THE JOURNAL OF FINANCE • VOL. LXII, NO. 3 • JUNE 2007

Stochastic Volatilities and Correlations of Bond Yields

BING HAN*

ABSTRACT

I develop an interest rate model with separate factors driving innovations in bond yields and their covariances. It features a flexible and tractable affine structure for bond covariances. Maximum likelihood estimation of the model with panel data on swaptions and discount bonds implies pricing errors for swaptions that are almost always lower than half of the bid–ask spread. Furthermore, market prices of interest rate caps do not deviate significantly from their no-arbitrage values implied by the swaptions under the model. These findings support the conjectures of Collin-Dufresne and Goldstein (2003), Dai and Singleton (2003), and Jagnnathan, Kaplin, and Sun (2003).

THE TURMOIL IN THE BOND MARKETS and increased interest rate volatility since the 1980s have provided a boost for the rapid growth of interest rate hedging vehicles such as swaptions and caps. As these interest rate derivatives become liquid, researchers start using their prices to evaluate term-structure models. There is a rich cross section of swaptions and caps. Their prices are sensitive to the volatilities and correlations of bond yields and contain valuable information regarding the market's view on the evolution of the yield curve beyond that contained in interest rate data such as Treasury bonds, Eurodollar futures, or swap rates.

Recent studies find that it is challenging to explain the market prices of swaptions and caps under many popular term-structure models, including those that by construction fit the bond prices exactly. For example, Jagnnathan, Kaplin, and Sun (2003) find that the multi-factor Cox-Ingersoll-Ross (CIR) models generate pricing errors for caps and swaptions that are very large relative to the bid–ask spread, although the fit to the swap rates is very good. Longstaff, Santa-Clara, and Schwartz (2001), who calibrate string-market models, find that short-dated and long-dated swaptions tend to be priced inconsistently, and cap prices periodically deviate significantly from their no-arbitrage values

*Han is from the McCombs School of Business at the University of Texas, Austin. I am grateful to Martin Dierker, Heber Farnsworth, Mark Grinblatt, Jean Helwege, Jason Hsu, Jingzhi Huang, Christopher Jones, Andrew Karolyi, Francis Longstaff, Monika Piazzesi, Pedro Santa-Clara, Robert Stambaugh, Kenneth Singleton, an anonymous referee, and seminar participants at the Ohio State University for many useful comments. All errors are my own. Financial support from the Dice Center for Financial Economics at the Ohio State University is acknowledged.

implied by the swaptions. Other studies that document large and systematic pricing errors for swaptions and caps include Driessen, Klaassen, and Melenberg (2003) and Fan, Gupta, and Ritchken (2001) for multi-factor Heath, Jarrow, and Morton (1992) (HJM) models and De Jong, Driessen, and Pelsser (2001) and Hull and White (1999) for Libor and swap market models.

The large mispricing of swaptions and caps may suggest that the existing term-structure models do not adequately describe the nature of the stochastic volatilities and correlations of bond yields (e.g., Jagnnathan et al. (2003)). In most previous studies, the covariances of interest rates are either deterministic or depend at most on interest rate levels. Furthermore, models are recalibrated each date, allowing constant model parameters to change over time. Note that this implicitly induces time variation in bond covariances. Consistent with the presence of stochastic volatilities and correlations, recalibrated models yield significantly better fit to interest rate derivatives data compared to models in which the parameters are held constant (e.g., Driessen et al. (2003)). However, it has not been enough to continuously recalibrate simplistic models.

While the affine framework of Duffie and Kan (1996) can accommodate stochastic covariances of bond yields, it implies strong restrictions on the covariance structure. For example, Dai and Singleton (2000) find that for the affine models to be admissible, there is an important trade-off between flexibility in modelling the factor volatilities and correlations. More importantly, under the affine framework, risk factors that drive bond covariances generally can be hedged by a portfolio that consists solely of bonds. However, Collin-Dufresne and Goldstein (2002) and Heidari and Wu (2003) show that interest rate options markets exhibit risk factors unspanned by, or independent of, the underlying yield curve. Further evidence of systematic unspanned factors related to stochastic volatility in interest rate derivatives markets can be found in Li and Zhao (2006).

Collin-Dufresne and Goldstein (2003) and Dai and Singleton (2003) conjecture that the ultimate resolution of the swaptions and caps valuation puzzle may require time-varying correlations and possibly factors affecting the volatility of yields that do not affect bond prices. Jagnnathan et al. (2003) also suggest that it may be necessary to consider models outside the affine class that are flexible in accommodating stochastic volatility of more general forms.

In this paper, I develop a string market model of interest rates with stochastic volatility and correlation that satisfy the properties in the conjectures above. Empirically, I find strong evidence that market prices of interest rate caps do not deviate significantly from their no-arbitrage values implied from the swaptions under my model. This paper is the first to extract the market's view on the dynamics of bond covariances from liquid interest rate derivatives. Such information is valuable for risk management and the valuation of exotic interest rate derivatives.

The rest of the paper is organized as follows. Section I briefly introduces swaptions and interest rate caps. Section II develops the model and derives closed-form pricing formulas for European swaptions and caps. Section III discusses the data and Section IV presents the econometric method used to estimate the model. Section V reports the empirical results. Finally, Section VI concludes the paper.

I. Swaptions and Interest Rate Caps

The swaptions studied in this paper are written on semiannually settled interest rate swaps. Every 6 months over the term of the swap, one counterparty receives a fixed annuity from and makes a floating payment tied to the 6-month Libor rate to the other counterparty of the swap contract. Thus, an interest rate swap can be viewed as an agreement to exchange a fixed rate bond for a floating rate bond. The coupon rate on the fixed leg of a swap, also known as the swap rate, is set so that the present value of the fixed and floating legs are equal at the start of the swap contract.

Fix two dates $T > \tau$. A European style τ by T (or "τ into $T-\tau$") year receiver swaption is a single option giving its holder the right, but not the obligation, to enter into a $T-\tau$ year interest rate swap at date τ and receive semiannual fixed payments at a pre-agreed coupon rate between date τ and T. Let $D(t, T)$ denote the time-t price of a discount Libor bond that matures at time T. Then at $t < \tau$, the value of a forward swap that starts at τ and matures at T with a coupon rate c is given by

$$V(t, \tau, T, c) = \frac{c}{2} \sum_{i=1}^{2(T-\tau)} D(t, \tau_i) + D(t, T) - D(t, \tau),$$

where $\tau_i = \tau + \frac{i}{2}$ years. It follows that the payoff to the holder of a European style τ by T receiver swaption at its maturity date τ is given by

$$\mathrm{Max}(V(\tau, \tau, T, c), 0) = \mathrm{Max}\left(\frac{c}{2} \sum_{i=1}^{2(T-\tau)} D(\tau, \tau_i) + D(\tau, T) - 1, 0\right).$$

Thus, a swaption is an option on a portfolio of discount bonds.

A European style τ by T swaption is said to be at-the-money forward when the coupon rate c equals the corresponding forward swap rate $FSR(0, \tau, T)$, where

$$FSR(0, \tau, T) = 2\left(\frac{D(0, \tau) - D(0, T)}{\sum_{i=1}^{2(T-\tau)} D(0, \tau_i)}\right). \tag{1}$$

European at-the-money forward swaptions are actively traded over the counter, and are quoted in terms of implied volatilities relative to the Black (1976) model as applied to the corresponding forward swap rate. The market price for a τ by

T at-the-money forward swaption is obtained by plugging the quoted Black implied volatility σ into the formula

$$(D(0,\tau) - D(0,T))\left(\mathcal{N}\left(\frac{\sigma\sqrt{\tau}}{2}\right) - \mathcal{N}\left(-\frac{\sigma\sqrt{\tau}}{2}\right)\right), \tag{2}$$

where $\mathcal{N}(\cdot)$ is the cumulative density function of a standard normal random variable.

An interest rate cap provides insurance against the rate of interest on a floating rate loan rising above the pre-specified cap rate. It gives its holder a series of European call options, or caplets, on the underlying Libor rates. Each caplet has the same strike level but a different expiration date. For example, a T-year cap on the 6-month Libor rate consists of $2T-1$ caplets;[1] the first caplet matures in 1 year, and the last caplet matures in T years. Let $t_i = \frac{i}{2}$ years, a_i be the actual number of days between t_i and t_{i+1}, and L_i denote the 6-month Libor rate that is applicable over the period $[t_i, t_{i+1}]$. Then the cash flow received at time t_{t+1} on the caplet maturing at t_i is $\frac{a_i}{360}\max(0, L_i - R)$, where R is the cap rate. A T-year cap is said to be at the money if the cap rate R equals the current T-year swap rate.

For date $t < \tau < T$, let $F(t,\tau,T)$ denote the time-t Libor forward rate that is applicable over the period from τ to T. The i^{th} caplet is an option on the forward rate $F(t, t_i, t_{i+1})$. Assuming that each forward Libor rate is lognormal with constant volatility σ_i, the Black model price of a cap with cap rate R is

$$\sum_{i=1}^{2T-1} \frac{a_i}{360} D(0,t_{i+1})(F(0,t_i,t_{i+1})\mathcal{N}(d_i) - R\mathcal{N}(d_i - \sigma_i\sqrt{t_i})),$$

where

$$d_i = \frac{\mathrm{In}(F(0,t_i,t_{i+1})/R) + \sigma_i^2 t_i/2}{\sigma_i\sqrt{t_i}}.$$

The market convention is to quote the price of a cap in terms of an implied volatility σ, which is the same across caplets, so that the Black model price at $\sigma_i = \sigma$ equals the market price of the cap.

A caplet can also be viewed as a put option on the corresponding Libor discount bond. Thus, an interest rate cap is a portfolio of options on discount bonds. In contrast, a swaption is an option on a portfolio of discount bonds. Although swaptions and interest rate caps are traded as separate products, they are linked by no-arbitrage relations through the correlation structure of bond yields. It is important to note that the relative valuation of swaptions and

[1] Note that although the cash flow of this caplet is paid at time t_{i+1}, the applicable Libor rate L_i is determined at t_i. For this reason, the cash flow for the first caplet maturing in 6 months is nonstochastic and thus omitted by market convention.

caps can only be judged under a term-structure model, and not by a simple comparison of their Black implied volatilities quoted in the market.[2]

II. The Valuation Framework

In this section, I develop a term-structure model with flexible and intuitive specifications for the stochastic volatilities and correlations of bond yields. In my model, separate factors drive the innovations in bond yields and their covariances, and thus bonds alone cannot hedge volatility risk. This modeling framework, which is reminiscent of the large literature on stochastic volatility that specifies the joint dynamics of a traded asset and its volatility (e.g., Heston (1993)), captures the empirical evidence of unspanned stochastic volatility. The model is also tractable, since just like the affine framework, bond covariances are affine in the volatility state variables. This property is key to deriving closed-form solutions for a variety of interest rate derivatives (e.g., Collin-Dufresne and Goldstein (2003)).

A. Model

Similar to the string market model of Longstaff et al. (2001), I directly model the dynamics of bond prices. The risk-neutral drifts of traded bond prices are determined by the no-arbitrage condition that their expected rates of return under the risk-neutral measure equal the spot risk-free rate. Thus, the focus of my model is on the dynamics of bond covariances. Given the contemporaneous bond prices, prices of European swaptions and interest rate caps are determined by the dynamics of bond covariances (see Section II.B).

By Girsanov's theorem, the instantaneous bond covariances are invariant with respect to an equivalent change of probability measure. Thus, when modeling bond covariances under the risk-neutral (equivalent martingale) measure, I can utilize information contained in the historical estimates of bond covariances. It is well known that most of the observed variation in historical bond prices is explained by a few common factors (e.g., Litterman and Scheinkman (1991), and Dai and Singleton (2000)). The explanatory power of these factors is stable and the factor loadings show a persistent pattern. However, there is significant time variation in the variances of the common yield factors (e.g., Bliss (1997), Perignon and Villa (2005)).

Motivated by these findings, I assume that the yield curve is driven by N common factors with time-invariant weights but possibly stochastic volatility. The stochastic volatilities of common yield factors lead to stochastic covariances

[2] The reason is that the Black implied volatilities apply to different underlying interest rates that are assumed to be lognormally distributed: forward swap rates in the case of swaptions and forward Libor rates in the case of interest rate caps. Each forward swap rate is approximately a linear combination of the underlying forward Libor rates. Thus, forward swap rates and forward Libor rates can not be simultaneously lognormally distributed.

of bond yields. The risk-neutral dynamics of the discount bond prices are given by

$$\frac{dD(t,T)}{D(t,T)} = r_t dt - \sum_{k=1}^{N} B_k(T-t)\sqrt{\nu_k(t)}\, dZ_k^Q(t), \tag{3}$$

where r_t is the instantaneous short rate. For each $k = 1, \ldots, N$, dZ_k^Q is a Brownian motion (under the risk-neutral measure) that represents shocks to the kth factor driving the yield curve, and $\nu_k(t)$ is the instantaneous variance of the k^{th} yield factor. Without loss of generality, the yield factors are orthogonal to each other. The function $B_k(T-t)$ describes the loadings of the bond with maturity T on the k^{th} yield factor at time t. It is a deterministic function of the time-to-maturity $T-t$ only. This ensures that the term-structure dynamics under the model are time homogeneous.

