Nonparametric Estimation of State-Price Densities Using Interest Rate Options

Haitao Li and Feng Zhao*

November 15, 2006

(Preliminary, Comments Welcome)

*Li is from the Stephen M. Ross School of Business, University of Michigan. Zhao is from the Rutgers Business School, Rutgers University. We thank Yacine Aït-Sahalia, Ivan Brick, Jefferson Duarte, Yongmiao Hong, Robert Jarrow and seminar participants at Rutgers University for helpful discussions. We are responsible for any remaining errors.

Nonparametric Estimation of State-Price Densities Using Interest Rate Options

ABSTRACT

We provide one of the first nonparametric estimates of the probability densities of LIBOR rates under forward martingale measures and the state-price densities (SPDs) using interest rate cap prices. We extend the constrained local polynomial approach of Aït-Sahalia and Duarte (2003) to a multivariate setting and estimate the forward densities and the SPDs conditional on the level, slope, and volatility of the LIBOR rates. The multivariate constrained local polynomial approach has excellent finite sample performances and guarantees that the nonparametric estimates satisfy necessary theoretical restrictions. We find that the forward densities deviate significantly from the log-normal distribution and are strongly negatively skewed. The SPDs exhibit a pronounced U-shape as a function of future LIBOR rates. This suggests that the state prices are high at low and high interest rates, which tend to correspond to economic recessions and high inflations, respectively. Both the forward densities and the SPDs depend significantly on the volatility of the LIBOR rates, suggesting that it is important to incorporate unspanned stochastic volatility into equilibrium term structure models.

Over-the-counter interest rate derivatives, such as caps and swaptions, are among the most widely traded derivatives in the world. According to the Bank for International Settlements, in recent years, the notional value of caps and swaptions exceeds 10 trillion dollars, which is many times larger than that of exchange-traded options. Although the extensive term structure literature of the last decade has mainly focused on explaining bond yields and swap rates,[1] prices of caps and swaptions are likely to contain richer information on term structure dynamics because their payoffs are nonlinear functions of underlying interest rates. As a result, in a recent survey of the term structure literature, Dai and Singleton (2003) suggest that there is an "enormous potential for new insights from using (interest rate) derivatives data in (term structure) model estimations."

In recent years, a fast-growing literature has been developed to understand term structure dynamics from the perspective of pricing and hedging interest rate derivatives. So far the literature has mainly focused on the following three important issues: (i) whether bonds span interest rate derivatives; (ii) relative pricing between caps and swaptions; and (iii) relative pricing of caps with different strike prices. This literature has provided important insights that otherwise would be difficult to obtain by studying only bond yields and swap rates. For example, Collin-Dufresne and Goldstein (2002), Heidari and Wu (2003), and Li and Zhao (2006) show that interest rate derivatives are not redundant securities and cannot be spanned by bond market factors. While Han (2002) shows that unspanned stochastic volatility helps to reconcile relative mispricing between caps and swaptions, Jarrow, Li, and Zhao (2006) show that significant negative jumps in interest rates are needed to capture the volatility skew they document in the cap markets.

Despite these interesting progresses, one common feature of most existing term structure models is that they rely on parametric assumptions to obtain closed-form pricing formulae for interest rate derivatives. For example, the popular LIBOR (Swap) market model assumes that LIBOR forward (swap) rates follow the log-normal distribution and prices caps (swaptions) using the Black (1976) formula. The models of Collin-Dufresne and Goldstein (2002), Han (2002), and Jarrow, Li, and Zhao (2006) rely on the affine jump-diffusion models of Duffie, Pan, and Singleton (2000), while the models of Li and Zhao (2006) rely on the quadratic term structure models of Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2003). As a result, these models are likely to be misspecified and therefore may not be able to fully capture the prices of interest rate derivatives.

The main objective of our paper is to extract the rich information on term structure dynamics contained in the prices of interest rate derivatives without imposing restrictive parametric assump-

[1] See Dai and Singleton (2003) and Piazzesi (2003) for surveys of the literature.

tions. Specifically, we provide nonparametric estimates of the probability densities of LIBOR rates under forward martingale measures and the state-price densities (SPDs) using LIBOR rates and interest rate caps with different strike prices and maturities.

The nonparametric densities of the LIBOR rates under different forward measures can be useful for many purposes. First, they contain the most important information for pricing interest rate derivatives due to the widely used forward-measure approach in the current term structure literature. Under a risk-neutral measure, the price of any security discounted by the money market account is a martingale. Similarly under a forward measure, the price of any security discounted by a zero-coupon bond associated with the forward measure is a martingale. Therefore, the forward densities of the LIBOR rates backed out from the interest rate caps, which are among the simplest and the most liquid OTC interest rate derivatives, allow consistent pricing of more exotic and/or less liquid OTC interest rate derivatives. Second, the nonparametric forward densities can reveal potential misspecifications of existing parametric term structure models, which could be helpful for the development of new parametric term structure models. Third, combining the physical and forward densities of the LIBOR rates, we can estimate the SPDs or the intertemporal marginal rate of substitutions of the representative investor implicit in interest rate derivative prices. This allows us to study investor preferences from a perspective that is different from most of those in the existing literature, which are mainly based on index option prices.[2]

Our paper makes several contributions to the fast-growing literature on interest rate derivatives. First, building on the insights of Breeden and Litzenberger (1978), we provide nonparametric estimates of the forward LIBOR densities using the constrained local polynomial approach of Aït-Sahalia and Duarte (2003). Compared to the nonparametric kernel regression approach of Aït-Sahalia and Lo (1998, 2000) and others for index options, the approach of Aït-Sahalia and Duarte (2003) has superior finite sample performances and guarantees that the estimated forward densities satisfy the necessary theoretical restrictions. One important innovation of our analysis is that we extend the approach of Aït-Sahalia and Duarte (2003) to a multivariate setting. The new extension makes it possible to estimate the forward densities and the SPDs conditional on important economic variables in addition to the usual inputs to the Black-Scholes-type of models considered in most existing studies.

Second, based on the newly extended approach of Aït-Sahalia and Duarte (2003), we conduct

[2] Beber and Brandt (2006) estimates investor preferences using interest rate options. However, as explained later, there are several important differences between their analysis and ours.

probably the first nonparametric estimation of the forward LIBOR densities conditional on the level, slope, and volatility factors of the LIBOR rates.[3] We include these three conditioning variables in our analysis because of the important roles they play for term structure modeling. For example, it has been widely documented that the level and slope factors of LIBOR rates can explain close to 99% of the variations of the LIBOR rates. Many studies, such as Jarrow, Li, and Zhao (2006), also have documented the importance of stochastic volatility for pricing and hedging LIBOR-based derivatives. While the level factor can be easily incorporated under existing methods, our new extension of Aït-Sahalia and Duarte (2003) is needed to incorporate the slope and volatility factors in our nonparametric estimation. In comparison, despite the overwhelming evidence of stochastic volatility in index returns, most existing nonparametric estimates of the SPDs using index options do not allow for stochastic volatility.

Third, we provide nonparametric estimates of the SPDs conditional on the level, slope, and volatility factors of the LIBOR rates over different horizons. The SPDs are estimated as the ratio between the forward and physical densities of the LIBOR rates conditional on the three state variables. The latter is estimated using the kernel method of Ait-Sahlaia and Lo (2000). Our results provide a new perspective on investor preferences that is different from that of most existing studies. While most existing studies estimate the SPDs as a function of the level of the equity market, our analysis documents the dependence of the SPDs on important term structure factors. Moreover, while the index options used in most existing studies tend to have very short maturities (less than one or two years), the interest rate caps we consider allow us to estimate the SPDs over longer horizons.

Finally, we provide interesting new empirical evidence on the forward and physical densities of the LIBOR rates as well as the SPDs. For example, we find that the forward densities deviate significantly from the log-normal distribution assumed by the LIBOR market models and are strongly negatively skewed. In addition, we find that both densities depend significantly on the slope and volatility factors of the LIBOR rates. Most interestingly, we find a pronounced U-shape of the SPDs as a function of future LIBOR rates, suggesting that investors attach high values to payoffs when interest rates are extremely high or low. This is consistent with the notion that extremely low interest rates tend to be associated with economic slowdowns or even recessions, while extremely high interest rates tend to be associated with hyper inflations. This pattern differs significantly from

[3] In this paper, the volatility factor is the first principal component of the filtered instantaneous volatilities of the LIBOR rates at all maturities via an EGARCH model. We obtain very similar results using GARCH models.

that estimated from index options, which is typically a declining function of the level of the equity market. The SPDs also depend significantly on the volatility of the LIBOR rates, which strongly suggests that unspanned stochastic volatility is an important state variable that affects the pricing kernel and should be incorporated into equilibrium term structure models.

Our paper is closely related to Aït-Sahalia (1996, 1998), which are the earliest studies on nonparametric pricing of interest rate derivatives. While Aït-Sahalia (1996, 1998) price interest rate derivatives based on nonparametrically estimated diffusion processes for spot interest rates, we provide nonparametric estimates of the forward densities of the LIBOR rates conditional on the slope and volatility of the LIBOR rates, and thus explicitly allows for unspanned stochastic volatility in interest rate derivatives markets. Beber and Brandt (2006) is one of the few papers that estimates investor preferences using interest rate option prices. Our paper, however, differs from Beber and Brandt (2006) in several important aspects. First, while Beber and Brandt (2006) use the Edgeworth expansion method of Jarrow and Rudd (1982) to estimate the SPDs, we use the extended local polynomial method of Aït-Sahalia and Duarte (2003). Second, the nonparametric method allows us to estimate the SPDs conditional on important term structure factors, such as the slope and volatility of the LIBOR rates. In contrast, Beber and Brandt (2006) do not account for stochastic volatility in interest rates. Third, while Beber and Brandt (2006) consider only short-term interest rate options with maturities typically less than one year, we consider interest rate caps with maturities up to ten years.

The rest of the paper is organized as follows. In Section 1, we discuss how to estimate forward densities from cap prices. In Section 2, we introduce the nonparametric methods used for estimating the forward densities and the SPDs. In Section 3, we present the data and document a volatility smile in interest rate cap markets. Section 4 reports the empirical findings and Section 5 concludes.

1. Forward Densities Implicit in Cap Prices

In this section, we discuss the general idea of estimating forward densities from cap prices. 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. For example, a five-year cap on three-month LIBOR struck at 6% represents a portfolio of 19 separately exercisable caplets with quarterly maturities ranging from six months to five years, where each caplet has a strike price of 6%.

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 we assume that the intervals $T_{k+1} - T_k$ are equally spaced by δ, a quarter of a year as in the U.S. cap markets. Let $L_k(t) = L(t, T_k)$ be the LIBOR forward rate for the actual period $[T_k, T_{k+1}]$, and let $D_k(t) = D(t, T_k)$ be the price of a zero-coupon bond maturing at T_k. We then have

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

A caplet for the period $[T_k, T_{k+1}]$ struck at X pays $\delta(L_k(T_k) - X)^+$ at T_{k+1}. Although the cash flow of this caplet is received at time T_{k+1}, the LIBOR rate is determined at time T_k and there is no uncertainty about the caplet's cash flow after T_k.

For LIBOR-based instruments such as caps, floors, and swaptions, it is convenient to consider pricing under the forward measure. We will therefore 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. Thus, the time-t price of a caplet maturing at T_{k+1} with a strike price of X is

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

where $E_t^{\mathbb{Q}^{k+1}}$ is taken with respect to $\mathbb{Q}^{k+1}$ given the information set at t and $\tau_k = T_k - t$, the time horizon over which $L_k(t)$ can randomly fluctuate. The key to valuation is the distribution of $L_k(t)$ under $\mathbb{Q}^{k+1}$. Once we know this distribution, we can price any security whose payoff on T_{k+1} depends only on $L_k(t)$ by discounting its expected payoff under $\mathbb{Q}^{k+1}$ using $D_{k+1}(t)$.

Existing term structure models rely on parametric assumptions on the distributions of $L_k(t)$ to obtain closed-form pricing formulae for caplets. For example, the standard LIBOR market models of Brace, Gatarek, and Musiela (1997) and Miltersen, Sandmann, and Sondermann (1997) assume that $L_k(t)$ follows a log-normal distribution and price caplet using the Black formula. The models of Jarrow, Li, and Zhao (2006) assume that $L_k(t)$ follows affine jump-diffusions of Duffie, Pan, and Singleton (2000).

