framework to characterize the evolution of the whole yield curve, the LIBOR market models typically make separate specifications of the dynamics of LIBOR rates with different maturities. As suggested by LSS (2001), the standard LIBOR models are "more appropriately viewed as a collection of different univariate models, where the relationship between the underlying factors is left unspecified." In contrast, the dynamics of LIBOR rates with different maturities under their related forward measures are internally consistent with each other given their dynamics under the physical measure and the market prices of risks. Once our models are estimated using one set of prices, they can be used to price and hedge other fixed-income securities.

B. Parameter Estimation and Model Comparison

In this section, we discuss the estimation of our new market model using prices form a wide cross section of *difference caps* with different strikes and maturities. One challenge we face is that in addition to the model parameters, we also need to deal with the latent stochastic volatility

[19] Andersen and Brotherton-Radcliff (2001) and Glasserman and Kou (2003) develop LIBOR models with stochastic volatility and jumps, respectively.

variables. We adopt the IS-GMM approach of Pan (2002), which is more suitable to our applications than other existing methods.[20] The IS-GMM approach is an important extension of the standard GMM to dynamic models with unobservable latent state variables. Using IS-GMM, Pan (2002) conducts a joint analysis of stochastic volatility and jump models using stock and option prices. She first backs out the volatility variables from short-term ATM options, then estimates model parameters based on the moment conditions implied by AJDs of spot price and stochastic volatility. Specializing to our models, this approach allows us to back out the latent volatility variables from observed *difference cap* prices given a parametric pricing formula. Model parameters can then be estimated by minimizing appropriately chosen moment conditions based on both observed and latent variables.

Our implementation of IS-GMM to the cross section of *difference caps* differs from Pan (2002) in several respects. First, instead of backing out volatility variables from short-term ATM options, we use all *difference caps* in estimating the volatility variables at each point in time. Specifically, every week, for a given set of model parameters, we minimize the RMSE of all *difference caps* to obtain the estimates of the volatility variables. This approach allows us to fully utilize the information in the prices of all *difference caps* to estimate the volatility variables. Second, instead of using the time series properties of state variables as moment conditions, we use the absolute percentage pricing errors (the absolute value of the difference between observed and theoretical prices of *difference caps* divided by observed prices of *difference caps*) of all *difference caps* as the moment conditions. This is motivated by our objective to explain the cap market smile. Unlike exchange-traded options studied in Pan (2002), the *difference caps* in our paper have fixed time-to-maturity and moneyness. This removes the time dependency in the contract variables and makes IS-GMM especially suitable for our situation.

Every week we observe prices of *difference caps* with ten moneyness and thirteen maturity. Theoretically, in total we have 130 moment conditions. However, due to changing interest rates, we do not have enough observations in all moneyness/maturity categories throughout the sample. Thus, we focus on the 53 moneyness/maturity categories that have less than ten percent of missing values over the whole sample period used in our estimation. The moneyness and maturity of all *difference caps* belong to the following sets $\{0.7, 0.8, 0.9, 1.0, 1.1\}$ and $\{1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0\}$ (unit in years), respectively. The *difference caps* with time-to-maturity less than or equal to five years represent portfolios of two caplets, while those longer than five years represent portfolios of four caplets. So in total, we use 53 moment conditions in the empirical estimation.

[20] In studies of stochastic volatility models using stock and option prices, Chernov and Ghysels (2000) use the efficient method of moments (EMM) of Gallant and Tauchen (1998), and Eraker (2003) uses the Bayesian Markov Chain Monte Carlo (MCMC) method. Both methods are based on simulations and can be challenging to apply to a wide cross section of options.

Suppose we have time series observations over $t = 1, ..., \mathcal{T}$, of the prices of 53 *difference caps* with moneyness m_i and time-to-maturity $\tau_i, i = 1, ..., M = 53$. Let θ represent the model parameters which remain constant over the whole sample period. Let $C(t, m_i, \tau_i)$ be the observed price of a *difference cap* with moneyness m_i and time-to-maturity τ_i and $\hat{C}(t, \tau_i, m_i, V_t(\theta), \theta)$ be the corresponding theoretical price under a given model, where $V_t(\theta)$ is the model implied instantaneous volatility at t given model parameters θ. For each i and t, denote the absolute relative pricing error as

$$u_{i,t}(\theta) = \left| \frac{C(t, m_i, \tau_i) - \hat{C}(t, m_i, \tau_i, V_t(\theta), \theta)}{C(t, m_i, \tau_i)} \right|. \tag{12}$$

To estimate the unobserved volatility variables V_t, for a specific parameter θ, we choose V_t to minimize the RMSE of all *difference caps* at t, $\varepsilon_t(\theta) = \sqrt{\frac{1}{M}\sum_{i=1}^{M}[u_{i,t}(\theta)]^2}$. That is,

$$V_t(\theta) = \arg\min_{\{V_t\}} \varepsilon_t(\theta).$$

Similar to Pan (2002), we make the assumption that under correct model specification and true model parameters, θ^0, $V_t\left(\theta^0\right) = V_t^0$, the true instantaneous stochastic volatility.

For given model parameters, θ, denote $u_t(\theta)$ as the 53-by-1 vector of absolute percentage pricing errors on date t of the *difference caps* in the 53 moneyness/maturity groups. Then the sample mean of u_t in a sample of size $\mathcal{T}$ equals

$$g_{\mathcal{T}}(\theta) = \frac{1}{\mathcal{T}}\sum_{t=1}^{\mathcal{T}} u_t(\theta).$$

We obtain parameter estimates by minimizing the moment conditions as in the traditional GMM framework. That is,

$$\hat{\theta} = \arg\min_{\{\theta\}} g_{\mathcal{T}}(\theta)' W_{\mathcal{T}} g_{\mathcal{T}}(\theta),$$

where $W_{\mathcal{T}}$ is a weighting matrix. As shown by Hansen (1982), the optimal weighting matrix is the inverse of the covariance matrix of pricing errors. We follow the standard GMM two-step estimation approach. In the first approach, we use the identity weighting matrix to obtain the first stage GMM estimators via the following minimization

$$\hat{\theta}_1 = \arg\min_{\{\theta\}} g_{\mathcal{T}}(\theta)' W_{\mathcal{T}} g_{\mathcal{T}}(\theta),$$

where $W_{\mathcal{T}} = I$. Using $\hat{\theta}_1$, we form an estimate of the covariance matrix of pricing errors $\hat{S}$ of

$$S = \sum_{j=-\infty}^{\infty} E\left[u_t\left(\hat{\theta}_1\right) u_{t-j}\left(\hat{\theta}_1\right)'\right].$$

We obtain a second-stage estimate $\hat{\theta}_2$ using the inverse of matrix $\hat{S}$ as the weighting matrix in the quadratic form

$$\hat{\theta}_2 = \arg\min_{\{\theta\}} g_{\mathcal{T}}(\theta)' \hat{S}^{-1} g_{\mathcal{T}}(\theta).$$

As shown by Hansen (1982), $\hat{\theta}_2$ is a consistent, asymptotically normal, and asymptotically efficient estimate of the parameter vector θ.

One important feature of our moment condition is that $g_{\mathcal{T}}(\theta)$ is always positive.[21] This violates the standard GMM assumption of Hansen (1982) that sample moments under the null hypothesis have a normal distribution with zero mean. Instead, our sample moments always have positive mean. While this issue does not affect parameter estimates, it renders the standard GMM χ^2 specification tests not directly applicable to our situation.

To compare model performance, especially to test whether one model has statistically smaller pricing errors than another, we adopt an approach developed by Diebold and Mariano (1995) in the time-series forecast literature. Consider two models whose associated weekly RMSEs $\{\varepsilon_1(t)\}_{t=1}^{\mathcal{T}}$ and $\{\varepsilon_2(t)\}_{t=1}^{\mathcal{T}}$, respectively . The null hypothesis that the two models have the same pricing errors is $E[\varepsilon_1(t)] = E[\varepsilon_2(t)]$, or $E[d(t)] = 0$, where $d(t) = \varepsilon_1(t) - \varepsilon_2(t)$. Diebold and Mariano (1995) show that if $\{d(t)\}_{t=1}^{\mathcal{T}}$ is covariance stationary and short memory, then

$$\sqrt{\mathcal{T}}\left(\overline{d} - \mu_d\right) \sim N\left(0, 2\pi f_d(0)\right), \tag{13}$$

where $\overline{d} = \frac{1}{\mathcal{T}}\sum_{t=1}^{\mathcal{T}}[\varepsilon_1(t) - \varepsilon_2(t)]$, $f_d(0) = \frac{1}{2\pi}\sum_{q=-\infty}^{\infty}\gamma_d(q)$ and $\gamma_d(q) = E[(d_t - \mu_d)(d_{t-q} - \mu_d)]$. In large samples, $\overline{d}$ is approximately normally distributed with mean μ_d and variance $2\pi f_d(0)/\mathcal{T}$. Thus under the null hypothesis of equal pricing errors, the following statistic

$$S = \frac{\overline{d}}{\sqrt{2\pi\widehat{f}_d(0)/\mathcal{T}}} \tag{14}$$

is distributed asymptotically as $N(0,1)$, where $\widehat{f}_d(0)$ is a consistent estimator of $f_d(0)$.[22] To compare the overall performance of the two models, we use the above statistic to measure whether one model has significantly smaller RMSEs than another. We can also use the above statistic to measure whether one model has smaller absolute percentage pricing errors than another for *difference caps* in a specific moneyness/maturity group.

[21] While moment conditions based on percentage pricing errors satisfy the standard GMM assumption, they yield parameter estimates that give large variability in pricing errors. For example, the estimated models can significantly underprice short-term caps and overprice long-term caps while still having an average pricing error that is close to zero. Moment conditions based on absolute percentage pricing errors eliminate such problems.

[22] We estimate the variance of the test statistic using the Bartlett estimate of Newey and West (1987). As the nonparametric estimator of the variance has a slower convergance rate than that of the parameter estimates, asymptotically parameter estimation uncertainty has no impact on the test statistic.

III. Empirical Results

In this section, we provide empirical evidence on the performance of six different models in capturing the cap volatility smile. The first three models, denoted as SV1, SV2 and SV3, allow one, two, and three principal components to drive the forward rate curve, respectively, each with its own stochastic volatility. The next three models, denoted as SVJ1, SVJ2 and SVJ3, introduce jumps in LIBOR rates in each of the previous SV models. SVJ3 is the most comprehensive model and nests all the others as special cases. We first examine the separate performance of each of the SV and SVJ models, then we compare performance across the two classes of models.