It follows from (3) that the date-t instantaneous covariances of log-bond prices for a set of bonds with maturity $\tau_1, \ldots, \tau_n$ can be written as a product of three matrices:

$$B_t Diag(\nu_t) B_t', \tag{4}$$

where B_t is an $n \times N$ matrix whose $(i,k)^{\text{th}}$ element is $B_k(\tau_i - t)$, and $Diag(\nu_t)$ is an $N \times N$ diagonal matrix whose diagonal elements are $\nu_k(t), k = 1, \ldots, N$. By construction, the covariance matrix (4) is positive semidefinite. Furthermore, the covariance between any two bonds is *linear* in the variances of the N yield factors $\nu_1(t), \ldots, \nu_N(t)$.

There are two sources of time variation in bond covariances under my model. One, as time passes by, the time-to-maturities of bonds decreases, and hence their loadings on the common yield factors change correspondingly. This leads to deterministic changes in bond covariances. Another source of movement in the covariances of bond yields is induced by the stochastic volatilities of the yield factors, which I model next.

Consider the general case in which K of the N yield factors ($K \leq N$), labeled by I_K, display stochastic volatility, and the remaining $N-K$ factors have constant volatility. The instantaneous variance of the i^{th} yield factor ($\forall i \in I_K$) follows an autonomous square root process:

$$d\nu_i(t) = \kappa_i(\theta_i - \nu_i(t))\,dt + \sigma_i\sqrt{\nu_i(t)}\, dW_i^Q(t). \tag{5}$$

Model parameters κ_i and θ_i are, respectively, the mean reversion speed and the long-run mean level for the variance of the i^{th} yield factor. For tractability, I assume that the stochastic volatilities of the yield factors (the covariance state variables in my model) are independent of each other. Furthermore, I assume that the Brownian motions dW's that drive bond yields covariances are uncorrelated with the dZ's that drive innovations in bond yields. This assumption is motivated by previous findings that innovations in interest rate levels are largely uncorrelated with innovations in the volatility of interest rates (e.g., Ball and Torous (1999), Chen and Scott (2001), and Heidari and Wu (2003)), and it implies that the dynamics for the variances of yield factors are the same under the risk-neutral measure and all forward measures (e.g., Goldstein (2000)).

In the rest of the paper, I denote by $GA_{N,K}$ $(K \leq N)$ the above model specification with N factors driving innovations in bond yields, the first K of which display unspanned stochastic volatility while the others have constant volatility.[3] The risk-neutral dynamics of bond prices are given by (3), and the unspanned stochastic volatilities satisfy (5). These volatility factors, together with the discount bond prices, form the state vector of the model.

Risk-neutral dynamics for bond prices and volatility state variables are sufficient for valuing swaptions and interest rate caps. To complete the model and to estimate it via maximum likelihood, I also need the dynamics of the bond prices and the volatility state variables under the empirical measure. Market prices of risk for the common yield factors are assumed to be proportional to the volatility of the yield factors, so that

$$dZ_k^P(t) = dZ_k^Q(t) + \gamma_k \sqrt{\nu_k(t)}\, dt,$$

where the γ_k's are constant model parameters. By this assumption and (3), the bond price dynamics under the empirical measure are given by

$$\frac{dD(t,T)}{D(t,T)} = \left(r_t + \sum_{k=1}^{N} \gamma_k B_k(T-t)\nu_k(t) \right) dt - \sum_{k=1}^{N} B_k(T-t)\sqrt{\nu_k(t)}\, dZ_k^P(t). \tag{6}$$

The risk premium for the i^{th} volatility factor $(i = 1, \dots, K)$ is modeled as $\lambda_i \sqrt{\nu_i}$, where the λ_i's are constant model parameters (see, e.g., Heston (1993) for an equilibrium justification of this volatility risk premium specification). Under the empirical measure, ν_i satisfies the following affine process:

$$d\,\nu_i(t) = \hat{\kappa}_i(\hat{\theta}_i - \nu_i(t))\, dt + \sigma_i \sqrt{\nu_i(t)}\, dW_i^P(t), \tag{7}$$

where dW_i^P is a standard Brownian motion under the empirical measure, and

$$\hat{\kappa}_i = \kappa_i - \lambda_i, \ \hat{\theta}_i = \frac{\kappa_i \theta_i}{\kappa_i - \lambda_i}.$$

B. Model Valuation of Swaptions and Caps

It is more convenient to use the forward risk-neutral measure (e.g., Jamshidian (1997)) to value swaptions and interest rate caps, since their payoffs are homogeneous of degree one in a finite number of discount bond prices. Let $D(t, \tau, T)$ denote the date-t price of a forward contract to buy at date τ a bond that matures at $T > \tau$. In the absence of arbitrage, the forward bond prices are

[3] The common yield factors are labeled in decreasing order according to their unconditional variance.

related to the discount bond prices by $D(t, \tau, T) = D(t, T)/D(t, \tau)$. The forward risk-neutral measure corresponding to date τ, denoted by Q^τ, uses the discount bond that matures at τ as the numeraire asset. By construction, forward bond prices $D(t, \tau, T)$ are martingales under Q^τ. Thus, the dynamics of the forward bond prices under the corresponding forward risk-neutral measure are determined by their covariances.

The payoff at maturity of a τ by T European at-the-money forward receiver swaption can be written in terms of the forward bond prices as

$$\mathrm{Max}(\tilde{A}(\tau) - 1, 0),$$

where

$$\tilde{A}(t) = \sum_{j=1}^{2(T-\tau)} \tilde{\omega}_j S_j(t),$$

$$S_j(t) = \frac{D(t, \tau, \tau_j)}{D(0, \tau, \tau_j)}, \;\; \tau_j = \tau + \frac{j}{2}, \tag{8}$$

$$\tilde{\omega}_j = \frac{\omega_j D(0, \tau, \tau_j)}{\sum_{k=1}^{2(T-\tau)} \omega_k D(0, \tau, \tau_k)}, \tag{9}$$

$$\omega_j = \frac{c}{2(1 + (T - \tau)c)}, \quad j = 1, \dots, 2(T - \tau) - 1,$$

$$\omega_{2(T-\tau)} = \frac{1 + c/2}{1 + (T - \tau)c},$$

and c is the strike rate, which equals the forward swap rate $FSR(0, \tau, T)$ given in (1). Note that the $\tilde{\omega}_j$'s are positive constants and $\sum_{j=1}^{2(T-\tau)} \tilde{\omega}_j = 1$.

The no-arbitrage price of a contingent claim, which settles at time τ, is given by first taking the expectation of its payoff under the forward risk-neutral measure, and then multiplying it by $D(0, \tau)$ (e.g., lemma 13.2.3 of Musiela and Rutkowski (1997)). It follows that the date-0 price of a τ by T at-the-money forward receiver swaption is

$$P(\tau, T) = D(0, \tau)\, \mathrm{E}^{Q^\tau}[\mathrm{Max}(\tilde{A}(\tau) - 1, 0)]. \tag{10}$$

Thus, the valuation of the swaption is reduced to computing the expectation under the forward risk-neutral measure of an arithmetic sum of a set of random variables $S_j(\tau)$, for $j = 1, \dots, 2(T - \tau)$. Since each $S_j(t)$ is just a constant multiple of $D(t, \tau, \tau_j)$, which is a martingale under Q^τ, the drift of $S_j(t)$ under Q^τ is also zero. Therefore, the covariances of $\{D(t, \tau, \tau_j)\}_{j=1,\dots,2(T-\tau)}$ determine their joint distributions under the forward risk-neutral measure Q^τ, and hence the price of a τ by T European swaption as given by (10).

To value European swaptions and interest rate caps, the covariances of forward bonds with fixed maturities are required. It is convenient to first model the covariances of bonds with fixed time to maturity (e.g., multiples of 6 months).

Let Ω_t be the date t instantaneous covariance matrix of changes in the logarithm of the 6-month forward Libor bond prices $\{D(t, t+t_i, t+t_{i+1})\}_{i=0}^{19}$, where $t_i = \frac{i}{2}$ years. (Only bonds with maturity up to 10 years are considered since the swaptions and caps used in my empirical study have maturity no greater than 10 years.) Let H be the corresponding unconditional covariance matrix estimated from the historical bond prices. This matrix can be decomposed as $H = U\Lambda_0 U'$, where Λ_0 is a diagonal matrix whose diagonal elements are the eigenvalues of H, and the columns of U are the corresponding eigenvectors.[4] Under the $GA_{N,K}$ model, the conditional covariance matrix of the forward bonds is[5]

$$\Omega_t = U\Lambda_t U', \tag{11}$$

where Λ_t is a diagonal matrix whose first N main diagonal elements are the instantaneous variances of the N yield factors. The variances of the first K yield factors follow the CIR processes specified in (5). The remaining $N-K$ factors have constant volatility.

The covariances of forward bonds with fixed maturities can be obtained from Ω_t. Note that

$$\log(D(t,\tau,\tau_j)) = \log(D(t,\tau,\tau_1)) + \log(D(t,\tau_1,\tau_2)) + \cdots + \log(D(t,\tau_{j-1},\tau_j)).$$

Every 6 months from date 0 to date τ, $\tau_j - t$'s are multiples of 6 months. On these dates, the instantaneous covariances of the bonds on the right-hand side of the last equation can be read off from $\Omega_t = \{c_{ij}(t)\}_{ij}$. On other dates, I linearly interpolate the covariances to preserve the continuity of the covariances as functions of time to maturity.[6] More precisely, for integers $1 \le i < j < 20$, at any time $t \le t_i = \frac{i}{2}$, let k be the integer such that $t_k \le t < t_{k+1}$. Assume

$$\mathrm{Cov}(D(t,t_i,t_{i+1}), D(t,t_j,t_{j+1})) = (1-2(t-t_k))c_{i-k,j-k}(t) + 2(t-t_k)c_{i-k-1,j-k-1}(t).$$

Now I continue with the valuation of a European τ by T swaption given by equation (10). Because of the assumption that the volatility state variables are instantaneously uncorrelated with innovations in the yield curve, each $S_j(\tau)$ is lognormal conditional on average values of ν_k ($k = 1, \ldots, N$) between date 0 and τ (see lemma 1 of Hull and White (1987)). Thus, their geometric sum

[4] I normalize U to have unit length for each column. The i^{th} column of U and the i^{th} main diagonal element of Λ_0 are the weights and the unconditional variance of the i^{th} common factor driving the yield curve. Furthermore, a rotation of the yield factors does not change their variances (the ν_i's).

[5] The covariances of discount bonds $\{D(t, t+t_i)\}_{i=1}^{20}$ ($t_i = \frac{i}{2}$) are easily recovered from the covariance of forward bonds modeled here. The matrix B in the covariances of discount bonds $\{D(t, t+t_i)\}_{i=1}^{20}$ written in the form of (4) is $B = TU_N$, where T is a 20×20 lower-triangular matrix whose entries on and below the main diagonal are all one, and U_N is a matrix consisting of the first N columns of U.

[6] I experiment with alternative interpolating schemes using functions that are exponentially decaying in time to maturity, and find that the interpolation scheme has a negligible influence on the prices of swaptions and caps.