In this paper, we estimate the distribution of $L_k(t)$ under $\mathbb{Q}^{k+1}$ using the prices of a cross section of caplets that mature at T_{k+1} and have different strike prices. Breeden and Litzenberger (1979) show that the density of $L_k(t)$ under $\mathbb{Q}^{k+1}$ is proportional to the second derivative of $Caplet(L_k(t), \tau_k, X)$ with respect to X. Specifically, define

$$Caplet(L_k(t), \tau_k, X) = \delta D_{k+1}(t)\, C(L_k(t), \tau_k, X), \tag{3}$$

where

$$
\begin{aligned}
C\left(L_k\left(t\right),\tau_k,X\right) &= E_t^{\mathbb{Q}^{k+1}}\left[\left(L_k\left(T_k\right)-X\right)^+\right] \\
&= \int_X^\infty \left(y-X\right)p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)=y|L_k\left(t\right)\right)dy.
\end{aligned}
$$

Then the conditional density of $L_k\left(T_k\right)$ under the forward measure $\mathbb{Q}^{k+1}$ is given by

$$
p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)|L_k\left(t\right)\right)=\frac{\partial^2 C\left(L_k\left(t\right),\tau_k,X\right)}{\partial X^2}|_{X=L_k(T_k)}. \tag{4}
$$

The formula in (4) makes the assumption that the conditional density of $L_k\left(T_k\right)$ depends only on the current LIBOR rate, i.e., $p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)|\mathcal{F}_t\right)=p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)|L_k\left(t\right)\right),$ where $\mathcal{F}_t$ represents the information set at t. This assumption, however, can be overly restrictive given the multifactor nature of term structure dynamics. For example, while the level factor can explain a large fraction (between 80-90%) of the variations of the LIBOR rates, the slope factor still has significant explanatory power of interest rate variations. Moreover, there is overwhelming evidence that the volatility of interest rates are stochastic, and these stochastic volatility factors are unspanned in the sense that they can not be fully explained by the yield curve factors such as the level and slope factors.

One important innovation of our study is that we allow the volatility of $L_k\left(t\right)$ to be stochastic and the conditional density of $L_k\left(T_k\right)$ to depend on not only the level, but also the slope and volatility factors of the LIBOR rates. That is, we assume that

$$
p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)|\mathcal{F}_t\right)=p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)|L_k\left(t\right),Z\left(t\right)\right),
$$

where $Z\left(t\right)=\{s(t),v\left(t\right)\}$, $s(t)$ is the slope factor of the LIBOR forward curve, and $v\left(t\right)$ is the common volatility factor of all the LIBOR rates. Specifically, $v\left(t\right)$ is the first principle component of the filtered instantaneous volatilities of the LIBOR rates across all maturities via an EGARCH model.

Under this generalization, the above analysis can be modified as

$$
Caplet\left(L_k\left(t\right),Z\left(t\right),\tau_k,X\right)=\delta D_{k+1}\left(t\right)C\left(L_k\left(t\right),Z\left(t\right),\tau_k,X\right),
$$

and

$$
\begin{aligned}
C\left(L_k\left(t\right),Z\left(t\right),\tau_k,X\right) &= E_t^{\mathbb{Q}^{k+1}}\left[\left(L_k\left(T_k\right)-X\right)^+\right] \\
&= \int_X^\infty \left(y-X\right)p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)=y|L_k\left(t\right),Z\left(t\right)\right)dy.
\end{aligned}
$$

And the conditional density of $L_k\left(T_k\right)$ under the forward measure $\mathbb{Q}^{k+1}$ is given by

$$
p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)|L_k\left(t\right),Z\left(t\right)\right)=\frac{\partial^2 C\left(L_k\left(t\right),Z\left(t\right),\tau_k,X\right)}{\partial X^2}|_{X=L_k(T_k)}.
$$

Similar to many existing studies, to reduce the dimensionality of the problem, we assume that the caplet price is homogeneous of degree 1 in the current LIBOR rate. That is,

$$\begin{aligned} C\left(L_k\left(t\right), Z\left(t\right), \tau_k, X\right) &= L_k\left(t\right) C\left(1, Z\left(t\right), \tau_k, M_k(t)\right) \\ &= L_k\left(t\right) C_M\left(M_k(t), Z\left(t\right), \tau_k\right), \end{aligned}$$

where the moneyness of the caplet $M_k(t) = X/L_k\left(t\right)$. Hence, we estimate the forward density of $L_k\left(T_k\right)/L_k\left(t\right)$ as the second derivative of the price function C_M with respect to M :

$$p^{\mathbb{Q}^{k+1}}\left(\frac{L_k\left(T_k\right)}{L_k\left(t\right)}|Z\left(t\right)\right) = \frac{\partial^2 C_M\left(M_k(t), Z\left(t\right), \tau_k\right)}{\partial M^2}|_{M=L_k(T_k)/L_k(t)}.$$

2. Nonparametric Estimation of Forward and State-Price Densities

In this section, we discuss nonparametric estimation of the forward densities and the SPDs using interest rate cap prices. We first provide a brief introduction to the local polynomial approach of Aït-Sahalia and Duarte (2003).[4] Then we discuss how to extend the method of Aït-Sahalia and Duarte (2003) to a multivariate setting to estimate the forward densities conditional on the level, slope, and volatility factors of the LIBOR rates. Finally, we discuss how to combine the forward and physical densities of the LIBOR rates to estimate the SPDs at different maturities.

2.1. A Brief Review of Local Polynomial Estimation

Most studies in the existing literature typically estimate the option pricing formula $C_M\left(M_k(t), Z\left(t\right), \tau_k\right)$ nonparametrically, and then differentiates it twice with respect to M to obtain $\partial^2 C/\partial M^2$ and $p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)|L_k\left(t\right), Z\left(t\right)\right)$. Nonparametric estimation of the derivatives of a regression function, however, requires much more data than estimating the regression function itself. The increase of the dimensionality of the problem due to the conditioning variables $Z\left(t\right)$ further worsens the problem. Therefore, for a given sample size, the choice of the estimation method is crucial.

In our analysis, we adopt and extend the constrained local polynomial method of Aït-Sahalia and Duarte (2003). Compared to the traditional kernel regression method, the local polynomial approach has several important advantages. First, it has been well-established in the statistics and econometrics literature that the local polynomial approach is superior to the kernel method in estimating the derivatives of nonlinear functions (See, for example, Fan and Gijbels (1996)). Second, the constrained local polynomial approach of Aït-Sahalia and Duarte (2003) guarantees that the nonparametric option pricing function $\hat{C}_M\left(\cdot\right)$ satisfies the necessary theoretical restrictions.

[4] The discussion here relies heavily on Aït-Sahalia and Duarte (2003), which contains much more detailed descriptions of the local polynomial approach.

For example, to guarantee the absence of the arbitrage across moneyness and the positivity of the density function, we must have $-1 \leq \partial\hat{C}/\partial M \leq 0$ and $\partial^2\hat{C}/\partial M^2 \geq 0$, which we refer to as the shape restrictions on $\hat{C}_M(\cdot)$. Finally, Aït-Sahalia and Duarte (2003) show that the local polynomial estimation coupled with constrained least square regression also has better finite sample performances than existing models. Below we provide a brief review of the local polynomial approach.

Suppose we have observations $\{(y_i, x_i)\}_{i=1}^n$ generated from the following relation

$$y = f(x) + \epsilon,$$

where $f(x)$ is an unknown nonlinear function and ϵ is a zero-mean error term. Suppose the $(p+1)^{\text{th}}$ derivative of $f(\cdot)$ at x exists. Then a Taylor expansion gives us an approximation of the unknown function $f(\cdot)$ in a neighborhood of x

$$\begin{aligned} f(z) &\approx f(x) + f'(x)(z-x) + ... + \frac{f^{(p)}(x)}{p!}(z-x)^p \\ &= \sum_{k=0}^{p} \beta_{k,p}(x) \times (z-x)^k, \end{aligned}$$

where $\beta_{k,p}(x) = f^{(k)}(x)/k!$ and $f^{(k)}(x) = \frac{\partial f^k(z)}{\partial z^k}|_{z=x}$.

This representation of $f(\cdot)$ suggests that we can model $f(z)$ around x by a polynomial in z, and to use the regression of $f(z)$ on powers of $(z-x)$ to estimate the coefficients $\beta_{k,p}$. To insure the local nature of the representation, we weight the observations by a kernel $K_h(x_i - x) = K((x_i - x)/h)/h$, where h is a bandwidth. Then the estimates of the coefficients $\hat{\beta}_{k,p}(x)$ are the minimizers of

$$\sum_{i=1}^{n} \left\{ y_i - \sum_{k=0}^{p} \beta_{k,p}(x) \times (x_i - x)^k \right\}^2 K_h(x_i - x).$$

At each fixed point x, this is a generalized least squares regression of the y_is on the powers of $(x_i - x)$s with diagonal weight matrix formed by the weights $K_h(x_i - x)$. This regression is "local" in the sense that the regression coefficients in equation are only valid in a neighborhood of each point x.

The optimal way to estimate $f^{(k)}(x)$ based on asymptotics is to choose $p = k+1$ and use the estimator

$$\hat{f}^{(k)}(x) = \hat{f}_p^{(k)}(x) = k!\hat{\beta}_{k,p}(x).$$

For example, a locally linear regression serves to estimate the regression function $\hat{f}_1^{(0)}(x)$, a locally quadratic regression for the first derivative $\hat{f}_2^{(1)}(x)$, and a locally cubic regression for the second derivative $\hat{f}_3^{(2)}(x)$.

Aït-Sahalia and Duarte (2003) show that in small samples the asymptotically optimal solution may not have the best performance. Instead, they consider the following alternative approach, which has superior finite sample performances than the asymptotic approach. Specifically, they estimate $f(\cdot)$ using a locally linear regression

$$\hat{f}(z) = \hat{\beta}_{0,1}(x) + \hat{\beta}_{1,1}(x)(z - x).$$

Hence, the regression function, the first and the second derivatives are estimated as, respectively,

$$\hat{f}(x) = \hat{\beta}_{0,1}(x), \ \hat{f}^{(1)}(x) = \hat{\beta}_{1,1}(x), \text{ and } \hat{f}^{(2)}(x) = \frac{\partial \hat{\beta}_{1,1}(x)}{\partial x} = \hat{\beta}'_{1,1}(x).$$

2.2. A Multivariate Extension of the Constrained Local Linear Estimator

In this section, we extend the constrained local polynomial method of Aït-Sahalia and Duarte (2003) to a multivariate setting to incorporate the conditioning variables $Z(t)$ in our analysis. The main objective of Aït-Sahalia and Duarte (2003) is to estimate risk-neutral densities using a small sample of option data, typically one day's observations. This means that they essentially estimate the option price as an univariate function of the strike price. In our analysis, we would like to estimate the caplet price as a function its moneyness conditional on the slope and volatility factors of the LIBOR rates. The basic idea behind our approach is that we group observations on different dates that share similar values of the conditioning variables, and then within each group we solve an univariate problem similar to that of Aït-Sahalia and Duarte (2003).

For ease of exposition, we describe the extended method in a bivariate setting, although the description can be easily generalized to higher dimensions. Suppose we have a set of n observations $y_1, y_2, ..., y_n$ and their corresponding explanatory variables $(x_{11}, x_{12}), (x_{21}, x_{22}), ..., (x_{n1}, x_{n2})$. Throughout the paper, we assume that the observations are ordered by the first explanatory variable, i.e., $x_{i1} \geq x_{j1}$ if $i > j$, $1 \leq i, j \leq n$. Suppose we want to estimate the y_i's as a function of x_{i1}'s for a fixed x_2 subject to the necessary shape restrictions. Define

$$D\left(x_2; \bar{h}\right) \triangleq \left\{i | x_{i2} \in \left[x_2 - \bar{h}, x_2 + \bar{h}\right], \text{ for } 1 \leq i \leq n\right\},$$

where $\bar{h} > 0$. Basically, $D\left(x_2; \bar{h}\right)$ contains the subsample of observations whose second explanatory variables are grouped around x_2.

Aït-Sahalia and Duarte (2003) point out that nonparametric estimates of $\hat{C}_M$ using the original data are not guaranteed to be arbitrage-free in finite samples. To address this problem, they first filter the data by a solving the following constrained optimization problem:

$$\min_{m \in R^{\rho}} \sum_{i \in D\left(x_2; \bar{h}\right)} (m_i - y_i)^2$$

subject to the slope and convexity constraints:

$$-1 \leq \frac{m_{i+1} - m_i}{x_{i+1,1} - x_{i1}} \leq 0 \text{ for all } i = 1, ..., d-1,$$
$$\frac{m_{i+2} - m_{i+1}}{x_{i+2,1} - x_{i+1,1}} \geq \frac{m_{i+1} - m_i}{x_{i+1,1} - x_{i,1}} \text{ for all } i = 1, ..., d-2,$$

where ρ is the number of elements in $D\left(x_2; \bar{h}\right)$. Note that the solution m depends on the fixed value x_2 and the window size $\bar{h}$. This means that the filtering has to be done for each grid point of x_2 for nonparametric estimation. The basic restriction on $\bar{h}$ is that it should be no smaller than the bandwidth used in the nonparametric smoothing. Aït-Sahalia and Duarte (2003) provide a fast computational algorithm for solving the above constrained optimization problem. The filtered data m closely resembles the original data y and satisfies the shape restrictions imposed by the theory. It is important to note that the filtering is done after obvious data errors have been removed.