The estimation of all models is based on the principal components extracted from historical LIBOR forward rates between June 1997 and July 2000.[23] Figure 4 shows that the three principal components can be interpreted as in Litterman and Scheinkman (1991). The first or the "level" factor represents a parallel shift of the forward rate curve. The second or the "slope" factor twists the forward rate curve by moving the short and long end of the curve in opposite directions. The third or the "curvature" factor increases the curvature of the curve by moving the short and long end of the curve in one direction and the middle range of the curve in the other direction. The three factors explain 77.78%, 14.35%, and 7.85% of the variations of LIBOR rates up to ten years, respectively.

A. Performance of Stochastic Volatility Models

The SV models contribute to cap pricing in four important ways. First, the three principal components capture variations in the levels of LIBOR rates caused by innovations in the "level", "slope", and "curvature" factors. Second, the stochastic volatility factors capture the fluctuations in the volatilities of LIBOR rates reflected in the Black implied volatilities of ATM caps.[24] Third, the stochastic volatility factors also introduce fatter tails in LIBOR rate distributions than implied by the log-normal model, which helps capture the volatility smile. Finally, given our model structure, innovations of stochastic volatility factors also affect the covariances between LIBOR rates with different maturities. The first three factors, however, are more important for our applications, because *difference caps* are much less sensitive to time varying correlations than swaptions.[25] Our discussion of the performance of the SV models focuses on the estimates of the model parameters and the latent volatility variables, and the time series and cross-sectional pricing errors of *difference caps*.

A comparison of the parameter estimates of the three SV models in Table 1 shows that the "level" factor has the most volatile stochastic volatility, followed, in decreasing order, by the

[23] The LIBOR forward curve is constructed from weekly LIBOR and swap rates from Datastream following the bootstrapping procedure of LSS (2001).

[24] Throughout our discussion, volatilities of LIBOR rates refer to market implied volatilities from cap prices and are different from volatilities estimated from historical data.

[25] See Han (2002) for more detailed discussions on the impact of time varying correlations for pricing swaptions.

"curvature" and "slope" factor.[26] The long-run mean ($\bar{v}_1$) and volatility of volatility (ξ_1) of the first volatility factor are much bigger than that of the other two factors. This suggests that the fluctuations in the volatilities of LIBOR rates are mainly due to the time varying volatility of the "level" factor. The estimates of the volatility risk premium of the three models are significantly negative, suggesting that the stochastic volatility factors of longer maturity LIBOR rates under the forward measure are less volatile with lower long-run mean and faster speed of mean reversion. This is consistent with the fact that the Black implied volatilities of longer maturity *difference caps* are less volatile than that of short-term *difference caps*. The contributions from additional volatility factors in capturing the volatility of LIBOR rates tend to increase the speed of mean reversion and to reduce the long-run mean and volatility of the existing volatility factors.

Our parameter estimates are consistent with the volatility variables inferred from the prices of *difference caps* in Figure 5. The volatility of the "level" factor is the highest among the three (although lower in the more sophisticated models). It starts at a low level and steadily increases and stabilizes at a high level in the later part of the sample period. The volatility of the "slope" factor is much lower and relatively stable during the whole sample period. The volatility of the "curvature" factor is generally between that of the first and second factors. The steady increase of the volatility of the "level" factor is consistent with the increase of Black implied volatilities of ATM *difference caps* throughout our sample period. In fact, the correlation between the Black implied volatilities of most *difference caps* and the implied volatility of the "level" factor are higher than 0.8. The correlation between Black implied volatilities and the other two volatility factors is much weaker. The importance of stochastic volatility is obvious: the fluctuations in Black implied volatilities show that a model with constant volatility simply would not be able to capture even the general level of cap prices.

The other aspects of model performance are the time series and cross-sectional pricing errors of *difference caps*. Figure 7 plots the time series of RMSEs of the three SV models over our sample period. The Diebold-Mariano statistics in Panel A of Table 2 show that SV2 and SV3 have significantly smaller RMSEs than SV1 and SV2, respectively, suggesting that the more sophisticated SV models improve the pricing of all caps. Except for two special periods where all models have extremely large pricing errors, the RMSEs of all models are rather uniform over the whole sample period, with the best model (SV3) having RMSEs slightly above 5%. The two special periods with high pricing errors cover the period between the second half of December of 2000 and the first half of January of 2001, and the first half of October 2001, and coincide

[26] The estimates are the first stage GMM estimates, which are very similar to the second stage GMM estimates. The objective function reported in Table 1 are the rescaled objective functions of the first GMM estimation and are essentially the RMSEs of each model. It is difficult to compare the objective functions of the second stage GMM because of the different weighting matrices used in the different models. While we do not use the second stage GMM estimates explicitly, they serve as a robustness check of the first stage estimates.

with high prepayments in mortgage-backed securities (MBS). Indeed, the MBAA refinancing index and prepayment speed (see Figure 3 of Duarte 2004) show that after a long period of low prepayments between the middle of 1999 and late 2000, prepayments dramatically increased at the end of 2000 and the beginning of 2001. There is also a dramatic increase of prepayments at the beginning of October 2001. As widely recognized in the fixed-income market,[27] excessive hedging demands for prepayment risk using interest rate derivatives may push derivative prices away from their equilibrium values, which could explain the failure of our models during these two special periods.[28]

In addition to overall model performance as measured by RMSEs, we also examine the cross-sectional pricing errors of *difference caps* with different moneyness and maturity. We first look at the absolute percentage pricing errors, which measure both the biasedness and variability of the pricing errors. Then we look at the average percentage pricing errors (the difference between market and model prices divided by the model price) to see whether SV models can on average capture the volatility smile in the cap market.

The Diebold-Mariano statistics of absolute percentage pricing errors between SV2 and SV1 in Panel B of Table 2 show that SV2 reduces the pricing errors of SV1 for most *difference caps* (bold means the difference is significant at 5% level). SV2 has the most significant reductions in pricing errors of SV1 for long and short term ATM and OTM *difference caps*, and mid-term slightly ITM ($m = 0.9$) *difference caps*. The improvements for most ITM caps are not significant. For some deep ITM ($m = 0.7$) caps, SV2 actually has larger, although not significant, pricing errors over SV1. The Diebold-Mariano statistics between SV3 and SV2 in Panel C of Table 2 show that SV3 significantly reduces the pricing errors of many short-term ATM, slightly ITM ($m = 0.9$) and OTM, and long-term ITM *difference caps*. For medium maturity range, while SV3 significantly reduces the pricing errors of ATM and OTM *difference caps*, it significantly increases the pricing errors of ITM caps.

Table 3 reports the average percentage pricing errors of all *difference caps* under the three SV models. Panel A of Table 3 shows that, on average, SV1 underprices short-term and overprices long-term ATM *difference caps*, and underprices ITM and overprices OTM *difference caps*. This suggests that SV1 cannot generate enough skewness in the implied volatilities to be consistent with the data. Panel B shows that SV2 has some improvements over SV1, mainly for some short term (less than 3.5 yr) ATM *difference caps* and most long-term (7-8 yr) slightly ITM ($m = 0.9$) *difference caps*. But SV2 has worse performance for most deep ITM ($m = 0.7$ and 0.8) and OTM *difference caps*: it worsens the underpricing of ITM and the overpricing of OTM caps. Panel C of

[27]We would like to thank Pierre Grellet Aumont from Deutsche Bank for his helpful discussions on the influence of MBS markets on OTC interest rate derivatives.

[28]While the prepayments rates were also high in later part of 2002 and for most of 2003, they might not have come as surprises to participants in the MBS markets given the two previous special periods.

Table 3 shows that relative to SV1 and SV2, SV3 has smaller average percentage pricing errors for most long-term (7-10 yr) ITM, mid-term (3.5-5yr) OTM, and short-term (2 and 2.5 yr) ATM *difference caps*, and bigger average percentage pricing errors for mid-term (3.5 to 6 year) ITM *difference caps*. There is still significant underpricing of ITM and overpricing of OTM *difference caps* under SV3.

Overall, the results show that stochastic volatility factors are essential for capturing the time varying volatilities of LIBOR rates. The Diebold-Mariano statistics in Table 2 show that in general more sophisticated SV models have smaller absolute percentage pricing errors than simpler models, although the improvements are more important for close-to-the-money *difference caps*. The average percentage pricing errors in Table 3 show that, however, even the most sophisticated SV model cannot generate enough volatility skew to be consistent with the data. While previous studies, such as Han (2002), have shown that a three-factor stochastic volatility model similar to ours performs well in pricing ATM caps and swaptions, our analysis shows that the model fails to completely capture the volatility smile in the cap markets. Our findings highlight the importance of studying the relative pricing of caps with different moneyness to reveal the inadequacies of existing term structure models, the same inadequacies cannot be obtained from studying only ATM options.

B. Performance of Stochastic Volatility and Jump Models

One important reason for the failure of SV models is that the stochastic volatility factors are independent of LIBOR rates. As a result, the SV models can only generate a symmetric volatility smile, but not the asymmetric smile or skew observed in the data. The pattern of the smile in the cap market is rather similar to that of index options: ITM calls (and OTM puts) are overpriced, and OTM calls (and ITM puts) are underpriced relative to the Black model. Similarly, the smile in the cap market could be due to a market expectation of dramatically declining LIBOR rates. In this section, we examine the contribution of jumps in LIBOR rates in capturing the volatility smile. Our discussion of the performance of the SVJ models parallels that of the SV models.

Parameter estimates in Table 4 show that the three stochastic volatility factors of the SVJ models resemble that of the SV models closely. The "level" factor still has the most volatile stochastic volatility, followed by the "curvature" and the "slope" factor. With the inclusion of jumps, the stochastic volatility factors in the SVJ models tend to be less volatile than that of the SV models (faster speed of mean reversion and lower long run mean and volatility of volatility). Negative estimates of the volatility risk premium show that the volatility of the longer maturity LIBOR rates under the forward measure have lower long-run mean and faster speed of mean-reversion. Figure 7 shows that the volatility of the "level" factor experiences a steady increase over the whole sample period, while the volatility of the other two factors are relatively stable over time.