$\tilde{G}(\tau) = \prod_{j=1}^{2(T-\tau)} S_j^{\tilde{\omega}_j}(\tau)$ is also conditionally lognormally distributed. Upon replacing the difference of the arithmetic sum $\tilde{A}(\tau)$ and the geometric sum $\tilde{G}(\tau)$ by the mean of the difference,[7] the price of a τ by T at-the-money forward European swaption approximately equals

$$P(\tau, T) = D(0, \tau)\, \mathrm{E}^{Q^\tau}[\mathrm{Max}(\tilde{G}(\tau) - g, 0)], \tag{12}$$

where $g = 1 + \mathrm{E}^{Q^\tau}[\tilde{G}(\tau) - \tilde{A}_\tau]$. I apply the law of iterative expectations to derive the following closed-form pricing formula (see the Appendix for details of the derivation):

PROPOSITION 1: *Under the $GA_{N,K}$ model, the price of a τ by T European at-the-money forward swaption is given by:*

$$P(\tau, T) = D(0, \tau)\left[2\mathcal{N}\left(\frac{1}{2}\sqrt{\tilde{\omega}'\ \hat{\Sigma}\ \tilde{\omega}\tau}) - 1\right)\right], \tag{13}$$

where

$$\hat{\Sigma} = \frac{1}{\tau}\sum_{l=0}^{2\tau-1}\left[A_l U Diag_1^l U' A_l' + A_{l+1} U Diag_2^l U' A_{l+1}' - A_l U Diag_2^l U' A_l'\right] \tag{14}$$

and $\mathcal{N}(\cdot)$ is the cumulative density function of a standard normal random variable. The weights $\omega = \{\tilde{\omega}_j\}$, $j = 1, \ldots, 2(T-\tau)$, are given by (9). Matrix U consists of the eigenvectors for the unconditional covariance matrix of changes in the logarithm of the 6-month forward Libor bonds with semiannual maturities ranging from 6 months to 10 years. For each non-negative integer $l \leq 2\tau - 1$, Diag_1^l and $Diag_2^l$ are 20×20 matrices whose entries are zero except the first N diagonal elements

$$Diag_1^l(i,i) = \frac{1}{2}\theta_i + \frac{e^{-\kappa_i l/2} - e^{-\kappa_i(l+1)/2}}{\kappa_i}(\nu_i(0) - \theta_i), \quad i = 1, \ldots, K,$$

$$Diag_1^l(i,i) = \frac{1}{2}\theta_i, \quad i = K+1, \ldots, N,$$

$$Diag_2^l(i,i) = \frac{1}{4}\theta_i + e^{-\kappa_i l/2}\left(\frac{2}{\kappa_i^2} - \frac{e^{-\frac{\kappa_i}{2}}}{\kappa_i} - \frac{2}{\kappa_i^2}e^{-\frac{\kappa_i}{2}}\right)(\nu_i(0) - \theta_i), \quad i = 1, \ldots, K,$$

$$Diag_2^l(i,i) = \frac{1}{4}\theta_i, \quad i = K+1, \ldots, N.$$

For $i = 1, \ldots, K$, $\nu_i(0)$ denotes the spot variance of the i^{th} common yield factor, which reverts to a long-run mean level of θ_i with mean reversion speed κ_i. For

[7] This approximation technique has been applied to price Asian options and basket options. It is known to be very accurate (e.g., Vorst (1992)). In my case, simulations show that the typical approximation error is less than 0.1%, even for maturities as long as $\tau = 10$, and for values of the volatility state variables that generate bond yield volatilities that are twice as high as those observed in the data.

each nonnegative integer $l \leq 2\tau - 1, A_l$ is a $2(T-\tau) \times 20$ matrix. The $2\tau - l + 1$ to $2T - l$ column of A_l forms a lower triangular matrix whose elements on and below the diagonal are one. The other elements of A_l are zero.

The price of an interest rate cap is the sum of the prices of its constituent caplets. A caplet on the 6-month Libor rate with maturity $t_i = i/2$ year is a put option on the Libor forward bond $D(t, t_i, t_{i+1})$ and can be valued the same way as a t_i by t_{i+1} swaption.

PROPOSITION 2: *Under the $GA_{N,K}$ model, the price of a T-year at-the-money interest rate cap on the 6-month Libor rate is*

$$\sum_{i=1}^{2T-1} D(0,t_i)\mathcal{N}(-d_{2i}) - \left(1+\frac{1}{2}R\right)D(0,t_{i+1})\mathcal{N}(-d_{1i}), \tag{15}$$

where $d_{1i} = \frac{\log(D(0,t_i,t_{i+1})(1+\frac{1}{2}R))+\frac{1}{2}\bar{\sigma}_i^2 t_i}{\sqrt{\bar{\sigma}_i^2 t_i}}$, $d_{2i} = d_{1i} - \sqrt{\bar{\sigma}_i^2 t_i}$, $t_i = \frac{i}{2}$ year, the cap rate R equals the T-year swap rate, and $\bar{\sigma_i}^2$ is the average expected variance over $[0, t_i]$ of changes in the logarithm of forward bond price $D(t, t_i, t_{i+1})$, given by equation (14) with $\tau = t_i$ and $T - \tau = \frac{1}{2}$.

C. Relation to Other Term Structure Models

The model I develop in this paper belongs to the string model/random field framework (e.g., Goldstein (2000), and Santa-Clara and Sornette (2001)),[8] which generalizes the HJM (1992) model. In particular, I generalize the constant covariance string market model of Longstaff et al. (2001) by explicitly modeling the dynamics of the stochastic covariances of bond yields. These dynamics are taken into account both in model valuation of swaptions and caps, as well as in model estimation. Below I show that incorporating stochastic covariances into the string market model is key to reconciling the relative valuation of swaptions and caps.

My model can be viewed as a "reduced-form" representation of an affine model with unspanned stochastic volatility. First, the dynamics of discount bond prices (3) under my model are consistent with the affine model of Duffie and Kan (1996). To see this, consider an affine model with N factors driving the short rate:

$$r_t = \delta_0 + \delta_1 X_t.$$

The risk-neutral dynamics of the risk factors X_t are given by

$$dX_t = \kappa(\theta - X_t)\,dt + \Sigma\sqrt{V(t)}\,dZ_t^Q,$$

[8] Strictly speaking, the covariance matrix of a set of forward rates or bonds is of full rank under the string/random field model. It has a rank of N under my $GA_{N,K}$ model because the diagonal matrix Λ_t in (11) only has N nonzero diagonal entries. However, it is trivial to generate a full-ranked bond covariance matrix under my model (e.g., by letting the other diagonal elements of Λ_t take some arbitrarily small positive values) without affecting the valuation of interest rate derivatives.

where κ and Σ are $N \times N$ matrices, and V is a diagonal matrix with components

$$V_{ii}(t) = \alpha_i + \beta_i X(t).$$

The discount bond prices under the above affine model are given by

$$D(t, T) = e^{A(T-t) - B(T-t)' X_t}.$$

By Ito's lemma and the dynamics of X_t, the risk-neutral dynamics for the discount bond prices under the affine model are

$$\frac{dD(t, T)}{D(t, T)} = r_t dt - B(T-t)' \Sigma \sqrt{V(t)}\, dZ_t^Q,$$

which are of the same form as the bond price dynamics (3) under my model. The difference is that under my model, the B functions in (3) are pre-specified and do not depend on model parameters, whereas under the affine models, the B functions satisfy a system of ordinary differential equations and depend on the model parameters. This difference reflects the fact that in my model, bond prices are part of the state vector rather than derived as functions of latent factors as in the affine models.

Second, the covariances of bond yields are affine in the volatility state variables under both my model and the affine model. The difference is that in my model, innovations in bond yields are not contemporaneously affected by volatility innovations. In contrast, the stochastic volatility factors in the affine framework typically enter into bond prices, and thus can be represented as linear combinations of bond yields. In other words, the volatility factors in the traditional affine framework drive both the cross-sectional differences in bond yields and the changes in the conditional volatility. My model, as well as affine models with unspanned stochastic volatility (e.g., Collin-Dufresne, Goldstein, and Jones (2004)), breaks this dual role of stochastic volatility.

Two contemporaneous theoretical papers present models similar to that I introduce here. First, Collin-Dufresne and Goldstein (2003) generalize the affine framework to HJM and random field models. Like my model, the generalized affine framework allows for unspanned stochastic volatility yet maintains the tractability of the affine framework. Collin-Dufresne and Goldstein (2003) illustrate how my model can be mapped into their framework. Second, Kimmel (2004) develops a class of random field models in which the volatilities of forward rates depend on a finite set of latent variables that follow diffusion processes. He focuses on conditions necessary for the existence and uniqueness of the forward rate process, and derives theoretic results for derivative pricing. Although the state vector in his model is infinite dimensional, each forward rate follows a low-dimensional diffusion process. My model shares this property. In particular, I show that under my model, it is still tractable to price interest rate derivatives and conduct econometric estimation.

III. Data

A panel data set of interest rates and Black implied volatilities for at-the-money forward European swaptions is used to estimate my model. The interest rates include 6-month and 1-year Libor rates as well as 2-, 3-, 4-, 5-, 7-, 10-, and 15-year swap rates. I use 34 swaptions whose total maturity is no greater than 10 years. The option maturity ranges from 6 months to 5 years and the tenor of the underlying swap is between 1 year and 7 years. Implied volatilities for Libor interest rate caps of 2-, 3-, 4-, 5-, 7- and 10-year maturities are used in the study of relative valuation of swaptions and caps. All data are collected from the Bloomberg system and represent the average of best bid and ask among many large swap and swap derivatives brokers. There are 220 weekly observations for each series, sampled every Friday from January 24, 1997 to April 6, 2001.

To compute both the market and model prices for swaptions and interest rate caps, I need prices of discount bonds with semiannual maturity ranging from 6 months to 10 years. Following Longstaff et al. (2001), I apply a least-square cubic spline approximation to the Libor and swap rates to get the par yield curve, and then bootstrap the discount bond prices from the par yield curve.[9] The unconditional covariance matrix of log forward bond prices is computed using historical data between January 17, 1992 and January 17, 1997. Its eigenvector matrix U is used in forming conditional bond covariances as specified in (11).

Table I reports the mean and standard deviation of Black implied volatilities for the swaptions. On average, the implied volatility is humped as a function of swaption maturity, with a maximum at 2 years. However, on high volatility dates, the swaption implied volatility tends to decrease monotonically with option maturity. Consistent with mean reversion in the interest rates, the swaption implied volatility is usually monotonically decreasing as a function of the tenor of the underlying swap. Furthermore, the standard deviation of swaption implied volatility decreases with option maturity as well as the tenor of the underlying swap. For example, although the implied volatilities of 6 month into 1 year and 5 year into 5 year swaptions have about the same sample mean, the first is almost three times as volatile as the latter.

Figure 1 confirms that there is a fair amount of time-series variation in the swaption implied volatilities, especially for short-dated swaptions. My sample period includes the Asian crisis, the Russian moratorium, the Long-Term Capital Management (LTCM) crisis, the crash of technology stocks, as well as several quiet periods of low interest rate volatility. These different volatility environments help me pin down the dynamics of bond covariances in model

[9] The least-squares cubic spline approximation fits the swap rates very well, with average absolute fitted error of about 0.76 basis points. Interpolation schemes that exactly fit observed Libor and swap rates tend to lead to unreasonably high estimates for the volatilities of long-term bonds and somewhat rugged correlations. An alternative method to generate reasonable estimates for the bond covariances is to put some smoothness condition on the shape of the forward rate curve (e.g., Driessen, Klaassen, and Melenberg (2003)).

Table I

Descriptive Statistics of European Swaption Volatilities

This table presents descriptive statistics for the mid-market Black implied volatilities for the 34 at-the-money forward European swaptions analyzed in the paper. The data consist of Friday closing quotes from January 24, 1997 to April 6, 2001, and are collected from Bloomberg. "Expiration" refers to the number of years till option expiration and "Tenor" refers to the maturity of the underlying swap. The last column reports the first-order serial correlation of each swaption implied volatility series. The swaption implied volatilities are annualized and expressed in percentage.

Expiration	Tenor	Mean	Median	Standard Deviation	Min	Max	Serial Correlation
0.5	1	14.77	13.85	3.29	9.70	27.00	0.93
1	1	16.12	15.45	2.79	11.80	26.00	0.95
2	1	16.85	16.50	2.11	13.10	23.20	0.94
3	1	16.66	16.40	1.87	13.00	21.80	0.93
4	1	16.35	16.10	1.69	13.00	20.70	0.92
5	1	16.05	15.90	1.50	12.90	20.25	0.93
0.5	2	15.50	14.55	3.02	10.40	27.00	0.94
1	2	16.10	15.50	2.48	12.20	24.00	0.94
2	2	16.37	16.05	1.95	13.00	22.50	0.93
3	2	16.12	15.90	1.70	12.90	21.00	0.93
4	2	15.82	15.73	1.51	12.80	20.10	0.92
5	2	15.51	15.50	1.37	12.70	19.75	0.92
0.5	3	15.48	14.58	2.89	10.60	27.00	0.93
1	3	15.87	15.30	2.28	12.20	23.00	0.94
2	3	16.04	15.70	1.81	12.90	21.50	0.93
3	3	15.79	15.70	1.57	12.80	20.20	0.93
4	3	15.50	15.50	1.41	12.70	19.85	0.92
5	3	15.18	15.20	1.29	12.60	19.50	0.91
0.5	4	15.42	14.60	2.73	10.80	26.50	0.93
1	4	15.65	15.30	2.12	12.10	22.50	0.93
2	4	15.74	15.55	1.67	12.80	20.30	0.93
3	4	15.50	15.40	1.48	12.70	19.95	0.93
4	4	15.20	15.20	1.33	12.60	19.50	0.91
5	4	14.85	14.90	1.22	12.40	19.05	0.91
0.5	5	15.36	14.60	2.60	11.00	26.50	0.92
1	5	15.40	15.10	2.00	11.30	21.50	0.93
2	5	15.46	15.30	1.57	12.70	19.80	0.92
3	5	15.22	15.18	1.40	12.60	19.55	0.92
4	5	14.89	14.90	1.27	12.50	19.10	0.91
5	5	14.50	14.53	1.15	12.10	18.55	0.90
0.5	7	15.18	14.50	2.45	11.00	26.00	0.92
1	7	15.12	14.80	1.80	12.00	20.50	0.92
2	7	15.07	14.90	1.49	12.40	19.45	0.92
3	7	14.82	14.70	1.35	12.30	19.20	0.91

estimation. Figure 1 also shows that implied volatilities for short-dated and long-dated swaptions do not always move in lockstep. A principal component analysis suggests that at least two factors drive innovations in the swaption implied volatilities.