To estimate of the option pricing function $\widehat{m}(x_1, x_2)$, and its first and second partial derivatives, $\dfrac{\partial \widehat{m}(x_1, x_2)}{\partial x_1}$ and $\dfrac{\partial \widehat{m}^2(x_1, x_2)}{\partial x_1^2}$, we minimize the following weighted sum of squared errors

$$\sum_{i=1}^{n} \{m_i - \beta_0(x_1, x_2) - \beta_1(x_1, x_2) \times (x_{1i} - x_1)\}^2 K_h(x_{i1} - x_1, x_{i2} - x_2),$$

where $K_h(x_{i1} - x_1, x_{i2} - x_2)$ is the joint kernel function. We have the following proposition which extends the analysis of Aït-Sahalia and Duarte (2003) to a bivariate setting.

Proposition 1 *Consider a set of n observations of the dependent variables, $y_1, y_2, ..., y_n$ and the corresponding bivariate independent variables, $(x_{11}, x_{12}), (x_{21}, x_{22}), ..., (x_{n1}, x_{n2})$. Without loss of generality, let $x_{i1} \geq x_{j1}$ if $i > j$, $1 \leq i, j \leq n$. For any given pair (x_{i1}, x_{i2}), the estimators $\dfrac{\partial \widehat{m}(x_1, x_2)}{\partial x_1}$ and $\dfrac{\partial \widehat{m}^2(x_1, x_2)}{\partial x_1^2}$ satisfy the required constraints in sample: $-1 \leq \dfrac{\partial \widehat{m}(x_1, x_2)}{\partial x_1} \leq 0$ and $\dfrac{\partial \widehat{m}^2(x_1, x_2)}{\partial x_1^2} \geq 0$, provided that*

1. *The transformed data $m_i(x_1, x_2)$, $i = 1, 2, ..., d$, are obtained through the constrained least squares algorithm based on the original data in $D\left(x_2; \bar{h}\right)$;*

2. *The joint kernel function can be written as a product of two univariate kernel functions*

$$K_h(x_{i1} - x_1, x_{i2} - x_2) = K_{1,h}(x_{i1} - x_1) K_{2,h}(x_{i2} - x_2);$$

3. *The kernel function* $K_{1,h}(\cdot)$ *is log-concave;*

4. *The bandwidth for the kernel function* $K_{2,h}(\cdot)$, h_2^*, *satisfies* $h_2^* \leq \bar{h}$, *and* $K_{2,h}(z) = 0$ *for* $|z| > \bar{h}$.

Proof. For a fixed (x_1, x_2), we first augment the sample to the size of the original observations by including the unfiltered observations that are not in $D(x_2; \bar{h})$. Next we define

$$\begin{aligned} k_{l,i} &= K_{l,h}(x_{il} - x_l), l = 1, 2, \\ k_{i,j} &= (x_{i1} - x_1)^2 k_{1,i} k_{1,j} k_{2,i} k_{2,j}, \\ M_{i,j} &= (m_i - m_j) / (x_{i1} - x_1). \end{aligned}$$

The estimator of the first-order derivative equals

$$\frac{\partial \widehat{m}(x_1, x_2)}{\partial x_1} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} M_{i,j} k_{i,j}}{\sum_{k=1}^{n-1} \sum_{l=k+1}^{n} k_{k,l}}.$$

Since $k_{2,i,j} = 0$ outside the set $D(x_2; \bar{h})$, we can reduce the sum on this set only, i.e.,

$$\begin{aligned} \frac{\partial \widehat{m}(x_1, x_2)}{\partial x_1} &= \frac{\sum_{i=1}^{d-1} \sum_{j=i+1}^{d} M_{i,j} k_{i,j}}{\sum_{k=1}^{d-1} \sum_{l=k+1}^{d} k_{k,l}}, \\ \frac{\partial \widehat{m}^2(x_1, x_2)}{\partial x_1^2} &= \frac{\left(\sum_{i=1}^{d-1} \sum_{j=i+1}^{d} M_{i,j} k'_{i,j}\right)\left(\sum_{k=1}^{d-1} \sum_{l=k+1}^{d} k_{k,l}\right) - \left(\sum_{i=1}^{d-1} \sum_{j=i+1}^{d} M_{i,j} k_{i,j}\right)\left(\sum_{k=1}^{d-1} \sum_{l=k+1}^{d} k'_{k,l}\right)}{\left(\sum_{k=1}^{d-1} \sum_{l=k+1}^{d} k_{k,l}\right)^2} \end{aligned}$$

Aït-Sahalia and Duarte (2003) show $-1 \leq \dfrac{\partial \widehat{m}(x_1, x_2)}{\partial x_1} \leq 0$ and $\dfrac{\partial \widehat{m}^2(x_1, x_2)}{\partial x_1^2} \geq 0$ for the univariate case. It can be easily shown from the definition that the same result applies here, .

■

This proposition can be easily extended to higher dimensions. The key difference from the univariate case is that for the bivariate case we need to apply the constrained least square procedure at each grid point of x_2. The ideal choice of $\bar{h}$ should be h_2^*. However, the choice of h_2^* often depends on the filtered observations, which depend on the choice of $\bar{h}$. Therefore, in practice a two-stage procedure should be used in estimation. We can first conduct an unconstrained estimation to obtain the bandwidth h_2^*, which will be used later as the filtering window width in the second-stage constrained estimation.

In our implementation, x_1 represents the moneyness of the option, while x_2 and x_3 represent the slope and volatility factors, respectively. Following Proposition 1, we choose Epanechnikov kernels for the slope and volatility factors and the Gaussian kernel for the moneyness.

While the choices of kernels are relatively straightforward, the choices of bandwidths are much more complicated. The data are relatively evenly distributed along the moneyness dimension. Therefore, a global bandwidth (constant across moneyness) is acceptable in terms of the mean squared error (MSE) criterion. On the other hand, the observations of the slope and volatility factors are rather unevenly distributed. This makes a simple global bandwidth an inappropriate choice, because a global bandwidth will likely to be too large and oversmooth the regions with dense observations and too small and undersmooth the regions with sparse observations. To address this problem, we perform a monotonic transformation of the slope and volatility data via their empirical distribution functions and apply a global bandwidth to the transformed data. We first start with a preliminary bandwidth selection from the following relation

$$h_j = c_j \sigma(x_j) n^{-1/(4+d)}, j = 1, 2, 3,$$

where c_j is a constant that is close to one, $\sigma(x_j)$ is the sample standard deviation of x_j, and n is the sample size. Then we fine tune c_j to minimize the MSE via cross validation procedures.

The asymptotic and finite sample distributions of the constrained local polynomial estimator are not known analytically. We obtain the finite sample distribution of the estimator using bootstrap. In our estimation, the sample is divided into cells around certain slope and volatility grids. The estimator for a particular cell does not depend on the data outside that cell. We generate bootstrap samples by randomly drawing observations with replacements from a cell.[5] Conducting the estimation on all bootstrap samples, we obtain the finite sample distribution of the estimator.

2.3. Nonparametric Estimation of State-Price Densities

In this section, we discuss nonparametric estimation of the SPDs by combining the forward and physical densities of the LIBOR rates. Given a SPD function H, the price of the caplet can be calculated as

$$Caplet\left(L_k\left(t\right), Z(t), \tau_k, X\right) = \delta E_t^{\mathbb{P}}\left[H \cdot \left(L_k\left(T_k\right) - X\right)^+\right],$$

where the expectation is taken under the physical measure. In general, H depends on different

[5] For identical observations, our estimator essentially keeps one among them.

economic factors, and it is impossible to estimate it using interest rate caps above. Given the available data, all we can estimate is the projection of H onto the future spot rate $L_k(T_k)$. Define

$$H_k\left(L_k(T_k); L_k(t), Z(t)\right) = E^{\mathbb{P}}\left[H | L_k(T_k); L_k(t), Z(t)\right],$$

then by iterated expectation,

$$\begin{aligned}
Caplet\left(L_k(t), Z(t), \tau_k, X\right) &= \delta E_t^{\mathbb{P}}\left[H_k\left(L_k(T_k); L_k(t), Z(t)\right)\left(L_k(T_k) - X\right)^+\right] \\
&= \delta \int_X^\infty H_k(y)(y - X)\, p^{\mathbb{P}}\left(L_k(T_k) = y | L_k(t), Z(t)\right) dy.
\end{aligned}$$

Comparing the above equation with the forward measure approach, we have

$$H_k\left(L_k(T_k); L_k(t), Z(t)\right) = D_{k+1}(t)\frac{p^{\mathbb{Q}^{k+1}}\left(L_k(T_k) | L_k(t), Z(t)\right)}{p^{\mathbb{P}}\left(L_k(T_k) | L_k(t), Z(t)\right)}.$$

Therefore, by combining the densities of $L_k(T_k)$ under $\mathbb{Q}^{k+1}$ and $\mathbb{P}$, we can estimate the projection of H onto $L_k(T_k)$. In our empirical analysis, we focus on the density ratio, which is defined as

$$\widetilde{H_k}\left(L_k(T_k); L_k(t), Z(t)\right) = \frac{1}{D_{k+1}(t)} H_k\left(L_k(T_k); L_k(t), Z(t)\right).$$

The density ratio equals to the well-known Randon-Nikodym derivative in the option pricing literature. Another interpretation of $\widetilde{H_k}$ is the intertemporal rate of substitution of consumption of the representative investor. For an economy where the representative agent has a time-additive utility function, we have

$$\widetilde{H_k}\left(L_k(T_k)\right) = \frac{1}{D_{T_{k+1}}} E_t^{\mathbb{P}}\left[\frac{U'\left(c_{T_{k+1}}\right)}{U'\left(c_t\right)} | L_k(T_k)\right]$$

where $U'(\cdot)$ is the marginal utility of consumption, and $c_{T_{k+1}}$ and c_t are optimal consumptions at T_{k+1} and t, respectively. Therefore, we can estimate the dependence of future consumption $c_{T_{k+1}}$ on future spot interest rate $L_k(T_k)$ using a model-free approach based on the physical and forward densities of $L_{T_k}(T_k)$.[6]

The SPDs contain rich information on how risks are priced in financial markets. While Aït-Sahalia and Lo (2000), Jackwerth (2000) and others estimate SPDs using index options (i.e., the projection of H onto index returns), our analysis based on interest rate options documents the dependence of the SPDs on term structure factors. For example, economic theory suggests that real interest rates have both a wealth effect where higher rates lead to higher wealth level and therefore higher

[6] Here we ignore the small horizon difference between the consumption and spot interest rate and treat them as contemporaneous. The time difference is the tenor of the interest rate caps, which is three month.

current consumptions, and a substitution effect where higher rate make current consumptions more expensive relative to future consumptions and therefore lower current consumptions. The interest rate caps we consider allow us to examine the relative importance of these two effects at different time horizons.

