Hedging in the Possible Presence of Unspanned Stochastic Volatility: Evidence from Swaption Markets*

Rong Fan† Anurag Gupta‡ Peter Ritchken§

August 13, 2002

*The authors would like to thank seminar participants at presentations made at the Twelfth Annual Derivatives Securities Conference in New York, The 2002 FMA European Meetings in Copenhagen, the Seventh Annual US Derivatives and Risk Management Congress in Boston, The 2002 Western Finance Association Meetings in Park City, The Second World Congress of the Bachelier Finance Society in Crete, and Georgetown University. They also acknowledge support for this research from the Federal Reserve Bank in Cleveland.

†Case Western Reserve University, WSOM, 10900 Euclid Ave., Cleveland, OH 44106-7235, Phone: (216) 368–3849, Fax: (216) 368–4776, E-mail: rxf10@po.cwru.edu

‡Case Western Reserve University, WSOM, 10900 Euclid Ave., Cleveland, OH 44106-7235, Phone: (216) 368–2938, Fax: (216) 368–4776, E-mail: axg77@po.cwru.edu

§Case Western Reserve University, WSOM, 10900 Euclid Ave., Cleveland, OH 44106-7235, Phone: (216) 368–3849, Fax: (216) 368–4776, E-mail: phr@po.cwru.edu

Hedging in the Possible Presence of Unspanned Stochastic Volatility: Evidence from Swaption Markets

Abstract

This paper examines whether higher order multifactor models, with state variables linked *solely* to the full set of underlying LIBOR-swap rates, are by themselves capable of explaining and hedging interest rate derivatives, or whether models explicitly exhibiting features such as unspanned stochastic volatility are necessary. Our research shows that swaptions and even swaption straddles can be well hedged with LIBOR bonds alone. We examine the potential benefits of looking outside the LIBOR market for factors that might impact swaption prices without impacting swap rates, and find them to be minor, indicating that the swaption market is well integrated with the underlying LIBOR-swap market.

JEL Classification: G12; G13; G19.

Keywords: Interest Rate Derivatives; Swaptions; Delta and Gamma Neutral Hedging; Unspanned Stochastic Volatility; Spanning.

Understanding the dynamics of the term structure of interest rates is crucial for pricing and hedging fixed income positions. It is therefore not surprising that this subject continues to attract keen interest. In the last decade, researchers have focused their attention on term structure models where the yield to maturity is an affine function of the underlying state variables.[1] The empirical literature has evolved from studying one-factor diffusions to multi-factor models that incorporate stochastic central tendency, stochastic volatility, jumps, and non trivial correlations among state variables. With respect to derivatives, Chacko and Das (2002) show that there are computational advantages in pricing certain claims in the affine class, where simple solutions often exist.

Recently, several authors have identified limitations of affine models.[2] Jagannathan, Kaplin and Sun (2001) investigate whether these limitations are severe enough to affect their use in valuing interest rate claims. In particular, they estimate one, two and three-factor Cox, Ingersoll and Ross (1985) models and show that as the number of factors increases, the fit of the model to LIBOR and swap rates improves. However, when they use their estimated models to price swaptions, the pricing errors are biased and large. They conclude that while more factors might be necessary to explain swaption prices, it might be useful to investigate alternative structures, possibly outside the affine class. Their findings also point out the need for evaluating term structure models using data on derivative prices.

The Jagannathan, Kaplin, and Sun study uses time series information on LIBOR and swap rates to estimate models under the data generating measure. An alternative approach is to use a set of observable derivative prices to directly calibrate the parameters of a model under the risk neutral pricing measure and then analyze any systematic biases in the pricing errors. This process was followed by Longstaff, Santa Clara and Schwartz (2001)(hereafter LSS), among others.[3] They use observable swaption prices to extract information on volatility structures for forward rates, and conclude that in order to price swaptions effectively, models with three or four factors are necessary. However, they find that models calibrated using swaption data produce theoretical cap prices that often deviate significantly from the no-arbitrage values implied by the swaptions market. Their results indicate that the swaption and cap markets may not be well integrated, or that their models do not capture all elements of risk.

[1]Duffie and Kan (1996) link these models to the affine structure of the underlying stochastic processes of the state variables, and Dai and Singleton (2000) identify the *maximally flexible* affine models, that, from an empirical perspective, nest all other affine models.

[2]For example, see Backus, Foresi, Mozumdar and Wu (2001), Ghysels and Ng (1998), and Duffee (2002). Ahn, Dittmar and Gallant (2002) develop a family of quadratic term structure models that nest several existing models, including the SAINTS model of Constantinides (1992) and the double square root model of Longstaff (1989), and provide empirical support that their models outperform affine models in explaining historical bond price behavior in the United States.

[3]Hull and White (1999) develop very similar models to the LSS models using the LIBOR market based model of Brace, Gatarek and Musiela (1997).

The possible lack of integration between price behavior in related markets, such as the cap and swaption markets, could arise from the fact that derivative prices may be affected by factors outside the LIBOR bond market. If this is the case, then the swap markets are incomplete and hence there is no reason for the variance-covariance matrix of forward rates, implied out from cap and swaption data, to be similar to the variance-covariance matrix implied from interest rates. Indeed, De Jong, Driessen and Pelsser (2002) show that the historical correlations are significantly higher than those implied from cap and swaption data using a full-factor lognormal market model. One of their explanations is that a volatility risk premium may be present. Heidari and Wu (2001) show that less than 60% of the variation in swaption volatilities can be explained by the three common interest rate factors identified by Litterman and Scheinkman (1991). In addition to the level, slope and curvature of the yield curve, they identify three additional volatility related factors, independent of interest rate factors, that increase the explained variability of swaption volatilities to over 95%. They use this evidence to suggest that swaptions cannot be hedged using LIBOR state variables alone.

Other studies have also confirmed the lack of integration between prices in related markets. Collin-Dufresne and Goldstein (2002) regress the returns on at-the-money cap straddles, which are portfolios mainly exposed to volatility risk, against changes in swap rates drawn from across the yield curve. On average, over all contracts, they find that the swap rate changes only explain about 25% of the straddle returns. Further, the residuals of these regressions are highly correlated over the different contracts, and a principal component analysis indicates that about 85% of the remaining variation can be explained by a single common factor. They conclude that there is at least one state variable which drives innovations in interest rate derivatives, but which does not affect innovations in LIBOR-swap rates. In other words, they suggest that caps cannot be priced relative to bonds alone, or equivalently, that the bond market by itself is incomplete. They term this feature *unspanned stochastic volatility*, and establish models within an appropriately curtailed trivariate Markov affine system that display this interesting property. Han (2001) establishes a term structure model that explicitly models the covariances of bond yields as being linear in a set of state variables. In his model, the two factors for covariances can be viewed as a level dependent and a slope dependent factor for the term structure of volatilities. Han finds empirical support for this model.

The above studies all *suggest* that there are sources of risk that affect fixed income derivatives which cannot be effectively hedged by portfolios consisting solely of LIBOR bonds or swaps, and that the set of hedging instruments may have to be broadened, possibly to include some options. Our paper investigates the relative importance of these other sources of risk in the swaption market. In particular, if unspanned stochastic volatility is important, then it should not be possible to hedge swaptions effectively using a model based on state variables limited to the set of swap rates. More importantly, such models should certainly not be able to hedge contracts which have extreme sensitivity to volatilities, such as straddles. If a model can be identified in which the entire matrix of swaptions and their straddles can be well hedged by bonds, then the need to explicitly incorporate additional factors that influence swaptions but not swaps appears to be minor. On the other hand, if swaptions cannot be well hedged, then either the models are misspecified or indeed there are other factors, beyond the term structure, that influence the behavior of swaption prices.

We construct simple multifactor models that hedge swaptions and swaption straddles very effectively, and, based on the empirical performance of these hedges, we conclude that the benefits of explicitly incorporating unspanned stochastic volatility into models are minor.

Our paper examines the importance of multifactor models from the perspective of *hedging* effectiveness. In contrast, most of the other studies focus on pricing performance.[4] For example, LSS use pricing accuracy as the criterion to conclude that four factors are necessary for pricing swaptions. However, the number of free parameters in their models equals the number of factors. As a result, it is unclear whether the improvement in pricing is attributable to the number of factors or to the number of free parameters. Indeed, when we redo their pricing tests on our data we reconfirm their findings. In particular, we find significant improvements in pricing accuracy when incorporating up to four factors. However, it is easy to construct a one-factor model, containing the same number of parameters as their four-factor model, that prices swaptions equally well.

Unfortunately, a pricing analysis, even conducted out-of-sample, tells us little about the true volatility structure of forward rates. It therefore does not assist us in understanding the true importance of multiple factors and the real capability of a model. In contrast, our study carefully evaluates models in terms of hedging precision. If a continuous time model is correct, the risk of carrying a hedged position over a time increment is entirely due to the fact that the volatility parameters are not estimated precisely and continuous revisions are not accomplished. If one model consistently produces hedges that are more effective than another model, then it must be the case that the first model, with its volatility structure, better captures the true dynamics of the term structure and the true sensitivity of options to movements in the underlying term structure. Not only do we consider the effectiveness of multi-factor delta neutral hedges, but, for the case of high gamma positions such as straddles, we also investigate the effectiveness of delta and gamma neutral hedges. To our knowledge, this is the first study to investigate the consequences of adopting these hedge strategies using one, two, three and four-factor models in what is unquestionably a multifactor environment.