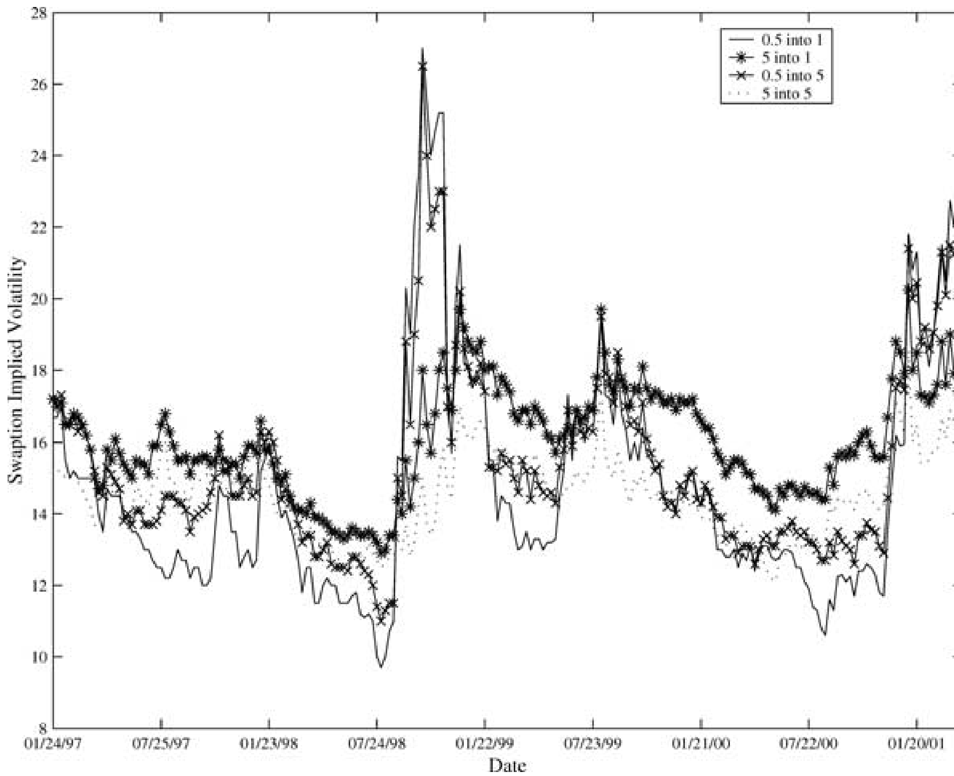

Figure 1. Time series of swaption implied volatilities. This figure plots the time series of implied volatilities for four at-the-money forward European swaptions. The data are mid-market quotes of annualized implied volatilities measured in percentage. The sample is weekly from January 24, 1997 to April 6, 2001.

IV. Econometric Method

The econometric exercise in my paper parallels many previous studies that use interest rate data to evaluate the performance of term structure models and infer dynamics of the yield curve factors. What is new here is that there are factors, besides the yield factors, that drive the evolution of the stochastic volatilities and correlations of bond yields. These volatility factors are not spanned by bonds. In general, there are parameters in models with unspanned stochastic volatility that are not identifiable from bond prices. Collin-Dufresne et al. (2004) isolate these parameters in the case of affine models. They find that interest rate volatility cannot be extracted from the cross section of bond prices. In this paper, I use panel data on both swaptions and bonds to estimate my model and infer the dynamics of the risk factors that drive the covariances of bond yields.

I estimate the model via the maximum likelihood approach. The estimated model parameters maximize the joint log-likelihood function of S_t, the prices of

34 European swaptions with total maturity no greater than 10 years, and D_t, the prices of 20 discount bonds with semiannual maturity between 6 months and 10 years. There are K latent volatility state variables ν_t that drive the covariances of bond yields. Since there are more swaptions than the number of latent state variables, I follow common practice (e.g., Chen and Scott (1993), Pearson and Sun (1994), and Duffie and Singleton (1997)) and assume that K swaptions S_t^1 are fitted exactly, while the other swaptions S_t^2 are observed with errors:[10]

$$S_t^1 = G(D_t, \nu_t; \tau_1, \Theta) \tag{16}$$

$$S_t^2 = G(D_t, \nu_t; \tau_2, \Theta)(1 + \epsilon_t), \tag{17}$$

where function G denotes the swaptions pricing formula given by (13), Θ is a vector of model parameters, and τ_1 and τ_2 are contract variables describing the exactly fitted swaptions and the remaining swaptions, respectively. The fitted swaption pricing errors ϵ_t in (17) depend on model parameters Θ, but they are assumed to be independent of the state vector and also to be independent over time.

By the assumptions above and the fact that (D_t, ν_t) follow jointly Markov processes in my model, the joint likelihood of bond prices and swaptions can be expressed as

$$\mathcal{L}(D_2, S_2, \ldots, D_T, S_T \mid D_1, S_1; \Theta) = \prod_{t=1}^{T-1} \mathcal{L}\big(D_{t+1}, S_{t+1}^1 \big| D_t, S_t^1; \Theta\big) \prod_{t=1}^{T-1} \mathcal{L}(\epsilon_{t+1}; \Theta).$$

Given a set of model parameters Θ, the volatility state variables can be recovered from the prices of the exactly fitted swaptions by inverting (16).[11] Through a change of variables from (D_{t+1}, S_{t+1}^1) to (D_{t+1}, ν_{t+1}), I obtain

$$\mathcal{L}(D_2, S_2, \ldots, D_T, S_T \mid D_1, S_1; \Theta) = \prod_{t=1}^{T-1} \mathcal{L}(D_{t+1}, \nu_{t+1} | D_t, \nu_t; \Theta) J_{t+1} \prod_{t=1}^{T-1} \mathcal{L}(\epsilon_{t+1}; \Theta), \tag{18}$$

where J_{t+1} is the Jacobian of the transformation. It is time dependent and a function of model parameters Θ. The Jacobian matrix of the transformation from (D_{t+1}, S_{t+1}^1) to (D_{t+1}, ν_{t+1}) is block lower-triangular, with an identity matrix in the upper left corner. Thus, J_{t+1} equals the inverse of the determinant of the matrix of first-order partial derivatives of S_{t+1}^1 with respect to each of the K volatility state variables ν_{t+1}.

[10] These measurement errors are included because without additional uncertainty, the model will imply deterministic relations for other swaptions that will almost surely be rejected by the data.

[11] It is straightforward to verify that swaption prices under my model are monotone functions of the volatility state variables. This guarantees a unique set of solutions for the inversion of volatility state variables from prices of exactly fitted swaptions.

Using the relations among the joint, conditional, and marginal densities, it follows that

$$\mathcal{L}(D_{t+1}, \nu_{t+1} \mid D_t, \nu_t; \Theta) = \mathcal{L}(D_{t+1}|\nu_{t+1}, D_t, \nu_t; \Theta)\, \mathcal{L}(\nu_{t+1} \mid D_t, \nu_t; \Theta). \tag{19}$$

I then take advantage of the known conditional densities of the volatility state variables as well as the closed-form expressions of bond price densities (conditional on the volatility state variables) to compute $\mathcal{L}(D_{t+1}, \nu_{t+1} \mid D_t, \nu_t; \Theta)$. The bond price dynamics in (6) imply that conditional on D_t and ν_t, D_{t+1} is well approximated by a lognormal distribution when the horizon of one period is short (e.g., an interval of 1 week in the data used to estimate the model). Thus,

$$\mathcal{L}(D_{t+1} \mid \nu_{t+1}, D_t, \nu_t; \Theta) = \frac{1}{(2\pi)^{10}\sqrt{\det(\Sigma_t)}\left(\prod_{k=1}^{20} D_{t+1}^k\right)} e^{-\frac{1}{2}x\Sigma_t^{-1}x'}, \tag{20}$$

where $\Sigma_t = \delta B \mathrm{Diag}(\nu_t) B'$, $\delta = 1/52$ year, B is the matrix of factor loadings for discount bonds, $\mathrm{Diag}(\nu_t)$ is a diagonal matrix whose main diagonal is ν_t, and

$$x = \log D_{t+1} - \log D_t - \left(r_t + B\gamma\nu_t - \frac{1}{2}\mathrm{diag}(B\mathrm{Diag}(\nu_t)B')\right)\delta,$$

where the operator diag extracts the main diagonal of a matrix.[12]

Under the $GA_{N,K}$ model, the K volatility state variables ν_t follow autonomous CIR-type dynamics. Therefore, the conditional density for each volatility state variable is noncentral chi-square, and

$$\begin{aligned}\mathcal{L}(\nu_{t+1} \mid D_t, \nu_t; \Theta) = \mathcal{L}(\nu_{t+1}|\nu_t; \Theta) &= \prod_{i=1}^{K} f(\nu_{i,t+1} \mid \nu_{i,t}) \\ &= \prod_{i=1}^{K} 2c_i e^{-u_i - w_i}\left(\frac{w_i}{u_i}\right)^{q_i/2} I_{q_i}(2\sqrt{u_i w_i}),\end{aligned} \tag{21}$$

where I_q is a qth-order modified Bessel function of the first kind,[13] and

$$c_i = \frac{2\hat{\kappa}_i}{\sigma_i^2(1 - e^{-\hat{\kappa}_i\delta})},\ u_i = c_i\nu_{i,t}e^{-\hat{\kappa}_i\delta},\ w_i = c_i\nu_{i,t+1},\ q_i = \frac{2\hat{\kappa}_i\hat{\theta}_i}{\sigma_i^2} - 1.$$

The model parameters $\hat{\kappa}$, $\hat{\theta}$, and σ govern the dynamics of the volatility state variables under the empirical measure (see equation (7)).

[12] Following Longstaff, Santa-Clara, and Schwartz (2001), I take r_t as the annualized yield on the 6-month discount bond (the shortest-maturity bond among the 20 discount bonds used in the model estimation).

[13] For the algorithms to compute Bessel functions of fractional order, I follow Chapter 6.7 in William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing* (ISBN 0-521-43108-5).

To summarize, by (18) and (19), the joint log-likelihood of swaptions and discount bonds can be written as

$$\begin{aligned}&\log \mathcal{L}(D_2, S_2, \ldots, D_T, S_T \mid D_1, S_1; \Theta)\\&\quad = \sum_{t=1}^{T-1} \{\log \mathcal{L}(D_{t+1} \mid \nu_{t+1}, D_t, \nu_t; \Theta) + \log \mathcal{L}(\nu_{t+1} \mid \nu_t; \Theta) + \log J_{t+1} + \log \mathcal{L}(\epsilon_{t+1})\}.\end{aligned}$$

The first two terms are given by (20) and (21), respectively. For the last term, I assume that the percent pricing errors of the non–exactly fitted swaptions ϵ are normally distributed with a covariance matrix Ω.[14] Thus,

$$\sum_{t=1}^{T-1} \log \mathcal{L}(\epsilon_{t+1}) = -\frac{(34-K)(T-1)}{2} \log(2\pi) - \frac{T-1}{2} \log|\Omega| - \frac{1}{2} \sum_{t=1}^{T-1} e'_{t+1} \Omega^{-1} e_{t+1}.$$

In carrying out the above maximum likelihood estimation, I need to choose the swaptions that are to be fitted exactly. Because each swaption provides information about the covariances of a particular segment of the yield curve, there is no good reason to choose one swaption over another to be fitted exactly. Thus, for all results reported in this paper, I invert the K volatility state variables under the $GA_{N,K}$ model by assuming that K portfolios of swaptions (corresponding to the first K principal components of the swaption implied volatilities) are fitted exactly, with measurement errors applying to other principal components. This approach not only circumvents the arbitrariness of fitting specific instruments exactly, it also approximately orthogonalizes the matrix of measurement errors.