While we estimate $p^{\mathbb{Q}^{k+1}}\left(L_k\left(T_k\right)|L_k\left(t\right),Z\left(t\right)\right)$ from caplet prices, we estimate $p^{\mathbb{P}}\left(L_k\left(T_k\right)|L_k\left(t\right),Z\left(t\right)\right)$ using the underlying LIBOR rates based on the kernel method of Aït-Sahalia and Lo (2000). Let $L\left(t,T\right)$ be the time-t three-month LIBOR forward rate with a maturity of T, and $L\left(T,T\right)$ be the three-month LIBOR spot rate at T. Suppose we have the following time series observations $\left\{L\left(t_i,T_i\right),L\left(T_i,T_i\right)\right\}_{i=1}^{n}$, where $T_i-t_i=T-t$. We define the log-return of the LIBOR rates as

$$u_{t_i,T_i}=\log\left(L(T_i,T_i)\right)-\log\left(L(t_i,T_i)\right).$$

We use Z_{t_i} to denote the conditioning variables at t_i. The joint distribution of the log-return and the state variables under the physical measure, $p^{\mathbb{P}}_{u,Z}\left(u_{t,T},Z_t\right)$, can be estimated as

$$\widehat{p}^{\mathbb{P}}_{u,Z}\left(u,z\right)=\frac{1}{n}\sum_{i=1}^{n}K_{h_u}\left(\frac{u_{t_i,T_i}-u}{h_u}\right)K_{h_z}\left(\frac{Z_{t_i}-z}{h_z}\right),$$

where $K_h\left(\cdot\right)=K(\cdot)/h$ and $K(\cdot)$ is a kernel function. In our estimation, we use the Gaussian kernel and assume the bivariate kernel is a product of one-dimensional kernels. The marginal density of the spot volatility under the physical measure can be estimated as

$$\widehat{p}^{\mathbb{P}}_{Z}\left(z\right)=\frac{1}{n}\sum_{i=1}^{n}K_{h_z}\left(\frac{Z_{t_i}-z}{h_z}\right),$$

and the density of the log-return conditional on the spot volatility is given by

$$\widehat{p}^{\mathbb{P}}_{u|Z}\left(u|Z\right)=\frac{\widehat{p}^{\mathbb{P}}_{u,Z}\left(u,z\right)}{\widehat{p}^{\mathbb{P}}_{Z}\left(z\right)}.$$

The bandwidths are chosen according to the following relation

$$h_j=c_j\sigma_j n^{-1/(4+d)},j=u,z$$

where σ_j is the unconditional standard deviation of the data, c_j is a constant, and d represents the dimension of the problem ($d=2$ for the joint distribution and $d=1$ for the marginal distribution). Since the joint density estimator converges slower than the marginal one, the speed of convergence for the conditional density estimator is the same as that of the joint distribution, i.e. $O(n^{-2/(4+d)})$. In practice, similar to Aït-Sahalia and Lo (1998), we let the bandwidth converge to zero slightly faster than the MSE-optimal rate stated above in order to eliminate the asymptotic bias. This makes the asymptotic variance slightly larger and the MSE convergence speed slightly slower. The asymptotic distribution for the kernel estimator based on this bandwidth selection method is provided below.

$$n^{1/2}\prod_{j=1}^{d}h_j^{1/2}\left[\widehat{p}_{u|Z}^{\mathbb{P}}(u|Z)-p_{u|Z}^{\mathbb{P}}(u|Z)\right]\xrightarrow{d}Normal\left(0,\frac{p_{u|Z}^{\mathbb{P}}(u|Z)}{p_Z^{\mathbb{P}}(z)}\left(\int K(z)^2dz\right)^d\right).$$

The estimator of the pricing kernel equals

$$\widehat{H_k}(u|Z)=\frac{\widehat{p}_{u|Z}^{\mathbb{Q}^{k+1}}(u|Z)}{\widehat{p}_{u|Z}^{\mathbb{P}}(u|Z)}.$$

Note that $\widehat{H_k}(u|Z)$ is the projection of H onto the log ratio between the future spot rate and the current forward rate. Based on similar arguments in Aït-Sahalia and Lo (2000), we obtain the distribution of the pricing kernel estimator. Intuitively, the estimator of the physical density $\widehat{p}_{u|Z}^{\mathbb{P}}(u|Z)$ converges faster than that of the forward density $\widehat{p}_{u|Z}^{\mathbb{Q}^{k+1}}(u|Z)$, since the latter involves estimation of second-order derivatives with the same dimensionality. So, the asymptotic distribution of $\frac{\widehat{p}_{u|Z}^{\mathbb{Q}^{k+1}}(u|Z)}{\widehat{p}_{u|Z}^{\mathbb{P}}(u|Z)}$ is identical to that of $\frac{\widehat{p}_{u|Z}^{\mathbb{Q}^{k+1}}(u|Z)}{p_{u|Z}^{\mathbb{P}}(u|Z)}$, where we replace the true physical density with the estimate. Since we do not have the asymptotic distribution of the estimator $\widehat{p}_{u|Z}^{\mathbb{Q}^{k+1}}(u|Z)$, we obtain finite sample distributions using bootstrap, which tends to give a larger confidence interval.

2. The Data

In this section, we introduce the two datasets used in our empirical analysis. The first one includes daily LIBOR and swap rates, and the second one includes daily prices of interest rate caps with different strike prices and maturities.

We obtain daily three-month LIBOR rates with maturities of three, six, and twelve months, as well as daily two-, three-, four-, five-, seven- and ten-year swap rates between August 13, 1990 and December 8, 2005 from Datastream. We bootstrap the swap rates to obtain daily three-month LIBOR forward rates with maturities beyond one year.

To capture the time-varying volatility of the LIBOR rates, we filter out spot volatilities of three-month LIBOR forward rates at different maturities using an EGARCH model. Figure 1 provides time series plots of the three-month LIBOR forward rates and their associated filtered spot volatilities at maturities of one, two, three, five, seven, and ten years. We observe strong negative correlations between the levels and the filtered spot volatilities of the LIBOR rates, especially toward the last few years of the sample period. For example, we find significant rises in the spot volatilities associated with dramatic declines in the LIBOR rates for most of the maturities between 2000 and 20005. The negative correlation between the LIBOR rates and the spot volatilities are different from those found in previous studies. For example, Andersen and Lund (1997), Ball and Torous (1999), and Chen and Scott (2001) show that correlations between spot rates and stochastic volatility are close to zero. We obtain different results, maybe because we consider different model specifications or different sample periods.

We also obtain daily prices of interest rate caps between August 1, 2000 and July 26, 2004 from SwapPX. 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, established in the early 1990s by the major U.S. fixed-income dealers in a response to regulators' demands for increased transparency in the fixed-income markets, 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 trades between dealers and large customers. The data are collected every day the market is open between 3:30 and 4 p.m. To our knowledge, our data set is the most comprehensive available for caps written on dollar LIBOR rates (see Gupta and Subrahmanyam (2005) and Deuskar, Gupta, and Subrahmanyam (2003) for the only other studies that we are aware of in this area).

One advantage of our data is that we observe prices of caps over a wide range of strike prices and maturities. For example, every day for each maturity, there are 10 different strike prices: 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 8.0, 9.0, and 10.0% between August 1, 2000 and October 17, 2001; 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0 and 5.5% between October 18 and November 1, 2001; 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, and 7.0% between November 2, 2001 and July 15, 2002; 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, and 6.5% between July 16, 2002 and April 14, 2003; and 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5 and 6.0% between April 15, 2003 and July 26, 2004. Moreover, caps have 15 different maturities throughout the whole sample period: 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.

Our analysis uses prices of caplets, although we only observe cap prices. To obtain caplet prices, we consider the difference between the prices of caps with the same strike and adjacent maturities, which we refer to as *difference caps*. A difference cap includes a few caplets between two neighboring maturities with the same strike. For example, 1.5-year difference caps with a specific strike represent the sum of the 1.25-year and 1.5-year caplets with the same strike. We assume all individual caplets of a difference cap share the same Black implied volatility and calculate the price of each individual caplet using the Black formula.

Due to daily changes in LIBOR rates, caplets realize different moneyness (defined as the ratio between the strike price and the LIBOR forward rate underlying the caplet) each day. Therefore, throughout our analysis, we focus on the prices of caplets at given fixed moneyness. That is, each day we interpolate caplet 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 while 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. We also eliminate observations with zero prices, and observations that violate either monotonicity or convexity with respect to the strikes.

Figure 2.A plots the average Black-implied volatilities of caplets across moneyness and maturity, while Figure 2.B plots the average implied volatilities of ATM caplets over the whole sample period. Consistent with the existing literature, the implied volatilities of caplets with a moneyness between 0.8 and 1.2 have a humped shape with a peak at around a maturity of two years. However, the implied volatilities of all other caplets decline with maturity. There is also a pronounced volatility skew for caplets at all maturities, with the skew being stronger for short-term caplets. The pattern is similar to that of equity options: In-the-money (ITM) caplets have higher implied volatilities than do out-of-the-money (OTM) caplets. The implied volatilities of the very short-term caplets are more like a symmetric smile than a skew. Figures 2.C, 2.D, and 2.E, respectively, plot the time series of Black-implied volatilities for the 2.5-, 5-, and 8-year caplets across moneyness, while Figure 2.F plots the time series of ATM implied volatilities of the three contracts. It is clear that the implied volatilities are time varying and they have increased dramatically (especially for the 2.5-year caplets) over our sample period. Due to changing interest rates and strike prices, there are more ITM caplets in the later part of our sample.

3. Empirical Results

3.1 Nonparametric Forward Densities

The volatility smile or skew documented in interest rate cap markets strongly suggests that the log-normal assumption of the LIBOR market models is violated in the data. Different extensions of the LIBOR market models have been developed to capture the smile. For example, Jarrow, Li, and Zhao (2006) introduce multi-factor Heath, Jarrow, and Morton (1992) models in which the LIBOR rates follow affine jump-diffusions to obtain closed-form cap pricing formula. Instead of making parametric assumptions on the LIBOR rate process, in this section, we estimate the probability densities of the LIBOR rates under forward measures nonparametrically.

Figure 3 provides the semi-nonparametric estimates of the Black-implied volatilities from cap prices as a nonlinear function of the moneyness and the spot volatility of the 1-, 2-, 3-, 5-, 7-, and10-year caplets. We see a clear dependence of the implied volatilities on the moneyness and the spot volatility for the caplets at all maturities. The dependence is especially pronounced for the caplets with shorter maturities. For the one-year caplets, the implied volatilities exhibit a rather symmetric smile when the spot volatility is low; and the implied volatilities of the ATM caplets increase monotonically with the level of the spot volatility. The implied volatilities of the 2-, 3-, and 5-year caplets exhibit a strong volatility skew, with the ITM caplets have higher implied volatilities than the ATM and OTM caplets; the implied volatilities of all caplets also increase with the level of their corresponding spot volatilities. The implied volatilities of the 7- and 10-year caplets exhibit similar behaviors as those of the 2- to 5-year caplets, although the general levels of the implied volatilities are lower and the volatility skews are not as dramatic. This is consistent with the fact that the longer maturity LIBOR rates have lower spot volatilities and the longer maturity caplets have less significant volatility skews.

Figure 4 contains the plots of the nonparametric densities of the log-returns of the LIBOR rates at the 2-, 3-, 4-, 5-, 7- and 10-year maturities for three different levels of the slope and volatility factors. The 95% confidence intervals are obtained through bootstrap. We measure the slope factor in percentage points and the three slope levels we consider represent the flat, average, and steeper than average term structures of the LIBOR rates. The volatility factor is normalized by its sample mean. Hence, if $v(t)$ is smaller (bigger) than 1, then the spot volatility is below (above) the sample mean. We consider volatility levels that are below, at, and above the sample mean. Under the forward measures, the LIBOR rates should be a martingale and the forward densities should have a mean that is close to zero. The expected log-returns of the LIBOR rates are slightly negative due to an adjustment from the Jensen's inequality. One common feature of the forward densities is that they all exhibit a negative skewness, suggesting that large negative moves of the LIBOR rates are more likely under the forward measures. Later results show that the physical densities of the LIBOR rates are not as negatively skewed. Therefore, the negative skewness in the forward densities is mostly due to the negative skewness in the SPDs.

Another important result is that all the forward densities depend significantly on the slope and volatility factors. For example, when the slope of the term structure becomes steeper than average, the forward densities across all maturities become much more dispersed. With a steep term structure, the current spot rate is low and is expected to rise. This coincides with periods when the Fed lowers the short rate to spur economic growth. This result reveals a positive relation between the volatility of future spot rate and the slope of the term structure.

The dependence of the forward densities on the volatility factor is not very transparent when the term structure is very steep. On the other hand, the volatility effect is very significant when the slope is around the average level: when the spot volatility is low, the forward densities are compact with high peaks; as the spot volatility rises, the forward densities become much more dispersed and negatively skewed. This pattern holds for all maturities, although the effect becomes weaker for longer maturities. This result suggests that the volatility process is very persistent because current high spot volatility leads to high future spot volatility. A flat term structure suggests that the short term interest rate can go up, down, or stay the same. Higher spot volatility probably means that the investors are less certain about the directions of future rate changes, and negative skewness suggests that either the investors care more about a downward move or the rate is more likely to go down. When the term structure is relatively flat, the effects of the volatility factor vary across maturities. The 2- and 3-year forward densities are more negatively skewed when the spot volatility is high. The seven- and ten-year forward densities, however, become more dispersed when the spot volatility is either below or above the sample mean.

The nonparametric forward densities provide important insights on term structure modelling. For example, we find that the log-normal assumption underlying the popular LIBOR market models is grossly violated in the data, and the forward densities across all maturities are significantly negatively skewed. Furthermore, we document significant nonlinear dependence of the forward densities on both the slope and volatility factors of the LIBOR rates. These results point out the importance of unspanned stochastic volatility and the challenges in modeling volatility dynamics due to their nonlinear impacts on the forward densities.

3.2 Implied State-Price Densities

In addition to the forward densities, we also estimate the physical densities of the LIBOR rates at different maturities. Combining the forward and physical densities, we provide nonparametric estimates of the SPDs over different horizons.

Figure 5 provides the plots of the physical densities of the LIBOR rates at different maturities at the three levels of the slope and volatility factors. The 95% confidence intervals are calculated based on the asymptotic distribution of the kernel density estimator. The most important result from Figure 5 is that the physical densities are not as negatively skewed and widely dispersed as the forward densities. This suggests that the high dispersion and negative skewness of the forward densities are caused by the SPDs rather than the physical densities.