Most importantly, we find overwhelming evidence of strong negative jumps in LIBOR rates under the forward measure. To the extend that cap prices reflect market expectations of future evolutions of LIBOR rates, the evidence suggests that the market expects a dramatic declining in LIBOR rates over our sample period. Such an expectation might be justifiable given that the economy has been in recession during a major part of our sample period. This is similar to the volatility skew in the index equity option market, which reflects investors fear of the stock market crash such as that of 1987. Compared to the estimates from index options (see, e.g., Pan 2002), we see lower estimates of jump intensity (between 2 to 6% per annual), but much higher estimates of jump size. The positive estimates of a jump risk premium suggest that the jump magnitude of longer maturity forward rates tend to be smaller. Under SVJ3, the mean relative jump size, $\exp\left(\mu_J + c_J\left(T_k - 1\right) + \sigma_J^2/2\right) - 1$, for one, five, and ten year LIBOR rates are -90%, -80%, and -56%, respectively. However, we do not find any incidents of negative moves in LIBOR rates under the physical measure with a size close to that under the forward measure. This big discrepancy between jump sizes under the physical and forward measures resembles that between the physical and risk-neutral measure for index options (see, e.g., Pan 2002). This could be a result of a huge jump risk premium.

Figure 8 plots the time series of RMSEs of the three SVJ models over our sample period. The Diebold-Mariano statistics in Panel A of Table 5 show that SVJ2 and SVJ3 have significantly smaller RMSEs than SVJ1 and SVJ2 respectively, suggesting that the more sophisticated SVJ models significantly improve the pricing of all *difference caps*. In addition to the two special periods in which the SVJ models have large pricing errors, the SVJ models have larger RMSEs than SV models during the first 20 weeks of the sample. This should not be surprising given the relatively stable forward rate curve and a less pronounced volatility smile. The RMSEs of all the SVJ models are rather uniform over the rest of the sample period.

The Diebold-Mariano statistics of the absolute percentage pricing errors in Panel B of Table 5 show that SVJ2 significantly improves the performance of SVJ1 for most *difference caps*. The most significant improvements occur for long and medium term ATM and ITM *difference caps*, and for some short-term ATM *difference caps*. The Diebold-Mariano statistics in Panel C of Table 5 show that the SVJ3 significantly reduces the pricing errors of SVJ2 for most *difference caps*, especially long-term ITM caps, and for some short-term ITM and mid-term OTM caps. But, the SVJ3 has bigger pricing errors than SVJ2 for some mid-term (3 and 3.5 yr) ITM caps.

The average percentage pricing errors in Table 6 show that the SVJ models capture the volatility smile much better than the SV models. Panel A of Table 6 shows that, although SVJ1 on average underprices short-term and overprices long-term ATM *difference caps*, the degree of mispricing is much smaller than that of SV1. While there is still an increasing degree of underpricing of *difference caps* that are deeper in the money (especially for 5 to 10 year caps), the magnitude of mispricing is again smaller than that of SV1. This suggests that with the introduction of negative jumps, SVJ1 can capture the volatility smile in the data much better than SV1. Panel B shows that in contrast to the SV models, SVJ2 significantly reduces the underpricing of deep ITM *difference caps* of SVJ1, especially with maturities between 5 and 10 years. Panel C of Table 3 shows that SVJ3 further improves SVJ2 in capturing the smile: for most *difference caps*, the average percentage pricing errors under SVJ3 are less than 1%, showing that the model can capture the smile well.

Table 7 compares the performance of the SVJ and SV models. As shown before, during the first 20 weeks of our sample, the SVJ models have much higher RMSEs than the SV models. As a result, the Diebold-Mariano statistics between the three pairs of SVJ and SV models are only significant at the 10% level. Excluding the first 20 weeks, the Diebold-Mariano statistics become overwhelmingly significant. The Diebold-Mariano statistics of individual *difference caps* in Panel B, C, and D show that the SVJ models significantly improve the performance of the SV models for most *difference caps* across moneyness and maturity. The most interesting results are in Panel D, which show that SVJ3 significantly reduces the pricing errors of most ITM *difference caps* of SV3, strongly suggesting that the negative jumps are essential for capturing the asymmetric smile in the cap market.

Our analysis shows that a low dimensional model with three principal components driving the forward rate curve, stochastic volatility of each component, and strong negative jumps captures the volatility smile in the cap markets reasonably well. The three yield factors capture the variations of the levels of LIBOR rates, while the stochastic volatility factors are essential to capture the time varying volatilities of LIBOR rates. Even though the SV models can price ATM caps reasonably well, they fail to capture the volatility smile in the cap market. Instead, significant negative jumps in LIBOR rates are needed to capture the smile. These results highlight the importance of studying the pricing of caps across moneyness: the importance of negative jumps is revealed only through the pricing of alway-from-the-money caps. Excluding the first 20 weeks and the two special periods, SVJ3 has a reasonably good pricing performance with an average RMSEs of 4.5%. Given that the bid-ask spread is about 2 to 5% in our sample for ATM caps, and because ITM and OTM caps tend to have even higher percentage spreads,[29] this cam be interpreted as a good performance.

Despite its good performance, there are several aspects of SVJ3 that deserve further analysis. First, the fact that the SVJ models have large pricing errors for the first 20 weeks shows that there might be a structural change in the data generating process. The expectation of negative jumps in LIBOR rates seem to be built into cap prices after this initial period. While this is similar to what happened after the 1987 crash to index option prices, there was not any single dramatic

[29] See, for example, Deuskar, Gupta, and M. Subrahmanyam (2003).

event that we can identify that caused such a change. Only additional research using independent data can determine whether this change is permanent or specific to our sample period. Second, while our model works reasonably well for most of the sample period, there are special segments coinciding with high prepayment activities in the MBS markets where our models have large pricing errors. A more careful analysis of the influence of MBS markets on cap prices would be interesting. Finally, there is evidence of model misspecification. Even though we assume that the stochastic volatility factors are independent of LIBOR rates and from each other, Table 8 shows strong negative correlations between the first stochastic volatility factor and the LIBOR rates, and a negative (positive) correlation between the first and second (third) stochastic volatility factor. Extending our model to incorporate these correlations is another future research project.

IV. Conclusion

In this paper, we have made significant theoretical and empirical contributions to the fast growing literature on LIBOR and swap-based interest rate derivatives. Theoretically, we develop multifactor HJM models that explicitly take into account the new empirical features of term structure data: unspanned stochastic volatility and jumps. Our models provide closed-form formula for caps which greatly simplifies an empirical implementation of the models. Empirically, we provide one of the first comprehensive analyses of the relative pricing of caps with different moneyness. Using a comprehensive dataset of three years of cap prices with different strike and maturity, we document a volatility smile in the cap market. Although previous studies show that multifactor stochastic volatility models can price ATM caps and swaptions well, we show that they fail to capture the volatility smile in the cap market. Instead, a three-factor model with stochastic volatility and significant negative jumps is needed to capture the smile. Our results show that the volatility smile indeed contains new information that is not available in ATM caps. Our paper is only one of the first attempts to explain the volatility smile in OTC interest rate derivatives markets. Even though our model exhibits reasonably good performance, there are several aspects of the model that are not completely satisfactory. Given that volatility smile has guided the development of equity option pricing literature since Black and Scholes (1973) and Merton (1973), we hope that the volatility smile documented here will help the development of term structure models in the years to come.

Mathematical Appendix

The solution to the characteristic function of $\log(L_k(T_k))$,

$$\psi(u_0, Y_t, t, T_k) = \exp\left[a(s) + u_0 \log(L_k(t)) + B(s)' V_t\right],$$

$a(s)$ and $B(s), 0 \leq s \leq T_k$ satisfy the following system of Ricatti equations:

$$\begin{aligned}
\frac{dB_j(s)}{ds} &= -\kappa_j^{k+1} B_j(s) + \frac{1}{2} B_j^2(s) \xi_j^2 + \frac{1}{2}\left[u_0^2 - u_0\right] U_{s,j}^2, \quad 1 \leq j \leq N, \\
\frac{da(s)}{ds} &= \sum_{j=1}^{N} \kappa_j^{k+1} \theta_j^{k+1} B_j(s) + \lambda_J \left[\Gamma(u_0) - 1 - u_0 (\Gamma(1) - 1)\right],
\end{aligned}$$

where the function Γ is

$$\Gamma(x) = \exp(\mu_J^{k+1} x + \frac{1}{2}\sigma_J^2 x^2).$$

The initial conditions are $B(0) = 0_{N\times 1}$, $a(0) = 0$, and κ_j^{k+1} and θ_j^{k+1} are the parameters of $V_j(t)$ process under $\mathbb{Q}^{k+1}$.

For any $l < k$, Given that $B(T_l) = B_0$ and $a(T_l) = a_0$, we have the closed-form solutions for $B(T_{l+1})$ and $a(T_{l+1})$. Define constants $p = \left[u_0^2 - u_0\right] U_{s,j}^2$, $q = \sqrt{\left(\kappa_j^{k+1}\right)^2 + p\xi_j^2}$, $c = \frac{p}{q - \kappa_j^{k+1}}$ and $d = \frac{p}{q + \kappa_j^{k+1}}$. Then we have

$$\begin{aligned}
B_j(T_{l+1}) &= c - \frac{(c+d)(c - B_{j0})}{(d + B_{j0})\exp(-q\delta) + (c - B_{j0})}, 1 \leq j \leq N, \\
a(T_{l+1}) &= a_0 - \sum_{j=1}^{N}\left[\kappa_j^{k+1}\theta_j^{k+1}\left(d\delta + \frac{2}{\xi_j^2}\ln\left(\frac{(d + B_{j0})\exp(-q\delta) + (c - B_{j0})}{c+d}\right)\right)\right] \\
&\quad + \lambda_J \delta \left[\Gamma(u_0) - 1 - u_0(\Gamma(1) - 1)\right],
\end{aligned}$$

if $p \neq 0$ and $B_j(T_{l+1}) = B_{j0}, a(T_{l+1}) = a_0$ otherwise. $B(T_k)$ and $a(T_k)$ can be computed via iteration.

REFERENCES

Andersen, L. and R. Brotherton-Ratcliffe, 2001, Extended LIBOR Market Models with Stochastic Volatility, working paper, Gen Re Securities.

Andersen, T.G. and J. Lund, 1997, Estimating Continuous Time Stochastic Volatility Models of the Short Term Interest Rate, *Journal of Econometrics* 77, 343-378.

Bakshi, G., C. Cao, and Z. Chen, 1997, Empirical Performance of Alternative Option Pricing Models, *Journal of Finance* 52, 2003-2049.

Black, F. and M. Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637-654.