V. Empirical Results

The empirical results presented below answer, among other things, the following questions: Can my model explain the relative valuation of swaptions and interest rate caps? How many yield factors and how many covariance factors are needed? How do the covariances of bond yields implied from the swaptions compare to their historical estimates based on bond prices alone?

A. Number of Factors Underlying the Swaptions Data

Table II reports the mean absolute swaption pricing errors under the estimated models for various specifications. The pricing error for a swaption is the difference between its fitted model price and its market price expressed as a percentage of the market price.

[14] Just like in the feasible General Least Squares, an estimate of Ω based on the fitted swaption errors is used in evaluating the log-likelihood function.

Table II

Swaption Pricing Errors

This table reports the mean absolute value of the percent pricing errors of the 34 at-the-money forward European swaptions under various model specifications. For the model corresponding to the column labeled by $(N, K), N$ factors drive the evolution of yield curve, of which the first K yield factors have unspanned stochastic volatility. All models are estimated using a panel data set consisting of 220 weekly observations on 34 swaptions from January 24, 1997 to April 6, 2001. The percent pricing error for a swaption is the difference between its fitted model price and its market price expressed as a percentage of the market price.

		Model					
Expiration	Tenor	(4,0)	(4,1)	(2,2)	(4,2)	(3,3)	(4,3)
0.5	1	13.84	7.96	11.79	4.40	4.92	4.08
1	1	11.18	4.31	7.64	3.28	2.91	2.85
2	1	10.29	2.96	4.13	3.15	3.23	3.30
3	1	10.00	2.95	3.87	2.77	2.76	2.71
4	1	10.13	3.57	4.45	3.28	3.09	3.06
5	1	10.18	4.15	3.59	3.09	2.92	2.89
0.5	2	12.56	5.02	3.21	2.05	3.01	1.69
1	2	10.10	2.25	3.04	2.74	3.81	3.10
2	2	9.05	1.47	1.80	1.72	1.90	1.95
3	2	8.82	1.73	1.47	1.34	1.37	1.32
4	2	8.83	2.33	1.46	1.47	1.32	1.32
5	2	8.78	3.05	1.21	1.33	1.31	1.25
0.5	3	12.14	4.15	7.08	2.52	2.88	2.38
1	3	9.69	1.59	4.36	2.30	2.50	2.54
2	3	8.77	1.59	1.34	1.05	1.06	1.07
3	3	8.66	2.03	0.92	1.09	1.02	1.03
4	3	8.70	2.63	0.72	1.08	0.93	0.97
5	3	8.45	3.31	2.01	1.23	1.23	1.20
0.5	4	11.44	3.63	5.73	2.30	2.05	1.95
1	4	9.38	1.66	2.93	1.67	1.66	1.72
2	4	8.72	2.32	0.94	1.23	1.35	1.25
3	4	8.74	2.81	1.03	1.39	1.38	1.40
4	4	8.58	3.29	1.00	1.22	1.20	1.20
5	4	8.14	3.98	3.39	2.00	1.96	1.98
0.5	5	10.82	3.62	3.48	2.85	2.65	2.55
1	5	9.48	2.17	1.97	1.82	1.82	1.80
2	5	8.81	3.09	1.43	1.59	1.66	1.61
3	5	8.76	3.60	1.29	1.55	1.57	1.58
4	5	8.36	4.14	2.02	1.82	1.83	1.82
5	5	7.80	4.65	5.09	3.09	2.93	3.07
0.5	7	10.73	4.12	4.83	4.75	4.70	4.72
1	7	9.28	2.83	2.97	2.63	2.64	2.66
2	7	9.00	3.83	2.10	2.05	2.07	2.08
3	7	8.66	4.30	1.86	1.94	1.95	1.93
Average		9.61	3.27	3.12	2.17	2.22	2.12

Under the $GA_{N,0}$ model, covariances of bond yields are deterministic. This model cannot fit the swaptions data well. Swaptions are systematically underpriced. Even when there are four factors driving the innovations in the yield curve ($N = 4$), the mean absolute swaption pricing error is 9.61% and the average root mean squared error (RMSE) is 10.32%. These errors are much higher than the typical bid–ask spread for at-the-money forward swaptions.[15]

Allowing stochastic covariances greatly improves models' fitting performance for swaptions. Table II shows that under the $GA_{4,1}$ model, the overall mean absolute swaption pricing error is reduced to 3.27% and the average (median) RMSE for the swaptions is 3.96% (3.57%). The swaption RMSE under $GA_{4,1}$ is smaller than that under $GA_{4,0}$ for all dates during the sample period. The difference is especially big in the months leading up to the LTCM crisis in the fall of 1998. Between January 1998 to the end of August 1998, the swaption RMSE under the constant covariance $GA_{4,0}$ model raises steadily from around 10% to over 30%, but it stays around 2.5% under the $GA_{4,1}$ model.

However, the $GA_{4,1}$ model, with only one factor driving the stochastic covariances of bond yields, has difficulty simultaneously fitting both short-maturity and long-maturity swaptions, especially during the LTCM crisis. The problem is that in the data, the implied volatilities for short-dated swaptions and long-dated swaptions usually move up and down by about the same amount, but sometimes there are periods during which this property breaks down. For example, Figure 1 shows that the implied volatility for the 0.5-year into 1-year swaption jumps from just above 10% before the LTCM crisis to more than 24% during the crisis, while the implied volatility for the 5-year into 1-year swaption only increases from about 13% to 17%. To match this feature of the data with only one volatility factor, its mean reversion speed must be high, implying that a volatility shock is expected to die out quickly, leading in turn to a much smaller impact on the implied volatilities for the long-dated swaptions than for the short-dated swaptions. Yet this would produce large swaption pricing errors during periods in which implied volatilities for short-dated and long-dated swaptions move up and down by about the same amount.

Table II shows that my models with multiple factors driving stochastic covariances of bond yields better fit the swaption data. The $GA_{4,2}$ and $GA_{4,3}$ models reduce the pricing errors of the $GA_{4,1}$ model for both short-dated and long-dated swaptions, especially for the short-dated swaptions. Under the $GA_{4,2}$ model, the mean absolute swaption pricing error is 2.17%, the median RMSE is 2.69%, and the maximum RMSE is 7.83%. Under the $GA_{4,3}$ model, the mean absolute swaption pricing error is 2.12%, the median RMSE is 2.59%, and the maximum RMSE is 7.73%. Under the $GA_{4,3}$ model, swaption RMSE stays below 4% except during the LTCM crisis (see Figure 2). For about two-thirds of the dates in my sample, the swaption root mean squared fitted error is lower than half of the bid–ask spread.

[15] The bid–ask spread for the swaption implied volatility is usually one volatility point (i.e., 1%) during my sample period. For a typical implied volatility of 16% for an at-the-money forward swaption, this translates into a bid–ask spread of about 6% for the swaption price, since the price of an at-the-money forward swaption is practically linear in its Black implied volatility.

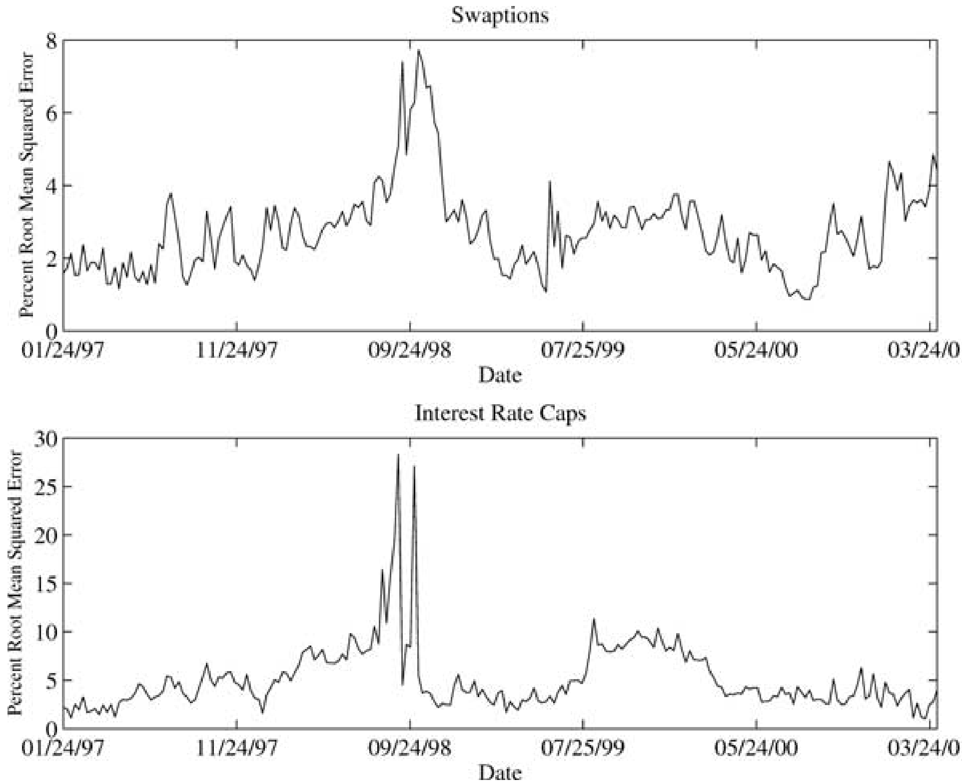

Figure 2. Time series of pricing errors for swaptions and interest rate caps. This graph plots the time series of root mean squared pricing errors (RMSE) for swaptions and interest rate caps under the $GA_{4,3}$ model. The model is estimated via maximum likelihood using 34 at-the-money forward European swaptions. Panel A plots the RMSE of these 34 swaptions. Panel B plots the RMSE for interest rate caps with maturities of 2, 3, 4, 5, 7, and 10 years. For each interest rate cap, the pricing error is calculated as the difference between its no-arbitrage value implied from the swaptions according to the $GA_{4,3}$ model and its market price, expressed as a percentage of the market price. The data set consists of 220 weekly observations on each series from January 24, 1997 to April 6, 2001.

The swaption pricing errors in previous studies are substantially higher, even when the models are frequently recalibrated in these studies in order to fit the data better. For example, De Jong et al. (2001) find that for a variety of Libor and swap market model specifications, the mean absolute pricing error ranges from about 5% to 14%. Driessen et al. (2003) report average absolute swaption pricing errors between 7.23% and 9.50% for several HJM-type models. Jagnnathan et al. (2003) report swaption pricing errors that are much larger than the bid–ask spread. For example, under the two-factor CIR model, the mean absolute pricing error is about 31.39% for the 2-year into 5-year swaption.

Among the previous studies, Longstaff et al.'s (2001) four-factor string market model generates the best fit to the swaptions data. The median swaption RMSE under their model is 3.10%, but it is over 15% during the LTCM crisis. Their model assumes constant covariances of bond yields (with constant time to maturities), but it is recalibrated each date. This effectively allows four factors to drive the covariances of bond yields. In contrast, I hold my model parameters fixed and explicitly model the dynamics of the risk factors that drive the covariances of bond yields. Although in my $GA_{4,2}$ model only two factors drive the covariances of bond yields, it is able to fit the swaption data better, both under normal market conditions and during market crises.[16]

Table III reports the Diebold and Mariano (1995) test statistics for pairwise comparisons of alternative specifications of my model.[17] Each model specification implies a time series of mean squared swaption pricing errors $SSE(t)$ corresponding to the optimized model parameters: $SSE(t) = \frac{1}{34}\sum_{i=1}^{34}\epsilon_{i,t}^2$, where $\epsilon_{i,t}$ is the percent pricing error for the i^{th} swaption on date t. Consider two models that give rise to $\{SSE_1(t)\}_{t=1}^T$ and $\{SSE_2(t)\}_{t=1}^T$, respectively ($T = 220$ in my case). The null hypothesis that the two models have the same pricing accuracy is equivalent to the null hypothesis that the population mean (μ) of the difference in pricing errors (d_t) is zero, where $d_t = SSE_1(t) - SSE_2(t)$. Diebold and Mariano (1995) show that if $\{d_t\}_{t=1}^T$ is covariance stationary and has short memory, then the asymptotic distribution of the sample mean of pricing error difference $\bar{d} = \frac{1}{T}\sum_{t=1}^T d_t$ is normally distributed as

$$\sqrt{T}(\bar{d} - \mu) \sim \mathcal{N}(0, 2\pi h_d(0)),$$

where $h_d(0)$ is the spectral density of the pricing error difference; that is,

$$h_d(0) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} r_d(\tau), \quad \text{where} \quad r_d(\tau) = \mathrm{E}[(d_t - \mu)(d_{t-\tau} - \mu)].$$

The formula for $h_d(0)$ shows correction for serial correlation in d_t. Diebold and Mariano (1995) obtain a consistent estimate of $2\pi h_d(0)$ by the sum of available autocovariances up to a certain truncation lag q:[18]

$$2\pi\hat{h}_d(0) = \sum_{\tau=-q}^{q}\hat{r}_d(\tau), \quad \text{where} \quad \hat{r}_d(\tau) = \frac{1}{T}\sum_{t=|\tau|+1}^{T}(d_t - \bar{d})(d_{t-|\tau|} - \bar{d}).$$

[16] The sample in Longstaff, Santa-Clara, and Schwartz (2001) coincides with the first 128 weeks of my sample. Over this period, the median and the maximum swaption RMSEs under my $GA_{4,2}$ model are, respectively, 2.61% and 7.83%.