We still see clear dependence of the physical densities on the slope and volatility factors. Different from the forward densities, the physical densities become more compact and closer to being normally distributed when the term structure is steeper than average. The dispersion of the physical densities is the largest when the slope is at the average level. This result could be due to mean reversion in interest rates. Also different from the forward densities, when the spot volatility is above the average, the physical densities become more compact and symmetric, which suggests a mean reverting volatility process. The 2- and 3-year LIBOR rates are more widely dispersed than the 4- and 5-year LIBOR rates.[7] A comparison between the forward and physical densities shows that the slope and volatility factors are more persistent under the forward measures.

Figure 6 plots the nonparametric estimates of the SPDs projected onto the LIBOR spot rates at four different maturities for the three different levels of the slope and volatility factors. The most important findings from Figure 6 is that the SPDs exhibit a pronounced U-shape as a function of the future LIBOR rates, especially at the 4- and 5-year maturities. This result suggests that the investors attach high values to payoffs received when the interest rates are either extremely high or low. This is consistent with the notion that low interest rates tend to be associated with economic slowdowns or even recessions, while high interest rates tend to be associated with high inflations. Investors with large bond portfolio holdings can hedge their potential losses due to rising interest rates by buying OTM caps. On the other hand, investors with large holdings in mortgage-backed securities can hedge their potential losses due to prepayments resulted from declining interest rates by buying OTM floors.[8] These effects could explain the high state prices at both low and high levels of interest rates. The of the SPDs estimated from interest rate options differ significantly from that estimated using index options, which is typically a declining function of the level of the equity market.

The SPDs at the 4- and 5-year maturities have much more pronounced U-shapes than those at the 2- and 3-year maturities. This is mainly driven by the fact that the physical densities at the 4- and 5-year maturities are more compact than those at the 2- and 3-year maturities, while the forward densities at all maturities have similar shapes. By the law of large numbers and mean-reversion in interest rates, fluctuations in interest rates should be canceled out over longer horizons. However, if interest rates indeed have gone up or down a lot over a longer horizon, it means that the rates

[7] We only report results up to fives because we do not have enough data on log-returns of LIBOR rates for longer maturities.

[8] See Duarte (2006) for excellent discussion on the impacts of mortgage prepayments on interest rate volatility.

probably have consistently gone up or down over the time period, respectively. This further means that extremely high or low interest rates correspond to really bad states of the economy, which lead to the high prices of risks in those states.

The SPDs at the 2- and 3-year maturities depend more significantly on the level and volatility factors. For example, at the 3-year maturity, when the slope is really steep, the left arm of the U-shape is more pronounced. In recessions, the Fed tends to lower rates to spur growth, which leads to an upward sloping yield curve. However, conditioning on a steep yield curve, if the rate keeps going down, then it means that the economy is sliding into recession and the SPDs for those states would go up. On the other hand, with a flat yield curve, the Fed is probably raising interest rate to slow the growth of the economy and inflation. Conditioning on that, if the rate increases a lot, then it means that the Fed has not been successful and the economy is probably suffering from hyper inflation. This explains why the right arm of the U-shape is more pronounced when the yield curve is flat.

The analysis in this section reveals some interesting features of the physical and forward densities of the LIBOR rates as well as the SPDs. They show that both the physical and forward densities depend significantly on the level, slope, and volatility factors of the LIBOR rates. The SPDs show that interest rate options allow us to study investor preferences from a different perspective than index options. Given that each market only contains a subset of information about the price kernel, our analysis shows that we should explore the implications of asset pricing models in different markets.

4. Conclusion

In this paper, we extract the rich information on term structure dynamics contained in the prices of interest rate caps using nonparametric methods. Methodologically, we extend the constrained local polynomial approach of Aït-Sahalia and Duarte (2003) to a multivariate setting and (for the first time) estimate the forward densities of the LIBOR rates and the SPDs conditional on the level, slope, and volatility of the LIBOR rates. The multivariate constrained local polynomial approach has excellent finite sample performances and guarantees that the nonparametric estimates satisfy necessary theoretical restrictions. Empirically, we provide interesting new evidence on the forward and physical densities of the LIBOR rates and the SPDs, which is important for future developments of term structure models. We find that the forward densities of the LIBOR rates deviate significantly from the log-normal distribution and are strongly negatively skewed. The SPDs exhibit a pronounced U-shape as a function of the future LIBOR rates. This suggests that the state prices are high at extremely low and high interest rates, which tend to correspond to economic recessions and high inflations, respectively. Both the forward densities and the SPDs depend significantly on the volatility of the LIBOR rates, suggesting that it is extremely important to incorporate unspanned stochastic volatility into equilibrium term structure models.

REFERENCES

Aït-Sahalia, Yacine, 1996a, Nonparametric pricing of interest rate derivative securities, *Econometrica* 64, 527-560.

Aït-Sahalia, Yacine, 1996b, Testing continuous-time models of the spot interest rate, *Review of Financial Studies* 9, 385-426.

Aït-Sahalia, Yacine, and Andrew Lo, 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance* 53, 499-547.

Aït-Sahalia, Yacine, and Andrew Lo, 2000, Nonparametric risk management and implied risk aversion, *Journal of Econometrics* 94, 9-51.

Andersen, Torben G., and Jesper Lund, 1997, Estimating continuous time stochastic volatility models of the short-term interest rate, *Journal of Econometrics* 77, 343-378.

Ball, Clifford A., and Walter Torous, 1999, The stochastic volatility of short-term interest rates: Some international evidence, *Journal of Finance* 54, 2339-2359.

Beber, Alessandro, and Michael Brandt, 2006, The effects of macroeconomic news on beliefs and preferences: Evidence from the options market, *Journal of Monetary Economics* forthcoming.

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

Black, Fischer, 1976, The pricing of commodity contracts, *Journal of Financial Economics* 3, 167-179.

Brace, Alan, Darius Gatarek, and Marek Musiela, 1997, The market model of interest rate dynamics, *Mathematical Finance* 7, 127-155.

Brenner, Robin J., Richard H. Harjes, and Kenneth F. Kroner, 1996, Another look at alternative models of short-term interest rate, *Journal of Financial and Quantitative Analysis* 31, 85-107.

Chen, Ren-Raw, and Louis Scott, 2001, Stochastic volatility and jumps in interest rates: An empirical analysis, Working paper, Rutgers University.

Collin-Dufresne, Pierre, and Robert S. Goldstein, 2002, Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility, *Journal of Finance* 57, 1685-1729.

Dai, Qiang, and Kenneth J. Singleton, 2003, Term structure dynamics in theory and reality, *Review of Financial Studies* 16, 631-678.

Deuskar, Prachi, Anurag Gupta, and Marti G. Subrahmanyam, 2003, Liquidity effects and volatility smiles in interest rate option markets, Working paper, New York University.

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

Duffie, Darrell, Jun Pan, and Kenneth J. Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343-1376.

Gupta, Anurag, and Marti Subrahmanyam, 2005, Pricing and hedging interest rate options: Evidence from cap-floor markets, *Journal of Banking and Finance* 29, 701-733.

Han, Bing, 2002, Stochastic volatilities and correlations of bond yields, Working paper, Ohio State University.

Heath, David, Robert Jarrow, and Andrew Morton, 1992, Bond pricing and the term structure of interest rates: A new methodology, *Econometrica* 60, 77-105.

Heidari, Massoud, and Liuren Wu, 2003, Are interest rate derivatives spanned by the term structure of interest rates? *Journal of Fixed Income* 13, 75-86.

Li, Haitao, and Feng Zhao, 2006, Unspanned stochastic volatility: Evidence from hedging interest rate derivatives, *Journal of Finance* 61, 341-378.

Litterman, Robert, and Jose Scheinkman, 1991, Common factors affecting bond returns, Journal of Fixed Income 1, 62–74.

Miltersen, Kristian, Klaus Sandmann, and Dieter Sondermann, 1997, Closed-form solutions for term structure derivatives with lognormal interest rates, *Journal of Finance* 52, 409-430.

Piazzesi, Monika, 2003, Affine term structure models, *Handbook of Financial Econometrics,* forthcoming.

Figure 1. Time Series of the Forward Rates and Spot Volatility
This figure plots and the time series of the Libor forward rates (solid) with the spot volatility (dotted) filtered through EGARCH model.

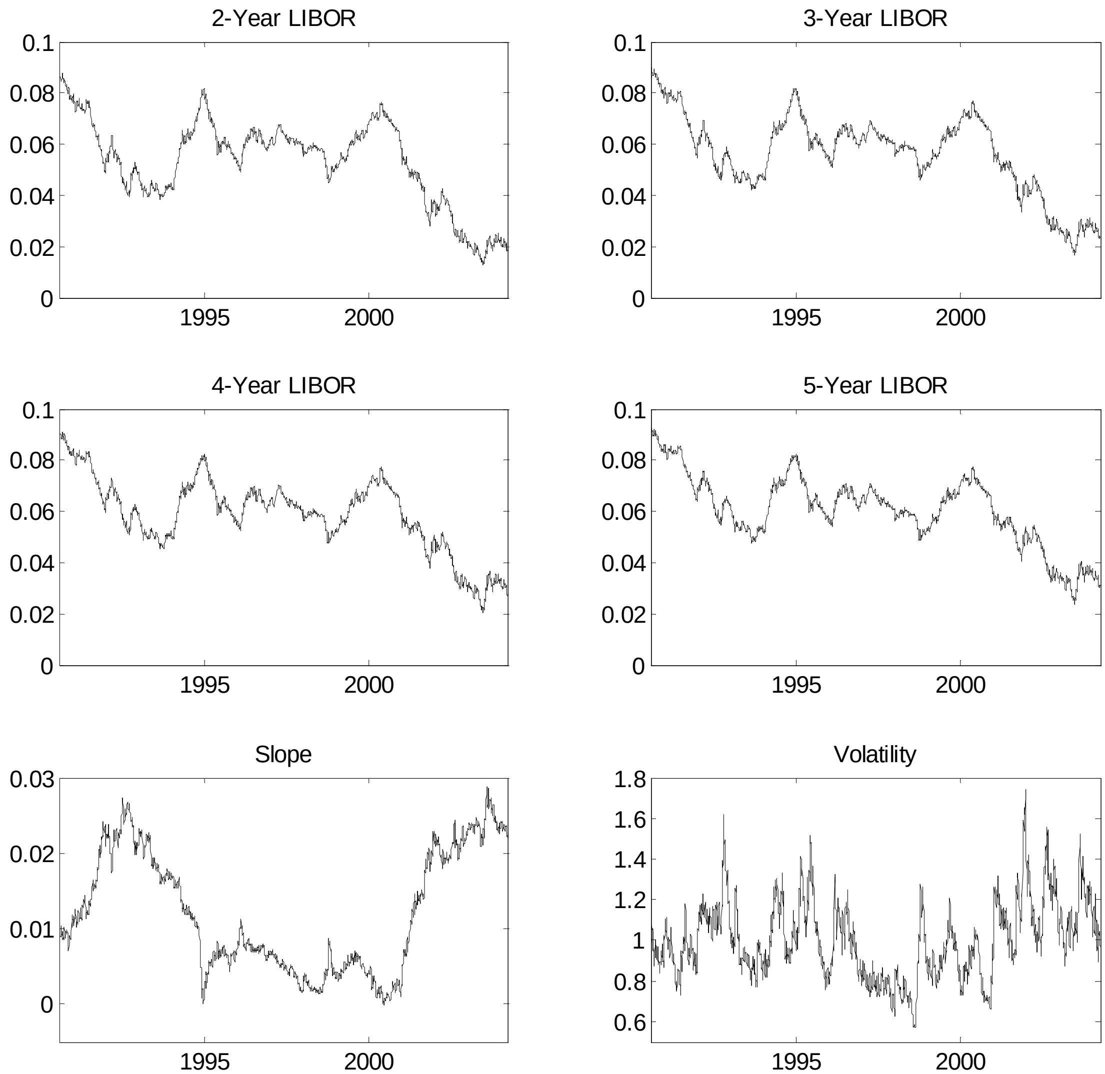
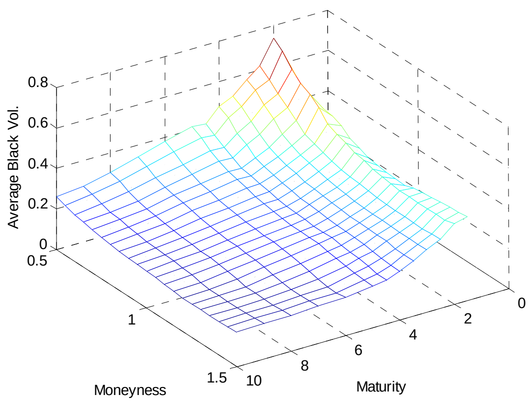
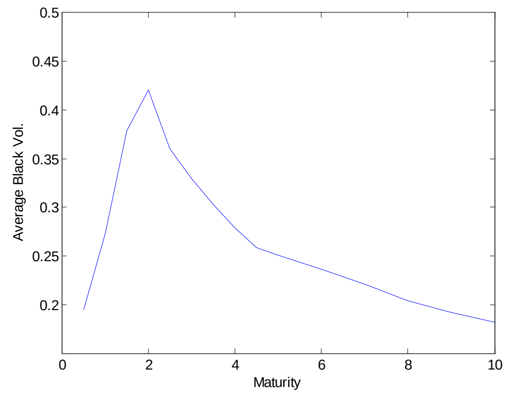