Black, F., 1976, The Pricing of Commodity Contracts, *Journal of Financial Economics* 3, 167-179.

Brace, A., D. Gatarek, and M. Musiela, 1997, The Market Model of Interest Rate Dynamics, *Mathematical Finance* 7, 127-155.

Brenner, R., R. Harjes, and K. Kroner, 1996, Another Look at Alternative Models of Short-Term Interest Rate, *Journal of Financial and Quantitative Analysis* 31, 85-107.

Campbell, J., A. Lo and C. MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, New Jersey).

Chernov, M., and E. Ghysels, 2000, A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation, *Journal of Financial Economics* 56, 407-458.

Collin-Dufresne, P. and R.S. Goldstein, 2002, Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility, *Journal of Finance* 57, 1685-1729.

Collin-Dufresne, P. and R.S. Goldstein, 2003, Stochastic Correlation and the Relative Pricing of Caps and Swaptions in a Generalized Affine Framework. Working paper, Carnegie Mellon University.

Cox, J.C., J.E. Ingersoll and S.A. Ross, 1985, A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385-407.

Dai, Q., and K. Singleton, 2003, Term Structure Dynamics in Theory and Reality, *Review of Financial Studies* 16, 631-678.

Das, S., 2002, The Surprise Element: Jumps in Interest Rates, *Journal of Econometrics* 106, 27-65.

Deuskar, P., Gupta, A. and M. Subrahmanyam, 2003, Liquidity Effects and Volatility Smiles in Interest Rate Option Markets, Working paper, New York University.

Diebold, F.X. and R.S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 253-265.

Duarte, J., 2004, Mortgage-Backed Securities Refinancing and the Arbitrage in the Swaption Market, working paper, University of Washington.

Duffie, D., 2002, *Dynamic Asset Pricing Theory* (Princeton University Press, New Jersey).

Duffie, D., J. Pan, K. Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* 68, 1343-1376.

Eraker, B., 2003, Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, *Journal of Finance* 59, 1367-1404.

Fan, R., A. Gupta, and P. Ritchken, 2003, Hedging in the Possible Presence of Unspanned Stochastic Volatility: Evidence from Swaption Markets, *Journal of Finance* 58, 2219-2248.

Gallant, A.R., and G. Tauchen, 1998, Reprojecting Partially Observed Systems with Applications to Interest Rate Diffusions, *Journal of American Statistical Association* 93, 10-24.

Glasserman, P. and S. Kou, 2002, The Term Structure of Simple Forward Rates with Jump Risk, *Mathematical Finance* forthcoming.

Gupta, A. and M. Subrahmanyam, 2001, An Examination of the Static and Dynamic Performance of Interest Rate Option Pricing Models in the Dollar Cap-Floor Markets, Working paper, Case Western Reserve University.

Goldstein, R.S., 2000, The Term Structure of Interest Rates as a Random Field, *Review of Financial Studies* 13, 365-384.

Han, Bing, 2002, Stochastic Volatilities and Correlations of Bond Yields, Working paper, Ohio State University.

Hansen, L.P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica* 50, 1029-1054.

Harris, L., 2003, *Trading and Exchanges: Market Microstructure for Practitioners*, Oxford University Press.

Heath, D., R. Jarrow, and A. Morton, 1992, Bond Pricing and the Term Structure of Interest Rates: A New Methodology, *Econometrica* 60, 77-105.

Heidari, M. and L. Wu, 2003, Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates?, *Journal of Fixed Income* 13, 75-86.

Heston, S., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 327-343.

Hull, J. and A. White, 1987, The Pricing of Options on Assets with Stochastic Volatilities, *Journal of Finance* 42, 281-300.

Hull, J. and A. White, 2000, Forward Rate Volatilities, Swap Rate Volatilities, and the Implementation of the LIBOR Market Models, *Journal of Fixed Income* 10, 46-62.

Jagannathan, R., A. Kaplin, and S. Sun, 2003, An Evaluation of Multi-factor CIR Models Using LIBOR, Swap Rates, and Cap and Swaption Prices, *Journal of Econometrics* 116, 113-146.

Johannes, M., 2004, The Statistical and Economic Role of Jumps in Interest Rates, *Journal of Finance* 59, 227-260.

Litterman, R., and J. Scheinkman, 1991, Common Factors Affecting Bond Returns, Journal of Fixed Income 1, 62-74.

Li, H., and F. Zhao, 2004, Hedging Interest Rate Derivatives Under Dynamic Term Structure Models: New Evidence of Unspanned Stochastic Volatility, working paper, Cornell University.

Longstaff, F., P. Santa-Clara, and E. Schwartz, 2001, The Relative Valuation of Caps and Swaptions: Theory and Evidence, *Journal of Finance* 56, 2067-2109.

Merton, R., 1973, The Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* 4, 141-183.

Miltersen, M., K. Sandmann, and D. Sondermann, 1997, Closed-form Solutions for Term Structure Derivatives with Lognormal Interest Rates, *Journal of Finance* 52, 409-430.

Newey, W., and K. West, 1987, A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.

Piazzesi, M. 2003, Affine Term Structure Models, *Handbook of Financial Econometrics,* forthcoming.

Piazzesi, M. 2004, Bond yields and the Federal Reserve, *Journal of Political Economy,* forthcoming.

Santa-Clara, P. and D. Sornette, 2001, The Dynamics of the Forward Interest Rate Curve with Stochastic String Shocks, *Review of Financial Studies* 14, 2001.

Table 1. Parameter Estimates of Stochastic Volatility Models

This table reports parameter estimates of the one, two, and three-factor stochastic volatility models. The estimates are based on the first stage GMM estimates with a identity weighting matrix and the standard errors are reported in the parentheses. The objective functions are the rescaled objective function of the first stage GMM and equal to the RMSE of each model. The volatility risk premium of the *i*th stochastic volatility factor is defined as $\eta_i = c_{iv}(T_k - 1)$.

	SV1		SV2		SV3	
Parameter	Estimate	Std. err	Estimate	Std. err	Estimate	Std. err
κ_1	0.0151	0.0091	0.0165	0.0136	0.0185	0.0150
κ_2			1.3995	0.0033	0.0002	0.0664
κ_3					0.0031	0.0174
$\bar{v}_1$	1.7297	0.9772	1.4310	0.0242	0.8447	0.8904
$\bar{v}_2$			0.0003	0.0001	0.0605	0.0069
$\bar{v}_3$					0.2106	2.5592
ζ_1	1.1012	0.0067	1.0356	0.0103	0.6948	0.0107
ζ_2			0.0451	0.0045	0.0282	0.0087
ζ_3					0.0565	0.0093
c_{1v}	-0.0024	0.00001	-0.0024	0.00002	-0.0021	0.0000
c_{2v}			-0.5661	0.0839	-0.0248	0.0070
c_{3v}					-0.0463	0.0081
Objective function	0.0638		0.0593		0.0543	

Table 2. Comparison of the Performance of Stochastic Volatility Models via Diebold-Mariano Statistics

This table reports comparison of model performance using Diebold-Mariano statistics, which measure whether a more sophisticated model has smaller pricing errors. A negative statistic means that the more sophisticated model has smaller pricing errors. The statistics are calculated according to equation (14) with a lag order q of 40 and follow an asymptotic standard Normal distribution under the null hypothesis of equal pricing errors.

Panel A. Diebold-Mariano statistics for overall model performance based on RMSEs.

Models	D-M Stats
SV2 – SV1	**-2.6779**
SV3 – SV2	**-3.5298**

Panel B. Diebold-Mariano statistics between SV2 and SV1 for individual difference caps based on absolute percentage pricing errors.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	0.67	0.58	-0.36	0.18	0.68	1.13	0.72	-0.84	-1.30
0.8	-	-	-0.03	-0.68	-1.04	-1.62	**-2.51**	-1.41	-0.41	0.38	-0.02	-1.30	-1.40
0.9	-	-1.49	-1.85	**-1.96**	**-2.06**	-1.81	**-2.15**	-1.67	-1.15	-0.77	-1.43	**-2.13**	-1.44
1.0	0.31	**-2.26**	-1.80	-1.52	-1.47	-1.29	-1.43	**-2.47**	**-3.58**	**-2.98**	**-2.76**	**-2.75**	-1.85
1.1	**-2.36**	**-2.65**	-1.32	-0.46	-0.45	-0.40	-0.90	**-1.97**	-	-	-	-	-

Panel C. Diebold-Mariano statistics between SV3 and SV2 for individual difference caps based on absolute percentage pricing errors.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	**3.23**	**3.73**	**3.44**	**4.23**	**2.71**	**-2.15**	**-2.25**	**-2.30**	**-3.44**
0.8	-	-	-1.02	1.87	**4.49**	**4.12**	**2.72**	**4.62**	**1.99**	**-4.11**	**-3.07**	**-2.89**	**-3.24**
0.9	-	**-2.31**	**-2.51**	-0.16	**7.11**	**4.10**	-0.86	1.43	1.13	**-6.26**	**-2.07**	**-2.31**	**-2.15**
1.0	-0.64	**-2.56**	**-2.65**	0.36	**1.99**	**-2.36**	**-2.76**	**-2.71**	1.04	0.62	1.82	-0.20	1.37
1.1	0.96	**-2.88**	1.89	**2.37**	**-2.58**	**-3.55**	**-2.95**	**-2.36**	-	-	-	-	-

Table 3. Average Percentage Pricing Errors of Stochastic Volatility Models

This table reports average percentage pricing errors of difference caps with different moneyness and maturity of three stochastic volatility models. Average percentage pricing errors are defined as the difference between market price and model price divided by market price.