[17] The Diebold and Mariano (1995) method allows the pricing errors to be non-Gaussian, nonzero mean, serially correlated, and contemporaneously correlated. It is applicable in nonstandard testing situations, such as when a nuisance parameter is not identified under the null (see page 262 of Diebold and Mariano (1995)). I verify that the same inferences as those reported in Table III also obtain based on the log-likelihood ratio tests.

[18] In Table III, lag order $q = 52$. I verify that my results are robust to alternative lag orders q = 5, 10, 26, and 102.

Table III

Tests of Alternative Model Specifications

This table reports pairwise comparisons of the pricing accuracy of alternative specifications of my model. All models are estimated using a panel data set consisting of 220 weekly observations on 34 swaptions and 20 discount bonds from January 24, 1997 to April 6, 2001. For the model labeled (N, K), N factors drive the evolution of the term structure of bond yields, of which the first K factors have unspanned stochastic volatility, and the other $N - K$ factors have constant volatility. The "DM" column reports the Diebold and Mariano (1995) statistic that tests the null hypothesis that the models in the first two columns have the same pricing accuracy for swaptions. Under the null hypothesis, the Diebold and Mariano statistic is distributed asymptotically as standard normal $\mathcal{N}(0, 1)$. Column four reports the number of weeks (out of a total of 220 weeks) that Model 1 implies a higher root mean squared percent pricing error (RMSE) for the cross section of swaptions, and column five reports the average reduction in swaption RMSE (reported in percent) implied by Model 2 during these weeks. Similarly, the sixth column reports the number of weeks that Model 2 implies a higher swaption RMSE, and the last column reports the average reduction in swaption RMSE (reported in percent) implied by Model 1 during these weeks.

			$RMSE_2 < RMSE_1$		$RMSE_2 > RMSE_1$	
Model 1	Model 2	*DM* Statistic	No. of weeks	Mean Difference	No. of weeks	Mean Difference
(1,1)	(2,1)	8.05	212	1.80	8	0.23
(2,1)	(3,1)	5.29	193	1.52	27	0.47
(3,1)	(4,1)	2.00	149	0.12	71	0.08
(2,1)	(2,2)	2.77	213	1.05	7	0.12
(2,2)	(3,2)	5.82	210	1.45	10	0.68
(3,2)	(4,2)	2.54	167	0.16	53	0.07
(3,1)	(3,2)	2.78	205	1.18	15	0.19
(3,2)	(3,3)	1.23	119	0.33	101	0.25
(3,3)	(4,3)	6.26	188	0.18	32	0.14
(4,0)	(4,1)	3.32	206	6.80	14	0.28
(4,1)	(4,2)	2.81	209	1.21	11	0.21
(4,2)	(4,3)	2.57	121	0.27	99	0.13
(4,3)	(4,4)	1.32	137	0.31	83	0.25

Thus, under the null hypothesis of equal pricing accuracy across the two models, the Diebold and Mariano statistic $\sqrt{T}\bar{d}/\sqrt{2\pi \hat{h}_d(0)}$ is distributed asymptotically as standard normal.

Table III shows that the largest improvement in model performance in terms of fit to the swaptions data comes from increasing the number of yield factors from one to two and from two to three, and also from increasing the number of stochastic volatility factors from zero to one and from one to two. In each pair of such comparisons, the Diebold and Mariano statistic is highly significant, and the model with more yield factors or volatility factors reduces the swaption RMSE of the alternative model by more than 1% at least 200 weeks out of a total of 220 weeks. Table III confirms the earlier result that at least two stochastic volatility factors are needed to explain the time-series and cross-sectional variations in the swaptions data. In addition, the table shows the importance of having a sufficient number of yield factors. In all comparisons of the $GA_{N_1,K}$ model with the $GA_{N_2,K}$ model, where $N_1 < N_2$ and $K = 1, 2, 3$, the $GA_{N_2,K}$ model performs significantly better. Overall, the $GA_{4,3}$ model provides the best fit to the swaptions data.

The results in Table III are consistent with Heidari and Wu (2003), who find by principal component analysis that one needs three volatility factors that are independent of the underlying yield curve to capture the movement of the swaption implied volatility surface. However, their principal component analysis relies purely on the unconditional covariance matrix of the swaption implied volatilities, completely ignoring the rich information contained in the cross section of swaptions that is crucial for the estimation of the covariance dynamics in my model. Furthermore, I obtain estimates of conditional covariances of bond yields and study the relative valuation of swaptions and interest rate caps. It is impossible to obtain these results using simple principal component analysis.

B. Relative Valuation of Swaptions and Caps

The relative valuation of swaptions and interest rate caps depends crucially on the covariances of bond yields, and thus provides a diagnostic test of my model. This is an out-of-sample test in the sense that I first estimate my models with the swaptions data, and I then compare the cap prices under the estimated models to the market prices of caps. Below, the pricing error for an interest rate cap is defined as the difference between its no-arbitrage model price implied from the swaptions and its market price expressed as a percentage of the market price.

Table IV reports the mean and mean absolute pricing errors of interest rate caps. The constant-covariance model $GA_{4,0}$ implies large mean absolute pricing error (11.41%). Just like the case for swaptions, introducing stochastic covariances of bond yields substantially reduces the pricing errors for caps. The mean absolute cap pricing error is 5.42% under $GA_{4,1}$, 5.07% under $GA_{4,2}$, and 4.59% under the $GA_{4,3}$ model. Again, it is important to include enough yield factors. For example, the average cap pricing errors are 0.62%, 1.34%, and 0.36% for the $GA_{4,1}$, $GA_{4,2}$ and $GA_{4,3}$ models, respectively. In contrast, the $GA_{2,2}$ and $GA_{3,3}$ model both generate large average pricing errors for the caps. Furthermore, the $GA_{4,1}$ model implies much smaller mean absolute cap pricing errors than the $GA_{2,2}$ model, although the latter has one more volatility factor and fits (in-sample) the swaptions data better. Overall, the $GA_{4,3}$ model, which is favored by the statistical tests based on the various models' fit to the swaptions, also implies the smallest pricing errors for the interest rate caps, and thus underlying the prices of swaptions and interest rate caps is a high-dimensional stochastic covariance matrix of bond yields.

Figure 2 shows that the differences between the no-arbitrage values of caps implied from the swaptions according to my $GA_{4,3}$ model and their market prices are typically less than 6% of the market prices (which is about the size of the bid–ask spread). There are only two brief periods during which cap RMSE

Table IV

Interest Rate Caps Pricing Errors

This table reports the average and mean absolute pricing errors for the interest rate caps of maturity 2, 3, 4, 5, 7, and 10 years under various model specifications. For the model labeled (N, K), N factors drive the evolution of bond yields, of which the first K factors have unspanned stochastic volatility, and the other $N-K$ factors have constant volatility. The pricing error for an interest rate cap is defined as the difference between its no-arbitrage model price implied from the swaptions and its market price expressed as a percentage of the market price. The model cap price is obtained from equation (15) using the model parameters (and implied volatility state variables) estimated from the swaption and bond prices data. The data consist of 220 weekly observations on each series from January 24, 1997 to April 6, 2001.

	Mean Percent Pricing Errors						Mean Absolute Percent Pricing Errors					
Maturity	(4,0)	(4,1)	(2,2)	(4,2)	(3,3)	(4,3)	(4,0)	(4,1)	(2,2)	(4,2)	(3,3)	(4,3)
2	−5.93	5.96	−12.60	6.99	−4.03	4.46	11.47	8.42	12.94	7.87	5.30	6.06
3	−7.25	0.77	−11.80	1.75	−5.18	0.35	11.79	5.20	12.03	4.94	5.78	4.33
4	−5.92	0.06	−9.89	0.83	−4.32	−0.09	11.73	4.75	10.14	4.70	5.30	4.42
5	−6.16	−1.33	−9.70	−0.77	−4.76	−1.38	11.80	4.60	9.91	4.49	5.53	4.42
7	−4.90	−1.30	−7.40	−0.92	−3.71	−1.22	11.36	4.79	7.66	4.39	5.00	4.33
10	−2.77	−0.41	−3.57	0.19	−1.98	0.04	10.29	4.78	4.66	4.06	4.20	3.99
Average	−5.49	0.62	−9.16	1.34	−4.00	0.36	11.41	5.42	9.56	5.07	5.19	4.59

exceeds 6%. The first occurs during the LTCM crisis.[19] The second starts at the beginning of August 1999, when, according to my $GA_{4,3}$ model, the no-arbitrage prices of interest rate caps implied from the swaptions are lower than the market prices of caps by about 8%. Supporting the view that the market overvalues interest rate caps relative to swaptions during this period, various news reports indicate that proprietary desks are short cap volatility and long swaption volatility.[20] The cap pricing errors revert to around 3% (about half of the bid–ask spread) by the end of 1999.

To summarize, under my $GA_{4,3}$ model, the market prices of interest rate caps do not deviate significantly from their no-arbitrage values implied from the swaptions. In contrast, Longstaff et al. (2001) find periodically large and somewhat bimodally distributed pricing errors for interest rate caps. The average cap pricing error is 10.38%, and the mean absolute cap pricing error is 14.19%. The smallest mean absolute pricing error for caplets among all the Libor and

[19] For example, the cap RMSE jumps to 19.9% on August 28, 1998 and 30.4% on September 4, 1998, right after *The Times* headlined "Meriwether fund plummets by 44% (http://www.thetimes.co.uk/)." For the next 4 weeks, the average cap pricing error stays below 6% until it shoots up again to 25.8% on October 2, 1998. During that week, many banks revealed huge exposure to or investment in the hedge fund, which led to a 3.5 billion bailout of LTCM.

[20] However, this position is risky. On August 5, 1999, Dow Jones Newswire reports that two major U.S. investment banks lost millions of dollars when they had to unwind big swaption positions and cover short cap positions. Interestingly, this correlation play on the forward rates was reported to be put back on just 3 weeks later.

swap market models tested in De Jong et al. (2001) is about 20%. Driessen et al. (2003) find that the average absolute pricing errors for several multifactor HJM models are all greater than 10%. Jagnnathan et al. (2003) find that even the smallest mean absolute pricing error for caps under one-, two-, or three-factor CIR models are still much larger than the bid–ask spreads. My model seems to realistically capture the dynamics of bond covariances and mostly eliminates the large and systematic pricing errors between swaptions and caps that obtain in the previous studies above.

As an additional test of model goodness of fit, I regress changes in the market prices of swaptions and interest rate caps on changes in their prices under the estimated $GA_{4,3}$ model. The regression is run individually for each swaption and interest rate cap series. If my model fits the data well, then the estimated intercept should be very close to zero, the slope coefficient should be indistinguishable from one, and the R^2's of the regressions should be high. These predictions are confirmed. I find that none of the intercept terms are even close to being statistically different from zero. The average estimated slope coefficient is not significantly different from one. The average R^2 is 0.91 for the swaptions and 0.78 for the interest rate caps.

C. Volatility Factors and Implied Bond Covariances

Through the model estimation, I extract the market's view about the dynamics of the volatility factors and obtain estimates of conditional bond covariances. Here I compare the implied volatilities and correlations of bond yields to their historical estimates (both unconditional and conditional). All results below are based on the estimated $GA_{4,3}$ model.

Three factors drive the covariances of bond yields under the $GA_{4,3}$ model. These factors can be interpreted as the variances of the first three common factors that drive innovations in the yield curve.[21] Based on the parameter estimates reported in Table V, Panel A, the volatility of the first yield factor seems quite persistent while the stochastic volatility of the second and third yield factors reverts to the mean much faster. The "half-life," defined as $\frac{\log 2}{\kappa}$, where κ is the mean reversion speed, measures the time when a mean-reverting variable is expected to reach a value that is halfway between the current level and its long-run mean. The half-life of the first volatility factor is about 6 years under the empirical measure, while it is only about 3 to 4 months for the second and third volatility factors.