The paper proceeds as follows. In section 1, we describe the models that we evaluate, the data, and the experimental design. In section 2, we examine the ability of different models in

[4]For exceptions to this rule see Driessen, Klaassen, and Melenberg (2001) and Gupta and Subrahmanyam (2001).

hedging swaptions. From a hedging perspective, the dangers in using lower order models are clearly revealed, and the benefits of using higher order models become apparent. In section 3, we address spanning issues in the swaption market. It is common to frequently recalibrate a model to observed option prices. This process may compensate for the inability of the model to capture the true term structure dynamics. Specifically, in the presence of unspanned stochastic volatility, the recalibrating process could incorporate swaption price information into a model that is based on swap rate state variables alone and mask the presence of missing state variables. In our analysis, we address this issue by measuring the hedging effectiveness when model recalibrations are not permitted. We analyze volatility sensitive straddles, which are hedged to be delta and gamma neutral. These hedges, maintained for one week, can explain about 85% of the unhedged variance. In contrast, when simply regressing swaption straddle returns on swap rate changes, we obtain R^2 values near 15%, comparable to the results Collin-Dufresne and Goldstein (2002) obtained for cap straddles. The consequences of our findings are explored. Section 4 concludes.

1 Models, Data, and Experimental Design

1.1 The Basic Models

Swaptions are actively traded, and, according to market convention, their prices are quoted in volatility form using the standard Black (1976) model, with instruments at different expiration dates, underlying swap maturities and strikes trading at different implied volatilities. The Black formula should be viewed only as a nonlinear transformation from prices into volatilities and vice-versa. This market convention provides a convenient way of communicating prices because volatilities tend to be more stable over time than actual dollar prices. The market convention does not imply that participants in this market view the Black model as being appropriate.

Let $f(t,T)$ denote the forward interest rate at time t for instantaneous riskless borrowing or lending at date T. The dynamics of forward rates are given by

$$df(t,T) = \mu_f(t,T)dt + \sum_{n=1}^{N} \sigma_{f_n}(t,T)dw_n(t), \text{ with } f(0,s) \text{ given for } s \geq 0. \tag{1}$$

where $\{dw_n(t)|n = 1, 2, \ldots, N\}$ are standard independent Wiener increments. The volatility structures, $\{\sigma_{f_n}(t,T)|n = 1, 2, \ldots, N\}$, could, in general, be functions of all path information up to date t.

Heath Jarrow and Morton (1992) show that to avoid riskless arbitrage, the drift term, under the equivalent martingale measure, is completely determined by the volatility functions in the above equation. Specifically:

$$\mu_f(t,T) = \sum_{n=1}^{N} \sigma_{f_n}(t,T) \int_t^T \sigma_{f_n}(t,u)du.$$

This implies that for pricing purposes, only the volatility structures need to be specified and estimated. The structures that we use are based on the loadings provided by the principal components of the historical correlation matrix of forward rates along the lines of the string models of LSS. In these models we consider a discrete set of M maturities say, $\{\tau_1, \tau_2, \ldots, \tau_M\}$ with $\tau_1 < \tau_2, \ldots, \tau_M$. Then:

$$\sigma_{f_j}(t, t+\tau_j) = g(\tau_j) f(t, t+\tau_j) \tag{2}$$

where $g(.)$ is a deterministic function of the maturity of the forward rate, that is estimated using the principal components extracted from the historical correlation matrix of weekly forward rate percentage changes. These forward rates are separated by three months for maturities less than a year, and six months thereafter (for up to ten years maturity). Specifically, twenty two forward rate maturities are used and a twenty two by twenty two correlation matrix is established. The matrix of eigenvectors (principal components) is computed, and the first four eigenvectors are retained. Let

$$T^* = \{\tau_1 = 0.25, \tau_2 = 0.5, \tau_3 = 0.75, \tau_4 = 1, \tau_5 = 1.5, \tau_6 = 2, \tau_7 = 2.5, ...\tau_{21} = 9.5, \tau_{22} = 10\}$$

represent the set of 22 forward rate maturities, and let h_i be a 22×1 vector representing the i^{th} eigenvector for $i = 1, 2, 3$ and 4. Then, define:

$$g_i(\tau_j) = \lambda_i h_{ij}. \text{ where } i = 1, 2, 3, 4 \text{ and } j = 1, 2, ..., 22.$$

where h_{ij} is the j^{th} element of the i^{th} eigenvector, and the λ_i values are the free parameters, the i^{th} one representing the scaling factor for all the elements of the i^{th} principal component, and is implied out at any date t using date t swaption data.

The principle behind such a procedure is simple. As shown by several researchers, including Litterman and Scheinkman (1991), the first four historical principal components identify the four most important types of orthogonal shocks to the forward rate curve. Since the exact contribution of each of these shocks may vary over time, the eigenvalues for the future period may be different from the eigenvalues over the historical period. Since, in an efficient market, the swaption data reflects all available information on the set of forward looking correlations among forward rates, this data should be used to establish the eigenvalues.

The above method, which we term an adapted Principal Component Analysis method has been used by LSS for models where forward rate volatilities are proportional to their levels, and by Driessen, Klaassen and Melenberg for models where forward rate volatilities do not depend on their levels.

1.2 Data

The data for this study consists of volatilities of USD swaptions of expiration 6 months, 1-, 2-, 3-, 4-, and 5-years, with the underlying swap maturities of 2-, 3-, 4, and 5-years each (in all, there are 24 swaption contracts). As per market convention, a swaption is considered at-the-money when the strike rate equals the forward swap rate for an equal maturity swap. The data consists of volatilities of at-the-money contracts over a 32 month period (March 1, 1998 - October 31, 2000), obtained from DataStream. Figure 1 presents the time series of implied Black volatilities for swaptions over the sample period.

Figure 1 Here

Each of the graphs corresponds to a time series for all swaption expirations for each underlying swap maturity. The figure clearly shows the changing variances over this time period, with the peaks occurring during the Long Term Capital Management crisis. Table 1 provides the time series properties for each of the contracts. The range of volatilities is largest for the short term contracts. The last column of the table indicates the typical bid-ask spread for each of the contracts. The typical bid-ask spread over this time period was approximately plus or minus a half Black vol. The spreads, in basis points, were obtained by repricing the swaption using the Black model, first with a volatility that exceeded the mean by 0.5 Black vols, and then with a volatility 0.5 units below the mean. These numbers indicate the coarseness of the data and serve as a benchmark for evaluating hedging precision.

Table 1 Here

For constructing the yield curve, we use futures and swap data. For the short end of the curve (up to 1 year maturity), we use the five nearest futures contracts on any given data. These futures rates are interpolated, and then convexity corrected to obtain the forward rates for 3, 6, 9, and 12-month maturities. The rest of the yield curve out to 5 years is estimated using the forward rates bootstrapped from market swap rates at 6 month intervals. The futures and swap data is obtained from DataStream. Eventually, we obtain weekly forward rate curves that start one year before our swaption data begins, and extend to the end of our swaption data period.

For the principal component analysis we use the one year history of forward rates that exists prior to the beginning of our swaption data, to estimate the correlation structure of forward rates. We then decompose the correlation matrix, R, into $U\Lambda^* U'$, where U is the matrix of eigenvectors and Λ^* is a diagonal matrix of eigenvalues. Finally, we retain the first four eigenvectors and assume a covariance structure for forward rates, Σ, given by $\Sigma = U\Lambda U'$ where Λ is a diagonal matrix with the first four diagonal elements positive, the others zero.

1.3 Model Implementation

We consider a discrete implementation of the multifactor HJM model. Towards this goal, we divide the time interval into trading intervals of length Δt, and label the periods with consecutive integers. Let $f^{\Delta t}(t,j)$ be the forward rate at period t, for the time interval $[j\Delta t,(j+1)\Delta t]$. Let $\Delta f^{\Delta t}(t,j)$ represent the change in the forward rate over a time increment Δt. That is

$$\Delta f^{\Delta t}(t,j) = f^{\Delta t}(t+1,j) - f^{\Delta t}(t,j)$$

The actual magnitude of this change could depend on the forward rate itself, its maturity date, and other factors.

We start with an initial forward rate curve, $\{f^{\Delta t}(0,j), j = 0,1,\ldots,m\}$ that is chosen to match the observed term structure at date 0 for all maturities up to date $m\Delta t$. Notice that $f^{\Delta t}(0,0)$ is just the spot rate for the immediate period, $[0,\Delta t]$. Over each time increment, the forward rates change as follows:

$$\Delta f^{\Delta t}(t,j) = \mu_f^{\Delta t}(t,j)\Delta t + \sum_{n=1}^{N} \sigma_{f_n}^{\Delta t}(t,j)\sqrt{\Delta t} Z_{t+1}^{(n)}. \tag{3}$$

where $Z_{t+1}^{(n)}$ is a standard normal random variable, j is an integer larger than the current date, t, $\mu_f^{\Delta t}(t,j)$ is the drift term, and $\sigma_{f_n}^{\Delta t}(t,j)$, is the volatility term associated with the n^{th} factor, $n = 1,2,...,N$, where the N standard normal random variables are independent. The discrete time equivalent of the Heath-Jarrow-Morton restriction is given by

$$\mu_f^{\Delta t}(t,j) = \sum_{n=1}^{N} \mu_{f_n}^{\Delta t}(t,j)$$

where

$$\mu_{f_n}^{\Delta t}(t,j) = \sigma_{f_n}^{2\ \Delta t}(t,j)\frac{\Delta t}{2} + \sigma_{f_n}^{\Delta t}(t,j)\sigma_{p_n}^{\Delta t}(t,j)$$

and

$$\sigma_{p_n}^{\Delta t}(t,j) = \sum_{i=t+1}^{j-1} \sigma_{f_n}^{\Delta t}(t,i)\Delta t$$

for $n = 1,2,....,N$.