Figure 2: Black-Implied Volatilities for Caplets

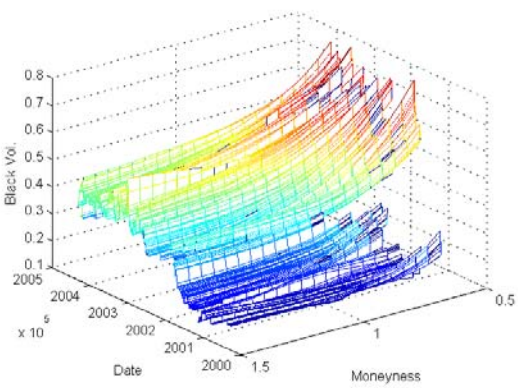
A. Average Black-Implied Volatilities



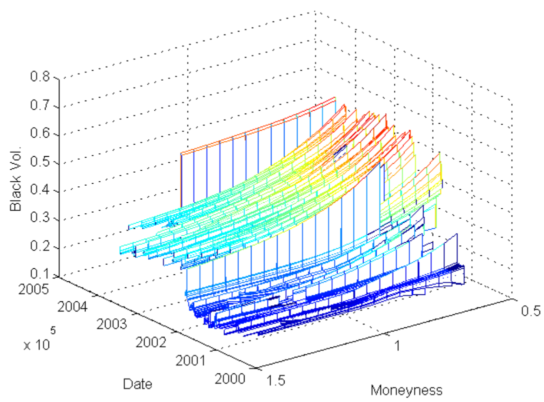
B. ATM Average Black-Implied Volatilities



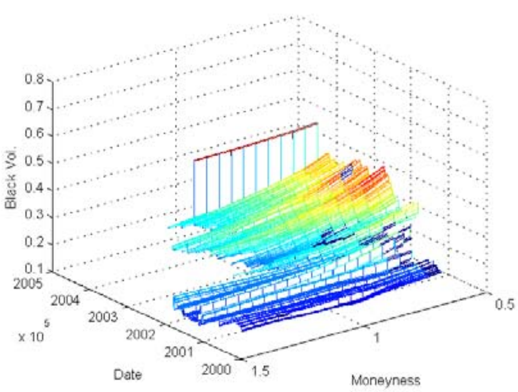
C. 2.5-Year



D. 5-Year



E. 8-Year



F. ATM

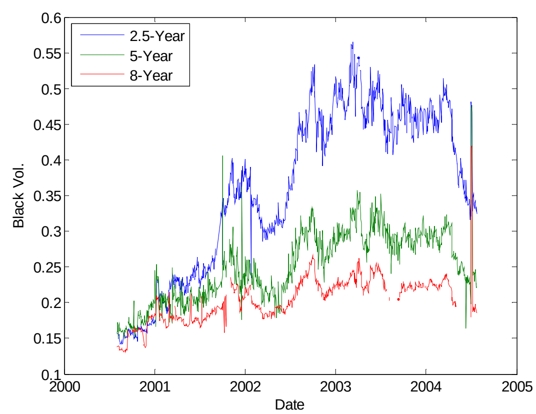
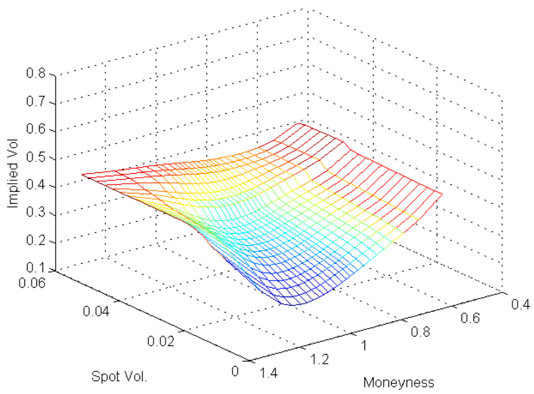


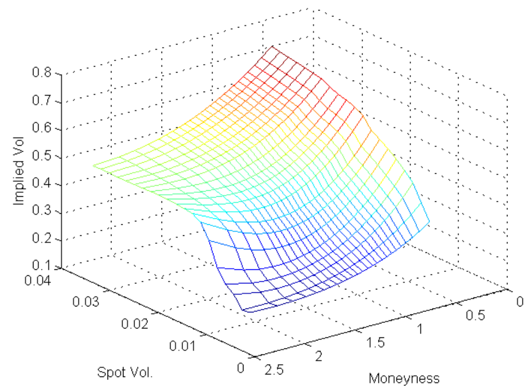
Figure 3: Semi-Nonparametric Estimation of the Implied Volatilities

This figures plot the semi-nonparametric estimates of the implied volatilities, maturing from 1 to 10 years. The estimates are conditional on the spot volatilities, filtered through EGARCH model.

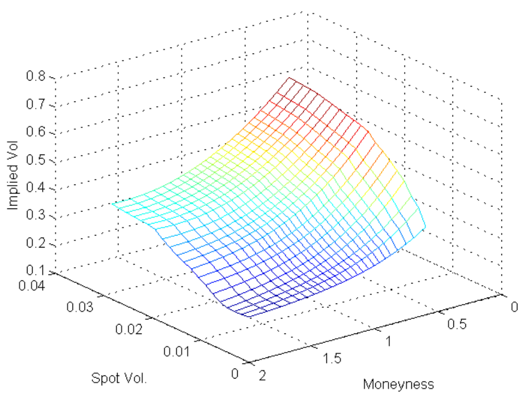
A. 1-Year



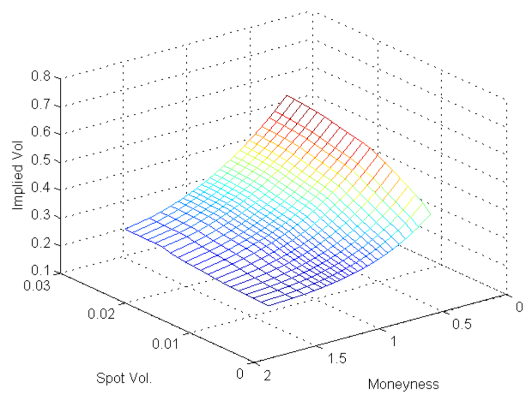
B. 2-Year



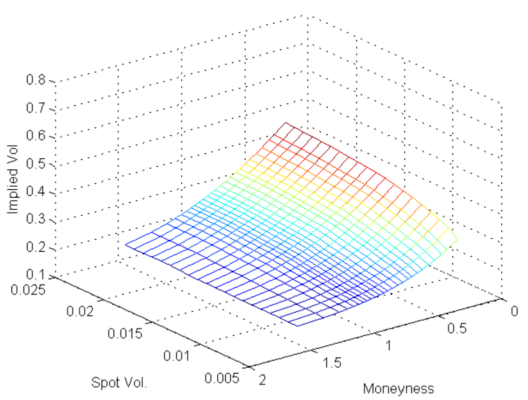
C. 3-Year



D. 5-Year



E. 7-Year



F. 10-Year

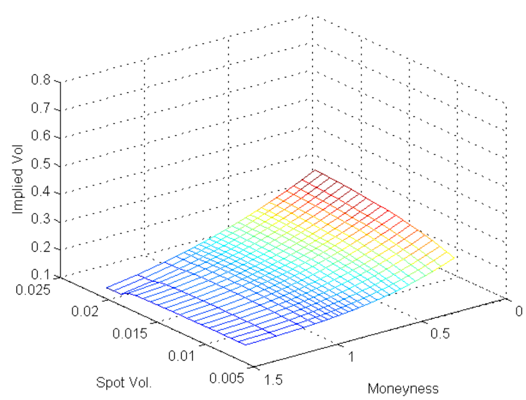
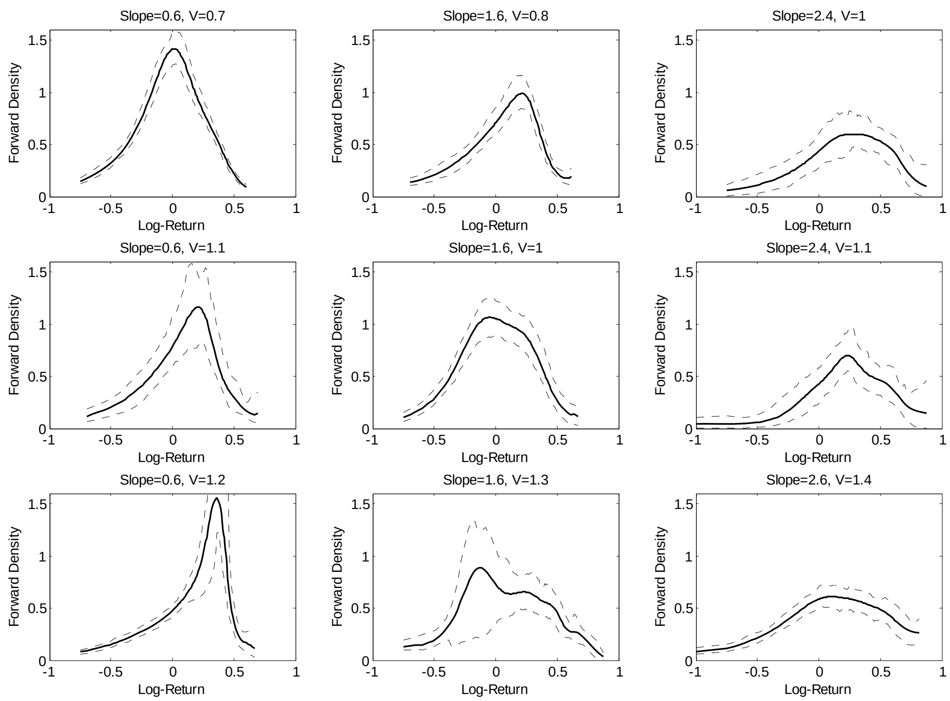


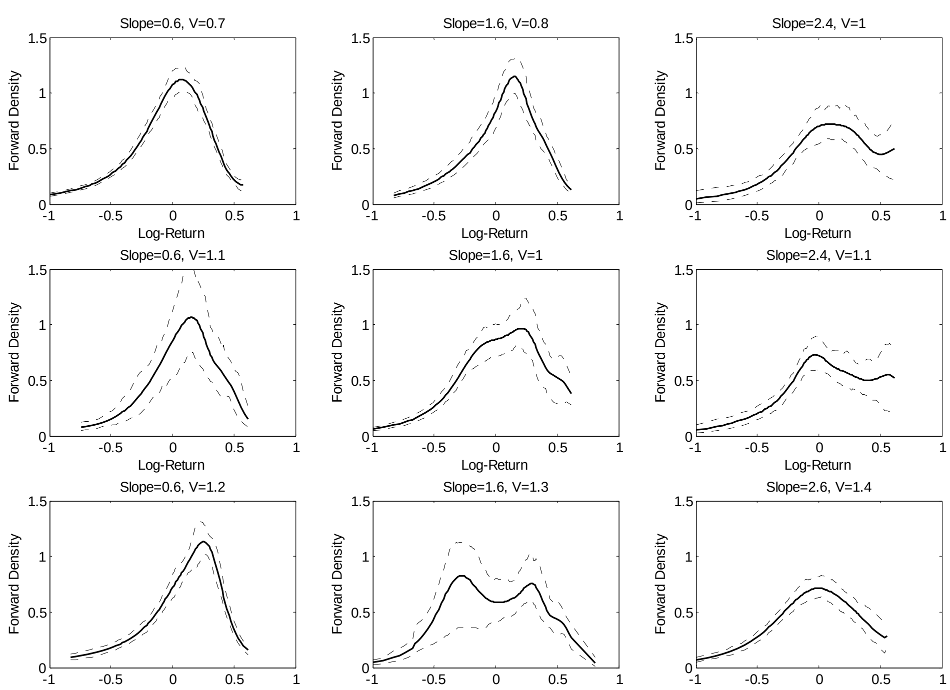
Figure 4: Nonparametric Estimation of Forward Density

This figures plot distribution densities of the Libor rates under forward measure, conditional on the slope of the forward rate curve and spot volatility levels. The slope is measured in percentage points. The volatility is normalized by the sample mean. The dotted lines are the 95% confidence interval.