Panel A. Average percentage pricing errors of SV1.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	0.0341	0.0265	0.0134	0.0356	0.0384	0.0532	0.0379	0.0336	0.0444
0.8	-	-	0.0416	0.0401	0.0320	0.0184	0.0116	0.0348	0.0346	0.0500	0.0339	0.0290	0.0376
0.9	-	0.1057	0.0503	0.0410	0.0301	0.0095	0.0007	0.0220	0.0208	0.0383	0.0146	0.0117	0.0221
1.0	0.0254	0.1160	0.0523	0.0336	0.0198	-0.0097	-0.0263	-0.0065	-0.0055	0.0127	-0.0061	-0.0125	0.0015
1.1	-0.1250	0.0530	-0.0097	-0.0281	-0.0376	-0.0730	-0.0816	-0.0549	-	-	-	-	-

Panel B. Average percentage pricing errors of SV2.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	0.0403	0.0322	0.0185	0.0399	0.0423	0.0564	0.0403	0.0339	0.0399
0.8	-	-	0.046	0.0461	0.0367	0.0221	0.0147	0.0374	0.0371	0.0522	0.0349	0.0272	0.0286
0.9	-	0.1109	0.0443	0.0418	0.0297	0.0086	-0.0005	0.0208	0.0206	0.038	0.0139	0.0073	0.0076
1.0	0.0222	0.1084	0.0288	0.0285	0.014	-0.0152	-0.0311	-0.0105	-0.0076	0.0107	-0.008	-0.0186	-0.0175
1.1	-0.0915	0.0623	-0.0266	-0.0274	-0.0391	-0.0753	-0.084	-0.0572	-	-	-	-	-

Panel C. Average percentage pricing errors of SV3.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	0.0521	0.0454	0.0307	0.0485	0.0442	0.0524	0.0353	0.0303	0.0304
0.8	-	-	0.0464	0.0499	0.048	0.038	0.0305	0.0483	0.0379	0.0448	0.0271	0.0226	0.015
0.9	-	0.0931	0.0363	0.0371	0.0386	0.0263	0.0186	0.0338	0.0195	0.0263	0.0023	0.0012	-0.0108
1.0	-0.0076	0.0779	0.0163	0.0157	0.0215	0.0059	-0.0073	0.0055	-0.0105	-0.0062	-0.0245	-0.0282	-0.0446
1.1	-0.1059	0.0339	-0.0347	-0.0365	-0.0245	-0.0447	-0.0513	-0.036	-	-	-	-	-

Table 4. Parameter Estimates of Stochastic Volatility and Jumps Models

This table reports parameter estimates of the one, two, and three-factor stochastic volatility and jumps models. The estimates are based on the first stage GMM estimates with a identity weighting matrix and the standard errors are reported in the parentheses. The objective functions are the rescaled objective function of the first stage GMM and equal to the RMSE of each model. The volatility risk premium of the *i*th stochastic volatility factor is defined as $\eta_i = c_{iv}(T_k - 1)$, and the jump risk premium is defined as $\eta_i = c_J(T_k - 1)$.

	SVJ1		SVJ2		SVJ3	
Parameter	Estimate	Std. err	Estimate	Std. err	Estimate	Std. Err
κ_1	0.0270	0.0098	0.0170	0.0188	0.0081	0.0002
κ_2			0.7660	0.0650	0.0004	0.0049
κ_3					0.0080	0.0216
$\bar{v}_1$	0.6112	0.3926	0.9336	0.1613	0.9272	2.1697
$\bar{v}_2$			0.0011	0.0001	0.1271	0.0328
$\bar{v}_3$					0.1989	0.8386
ζ_1	0.9719	0.0078	0.9552	0.0127	0.7713	0.0109
ζ_2			0.0129	0.0072	0.0292	0.0087
ζ_3					0.0577	0.0086
c_{1v}	-0.0037	0.00004	-0.0036	0.0006	-0.0020	0.0003
c_{2v}			-0.2670	0.1103	-0.0258	0.0037
c_{3v}					-0.0444	0.0011
λ	0.0597	0.0013	0.0208	0.0006	0.0218	0.0005
μ_J	-0.6687	0.0074	-1.7004	0.0480	-2.1273	0.0391
c_J	0.0277	0.0010	0.0631	0.0041	0.1298	0.0030
σ_J	0.4719	0.0058	0.3344	0.0817	-0.0100	0.1992
Objective Function	0.0575		0.0529		0.0482	

Table 5. Comparison of the Performance of Stochastic Volatility and Jump Models via Diebold-Mariano Statistics

This table reports comparison of model performance using Diebold-Mariano statistics, which measure whether a more sophisticated model has smaller pricing errors. A negative statistic means that the more sophisticated model has smaller pricing errors. The statistics are calculated according to equation (14) with a lag order q of 40 and follow an asymptotic standard Normal distribution under the null hypothesis of equal pricing errors.

Panel A. Diebold-Mariano statistics for overall model performance based on RMSEs.

Models	D-M Stats
SVJ2 – SVJ1	**-4.4229**
SVJ3 – SVJ2	**-3.6345**

Panel B. Diebold-Mariano statistics between SVJ2 and SVJ1 for individual difference caps based on absolute percentage pricing errors.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	-1.73	-0.98	0.25	**-2.19**	-1.94	**-2.76**	**-2.14**	-1.86	**-2.25**
0.8	-	-	-0.94	-1.71	-0.65	-0.76	-1.73	**-2.13**	-1.70	**-2.63**	**-2.55**	**-3.25**	**-1.98**
0.9	-	-0.03	-1.68	-1.80	-0.93	-1.38	**-2.08**	**-2.10**	-0.90	-1.47	-1.77	**-3.28**	-1.44
1.0	0.37	-1.91	**-3.39**	-1.13	-1.33	**-2.39**	**-2.65**	**-4.46**	-1.58	-1.15	**-2.08**	**-2.70**	**-1.98**
1.1	-0.65	**-2.07**	-0.81	-0.46	0.07	-0.77	-1.36	-1.93	-	-	-	-	-

Panel C. Diebold-Mariano statistics between SVJ3 and SVJ2 for individual difference caps based on absolute percentage pricing errors.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	**-**	**2.05**	1.21	-0.85	-1.25	-1.81	**-2.65**	**-2.55**	**-2.70**	-1.26
0.8	-	-	-0.04	**2.04**	**2.10**	0.82	-1.05	-1.41	-1.73	**-2.45**	**-2.58**	**-3.10**	**-3.47**
0.9	-	**-2.06**	**-2.07**	0.62	1.77	1.11	-1.44	-0.80	-1.31	-1.76	-1.14	**-2.91**	**-2.97**
1.0	-1.67	-1.87	-1.16	0.86	1.74	-0.82	-1.82	-0.74	0.11	-0.66	0.58	-0.44	-0.21
1.1	-1.63	-1.64	1.54	0.13	-1.93	**-3.40**	**-2.57**	-0.52	-	-	-	-	-

Table 6. Average Percentage Pricing Errors of Stochastic Volatility and Jump Models

This table reports average percentage pricing errors of difference caps with different moneyness and maturity of three stochastic volatility and jump models. Average percentage pricing errors are defined as the difference between market price and model price divided by market price.

Panel A. Average percentage pricing errors of SVJ1.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	0.0181	0.0118	-0.0006	0.0218	0.0241	0.0374	0.0194	0.0124	0.021
0.8	-	-	0.0082	0.0105	0.0056	-0.0053	-0.01	0.0146	0.0146	0.0293	0.0104	0.0031	0.0092
0.9	-	0.0579	0.0063	0.0044	-0.0002	-0.0154	-0.0202	0.0039	0.004	0.0211	-0.0054	-0.0108	-0.0036
1.0	-0.0272	0.0769	0.0227	0.0154	0.0106	-0.0117	-0.0238	-0.0015	0.0004	0.0157	-0.0078	-0.0192	-0.011
1.1	-0.119	0.0676	0.0166	0.0083	0.0081	-0.0208	-0.0263	-0.0006	-	-	-	-	-

Panel B. Average percentage pricing errors of SVJ2.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	0.0168	0.0089	-0.0048	0.0164	0.0174	0.0296	0.0107	0.0026	0.0091
0.8	-	-	0.0143	0.0113	0.0094	-0.003	-0.009	0.0146	0.0135	0.0272	0.007	-0.002	-0.0002
0.9	-	0.0706	0.013	0.002	0.0074	-0.0089	-0.0145	0.0091	0.0093	0.0258	-0.0008	-0.0088	-0.0085
1.0	-0.0255	0.0774	0.0139	-0.006	0.0143	-0.007 2	-0.0178	0.0057	0.0102	0.0269	0.0053	-0.0077	-0.009
1.1	-0.1062	0.0605	-0.009	-0.0358	0.0019	-0.0238	-0.0254	0.0039	-	-	-	-	-

Panel C. Average percentage pricing errors of SVJ3.

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	0.0173	0.008	-0.0083	0.0096	0.0055	0.0151	-0.0002	0.001	0.0145
0.8	-	-	0.0138	0.0144	0.0102	-0.0013	-0.0089	0.0105	0.0013	0.0102	-0.0055	-0.0027	0.0053
0.9	-	0.0632	0.0041	0.0045	0.0064	-0.0039	-0.0087	0.0098	-0.0019	0.0075	-0.0136	-0.0059	-0.0022
1.0	-0.0311	0.0558	-0.0034	-0.0002	0.0106	0.0017	-0.0045	0.0137	0.0016	0.0078	-0.006	0.0002	-0.0032
1.1	-0.089	0.0434	-0.023	-0.0184	0.0033	-0.0052	-0.0004	0.0216	-	-	-	-	-

Table 7. Comparison of the Performance of SV and SVJ Models via Diebold-Mariano Statistics

This table reports comparison of model performance using Diebold-Mariano statistics, which measure whether a more sophisticated model has smaller pricing errors. A negative statistic means that the more sophisticated model has smaller pricing errors. The statistics are calculated according to equation (14) with a lag order q of 40 and follow an asymptotic standard Normal distribution under the null hypothesis of equal pricing errors.

Panel A. Diebold-Mariano statistics for overall model performance based on RMSEs (with and without the first 20 weeks).

Models	D-M Stats (whole sample)	D-M Stats (without first 20 weeks)
SVJ1 - SV1	-1.7241	**-4.8358**
SVJ2 – SV2	-1.757	**-5.8541**
SVJ3 – SV3	-1.7054	**-10.039**

Panel B. Diebold-Mariano statistics between SVJ1 and SV1 for individual difference caps based on absolute percentage pricing errors (without first 20 weeks).

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	-	-	**-2.27**	-1.73	-0.33	-1.72	**-2.19**	**-2.72**	**-3.09**	**-2.99**	**-4.23**
0.8	-	-	**-3.16**	**-5.21**	**-3.64**	-1.51	-0.46	-1.40	**-2.10**	**-2.14**	**-2.59**	**-2.02**	**-2.27**
0.9	-	**-6.06**	**-8.94**	**-4.30**	**-2.78**	-0.79	-0.66	-0.98	-0.57	-0.79	-1.59	-1.35	-1.43
1.0	0.92	**-8.33**	**-5.42**	**-2.00**	-0.99	**-3.39**	**-4.83**	**-3.61**	-1.98	-0.66	-1.49	-1.86	-1.93
1.1	0.09	-1.50	0.00	0.18	-1.82	**-3.75**	**-3.24**	-1.65	-	-	-	-	-

Panel C. Diebold-Mariano statistics between SVJ2 and SV2 for individual difference caps based on absolute percentage pricing errors (without first 20 weeks).