Prices of European interest rate claims can be computed using Monte Carlo simulation. Specifically, assume K different paths, are simulated, each path initiated at date 0 where the initial term structure is given. Consider the k^{th} simulation. Given the date 0 term structure, forward rates are updated recursively using equation (3). This gives the k^{th} path of the term structure of forward rates. At date 0, \$1.0 is placed in a fund that rolls over at the short rate. At date $T\Delta t$ the value of the money fund, $M(T;k)$, is given by:

$$M(T;k) = \prod_{i=0}^{T-1} e^{f^{\Delta t}(i,i)\Delta t}.$$

Consider a claim that pays out in period T_E. Using simulation, a set of forward rates at this date can be computed, and hence all bond prices and swap rates can be recovered. In addition, the accumulated money fund, $M(T_E;k)$, along this path is known. The terminal value of this claim for this path can then be computed. Let $C(T_E;k)$ be this value. The date 0 value of the claim, for this path, is approximated by

$$C(0;k) = \frac{C(T_E;k)}{M(T_E;k)}.$$

The value of the claim at date 0 is then given by the average of all these values obtained over the K paths. Specifically:

$$C(0) = \frac{\sum_{k=1}^{K} C(0;k)}{K}.$$

Since repeated calls are used to estimate the parameters of the process, it is important that the pricing algorithms be as efficient as possible. Hence we use $\Delta t = 0.125$ years. Rather than price all the contracts separately, we simulate the money fund and forward rates along paths for a ten year period, and at each relevant maturity date along the path, all the appropriate caplet and swaption prices are computed. We repeat this procedure $K = 10,000$ times, and use antithetic variance reduction techniques, to establish the fair prices of all our contracts. We ran extensive robustness checks to ensure that the benefits of increasing the sample size and decreasing the time partition were negligible. Further, to the extent possible, we use the same stream of random numbers to price the same contracts with different volatility structures. This ensures that the difference in prices of the contracts is more tightly attributable to the different volatility structures rather than to sampling error.

Like Driessen, Klassen and Melenberg (2001), Moraleda and Pelsser (2000), LSS, and others, we estimate model parameters from cross sectional options data. At any date we fit models to the prices of swaptions for different expiry dates and underlying swap maturities. Our objective function is to minimize the sum of squared percentage errors between theoretical and actual prices using a non-linear least squares procedure. An alternative objective would be to minimize the sum of squared errors in prices. However, since prices of swaptions can range from a couple of basis points to a thousand basis points, which is almost three orders of magnitude apart, such a minimization would place more weight on the expensive contracts. In all these studies, the typical metric is to minimize root mean squared percentage pricing error at each optimization date.

Using mid-week data, for each odd week, we establish the best fit for the prices of all swaptions. Figure 2 presents the time series of forward rate volatilities estimated for each of the models. The forward rate volatility surfaces for all the models are clearly humped and they fluctuate over time, as the term structure moves. In general, forward rate volatilities increase over the time period with a spike during the Long Term Capital Management crisis.

Figure 2 Here

The implied correlations produced by the models are quite different from one another. The correlations among rates for the models generally decrease as the number of factors increase, with the four-factor model producing correlations that typically are the closest to the historical correlations. For example, the average correlations of the short 6 month rate with the 5 and 10 year six-month forward rates are 0.689 and 0.410 respectively, while the historical values were 0.695 and 0.370.[5]

2 Hedging Performance of Swaption Models

Our hedging experiments were conducted as follows. Given any calibrated n-factor model, we can establish a hedge position for a particular swaption using n different LIBOR discount bonds. For example, for a four-factor model, four price changes for each swaption are recorded, each price change arising after a small shock is applied to a single factor. In addition, the four price changes to a set of discount bonds are computed. The unique portfolio of the four bonds is then established that hedges the swaption against instantaneous shocks consistent with the model. The construction of the hedged position at any date t, only uses information available at date t. This analysis is repeated for all contracts and for all models. The hedge position is maintained unchanged for one week, and the hedged and unhedged residuals are obtained and stored. We then repeat this analysis for holding periods of two, three and four weeks.

Unfortunately, one week later, the prices of the old swaptions are unavailable, so the raw residuals of the unhedged positions cannot be directly observed. These prices, however, can be estimated using the observed set of fresh at-the-money swaption prices. We assume that the new volatility of each old swaption equals the volatility of the new at-the-money contract. Implicit in this analysis is the assumption that the volatility skew effect is insignificant over the region of strike price changes during a week. Indeed, we find that in more than 90% of the data, the strike changes by less than 30 basis points over a week, and in more than half the cases, the change is less than 10 basis points. Therefore, we really only require that contracts, with strikes within 30 basis points of the at-the-money contracts, have the same Black vols.[6] In contrast to the swaptions, the change in value of the bonds in the hedge portfolio is directly observable.

[5]For a detailed empirical investigation of the relationship between historical correlations and implied correlations, see De Jong, Driessen and Pelsser (2002). Additional discussions on the importance of matching model correlations to historical correlations is provided by Collin-Dufresne and Goldstein (2002b), Han (2001), LSS, Radhakrishnan (1998), Rebonato (1999), and others.

[6]LSS use the exact same procedure to compute the value of one week old swaptions, and conduct a variety of tests to indicate the viability of this procedure. Other studies that adopt the same assumptions include Driessen, Klaassen and Melenberg (2001) and Moraleda and Pelsser (2000). Until reliable prices of away-from-the-money swaptions are available, there is no objective way to adjust the Black vols. for any possible strike price bias.

For the one-factor model, we take the discount bond corresponding to the maturity date of the underlying swap as the hedging instrument. For the two-factor model, the two hedging instruments correspond to the discount bonds with maturities corresponding to the expiration date of the swaption and the maturity of the underlying swap. For the three-factor model we use these two bonds plus the bond with a maturity between the two. Finally, for the four-factor model, the hedging instruments are taken to be these two discount bonds, plus two additional bonds that have maturities equally spaced between the expiry and underlying swap maturity date. Given the swaption contracts, then, the hedging instruments are uniquely determined. Our analysis is limited to contracts with at least two years between the expiry date and swap maturity date. This is necessary, since four distinct instruments are needed for hedging within the four-factor model, and we want the hedging instruments to be separated by a minimum of six months. This simplifies the analysis since all swap rates are observed at six month increments, hence no interpolated rates are needed to estimate the prices of discount bonds.

For each swaption contract, the hedging analysis is conducted every second week. As a result, for each contract we have a time series of 70 nonoverlapping weekly unhedged and hedged residuals.

Figure 3 compares the box and whisker plots of hedging errors for each contract type for the four models with the unhedged pricing errors.

Figure 3 Here

The figure illustrates that all the hedges reduce the variance of the residuals. Indeed, the average variance reduction over all contracts achieved by the one-factor model is over 88%, with only marginal improvements for the higher order models. Graphically, this can be seen by the fact that the interquartile ranges of the hedging errors, as indicated by the width of the inner boxes, are of similar sizes for the four models. However, while the variances are similar for all the models, the biases are quite distinct, with the one-factor model, in particular, displaying the largest average deviations from zero.

As an example, consider the six month expiry contracts. The biases in the hedging errors, as indicated by the difference between the median error and zero, are large and positive for the one-factor model and negligible for the highest order model. In contrast, the interquartile ranges are somewhat similar over the models. This phenomenon holds true for almost all expiry dates, with the exception of the long term contracts, where the bias for the one-factor model is smaller, but comes at a cost of increased variance relative to the higher order multi-factor models. Figure 3 indicates that the common practice of measuring hedging effectiveness using variance reduction as the criterion may be flawed. It is only meaningful if the models produce average hedging errors close to zero. If average hedging errors are not near zero (i.e. the model is biased), then a better metric to use is the root mean squared error.

Table 2 presents the root mean squared error (multiplied by 10000) for each contract type for all 4 models. As a result, each entry can be interpreted in basis points.

Table 2 Here

As an example, consider the six month maturity contract on a swap of two years. The unhedged root mean squared error is 12.1 basis points, while the one-factor hedged position has a root mean squared error of 6.2 basis points. The four-factor model, however, has a root mean squared error of 3.0 basis points, indicating it is almost twice as effective. In comparing the root mean squared errors, contract by contract, the benefits of the higher order (three and four-factor) models become apparent. The bottom of the table reports the average effectiveness of the hedges. The higher order models account for over 90% of the variance of the unhedged error.[7]

From the last column in Table 1, we can see that the root mean squared error for the higher order models is of the same magnitude as the potential error in the change in swaption prices due to the bid-ask spread. This indicates that the hedges for the four-factor model are extremely precise given the coarseness of the data.

As a more formal test of comparing the hedging effectiveness among the different models, we conduct pairwise comparisons of the hedging residuals produced by each model for each of the 24 contracts. For each contract and for each week, the hedging residual is computed and the model with the smallest absolute value of hedging error is identified. The results are shown in Table 3. Simple proportion tests, at the one percent level of significance, reveal that the two-factor model outperforms the one-factor model, the three-factor model outperforms the two-factor model, and the four-factor model produces results indistinguishable from the three-factor model. Further, if the hedges are maintained unchanged over periods longer than one week, the relative advantage of multi-factor models, over one and two-factor models, increases, and there is still no significant advantage in moving beyond a three-factor model.