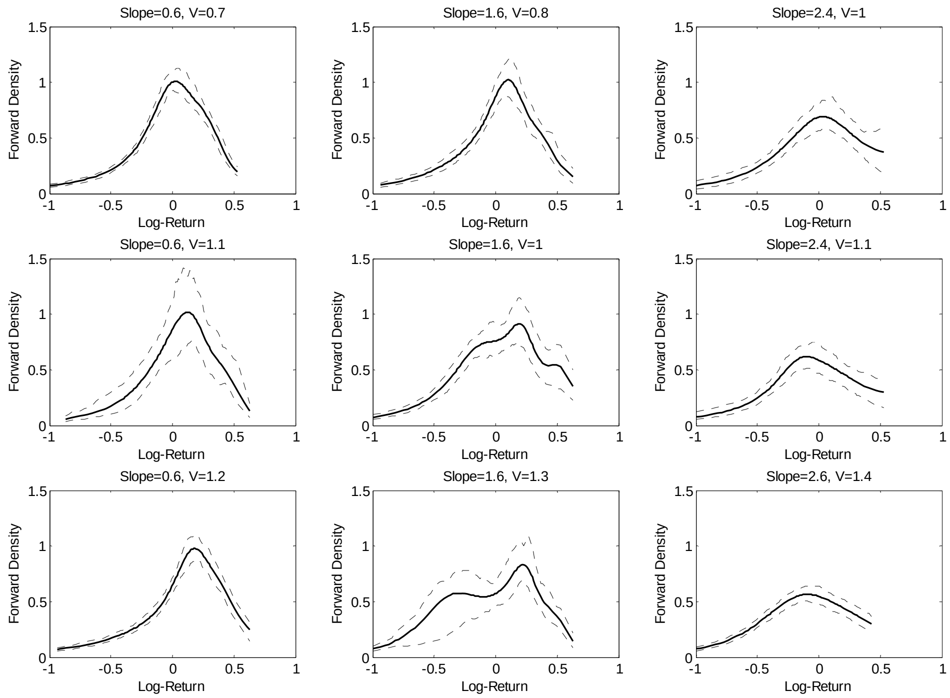
A. 2-Year



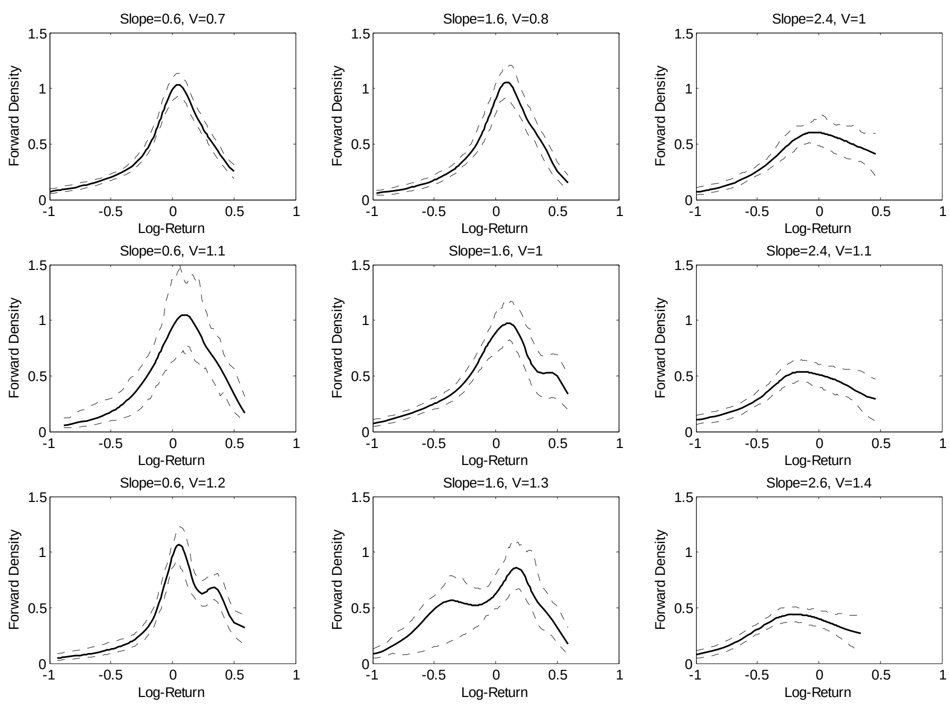
B. 3-Year



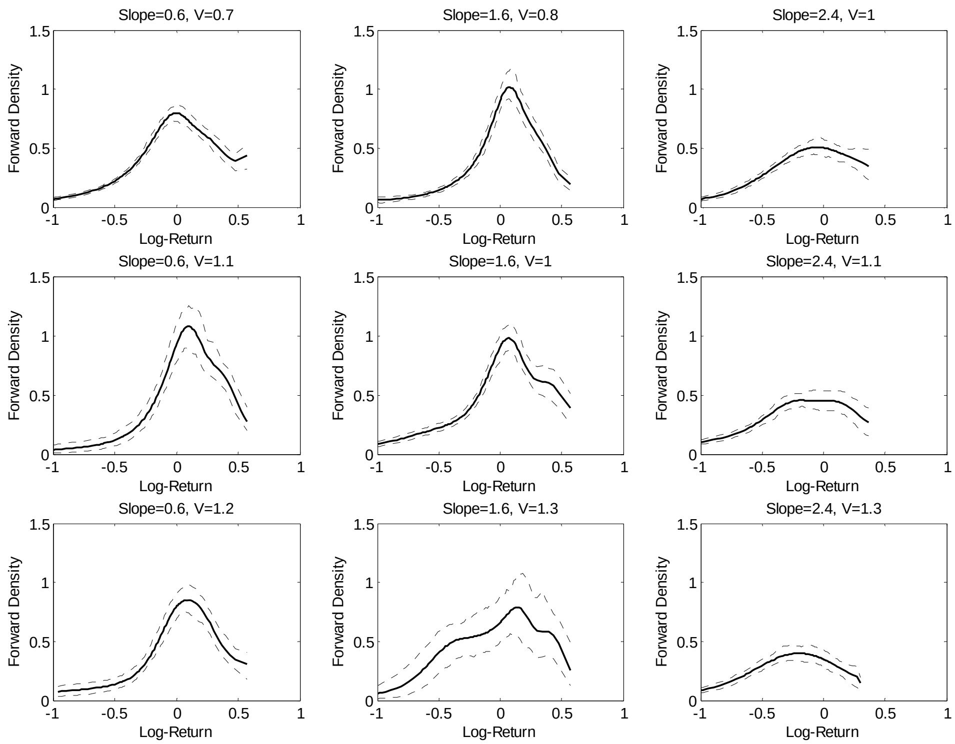
C. 4-Year



D. 5-Year



E. 7-Year



F. 10-Year

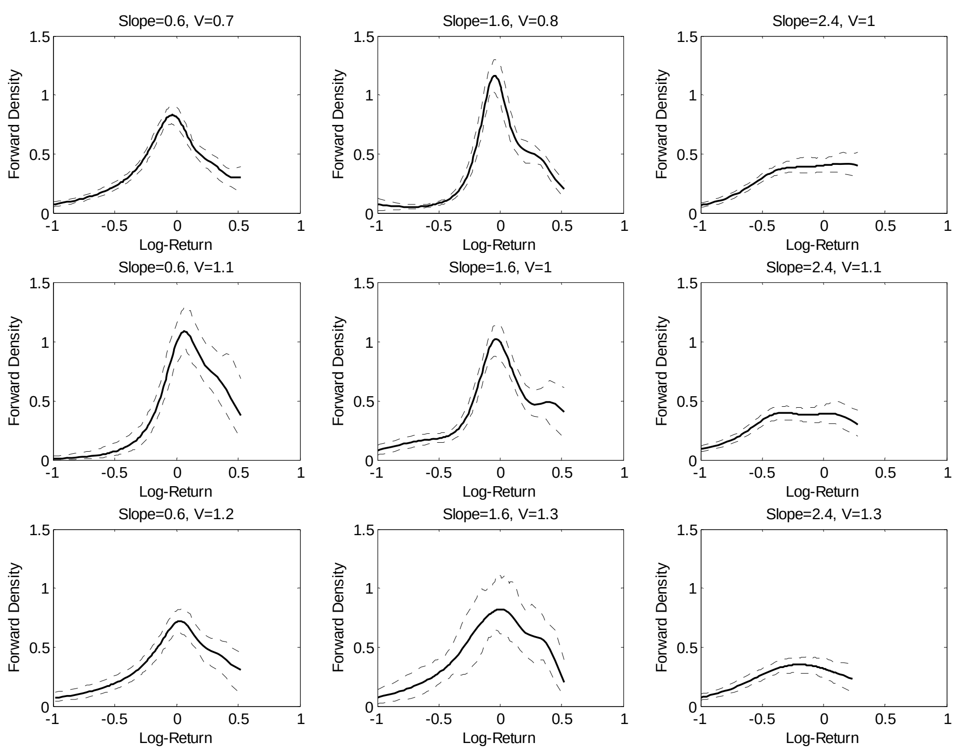
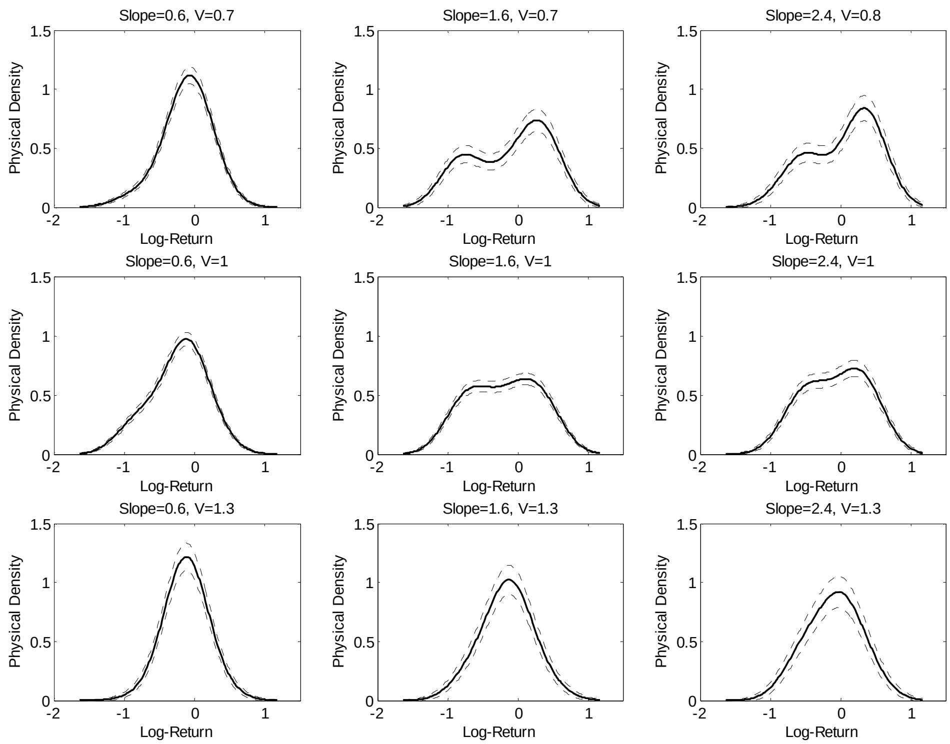
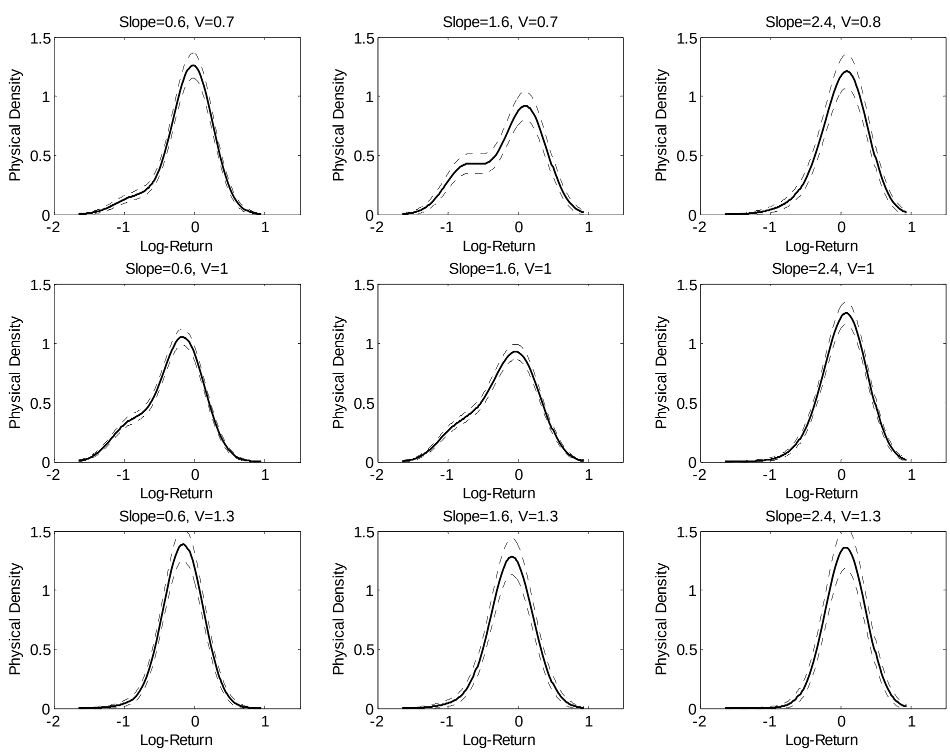


Figure 5: Nonparametric Estimation of Physical Density
This figures plot distribution densities of the Libor rates under physical measure, conditional on the slope of the forward rate curve and spot volatility levels. The slope is measured in percentage points. The volatility is normalized by the sample mean. The dotted lines are the 95% confidence interval.