Moneyness	1.5yr	2yr	2.5yr	3yr	3.5yr	4yr	4.5yr	5yr	6yr	7yr	8yr	9yr	10yr
0.7	-	-	**-**	**-**	**-2.58**	**-2.08**	-0.49	**-2.63**	**-3.43**	**-4.51**	**-4.63**	**-3.71**	**-5.82**
0.8	-	-	**-3.74**	**-4.68**	**-2.83**	-1.14	-0.26	-1.58	**-2.63**	**-2.99**	**-3.57**	**-2.37**	**-2.55**
0.9	-	**-8.42**	**-7.32**	**-4.74**	-1.91	0.00	-0.04	-0.70	-0.04	-0.73	-1.78	-1.88	-1.20
1.0	0.38	**-5.44**	**-4.28**	-1.10	1.39	**-3.31**	**-3.99**	**-3.24**	-0.56	1.00	-0.73	-1.41	-1.23
1.1	1.79	-1.35	0.33	**3.14**	**-3.51**	**-7.19**	**-5.30**	**-2.86**	-	-	-	-	-

Panel D. Diebold-Mariano statistics between SVJ3 and SV3 for individual difference caps based on absolute percentage pricing errors (without first 20 weeks).

Moneyness	**1.5yr**	**2yr**	**2.5yr**	**3yr**	**3.5yr**	**4yr**	**4.5yr**	**5yr**	**6yr**	**7yr**	**8yr**	**9yr**	**10yr**
0.7	-	-	-	-	**-4.54**	**-5.48**	**-4.32**	**-9.99**	**-8.02**	**-10.97**	**-7.10**	**-3.99**	**-5.35**
0.8	-	-	**-3.59**	**-5.44**	**-6.09**	**-7.72**	**-3.66**	**-5.65**	**-9.07**	**-6.67**	**-5.05**	**-2.27**	**-2.01**
0.9	-	**-3.23**	**-4.03**	**-3.41**	**-6.03**	**-3.93**	-0.90	**-2.58**	**-3.11**	-1.95	-1.38	**-2.53**	**-2.23**
1.0	1.09	**-2.40**	-1.83	-1.42	-0.29	0.16	**-3.11**	-0.84	-1.83	-1.49	**-2.82**	-1.41	**-2.94**
1.1	-1.85	**2.26**	**-2.96**	**-2.06**	-1.86	**-3.60**	**-3.48**	-1.27	-	-	-	-	-

Table 8: Correlation Between LIBOR Rates and Implied Volatility Variables

This table reports the correlations between LIBOR rates and implied volatility variables from SVJ3. Given the parameter estimates of SVJ model in Table 3, the implied volatility variables are estimated each week by minimizing the RMSEs of all difference caps.

	L(t,1)	**L(t,3)**	**L(t,5)**	**L(t,7)**	**L(t,9)**	**V1(t)**	**V2(t)**	**V3(t)**
V1(t)	-0.9415	-0.8969	-0.8416	-0.7948	-0.749	1	-0.3557	0.5942
V2(t)	0.1687	0.2109	0.1671	0.1203	0.0564	-0.3557	1	-0.0509
V3(t)	-0.6654	-0.5543	-0.4879	-0.4293	-0.3987	0.5942	-0.0509	1

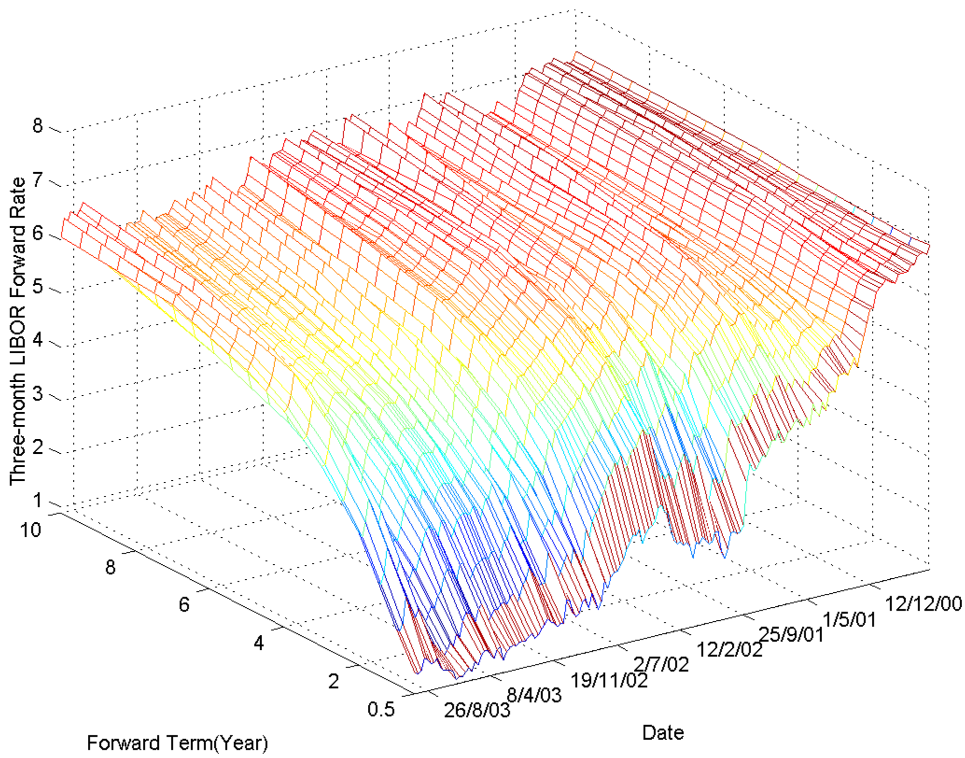

Figure 1. Term structure of three-month LIBOR forward rates between 1/8/2000 and 23/9/2003.

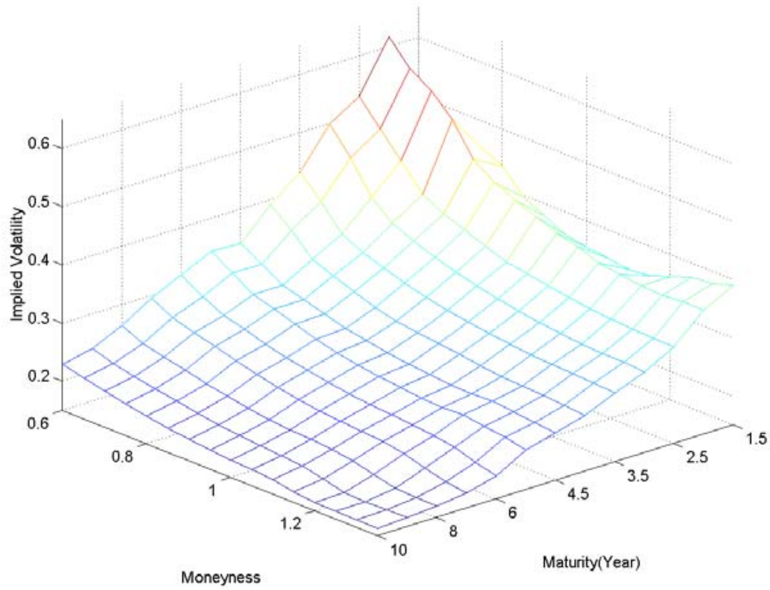

Figure 2.a. Average Black implied volatilities of difference caps between 1/8/2000 and 23/9/2003.

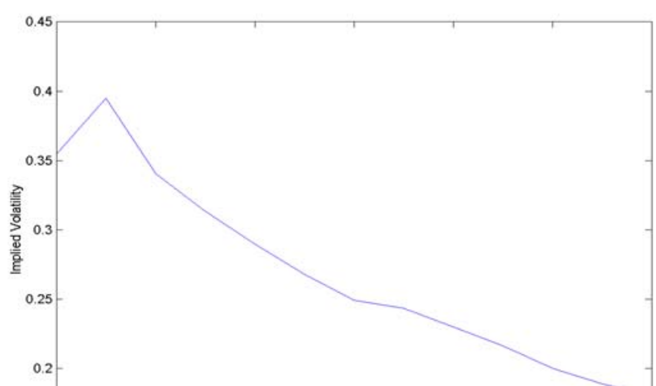

Figure 2.b. Average Black implied volatilities of ATM difference caps between 1/8/2000 and 23/9/2003.

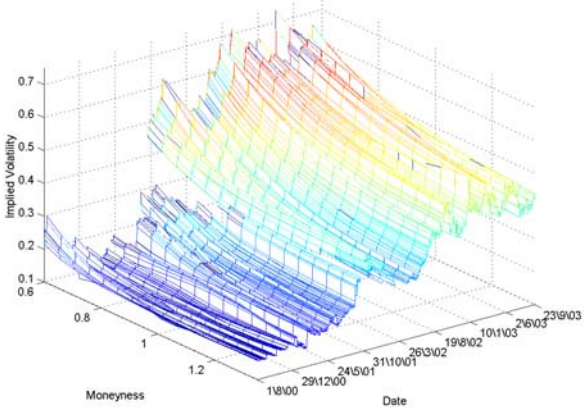

Figure 3.a. Black implied volatilities of 2.5-year difference caps.

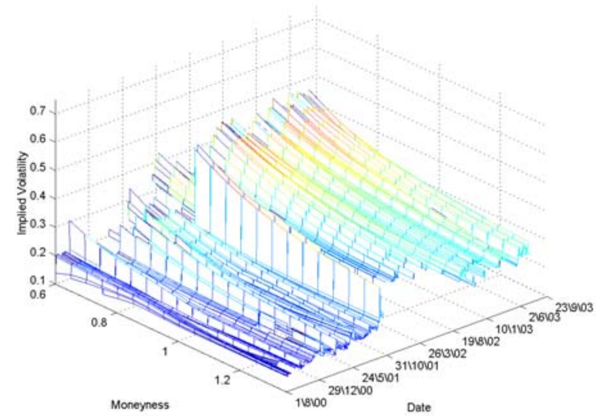

Figure 3.b. Black implied volatilities of 5-year difference caps.