Table 3 Here

The construction of hedge ratios for all the models is based on 70 separate cross sectional estimations. For the higher order models, the hedge ratios are remarkably stable over time, indicating that the models are capturing a stable volatility structure. To examine this issue more carefully, we freeze the parameter estimates at their values obtained from the first week, and redo all the hedging tests. The corresponding root mean squared hedging errors are reported,

[7] All R^2 values that we report are unadjusted for the means (hence they include the impact of the bias, if any). That is, the denominator is the sum of unhedged squared residuals, and the numerator is this value less the sum of hedged squared residuals.

by contract, in the last four columns of Table 2. These numbers are a bit higher than the corresponding numbers generated when the parameter estimates were updated every week, but overall, are remarkably similar. Indeed, over all contracts, the hedging effectiveness of the four-factor model decreases only from 91% to 90%, and the hedging errors, as measured by root mean squared error, typically only increase by less than one basis point. This analysis provides further support that our multifactor models are capturing the actual dynamics or shocks that ripple across the yield curve and that the performance of the hedges is not significantly influenced by any incremental information contained in the concurrent swaption prices.

3 Do LIBOR Bonds Span the Swaption Market?

Collin-Dufresne and Goldstein (2002) present empirical evidence that may suggest that interest rate volatility risk cannot be hedged by portfolios consisting of bonds alone. Their empirical support is based on analyzing straddles constructed using caps and floors, and regressing their monthly returns against linear combinations of swap rate changes of differing maturities. Creating straddle portfolios allows them to focus on stochastic volatility issues because straddles are insensitive to small changes in interest rates, but extremely sensitive to changes in volatility. In their analysis, between 8% and 39% of the variability of straddle returns could be explained by the swap rates, depending on the cap maturity. Although it could be argued that straddles will suffer from significant gamma slippage which could explain the poor regression results, it is interesting to note that a principal component analysis of the residuals shows that 85% of the remaining variability is explained by a single additional state variable. A consequence of their findings is that additional state variables beyond those driving bond prices may be needed to hedge interest rate claims. In a similar vein, Heidari and Wu (2001) conclude that a six-factor model, with three state variables linked solely to the volatility surface, is necessary for swaptions.

In this section we examine whether at-the-money swaption straddles can be effectively hedged using our models. If swaption straddles can be effectively hedged using just the underlying LIBOR bonds, then the need to incorporate unspanned stochastic volatility factors in these markets is diminished.

A straddle consists of a long position in a payer swaption and a long position in an otherwise identical receiver swaption. Since the delta value of the at-the-money straddle is zero, the previous delta hedge strategy will not be effective. In order to hedge the curvature of the straddle we construct delta-gamma neutral positions. If the models truly describe the shock processes that can occur to the term structure, then the gamma hedged position should be fairly effective in reducing the volatility of the payouts of the straddle.

3.1 Constructing the Delta Neutral-Gamma Neutral Positions

In order to set up a straddle at date t that expires at date T, on an n-period swap, we require data on receiver swaptions. At the initiation date, t, the price of a receiver swaption, $RS_t(tT, n)$, is linked to a payer swaption, $PS_t(t; T, n)$, by the no-arbitrage relationship:

$$RS_t(t; T, n) - PS_t(t; T, n) = \textit{ Value of a New Forward Starting Swap},$$

where the forward starting swap is an n period swap to be initiated at date T. Since new forward swaps have zero value, the receiver swaption has the same price as the payer swaption.

Now consider date $t + \Delta t$. Let $RS_{t+\Delta t}(t; T, n)$ and $PS_{t+\Delta t}(t; T, n)$ be the date $t + \Delta t$ value of the original swaptions entered into at date t. The value of these swaptions are linked as:

$$RS_{t+\Delta t}(t; T, n) - PS_{t+\Delta t}(t; T, n) = \textit{New Value of the date t Forward Starting Swap.}$$

Unfortunately, at date $t + \Delta t$, we only have the price of "new" at-the-money payer swaptions, in the form of Black vols. We assume the new volatility of the old payer swaption is equal to the volatility of a new payer swaption. Given this Black vol, we can price the payer swaption, $PS_{t+\Delta t}(t; T, n)$. Further, since we have the term structure of LIBOR rates at date $t+\Delta t$, we can easily compute the new forward starting n-period swap rate, and then, using the appropriate date $t + \Delta t$ n-period annuity factor and the old forward starting swap rate, we can compute the date $t + \Delta t$ value of the swap entered into at date t. Given the value, the above arbitrage relationship links the price of the old receiver swaption to the price of the old payer swaption. Once the receiver swaptions are priced we can compute the value and returns over the period $[t, t + \Delta t]$ on a straddle contract with expiry date T, for an underlying n-period swap.

To estimate the gamma values for all the swaptions and bonds, each factor is shifted upward and downward in a way consistent with the model. For each shock all the swaptions and bonds are repriced. Given these shocks, the delta and gamma values for each instrument can be estimated. This procedure is then repeated for each principal component. For the n factor model, n delta and n gamma values are computed for each security.

For the one-factor model, two bonds are needed to establish the hedge. We use a 1 year bond and a 10 year bond as the hedging instruments. For the two-factor model we need four bonds. We use the same two bonds as the one-factor model and add the 3.5 year and 7.5 year bonds. For the three factor model, we need six hedging instruments, so we add the 5 and 6 year bonds. Finally, for the 4 factor model we use these 6 bonds plus the 2.5 and 8.5 year maturities. The selection of these maturities is done so as to span the maturity set with equally spaced intervals.

3.2 Empirical Results for Hedging Swaption Straddles

Table 4 presents the root mean squared delta-gamma neutral hedging errors in basis points of the hedged and unhedged straddles for all the models, using a one week delta-gamma neutral hedge. For example, consider the first contract, a six month swaption on a two year swap. The root mean squared error is reduced from 5.9 to 3.1 basis points when a four-factor model is used to establish the hedge position. The hedge has accounted for 72% of the variance of the unhedged residuals. The results show that the four-factor model is very effective in reducing volatility. Indeed, on average, over all contracts, the four-factor model accounts for 85.3% of the variance of the unhedged straddle returns.

Table 4 Here

Table 4 also shows that the effectiveness of hedging straddles increases as the number of factors increases. While, on average, the one and two-factor models only account for 23% and 48% of the variance of straddle returns respectively, the three and four-factor models account for over 80% of the variability. The results are similar when the performance of the hedge is evaluated two, three and four weeks out-of-sample. Even four weeks out-of-sample, the higher order models hedge swaption straddles very well, removing nearly 80% of the unhedged variance.

Figure 4 presents the box and whiskers plots of the hedging errors, by contract, for each model. These plots show that there is very little difference between the three and four-factor models. However, relative to the lower order models, these models consistently produce unbiased hedging errors with smaller variance.

Figure 4 Here

Table 5 presents the results of pairwise comparisons of the residuals over the 70 weeks. These results confirm what we uncovered in the analysis of swaptions, namely that one and two-factor models are not as effective as higher order models and that there is little advantage, if any, in moving beyond a three-factor model.

Over a week the slippage in the hedges is small. From the last column of Table 1 we see the typical bid-ask spread in swaption prices, which reveals the inherent coarseness of price information in this market. In general, the root mean squared error is smaller than this range. Our results now provide a rather high hurdle which models that incorporate unspanned stochastic volatility must clear in order to demonstrate that swaptions are not spanned by LIBOR bonds. Indeed, given the coarseness of the data, any further improvements in hedging might be difficult to assess unless the actual swaption price data is recorded with more precision.

Table 5 Here

The bottom rows of Table 5 compares the performance of the hedges across models when the period between rebalancing is increased from one to four weeks. The results indicate that if less frequent rebalances are done, then the benefit in using multifactor models increases further. For example, a four-factor model outperforms a one-factor model 82% of the time over one week, but 97% of the time over four weeks.

In performing this analysis, however, we need to be careful to ensure that our hedges only use data from the underlying LIBOR-swap markets. The above analysis indirectly incorporates swaption information in the sense that the eigenvalues are periodically reestimated using swaption data. To ensure that we do not incorporate swaption information, even indirectly, we do further tests where we do not update the eigenvalues; we freeze them at their initial values and redo the entire set of hedging experiments for the swaption straddles. Specifically, the delta-gamma neutral hedging analysis is redone when the eigenvalues are not updated. These results are shown in the last four columns of Table 4. In this case, the performance of the hedges deteriorates a little, but overall, the hedging effectiveness is still very good. In particular, the three and four-factor hedges account for over 70% of the unhedged variance.

Figure 5 presents the time series of the percentage of total sums of squared errors over each quarter for the four-factor model, with and without recalibration of parameters using option information. The figure shows that delta-gamma hedging of swaption straddles is very effective in all time periods. The figure also shows that periodic recalibration provides only a modest improvement in hedging performance. This indicates that there may be some information contained in swaption prices, relating to volatility, which may not be impounded into the LIBOR bonds. However, relative to the empirical findings of Collin-Dufresne and Goldstein (2002) and others, the impact of unspanned stochastic volatility in hedging volatility sensitive straddles is very much diminished.