The following observations shed light on the economic meaning and impact of the volatility factors. First, the long-dated swaptions are mainly influenced by the first volatility factor. Second, swaptions of different maturities and tenors

[21] Consistent with many previous studies, I find that the first factor represents an approximately parallel shift of the yield curve. The second factor twists the yield curve by moving the short-term yield and the longer-term yield in opposite directions. The third factor bows the yield curve and increases its curvature.

Table V

Maximum Likelihood Estimation of the $GA_{4,3}$ Model

This table reports the results of maximum likelihood estimation of the $GA_{4,3}$ model. Four factors drive the innovation of bond yields. The first three factors display unspanned stochastic volatility, and the fourth factor has constant volatility. The model is estimated with a panel data set of 220 weekly observations (from January 24, 1997 to April 6, 2001) on 34 at-the-money forward European swaptions and 20 discount bonds with semiannual maturity no greater than 10 years. Panel A reports parameter estimates for the $GA_{4,3}$ model. Parameters κ_i, θ_i, and $\sigma_i (i = 1, 2, 3)$ govern the risk-neutral dynamics of the stochastic volatility state variables, and the λ_i's are the market prices of volatility risk. The γ_j's ($j = 1, 2, 3, 4$) are the market prices of risk for the yield factors. Standard errors of parameter estimates (reported in the parentheses) are obtained by calculating the inverse of the information matrix based on the Hessian of the likelihood function. Panel B reports the sample correlations between the yield factors and each volatility state variable, as well as the correlations between the implied volatility state variables. Panel C reports the sample mean of the implied variances of the first three yield factors and their unconditional variances.

Panel A: Parameter Estimates

Factor	κ	$10^4\theta$	σ	λ	$\gamma/100$
1	0.0516	3.5540	0.0059	−0.0648	−0.0156
	(0.0030)	(0.6098)	(0.0031)	(0.0631)	(0.0691)
2	1.7482	0.4347	0.0123	−0.0746	0.3084
	(0.0477)	(0.0241)	(0.0053)	(0.0569)	(0.2963)
3	2.4374	0.2030	0.0096	−0.2227	−5.7150
	(0.1853)	(0.0195)	(0.0046)	(0.1598)	(1.3510)
4	–	0.2617	–	–	−24.3746
	–	(0.0095)	–	–	(13.0733)

Panel B: Correlations between Yield Factors and Volatility Factors

	Yield Factors				Volatility Factors		
	z_1	z_2	z_3	z_4	ν_1	ν_2	ν_3
z_1	1						
z_2	−0.0000	1					
z_3	−0.0000	0.0000	1				
z_4	−0.0000	−0.0000	0.0000	1			
ν_1	−0.0288	0.0065	−0.0162	−0.0573	1		
ν_2	−0.0095	−0.0310	0.0595	0.0468	0.0868	1	
ν_3	0.0153	−0.0174	−0.1062	−0.0292	0.1083	−0.0956	1

Panel C: Unconditional and Average Implied Factor Variance

	$10^4\nu_1$	$10^4\nu_2$	$10^4\nu_3$
Historical	4.5837	0.6165	0.1594
Average Implied	4.4493	0.6411	0.1778

are about equally sensitive to innovations in the first volatility factor. The first volatility factor therefore proxies for the average level of the surface of swaption implied volatilities. Third, the second volatility factor proxies for the spread between the implied volatilities of short- and long-dated swaptions of the same

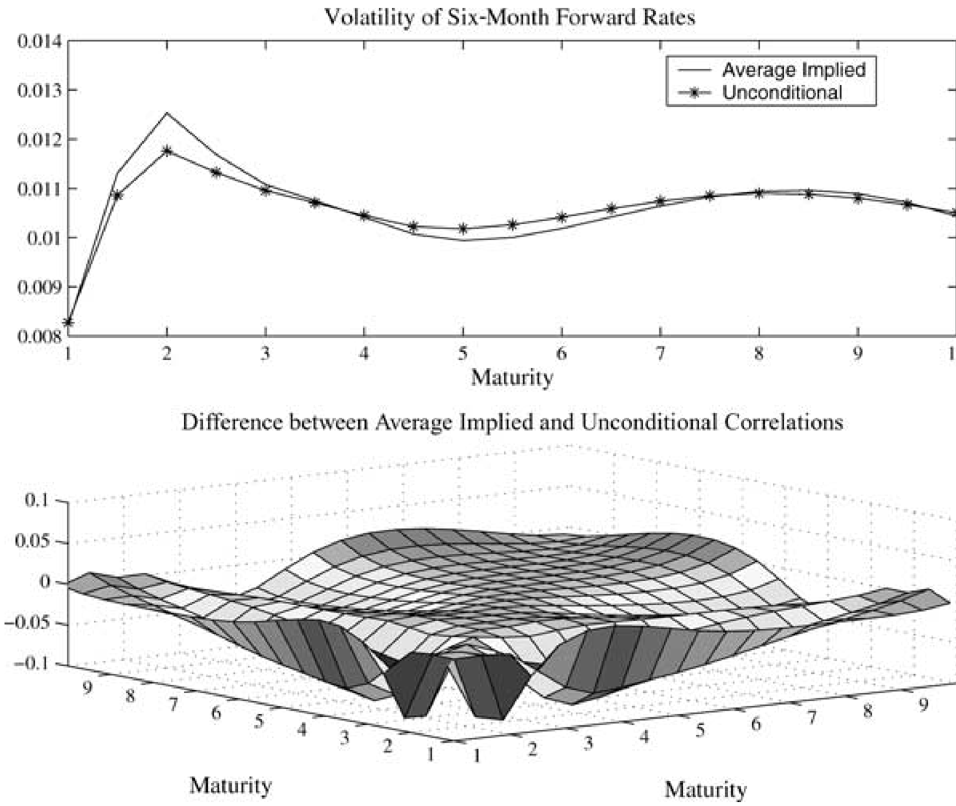

Figure 3. Average implied covariances versus unconditional covariances. This graph compares the time-series average of the swaption-implied conditional covariances of the 6-month forward rates (up to 10 years in maturity) under the $GA_{4,3}$ model to their unconditional estimates. Panel A plots the average implied volatility and the sample standard deviation of the 6-month forward rates (both annualized). Panel B plots the difference between the average implied correlations and the sample correlations based on weekly data from January 24, 1997 to April 6, 2001. The implied volatilities and implied correlations are calculated according to the covariance structure specified in equation (11) under the estimated $GA_{4,3}$ model.

tenor.[22] Finally, only short-dated swaptions on 1-year and 2-year swaps are sensitive to the third volatility factor.

Panel B of Table V reports the sample correlations between the yield factors and the volatility factors under the $GA_{4,3}$ model. It strongly supports my model's assumption that shocks to the yield factors are uncorrelated with their stochastic volatilities. Panel C of Table V shows that the time-series average of the implied variances of the yield factors are close to the unconditional variances of the yield factors. Figure 3 shows that the average implied covariances

[22] The sample correlations between implied ν_2 and the difference between the Black implied volatilities for 0.5-into T- and 5-into T-year swaptions are all about 0.95, for $T = 1, \ldots, 5$.

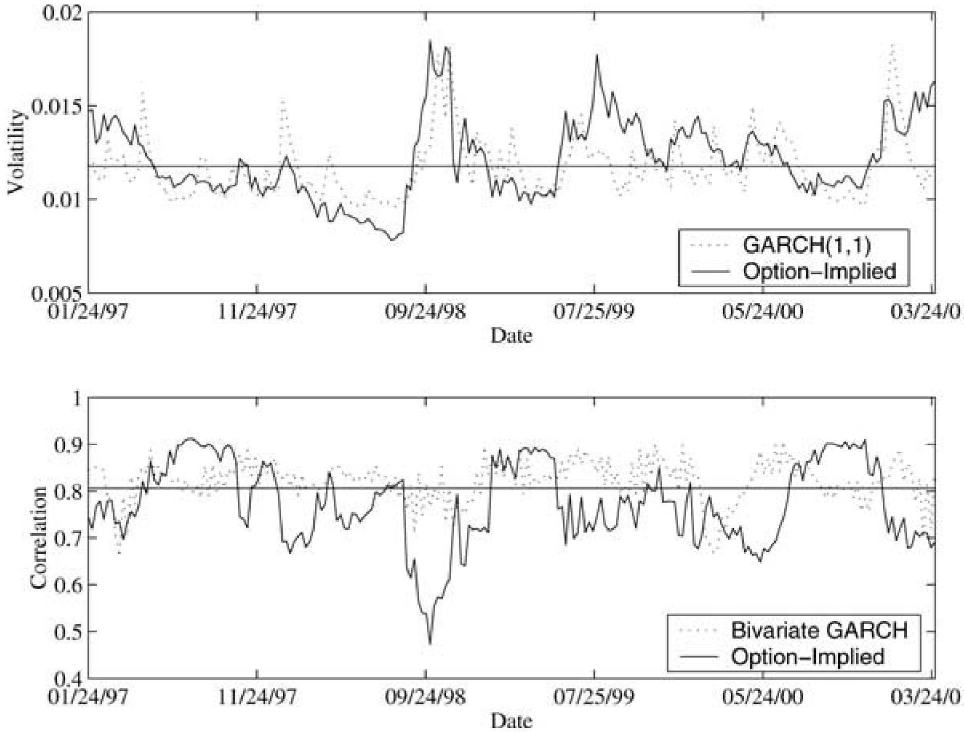

Figure 4. Conditional covariances: GARCH versus swaption-implied. This graph compares the conditional volatilities and correlations of the 6-month Libor forward rates implied from the swaptions according to my $GA_{4,3}$ model to those estimated from the historical bond prices according to Engle (2002)'s GARCH model with dynamic conditional correlations. Panel A illustrates two estimates of the volatility of the 6-month forward rate with 2-year maturity. Panel B plots two estimates of the correlation between the 2-year maturity and 5-year maturity 6-month forward rates. In both panels, the horizonal line marks the unconditional estimate. The data sample is weekly from January 24, 1997 to April 6, 2001.

of bond yields are also very close to their unconditional historical estimates.[23] These results are consistent with the insignificant risk premia for the stochastic volatilities of the common yield factors (Table V, Panel A).

Finally, Figure 4 compares the implied covariances of forward bonds to the conditional covariances obtained from the Dynamic Conditional Correlation (DCC) model, a new class of multivariate GARCH models proposed by Engle (2002). This model provides a very good approximation to a variety of time-varying correlation processes, yet it still has the flexibility of univariate GARCH model. I apply this model to changes in the logarithm of 6-month

[23] The only noticeable difference occurs for the correlations of bonds with maturity below 3 years, for which the implied correlations are lower than the historical correlations. But most of the differences are smaller than 0.05. The maximum difference is 0.08.

forward bonds. Conditional covariances are estimated following a two-step procedure as described in detail by Engle and Sheppard (2001).

Empirically, the instantaneous volatilities of 6-month forward rates implied from the swaptions closely track their GARCH(1,1) estimate. For example, in the case of 6-month forward rate with 2-year maturity plotted in the top panel of Figure 4, the correlation between the two conditional volatility estimates is 0.62. There is no systematic difference between the two conditional volatility estimates. My model with unspanned stochastic volatility captures interest rate volatility quite well. In contrast, Collin-Dufresne et al. (2004) show that under certain affine models, which do not allow for unspanned stochastic volatility, the interest rate volatility extracted from the cross section of bond prices is strongly negatively correlated with the GARCH estimate.

The bottom panel in Figure 4 plots two estimates for the correlation between the 6-month forward rates with 2- and 5-year maturities. The correlation implied from the swaptions is more volatile than that estimated from the bond prices under the DCC model. The sample correlation between the two correlation series is only 0.15. During periods of market crises, the implied correlation drops precipitously but this is not reflected in the DCC estimate. The information content of implied correlations deserves further investigation in future studies.

VI. Conclusion

Recent studies find that it is difficult for many popular term structure models to explain the valuation of swaptions and interest rate caps, although these models can price the underlying bonds very well or even fit them exactly by design. A potential limitation of these models is that the covariances of bond yields are either deterministic, or can be completely hedged by the underlying bonds. Collin-Dufresne and Goldstein (2003), Dai and Singleton (2003), and Jagnnathan et al. (2003) conjecture that the resolution of the swaptions and caps valuation puzzle may require models that can accommodate more general forms of stochastic volatility and time-varying correlations, for instance, by allowing for volatility factors that do not affect bond prices.