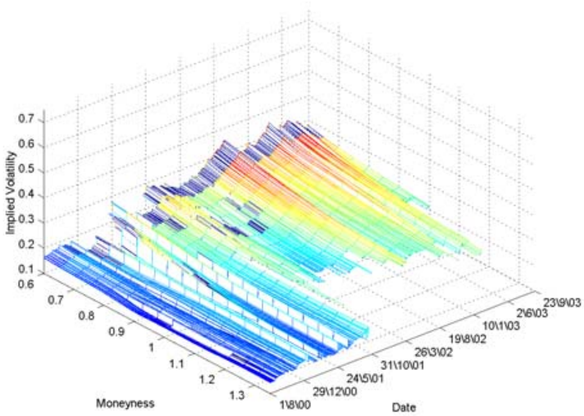

Figure 3.c. Black implied volatilities of 8-year difference caps.

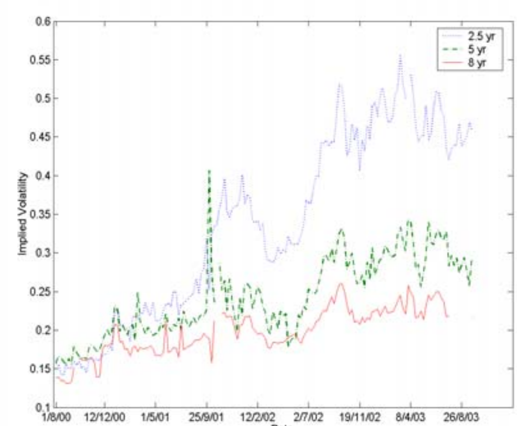

Figure 3.d. Black implied volatilities of 2.5-, 5-, and 8-year ATM difference caps.

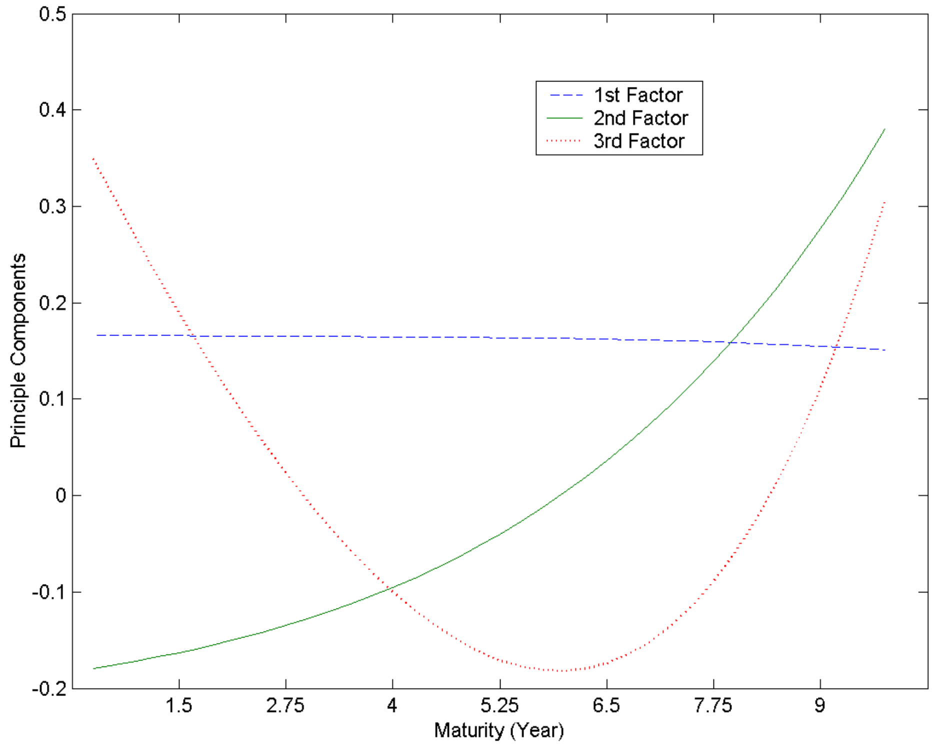

Figure 4. The first three principal components of weekly percentage changes of three-month LIBOR rates between 2/6/1997 and 31/7/2000.

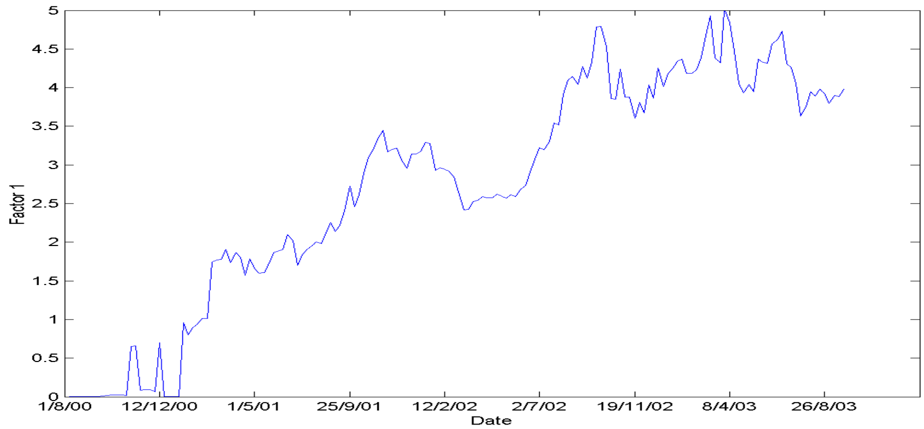

Figure 5.a. The implied stochastic volatility variables of SV1 between 1/8/2000 and 23/9/2003.

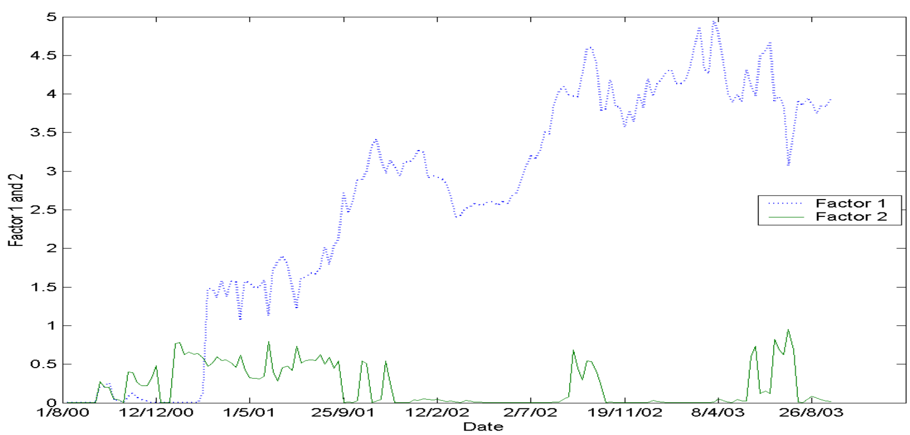

Figure 5.b. The implied stochastic volatility variables of SV2 between 1/8/2000 and 23/9/2003.

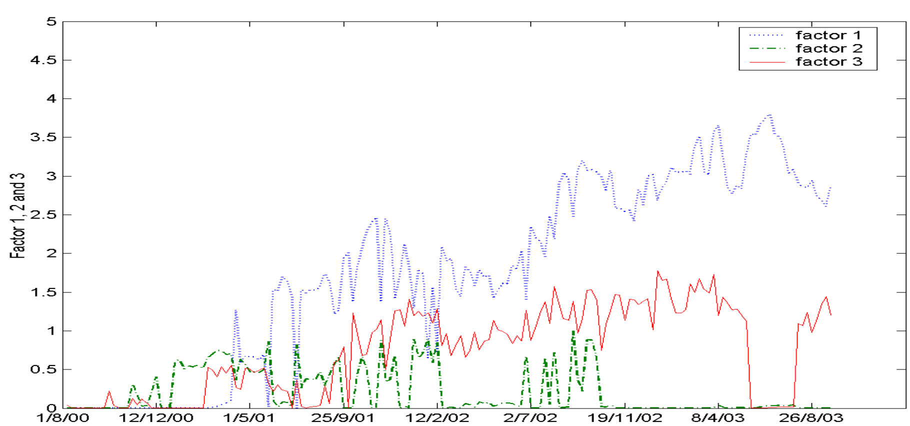

Figure 5.c. The implied stochastic volatility variables of SV3 between 1/8/2000 and 23/9/2003.

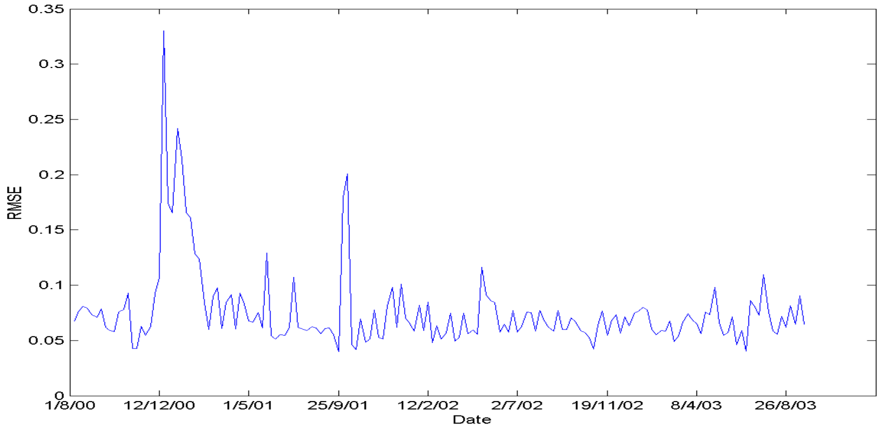

Figure 6.a. The RMSEs of SV1 between 1/8/2000 and 23/9/2003.

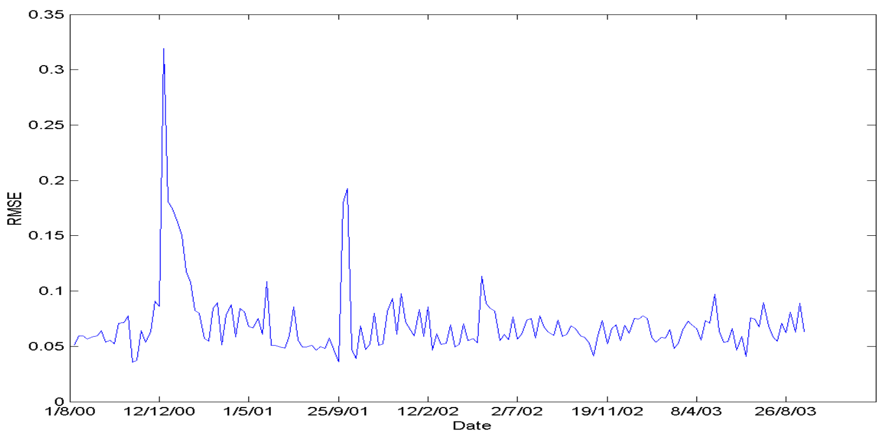

Figure 6.b. The RMSEs of SV2 between 1/8/2000 and 23/9/2003.