Figure 5 Here

Bakshi and Kapadia (2001) have recently investigated the performance of delta hedged positions in the equity market. Their analysis is based on the concept that if option prices incorporate a non zero volatility risk premium, then its existence can be inferred from the returns of an option portfolio that is hedged from all risks except volatility risk. They show that as long as volatility risk is stochastic, but unpriced, the expected gain from a delta hedged strategy should be zero. By analyzing the average gains from their delta hedges, they conclude that the market price of volatility risk is negative. Our straddles are delta and gamma hedged, and, if the model is correct, the resulting positions are only exposed to volatility risk. The delta-gamma neutral hedged straddle errors for the entire period reveal a symmetric distribution closely centered near zero. Indeed, on average, over the entire period and over all straddles, going short a straddle and hedging the position using delta-gamma neutral positions in LIBOR bonds, leads to an average loss of less than 0.50 basis point loss per contract per week. The average gain on an unhedged straddle for one week was 1.8 basis points. This positive return reflects the fact that over our data period, Black volatilities did rise. If all the risks of the straddle except volatility risk were hedged out, then, over this time period, one would expect a positive return, unless the risk premium for volatility risk is sufficiently negative. Since the average of the delta-gamma neutral hedging errors are closer to zero, relative to the unhedged positions, and since the distributions are almost centered around zero, this indicates that the risk premium for bearing volatility risk is not positive.

3.3 Unspanned Stochastic Volatility in Swaption Markets

Collin-Dufresne and Goldstein (2002) use the low R^2 values in their straddle regressions as a motivating point for their models. When we repeat their analysis, using swaption straddles, we get results similar to those for their cap straddles. Table 6, Panel A, presents the R^2 values for all our 24 contracts. These results are contrasted with the R^2 values from our four-factor model, with updating permitted for the eigenvalues, in panel B, and no updating, in panel C. As can be seen, over all contracts, the four-factor model accounts for a significant fraction of the straddle variability.

Table 6 Here

The poor results obtained using a linear regression methodology are not surprising. A straddle has a highly convex payout structure, and being at-the-money, its delta values are zero. As a consequence, given a small shock to the k^{th} principal component, $k = 1, 2, 3, 4$, the change in the straddle price should be zero. If the four-factor model is correct, then these are the only shocks that can occur to the yield curve and the straddle should be immunized against them. If this is the case, then there is no reason to expect straddle returns to be linearly related to changes in swap rates. Our regression results tend to confirm this feature. Indeed, the adjusted R^2 values from the regression are in line with these ex-ante expectations.

The regression analysis also tells us little about market completeness. To focus on this issue we need to establish a replicating portfolio of *traded* securities that mimics the straddle. Changes in swap rates do not correspond to changes in traded securities.[8] In contrast, our replicating portfolio consists of traded bonds and is constructed to have the same delta and gamma values as the straddle. Since the unhedged straddle is delta neutral, the significant improvement is

[8] A close substitute might be a swap contract. Unfortunately, this contract does not profit by the change in the swap rate, but by this amount multiplied by a time-varying annuity factor.

due to the ability of the hedge to match all the gamma exposures. Indeed, the fact that we can hedge this curvature risk is a great testimony to the four-factor model.

There are two additional reasons why the regressions may not be appropriate for evaluating whether bonds span swaptions. First, when the volatility structure of forward rates is level dependent, the factor loadings should be level dependent as well, implying that the coefficients in the regression equation should not be constant. Indeed, even if forward rates are not level dependent, since swap rates are weighted averages of forward rates, where the weights change as the term structure changes, the swap rates will have changing volatilities. Second, since the composition of the hedge changes as the term structure changes, the sensitivities of straddle returns to movements in particular swap rates are time varying. The regression methodology ignores both these effects, while our hedging strategies explicitly incorporate them.

Given that we can hedge swaption straddles very effectively, there appears to be very little advantage in extending our models to incorporate unspanned stochastic volatility features.

We further analyze our unexplained straddle hedging residuals, with the goal of identifying whether these residuals could be explained by a missing common factor. In particular, we construct a 24×24 variance covariance matrix of the residuals, and examine the principal components. Collin-Dufresne and Goldstein find a missing factor that accounts for 85% of the residual variability. In contrast, our first three principal components collectively only account for 70% of the remaining variance, with the first factor explaining only 42%.[9] Finally, when the unexplained residuals were regressed against the actual changes in all 24 Black volatilities in the same week, less than 60% of the variance could be explained, with the adjusted R^2 values averaging under 40%. While it is possible that models which incorporate unspanned stochastic volatility might account for some of this variance, even if they did, the additional contribution to hedging effectiveness, at best, will be very minor.[10]

We repeat the tests using the straddle residuals from the model when the parameter estimates were never updated. In this case, our first three principal components collectively account for 75% of the remaining variance, hardly different from the previous analysis. However, in this case, the first principal component accounts for 57% of the variability. This is higher than the 42% contribution from the first principal component that we had obtained earlier. This suggests that ignoring swaption information may lead to a loss of efficiency in the hedge performance, and that swaption prices may contain some information that may not be reflected in the LIBOR term structure. However, the role of this incremental information in hedging swaptions is minor, and is much less than that suggested in studies by Collin-Dufresne and Goldstein, and by Heidari and Wu.

[9] The results were almost unchanged when the correlation matrix was used.

[10] Explaining say 60% of the remaining 15% of variance, for example, might account for an additional one or two basis point improvement, which is hardly relevant.

4 Conclusion

We explore the consequences of using lower order models to price and hedge swaption contracts. We show that the higher order models hedge better than lower order models for all contracts, regardless of their expiration date and underlying swap maturity, and that decreasing the number of factors below three can result in significant biases and hedging errors. The main question addressed is not whether the cost of using one or two-factor models to hedge swaptions is acceptable, but whether models which incorporate three or four stochastic drivers, where the state variables relate solely to LIBOR swap rates, are sufficient for hedging. To investigate this issue, we use the models to hedge swaption straddles which have extreme curvature risk and are sensitive to volatility changes. If unspanned stochastic volatility and correlation are important factors, then it should be difficult to hedge these contracts using LIBOR bonds alone. We construct delta and gamma neutral hedges based on shocks that are consistent with our proposed volatility structures on forward rates. While the lower order models are not capable of effectively hedging these straddles, using the higher order (three and four-factor) models, the resulting hedges are found to explain over 80% of the variance of the straddle returns. Given the bid-ask spread of swaptions, the slippage in the hedges, held unchanged for one week, appear to be very modest.

A principal component analysis reveals that there is no dominant missing common factor in the residuals for the straddle hedges. This indicates that our models capture the majority of the risk in the straddle returns. We also investigate whether the act of recalibrating the models allows swaption information to indirectly affect the results. When we freeze the eigenvalues, the hedges are still effective, explaining over 70% of the variance. Of the remaining hedging error, a principal component analysis does reveal some weak structure in the residuals, but no dominant single factor that other studies suggest. In the context of the swaption market, our results for straddles, arguably one of the most volatility sensitive products, indicate that the benefits one could expect by explicitly introducing state variables that incorporate unspanned stochastic volatility appear to be minor. Our hedging results provide a much higher hurdle that models with unspanned stochastic volatility must clear, in order for them to conclude that LIBOR bond markets are incapable of hedging swaptions.

Our hedging analysis also reveals that the popular criterion of assessing hedging performance based on the percentage of unhedged *variance* explained by the hedging instrument could lead to flawed conclusions. Specifically, lower order models can produce hedging residuals with low variance, but the results may be an illusion since the low variance comes at the cost of large systematic biases that can lead to significant losses in hedging. Using *root mean squared error* as the criterion, which accounts for variance *and* bias, leads to more precise conclusions.

Recently, Collin-Dufresne and Goldstein (2002) have argued that the high residuals for caps, found in many studies, are due to the absence of state variables driving correlation and volatility risk, and that caps might be more sensitive to these factors than swaptions. It remains for future research to evaluate the performance of our models in the cap market. Indeed, it may be the case that unspanned stochastic volatility is more important in the cap market. Such a result would explain why models calibrated in the swaption market perform poorly when pricing caps.

Research has also been done regarding the importance of factors for pricing Bermudan swaptions. Longstaff, Santa-Clara and Schwartz (2001b) show that exercise strategies based on one-factor models understate the true option value for Bermudans. They contend that the current market practice of using one-factor models leads to suboptimal exercise policies and a significant loss of value for the holders of these contracts. However, Andersen and Andreasen (2001) conclude that the standard market practice of recalibrating one-factor models does not necessarily understate the price of Bermudan swaptions. The latter study is useful since it suggests that practitioners are not making systematic errors in marking their Bermudan swaptions to market. Yet, since they do not investigate any hedging issues, it does not resolve the issue of whether lower order models can be used for Bermudan swaptions. Our study implies that multifactor models are necessary for these products. It remains for future research to extend our analysis to include the Bermudan swaption market, and to establish whether our higher order models alleviate the need to incorporate unspanned stochastic volatility there.

Finally, it remains for future research to follow Bakshi and Kapadia (2001) and extract information that is helpful in identifying the market price of interest rate volatility risk. Continuing to extract information from interest rate derivative products, no doubt will lead us to an increased understanding of the nature of the volatility structure of forward rates, the linkage of markets, and the need to incorporate factors outside LIBOR swap rates. A deeper understanding of these factors is crucial for accurate hedging, interest rate risk management and value-at-risk.

References

Ahn D, R. Dittmar and A. Gallant, 2002, "Quadratic Term Structure Models: Theory and Evidence," *The Review of Financial Studies*,15, 243-288.

Andersen, L. and J. Andreasen, 2001, "Factor Dependence of Bermudan Swaptions: Fact or Fiction?," *Journal of Financial Economics*, 62, 3-37.