In this paper, I develop and estimate a term structure model that allows separate factors to drive innovations in bond yields and their stochastic volatilities and correlations. The covariances of bond yields are affine in a set of volatility state variables that are not spanned by bonds. Empirically, my model explains the cross-section and the time-series variation of the swaption implied volatilities very well, and reconciles the relative valuation of swaptions and interest rate caps. These results support the conjectures above.

My findings suggest that further exploration of the class of term structure models proposed here as well as by Collin-Dufresne and Goldstein (2003) and Kimmel (2004) is likely to contribute to a better understanding of the rich dynamics of bond volatilities and correlations as well as the valuation of interest rate derivatives.

Appendix: Model Valuation of Swaptions and Caps

Proof of Proposition 1 (valuation of a τ-year by T-year European swaption under the $GA_{N,K}$ model): (I continue to use the same notation as in Section II.B.) Under the forward risk-neutral measure Q^{τ}, $S_j(t)$ defined in equation (8) is a martingale, since it is a constant multiple of the forward discount bond price $D(t, \tau, \tau_j)$, $\tau_j = \tau + \frac{j}{2}$. Let $S(t)$ be a vector of length $2(T - \tau)$ that stacks $S_j(t)$, $j = 1, \ldots, 2(T - \tau)$. Denote by Σ_t the instantaneous covariance matrix of $d\log S(t)$; Σ_t can be expressed in terms of both U, the historical eigenvector matrix of 20 forward Libor bonds with semiannual maturities over the term of 10 years, and ν_k ($k = 1, \ldots, N$), the instantaneous variances of the common yield factors. At date $t < \tau = i/2$ year, let integer m be such that $m/2 \leq t < (m + 1)/2$. Then

$$\Sigma_t = (1 - 2(t - m/2))A_1 U\,Diag(\nu_t)U' A_1' + 2(t - m/2)A_2 U\,Diag(\nu_t)U' A_2', \tag{A1}$$

where $Diag(\nu_t)$ is a diagonal matrix whose main diagonal elements are the ν_t's, and A_1 and A_2 are $2(T - \tau) \times 20$ matrix. All elements of A_1 (resp., A_2) are zero, except that the $2(T - \tau) \times 2(T - \tau)$ submatrix consisting of the $i - m + 1$ column to the $i - m + 2(T - \tau)$ column (resp., the $i - m$ column to the $i + 2(T - \tau) - 1 - m$ column) is a lower triangular matrix whose elements on and below the main diagonal are one. Each covariance term in Σ_t is linear in the volatility state variables.

Let R_t be the Cholesky decomposition of Σ_t satisfying $R_t R_t' = \Sigma$, and let $dZ_t^{Q^{\tau}}$ be a standard Brownian motion of dimension $2(T - \tau)$ under the forward measure Q^{τ}. Then

$$\frac{dS(t)}{S(t)} = R_t dZ_t^{Q^{\tau}}.$$

By Ito's lemma, $\tilde{G}_t = \prod_{j=1}^{2(T-\tau)} S_j^{\tilde{\omega}_j}(t)$ satisfies

$$d\log\tilde{G}_t = \sum_{j=1}^{2(T-\tau)} \tilde{\omega}_j d\log S_j = -\frac{1}{2}\tilde{\omega}' diag(\Sigma)\,dt + \tilde{\omega}' R_t dZ_t^{Q^{\tau}},$$

where *diag* is an operator that maps a matrix to a column vector consisting of the main diagonal elements of the matrix. Recall that by equation (12), the price of the τ-year by T-year European swaption is determined by

$$\mathrm{E}^{Q^{\tau}}[\mathrm{Max}\,(\tilde{G}_{\tau} - g, 0)], \tag{A2}$$

where $g = 1 + \mathrm{E}^{Q^{\tau}}[\tilde{G}_{\tau} - \tilde{A}_{\tau}]$, and $\tilde{A}_{\tau} = \sum_{j=1}^{2(T-\tau)} \tilde{\omega}_j S_j(\tau)$ is an arithmetic sum of the $S_j(\tau)$'s that correspond to their geometric sum $\tilde{G}_{\tau}$. Note that $\mathrm{E}^{Q^{\tau}}[\tilde{A}_{\tau}] = 1$, since each S_j is a martingale under Q^{τ} and $S_j(0) = 1$ by definition. Hence, $g = \mathrm{E}^{Q^{\tau}}[\tilde{G}_{\tau}]$.

I compute (A2) via the law of iterated expectations by first conditioning on a vector $\bar{V}$ that stacks the average values of $\nu_k(t)$ over the time interval $[0, \tau]$, $k = 1, \dots, N$. Based on the assumption that the volatility state variables are instantaneously uncorrelated with innovations in the yield curve, and following an argument of Hull and White (1987) (see their lemma 1), $\tilde{G}_\tau$ is lognormally distributed conditional on $\bar{V}$:

$$\tilde{G}_\tau = \exp(\varepsilon_\tau - \mu\tau), \tag{A3}$$

where $\mu = -\frac{1}{2}\tilde{\omega}' diag(\bar{\Sigma})$, ε_τ is a normal random variable with mean zero and variance $(\tilde{\omega}'\bar{\Sigma}\tilde{\omega})\tau$, and $\bar{\Sigma}$ is the average covariance matrix of the forward bonds $\{D(t, \tau, \tau_j)\}_{j=1,\dots,2(T-\tau)}$ over horizon $[0, \tau]$. To finish the evaluation of (A2), I use the following lemma:

LEMMA 1: *Let ξ be a Gaussian random variable with zero mean and variance σ^2. For any strictly positive numbers a and b,*

$$E(\mathrm{Max}(ae^{\xi - 0.5\sigma^2} - b, 0)) = a\mathcal{N}(h) - b\mathcal{N}(h - \sigma),$$

where $h = \sigma^{-1} ln(a/b) + 0.5\sigma$ and $\mathcal{N}(-)$ is the cumulative density function of the standard normal random variable.

Applying lemma 1 to $\tilde{G}(\tau)$ as in (A3), the price of a τ by T at-the-money forward European swaption is

$$\begin{aligned} P &= D(0,\tau)\, \mathrm{E}_{\bar{V}}^{Q^\tau}(\mathrm{E}[\mathrm{Max}(\tilde{G}_\tau - g, 0)|_{\bar{V}}]) \\ &= D(0,\tau)\, \mathrm{E}_{\bar{V}}^{Q^\tau}\left[\mathcal{N}\left(\frac{1}{2}\sqrt{\tilde{\omega}'\bar{\Sigma}\tilde{\omega}\tau}\right) - \mathcal{N}\left(-\frac{1}{2}\sqrt{\tilde{\omega}'\bar{\Sigma}\tilde{\omega}\tau}\right)\right] \\ &= D(0,\tau)\left[2\mathcal{N}\left(\frac{1}{2}\sqrt{\tilde{\omega}'\, \mathrm{E}_{\bar{V}}^{Q^\tau}[\bar{\Sigma}]\, \tilde{\omega}\tau}\right) - 1)\right]. \end{aligned}$$

In the last equality, I use the well-known fact that the price of an at-the-money option is practically linear in its Black implied volatility. Finally, to compute $\mathrm{E}_{\bar{V}}^{Q^\tau}[\bar{\Sigma}]$, I use equation (A1) and the following formula for the expected value of the volatility state variables ν_j:

$$\mathrm{E}_0[\nu_j(t)] = e^{-\kappa_j t}\nu_j(0) + (1 - e^{-\kappa_j t})\theta_j.$$

Proposition 1 follows immediately after some algebra. Q.E.D.

Proof of Proposition 2 (valuation of a T-year interest rate cap under the $GA_{N,K}$ model): A T-year interest rate cap on the 6-month Libor rate with cap rate R consists of $2T - 1$ caplets C_i that mature at year $t_i = \frac{i}{2}$, $i = 1, \dots, 2T - 1$. Each caplet C_i protects the 6-month Libor rate L_i applicable over the horizon $[t_i, t_{i+1}]$

from rising above R. Its pays off $\frac{1}{2}\max(L_i - R, 0)$ per unit dollar of principal at date t_{i+1}. The payoff of caplet C_i discounted to date t_i is

$$\begin{aligned}\frac{1}{2}\frac{1}{1+0.5L_i}\mathrm{Max}(L_i - R, 0) &= (1+0.5R)\mathrm{Max}\left(\frac{1}{1+0.5R} - \frac{1}{1+0.5L_i}, 0\right)\\ &= (1+0.5R)\mathrm{Max}\left(\frac{1}{1+0.5R} - D(t_i, t_i, t_{i+1}), 0\right).\end{aligned}$$

Hence, each caplet C_i is a European put option on the forward Libor bond $D(t, t_i, t_{i+1})$ with a strike price of $\frac{1}{1+0.5R}$. I value each caplet C_i under the corresponding forward measure Q^{t_i} following the same procedures as in the proof of Proposition 1. The derivation is easier since each caplet is a European option on a single forward bond, rather than an option on a portfolio of forward bonds as in the case of a swaption. Q.E.D.

REFERENCES

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

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

Bliss, Robert R., 1997, Movements in the term structure of interest rates, *Federal Reserve Bank of Atlanta Economic Review*, Fourth Quarter, 16–33.

Chen, Ren-Raw, and Louis Scott, 1993, Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates, *Journal of Fixed Income* 3, 14–31.

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 Goldstein, 2002, Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility, *Journal of Finance* 57, 1685–1730.

Collin-Dufresne, Pierre, and Robert Goldstein, 2003, Generalizing the affine framework to HJM and random field models, Working paper, Carnegie Mellon University and Washington University.

Collin-Dufresne, Pierre, Robert Goldstein, and Christopher S. Jones, 2004, Can interest rate volatility be extracted from the cross section of bond yields? An investigation of unspanned stochastic volatility, NBER Working Paper 10756.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure of interest rates, *Journal of Finance* 55, 1943–1978.

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

De Jong, Frank, Joost Driessen, and Antoon Pelsser, 2001, Libor and swap market models for the pricing of interest rate derivatives: An empirical analysis, *European Finance Review* 5, 201–237.

Diebold, Francis X., and Roberto Mariano, 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics* 13, 253–263.

Driessen, Joost, Pieter Klaassen, and Bertrand Melenberg, 2003, The performance of multi-factor term structure models for pricing and hedging caps and swaptions, *Journal of Financial and Quantitative Analysis* 38, 635–672.

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 51, 921–949.

Duffie, Darrell, and Kenneth J. Singleton, 1997, An econometric model of the term structure of interest rate swap yields, *Journal of Finance* 52, 1287–1321.

Engle, Robert F., 2002, Dynamic conditional correlation—a simple class of multivariate GARCH models, *Journal of Business and Economic Statistics* 20, 339–350.

Engle, Robert F., and Kevin Sheppard, 2001, Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, Working paper, NYU Stern School.

Fan, Rong, Anurag Gupta, and Peter Ritchken, 2001, On the performance of multi factor term-structure models for pricing caps and swaptions, Working Paper, Case Western Reserve University.

Goldstein, Robert, 2000, The term structure of interest rates as a random field, *Review of Financial Studies* 13, 365–384.

Heath, David, Robert Jarrow, and Andrew Morton, 1992, Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation, *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.

Heston, Steven L., 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, John C., and Alan White, 1987, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42, 281–300.

Hull, John C., and Alan White, 1999, Forward rate volatilities, swap rate volatilities, and the implementation of the Libor market model, Working paper, University of Toronto.

Jagnnathan, Ravi, Andrew Kaplin, and Guoqiang Sun, 2003, An evaluation of multi-factor CIR models using LIBOR, swap rates, and cap and swaption prices, *Journal of Econometrics* 116, 113–146.

Jamshidian, Farshid, 1997, Libor and swap market models and measures, *Finance and Stochastics* 1, 293–330.

Kimmel, Robert, 2004, Modeling the term structure of interest rates: A new approach, *Journal of Financial Economics* 72, 143–183.

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

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

Longstaff, Francis A., Pedro Santa-Clara, and Eduardo S. Schwartz, 2001, The relative valuation of interest rate caps and swaption: Theory and empirical evidence, *Journal of Finance* 56, 2067–2169.

Musiela, Marek, and Marek Rutkowski, 1997, *Martingale Methods in Financial Modeling* (Springer, Berlin, Heidelberg, New York).

Pearson, Neil D., and Tong-Sheng Sun, 1994, Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model, *Journal of Finance* 49, 1279–1304.

Perignon, Christophe, and Christophe Villa, 2005, Sources of time variation in the covariance matrix of interest rates, *Journal of Business* 79, 1535–1550.

Santa-Clara, Pedro, and Didier Sornette, 2001, The dynamics of forward interest rate curve with stochastic string shocks, *Review of Financial Studies* 14, 149–185.

Vorst, Ton, 1992, Prices and hedge ratios of average exchange rate options, *International Review of Financial Studies* 1, 179–193.