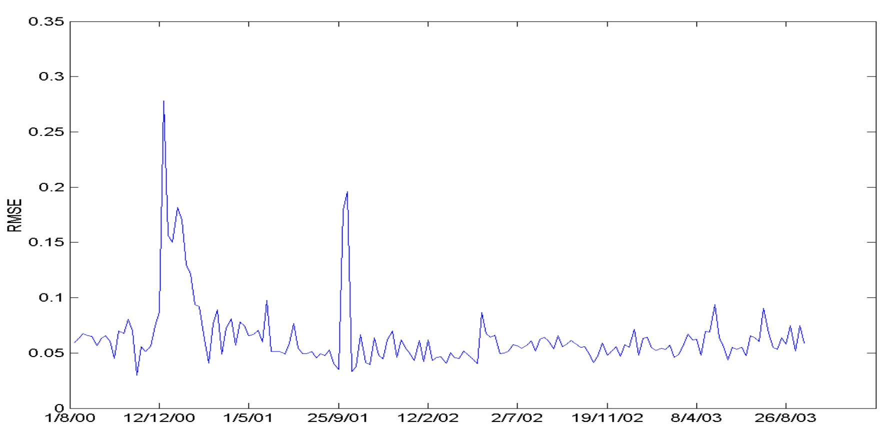

Figure 6.c. The RMSEs of SV3 between 1/8/2000 and 23/9/2003.

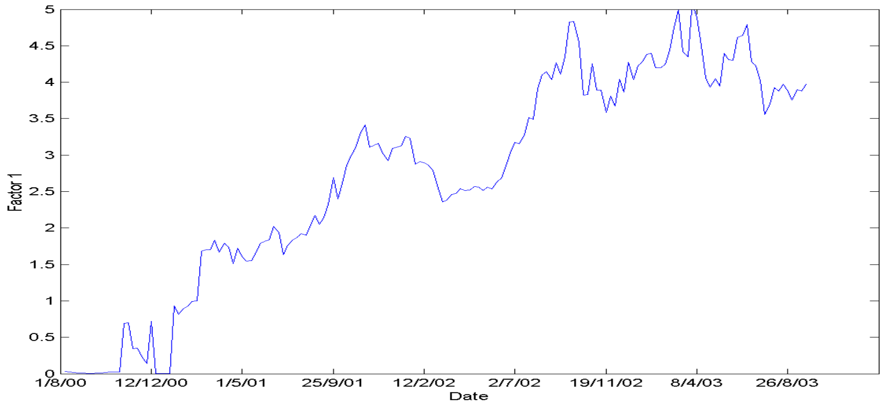

Figure 7.a. The implied stochastic volatility variables of SVJ1 between 1/8/2000 and 23/9/2003.

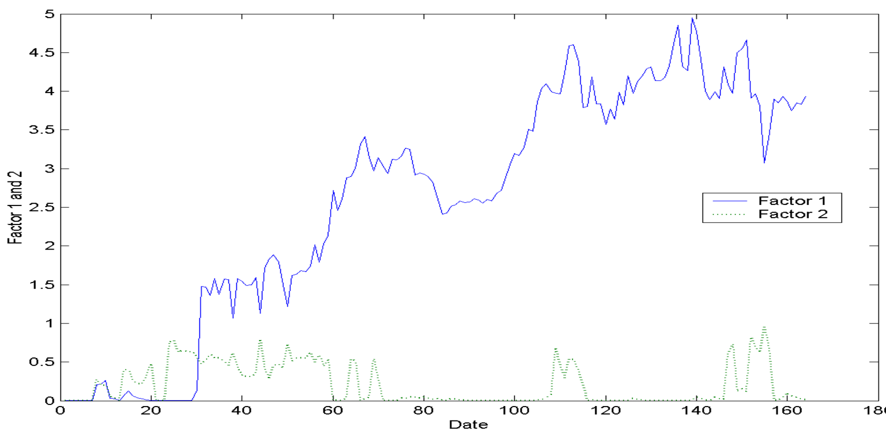

Figure 7.b. The implied stochastic volatility variables of SVJ2 between 1/8/2000 and 23/9/2003.

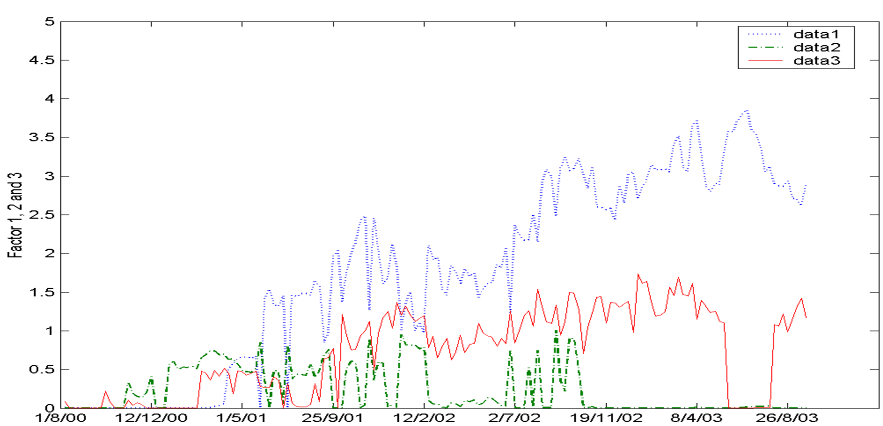

Figure 7.c. The implied stochastic volatility variables of SVJ3 between 1/8/2000 and 23/9/2003.

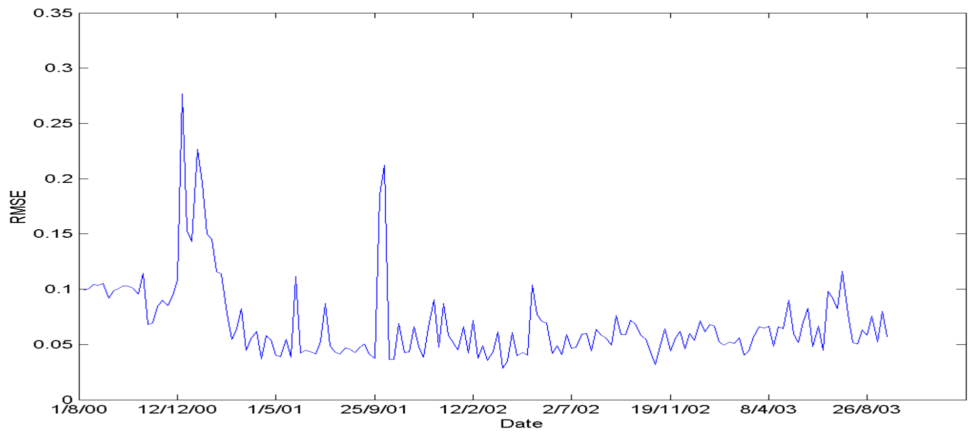

Figure 8.a. The RMSEs of SVJ1 between 1/8/2000 and 23/9/2003.

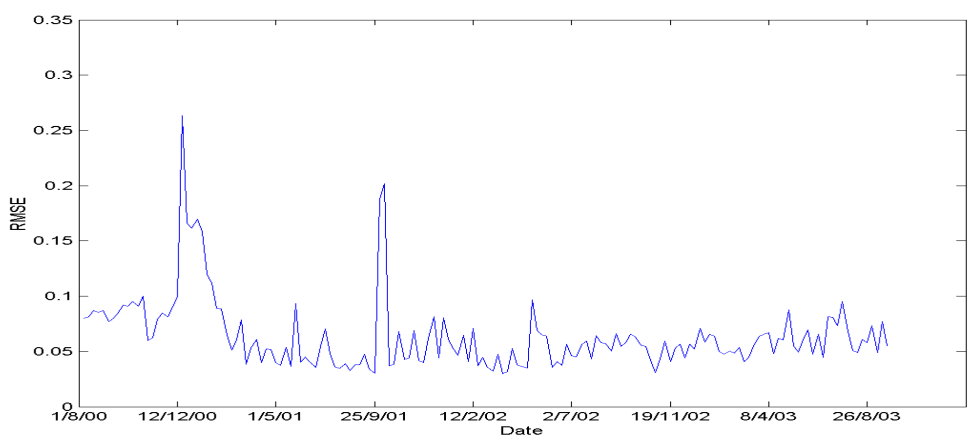

Figure 8.b. The RMSEs of SVJ2 between 1/8/2000 and 23/9/2003.

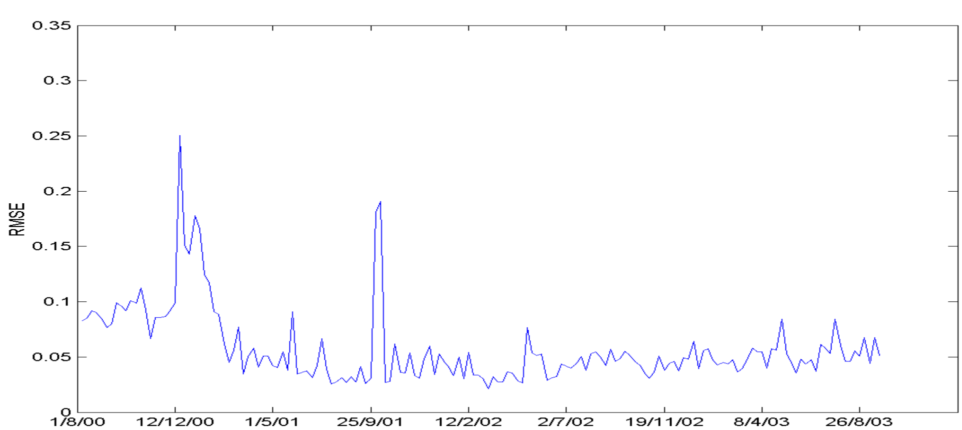

Figure 8.c. The RMSEs of SVJ3 between 1/8/2000 and 23/9/2003.