Backus, D., S. Foresi, A. Mozumbar, and L. Wu, 2001, "Predictable Changes in Yields and Forward Rates, " *Journal of Financial Economics*, 59 281-311.

Bakshi, G., and N. Kapadia, 2001, "Delta-Hedged Gains and the Negative Market Volatility Risk Premium," forthcoming, *Review of Financial Studies.*

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

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

Chacko, G. and S. Das, 2002, "Pricing Interest Rate Derivatives: A General Approach,", *The Review of Financial Studies*, 15, 195-241.

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

Collin-Dufresne, P. and R. S. Goldstein, 2002b, "Stochastic Correlation and the Relative Pricing of Caps and Swaptions in a Generalized-Affine Framework," working paper, Carnegie Mellon University.

Constantinides, G., 1992, "A Theory of the Nominal Structure of Interest Rates," *The Review of Financial Studies*, 5, 531-552.

Cox, J., J. Ingersoll, and S. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 363-384.

Dai, Q., and K. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance,* 55, 1943-1978.

De Jong, F., J. Driessen, and A. Pelsser, 2002, "On the Information in the Interest Rate Term Structure and Option Prices," Working Paper, University of Amsterdam.

Driessen, J., P. Klaassen, and B. Melenberg, 2001, "The Performance of Multi-factor Term Structure Models for Pricing and Hedging Caps and Swaptions", Working Paper, Department of Econometrics, Tilburg University.

Duffee, G., 2002, "Term Premia and Interest Rate Forecasts in Affine Models, " *The Journal of*

Finance, 57, 405-438

Duffie, D., and R. Kan, 1996, "A Yield Factor Model of Interest Rates," *Mathematical Finance*, 6, 379-406.

Ghysels, E. and S. Ng, 1998, "A Semi-Parametric Factor Model of Interest Rates and Tests of the Affine Term Structure," *Review of Economics and Statistics*, 80, 535-548.

Gupta, A., and M. Subrahmanyam, 2001, "Pricing and Hedging Interest Rate Options: Evidence from Cap-Floor Markets", Working Paper, New York University.

Han, B., 2001, "Stochastic Volatilities and Correlations of Bond Yields," working paper, Anderson Graduate School of management, University of California, Los Angeles.

Heath, D., R. Jarrow, and A. Morton, 1992, "Bond pricing and The Term Structure of Interest Rates,"*Economtrica*, 60, 77-105.

Heidari, M., and L. Wu, 2001, "Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates?, " working paper, Fordham University.

Hull, J., and A. White, 1999, "Forward Rate Volatilities, Swap Rate Volatilities, and the Implementation of the LIBOR Market Model", Working Paper, University of Toronto.

Jagannathan, R., A. Kaplin, and S. Sun, 2001, "An Evaluation of Multi-Factor CIR Models Using LIBOR, Swap Rates, and Cap and Swaption Prices", forthcoming, *Journal of Econometrics*

Litterman, T. and J. Scheinkman, 1991, "Common Factors Affecting Bond Returns", *Journal of Fixed Income* June, 54-61

Longstaff, F., 1989, "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates," *Journal of Financial Economics,* 23 195-224.

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

Longstaff, F., P. Santa-Clara, and E. Schwartz, 2001b,"Throwing Away a Billion Dollars: The Cost of Suboptimal Exercise in the Swaption Market", *Journal of Financial Economics*, 62, 39-66.

Moraleda, J. and A. Pelsser, 2000, " Forward versus Spot Interest Rate Models of the Term Structure: An Empirical Comparison", *The Journal of Derivatives* Spring, 9-21.

Radhakrishnan, A. R., 1998, Does correlation matter in swaption pricing, Working Paper, New York University.

Rebonato, R., 1999, "On the Simultaneous Calibration of Multifactor Lognormal Interest Rate Models to Black Volatilities and to the Correlation Matrix", *Journal of Computational Finance,* 4, 5-27.

Table 1

Descriptive Statistics for European Swaption Volatilities

This table presents descriptive statistics for the mid-market implied Black-model volatilities for European swaptions of different expirations and underlying swap maturities analyzed in the paper, from March 1, 1998 to October 31, 2000. The table also presents representative bid-ask spreads of each swaption, in basis points. The bid-ask spreads are estimated based on the mean volatility during the sample period, using the term structure at the start of the period. The presented spread is based on plus or minus half a Black vol., i.e., a total spread of one Black vol.

Expiration	Swap Maturity	Mean	Median	Standard Deviation	Minimum	Maximum	Bid-Ask Spread
0.5	2	15.2	14.3	3.1	10.5	27.0	3.0
	3	15.1	14.4	2.9	10.5	26.0	4.4
	4	15.1	14.4	2.8	10.3	25.0	5.8
	5	15.0	14.3	2.7	10.4	24.5	7.1
1	2	15.8	15.5	2.7	11.8	24.5	4.2
	3	15.5	15.3	2.4	12.0	23.0	6.2
	4	15.3	15.0	2.2	12.0	22.2	8.0
	5	15.1	14.8	2.0	12.0	21.0	9.8
2	2	16.1	16.0	2.1	13.0	23.0	5.7
	3	15.7	15.5	1.9	12.9	22.0	8.3
	4	15.4	15.3	1.8	12.8	21.0	10.8
	5	15.1	14.9	1.6	12.7	20.0	13.2
3	2	15.9	15.8	1.8	12.9	21.0	6.6
	3	15.5	15.5	1.7	12.8	20.0	9.7
	4	15.2	15.2	1.5	12.7	19.5	12.6
	5	14.9	14.9	1.4	12.5	19.5	15.3
4	2	15.6	15.6	1.6	12.8	19.9	7.3
	3	15.2	15.2	1.5	12.7	19.3	10.6
	4	14.9	14.9	1.4	12.6	18.6	13.7
	5	14.5	14.5	1.3	12.4	18.0	16.7
5	2	15.2	15.2	1.4	12.7	18.8	7.7
	3	14.8	14.9	1.3	12.6	18.2	11.2
	4	14.5	14.5	1.2	12.4	17.5	14.5
	5	14.2	14.2	1.1	12.1	17.2	17.6

Table 2

Absolute Hedging Errors for Swaptions

This table presents the root mean squared errors (in basis points) of the hedged and unhedged portfolios for all the models, one week out-of-sample, with and without recalibration of the model using option information. In the hedge portfolios, the number of hedging instruments used equals the number of factors in the model. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. The root mean square of the hedging errors for a contract, across all dates, is multiplied by 10,000 so that it can be interpreted as a basis point error. The corresponding root mean squared errors for the unhedged swaptions are also presented, for comparison.

Expiry	Swap Mat.	Unhedged Swaption	Number of Factors in the Model							
			With Recalibration				Without Recalibration			
			1	2	3	4	1	2	3	4
0.5	2	12.1	6.0	3.1	3.0	3.0	6.7	3.5	3.5	3.4
	3	18.1	7.0	4.3	4.1	4.2	7.5	4.7	4.5	4.5
	4	23.3	7.8	5.6	5.1	5.2	8.8	6.3	5.9	5.9
	5	28.2	8.4	6.8	5.9	5.9	9.2	7.5	6.7	6.7
1	2	13.1	5.5	3.3	3.3	3.2	6.1	3.9	3.9	3.8
	3	18.9	6.6	4.4	4.1	4.6	7.1	4.8	4.7	4.8
	4	24.1	7.8	5.7	5.4	5.7	8.5	6.4	6.3	6.3
	5	28.9	8.4	6.5	6.3	6.2	8.8	7.0	6.8	6.7
2	2	13.1	5.6	3.3	3.3	4.3	6.2	3.9	3.9	4.0
	3	18.7	7.3	4.8	4.7	4.6	7.9	5.5	5.4	5.7
	4	23.0	8.5	5.9	5.8	5.8	9.2	6.7	6.6	6.6
	5	27.5	9.7	7.3	6.8	7.0	10.3	8.0	7.7	7.8
3	2	12.6	5.6	3.4	3.3	5.4	6.0	3.8	3.9	3.8
	3	17.3	7.4	5.0	4.9	5.0	7.8	5.5	5.5	5.5
	4	21.9	8.7	6.4	6.2	6.3	9.1	6.9	6.8	6.9
	5	26.9	10.7	8.1	7.5	8.4	11.0	8.6	8.5	8.8
4	2	11.4	5.6	3.8	3.7	3.7	5.9	4.1	4.1	4.1
	3	15.8	7.1	5.3	5.3	5.3	7.5	5.9	5.9	5.9
	4	20.9	8.5	6.2	6.2	6.3	9.0	6.8	6.8	6.8
	5	25.9	10.3	7.6	7.3	7.6	10.7	8.2	8.2	8.2
5	2	10.8	5.5	4.0	3.9	4.5	5.9	4.3	4.3	4.5
	3	16.4	7.9	5.7	5.8	5.7	8.4	6.2	6.2	6.2
	4	21.5	10.1	7.2	7.2	7.1	10.7	7.9	7.9	7.9
	5	26.3	12.0	8.6	8.6	9.5	12.6	9.3	9.3	11.2
R^2 - 1 week out-of-sample			0.83	0.92	0.92	0.91	0.80	0.90	0.90	0.90
R^2 - 2 week out-of-sample			0.79	0.89	0.92	0.92	0.72	0.85	0.89	0.88
R^2 - 3 week out-of-sample			0.72	0.87	0.91	0.90	0.65	0.82	0.87	0.86
R^2 - 4 week out-of-sample			0.67	0.86	0.92	0.91	0.62	0.80	0.86	0.86

Table 3

Comparison of Hedging Errors for Swaptions

This table presents the fraction of times one model outperforms the other model in hedging forecasts one, two, three and four weeks out-of-sample, for swaptions. In the hedge portfolios, the number of hedging instruments equals the number of factors in the model. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. Therefore, for each contract, the proportions are computed from a comparison of 70 hedging errors.