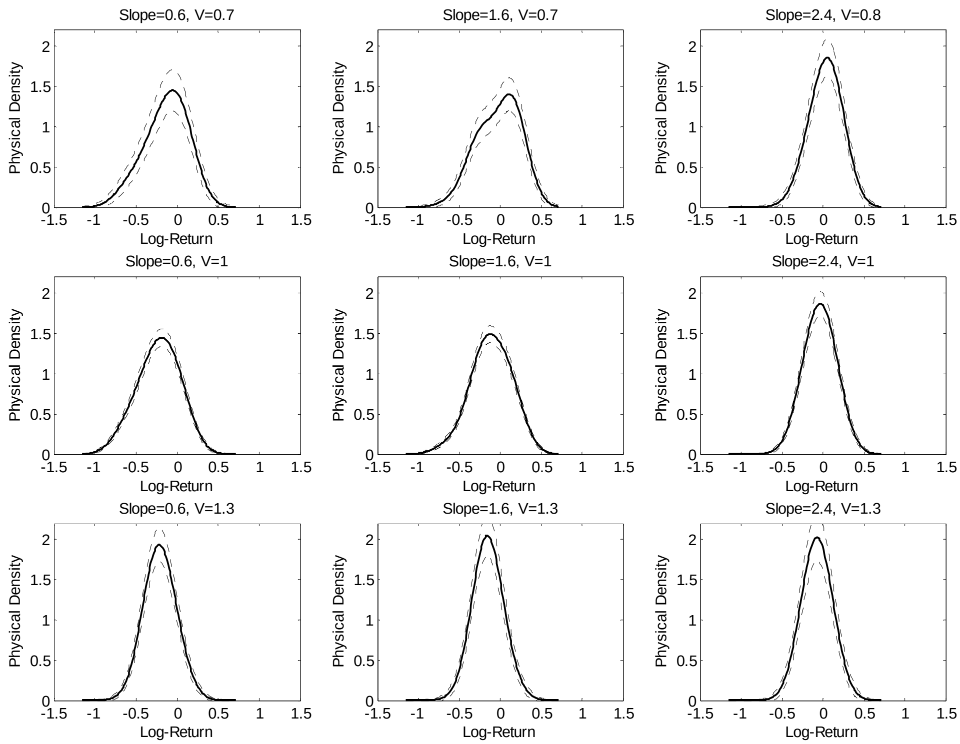
A. 2-Year



B 3-Year



C. 4-Year



D. 5-Year

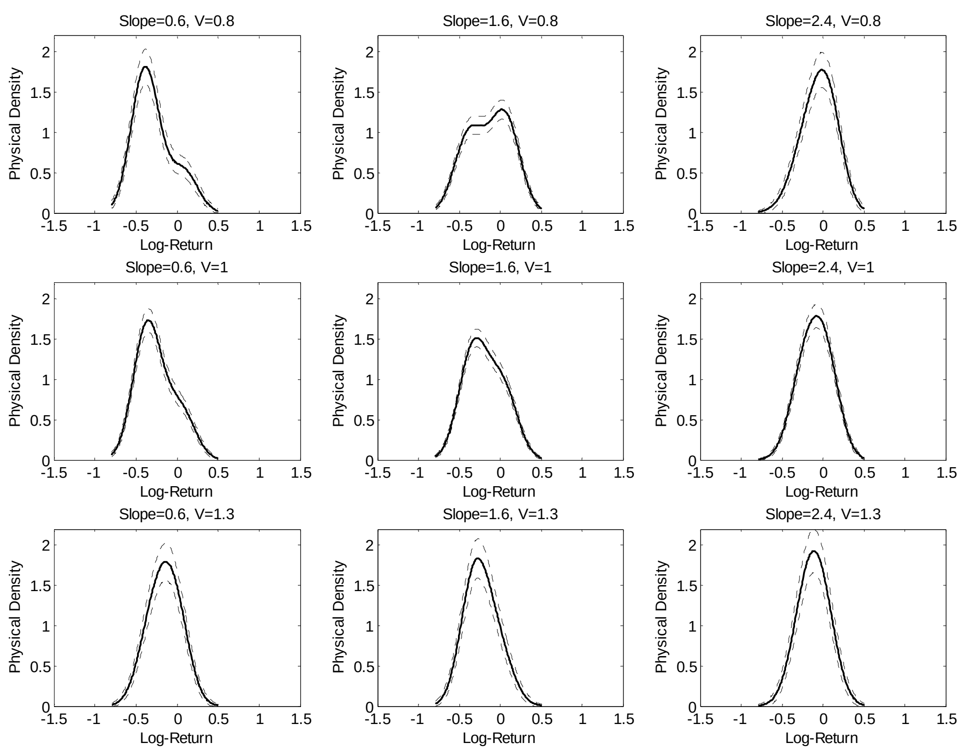
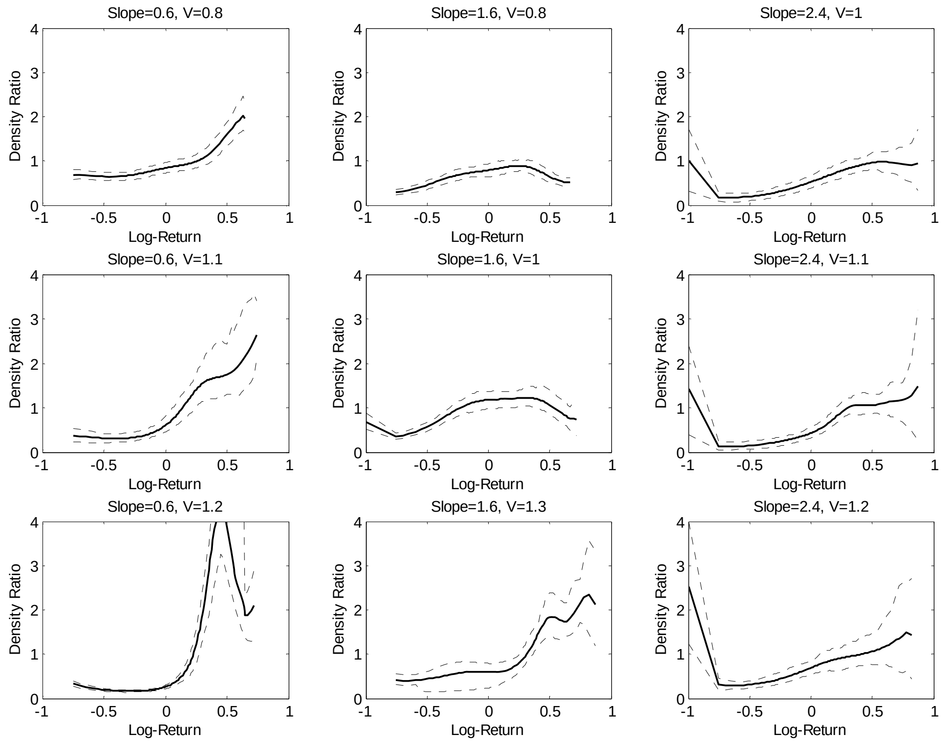
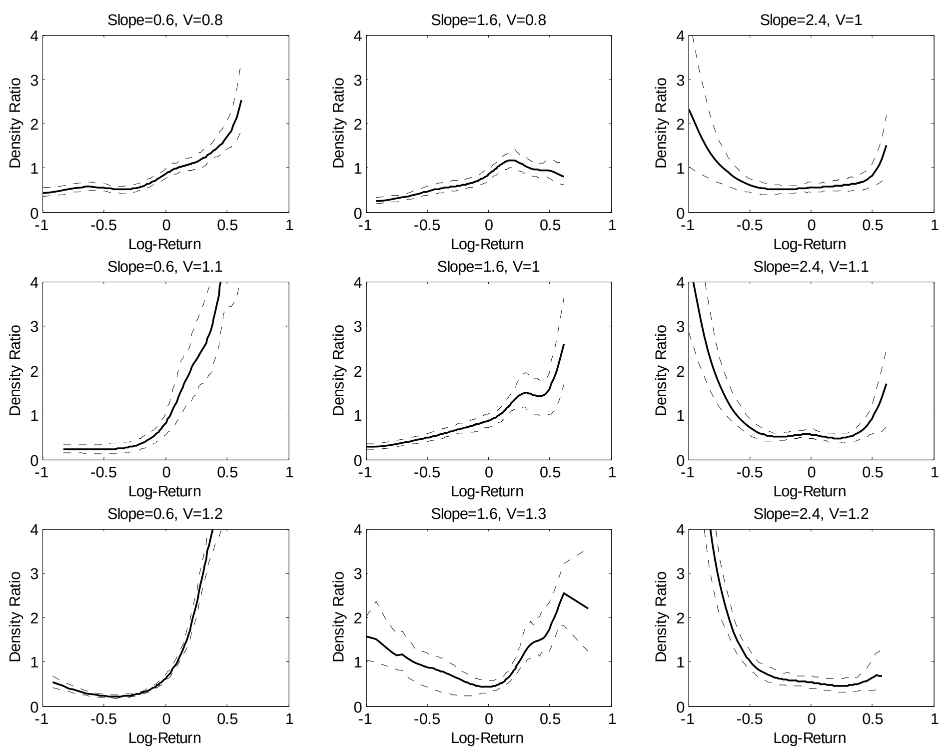


Figure 6: State Price Density
This figures plot the ratio of the distribution densities of the forward rates, forward over physical, conditional on the slope of the forward curve and the spot volatility levels.

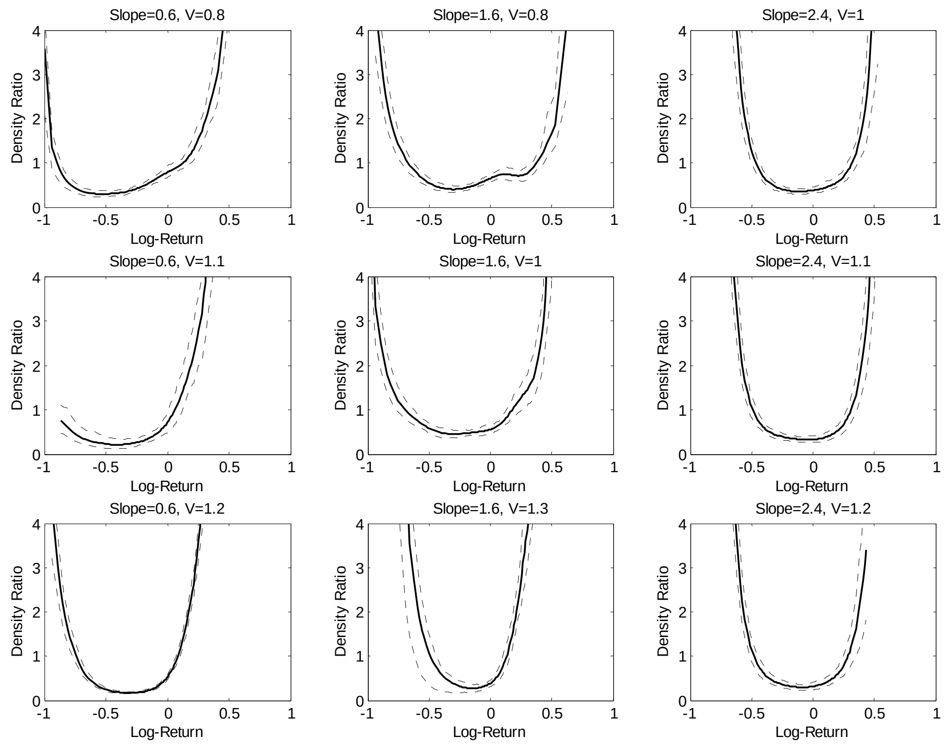
A. 2-Year



B. 3-Year



C. 4-Year



D. 5-Year