Expiration	Swap Maturity	1 vs Unhedged	2 vs 1	3 vs 1	4 vs 1	3 vs 2	4 vs 2	4 vs 3
0.5	2	0.67	0.91	0.91	0.91	0.59	0.61	0.61
	3	0.81	0.90	0.90	0.89	0.71	0.70	0.67
	4	0.86	0.90	0.89	0.89	0.79	0.77	0.40
	5	0.89	0.90	0.77	0.81	0.73	0.74	0.49
1	2	0.83	0.80	0.80	0.79	0.53	0.60	0.61
	3	0.83	0.73	0.76	0.73	0.59	0.49	0.39
	4	0.84	0.70	0.71	0.70	0.59	0.57	0.43
	5	0.83	0.70	0.54	0.73	0.34	0.59	0.73
2	2	0.79	0.73	0.71	0.67	0.57	0.38	0.39
	3	0.80	0.66	0.69	0.53	0.60	0.37	0.40
	4	0.79	0.69	0.73	0.70	0.54	0.63	0.54
	5	0.80	0.73	0.69	0.67	0.57	0.61	0.50
3	2	0.76	0.66	0.70	0.57	0.61	0.39	0.39
	3	0.79	0.64	0.67	0.66	0.66	0.57	0.39
	4	0.79	0.63	0.66	0.66	0.47	0.57	0.51
	5	0.81	0.70	0.71	0.63	0.59	0.40	0.39
4	2	0.77	0.70	0.74	0.63	0.49	0.46	0.49
	3	0.81	0.64	0.64	0.63	0.50	0.43	0.56
	4	0.77	0.70	0.71	0.71	0.54	0.47	0.44
	5	0.81	0.64	0.66	0.66	0.49	0.49	0.44
5	2	0.77	0.70	0.70	0.67	0.49	0.40	0.40
	3	0.77	0.73	0.69	0.70	0.46	0.50	0.59
	4	0.79	0.77	0.77	0.77	0.47	0.53	0.57
	5	0.83	0.73	0.76	0.60	0.60	0.40	0.40
Average (1 week out)		0.80	0.73	0.73	0.70	0.56	0.53	0.49
Average (2 weeks out)		0.84	0.76	0.74	0.74	0.56	0.54	0.49
Average (3 weeks out)		0.86	0.80	0.77	0.77	0.55	0.54	0.49
Average (4 weeks out)		0.90	0.82	0.81	0.79	0.55	0.54	0.49

Table 4

Absolute Hedging Errors for Swaption Straddles

This table presents the root mean squared delta-gamma hedging errors (in basis points) of the hedged and unhedged straddle portfolios for all the models, one week out-of-sample, with and without recalibration of the model using option information. In the hedge portfolios, the number of hedging instruments used equals twice the number of factors in the model. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. The root mean squared hedging errors for a contract is multiplied by 10,000 so that it can be interpreted as a basis point error. The corresponding root mean squared errors for the unhedged swaption straddles are also presented, for comparison.

Expiry	Swap Mat.	Unhedged Straddle	Number of Factors in the Model							
			With Recalibration				Without Recalibration			
			1	2	3	4	1	2	3	4
0.5	2	5.9	6.5	4.2	3.9	3.1	6.2	4.8	4.7	3.8
	3	8.2	9.0	6.7	4.0	3.2	8.2	7.6	5.1	4.2
	4	10.2	10.3	5.3	3.7	4.1	9.6	5.9	4.8	5.6
	5	12.0	14.1	7.7	4.7	4.5	12.0	8.7	5.9	6.2
1	2	6.4	6.8	5.4	3.6	2.7	7.1	6.0	4.7	3.6
	3	8.6	8.9	8.2	4.8	4.6	9.3	8.6	6.0	6.4
	4	11.2	11.5	8.8	4.7	4.4	12.1	9.0	6.0	6.0
	5	12.6	11.4	9.7	4.3	4.8	12.0	10.9	5.5	6.4
2	2	7.0	6.2	5.9	2.6	2.4	6.4	6.6	3.4	3.3
	3	10.0	8.3	9.6	3.3	3.3	8.9	10.1	4.2	4.5
	4	11.9	11.3	7.5	4.3	4.6	12.0	8.4	5.3	6.1
	5	14.6	10.2	7.6	5.2	5.5	10.8	8.5	6.6	7.4
3	2	7.3	6.8	5.9	2.4	2.5	7.4	6.7	3.1	3.5
	3	10.1	9.1	5.6	3.4	3.4	9.2	6.4	4.3	4.7
	4	13.0	10.5	8.0	4.4	4.9	10.6	9.0	5.6	5.9
	5	16.4	13.6	11.5	5.6	5.5	13.4	12.8	7.1	7.4
4	2	7.7	5.6	6.4	2.7	2.6	6.0	7.3	3.4	3.5
	3	10.7	7.4	6.5	4.0	3.8	7.6	7.4	4.9	5.2
	4	12.7	9.3	8.6	4.5	4.2	9.9	9.7	5.8	5.8
	5	15.6	11.4	10.4	6.5	5.4	12.3	11.5	7.4	7.3
5	2	7.9	5.1	4.7	2.7	2.6	5.4	5.2	3.4	3.6
	3	11.6	7.7	9.0	4.1	3.9	8.2	10.1	5.3	5.2
	4	14.8	9.8	8.8	4.9	6.9	10.5	9.8	6.3	9.7
	5	17.5	11.3	10.6	6.0	6.3	12.1	11.6	7.7	8.7
R^2 - 1 week out-of-sample			0.23	0.48	0.84	0.85	0.20	0.36	0.74	0.73
R^2 - 2 week out-of-sample			0.21	0.47	0.80	0.82	0.19	0.35	0.69	0.72
R^2 - 3 week out-of-sample			0.20	0.45	0.81	0.79	0.19	0.33	0.71	0.70
R^2 - 4 week out-of-sample			0.20	0.44	0.78	0.80	0.17	0.34	0.70	0.70

Table 5

Comparison of Hedging Errors for Swaption Straddles

This table presents the fraction of times one model outperforms the other model in delta-gamma hedging forecasts, one, two, three and four weeks out-of-sample, for swaption straddles. In the hedge portfolios, the number of hedging instruments equals twice the number of factors in the model. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. Therefore, for each contract, the proportions are computed from a comparison of 70 hedging errors.

Expiration	Swap Maturity	1 vs Unhedged	2 vs 1	3 vs 1	4 vs 1	3 vs 2	4 vs 2	4 vs 3
0.5	2	0.73	0.79	0.80	0.83	0.70	0.64	0.41
	3	0.66	0.79	0.89	0.94	0.73	0.73	0.46
	4	0.66	0.81	0.94	0.86	0.70	0.53	0.44
	5	0.64	0.83	0.89	0.91	0.71	0.63	0.57
1	2	0.80	0.59	0.80	0.83	0.80	0.83	0.53
	3	0.77	0.29	0.81	0.80	0.83	0.83	0.60
	4	0.73	0.34	0.89	0.86	0.90	0.90	0.43
	5	0.76	0.54	0.89	0.76	0.89	0.79	0.37
2	2	0.79	0.49	0.79	0.81	0.81	0.81	0.56
	3	0.80	0.30	0.84	0.76	0.94	0.91	0.39
	4	0.74	0.59	0.80	0.77	0.77	0.77	0.57
	5	0.73	0.56	0.79	0.74	0.79	0.73	0.49
3	2	0.83	0.54	0.84	0.80	0.86	0.84	0.49
	3	0.79	0.60	0.81	0.84	0.76	0.73	0.49
	4	0.80	0.54	0.79	0.80	0.77	0.79	0.59
	5	0.71	0.60	0.79	0.81	0.80	0.81	0.51
4	2	0.84	0.56	0.77	0.79	0.84	0.89	0.56
	3	0.80	0.47	0.80	0.81	0.79	0.81	0.64
	4	0.71	0.56	0.90	0.91	0.76	0.77	0.60
	5	0.83	0.51	0.70	0.74	0.73	0.76	0.60
5	2	0.80	0.60	0.90	0.87	0.81	0.84	0.57
	3	0.83	0.49	0.87	0.89	0.90	0.90	0.53
	4	0.80	0.56	0.87	0.84	0.80	0.81	0.46
	5	0.81	0.49	0.81	0.76	0.80	0.76	0.47
Average (1 week out)		0.76	0.56	0.83	0.82	0.80	0.78	0.51
Average (2 weeks out)		0.90	0.70	0.95	0.95	0.90	0.90	0.52
Average (3 weeks out)		0.89	0.74	0.97	0.97	0.88	0.88	0.52
Average (4 weeks out)		0.89	0.79	0.98	0.97	0.87	0.88	0.53

Table 6

Regression of Straddle Returns on Swap Rate Changes

Panel A presents the unadjusted R^2 values when weekly returns on the swaption straddles are regressed against the weekly changes in swap rates. The independent variables correspond to changes in 10 swap rates for maturities of 1 through 10 years. Panel B presents the percentage of error explained by the hedges formed using the four-factor model. Panel C presents the percentage of error explained when the four-factor model is not recalibrated based on swaption data. The data covers the period from March 1st 1998 to October 31st, 2000.

Swaption Expiration	Swap Maturity			
	2	3	4	5
Panel A				
0.5	0.120	0.120	0.136	0.148
1	0.052	0.079	0.086	0.094
2	0.149	0.165	0.143	0.169
3	0.195	0.117	0.143	0.159
4	0.164	0.123	0.184	0.217
5	0.219	0.245	0.250	0.238
Average R^2		15.5%		
Panel B				
0.5	0.724	0.848	0.838	0.859
1	0.822	0.714	0.846	0.855
2	0.882	0.891	0.851	0.858
3	0.883	0.887	0.858	0.888
4	0.886	0.874	0.891	0.880
5	0.892	0.887	0.783	0.870
Average R^2		85.3%		
Panel C				
0.5	0.585	0.735	0.704	0.735
1	0.692	0.460	0.707	0.741
2	0.780	0.803	0.735	0.743
3	0.767	0.788	0.792	0.795
4	0.791	0.766	0.795	0.778
5	0.794	0.796	0.568	0.753
Average R^2		73.3%		

Figure 1

Time Series of Swaption Volatilities

This figure presents the time series of the swaption volatility term structures for each underlying swap maturity, over the data sample period, from March 1, 1998 to October 31, 2000. The volatilities shown are Black volatilities of the underlying forward swap rates.

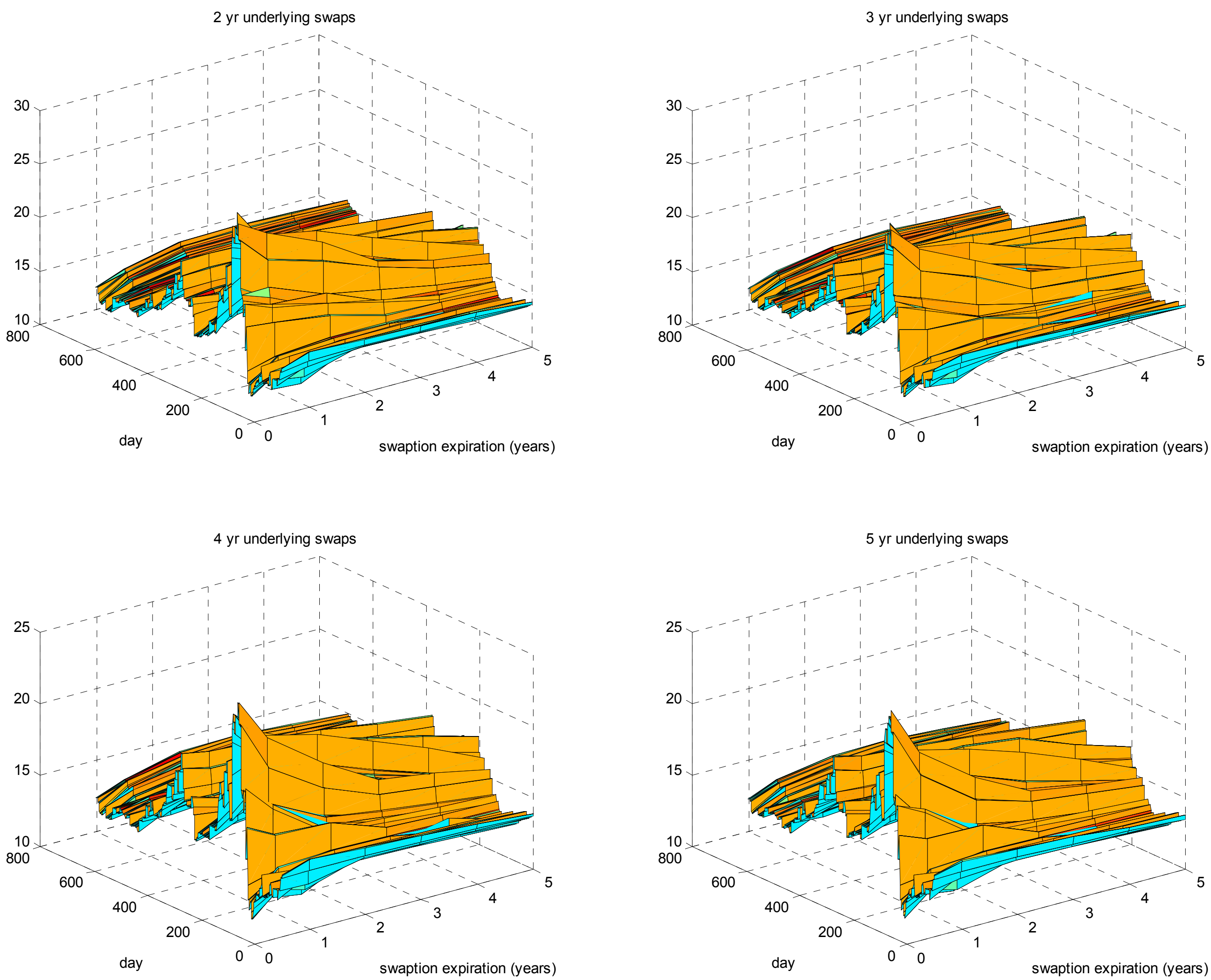

Figure 2

Estimated Volatilities of Forward Rates Implied from Swaption Prices

The figures presents the time series of estimated forward rate volatilities for all the models. There are 70 volatility curves on each figure, each curve separated by two weeks. The data for these curves are derived from the 70 optimization problems for each model.

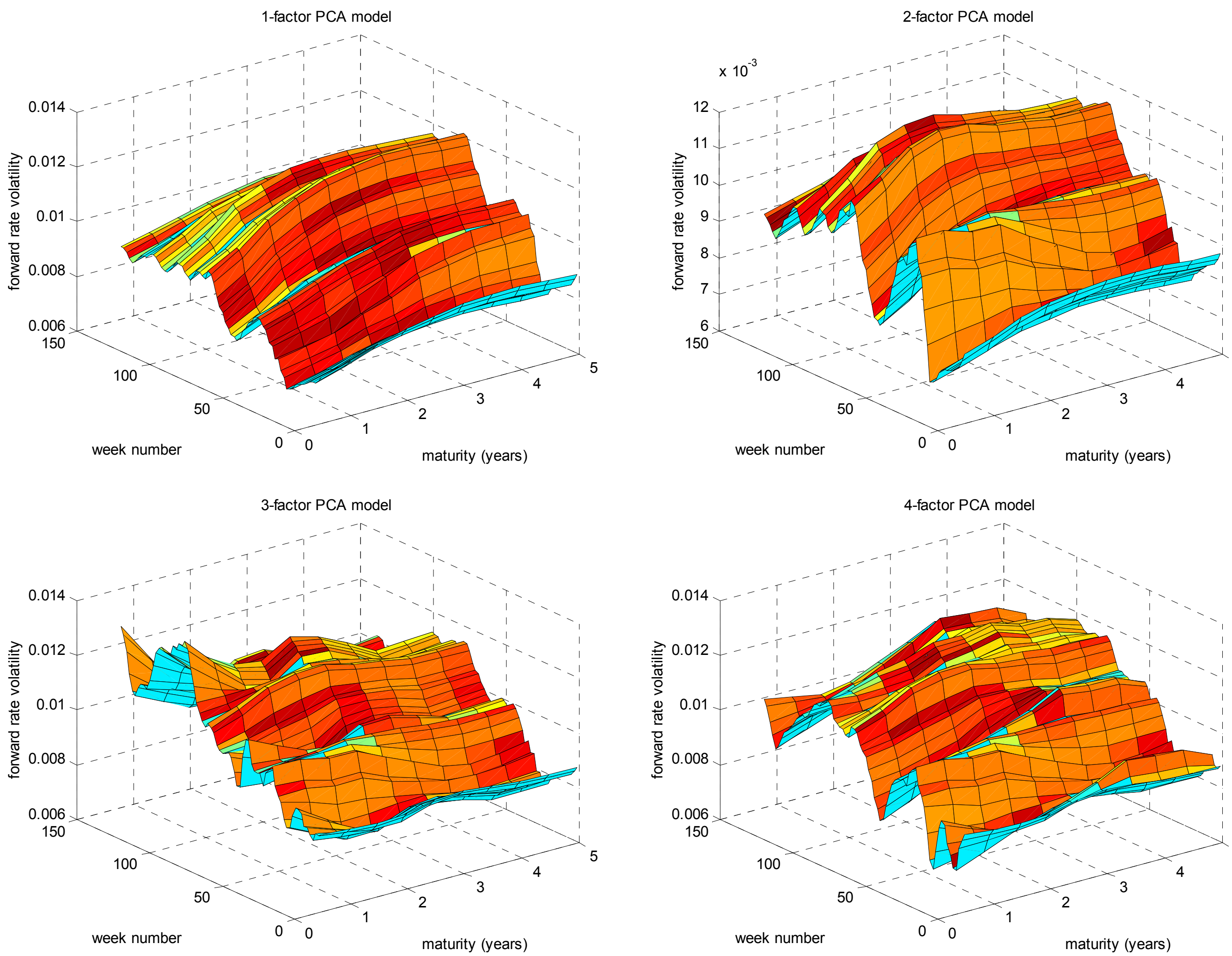

Figure 3

Box and Whiskers Plots of Swaption Hedging Errors

This figure presents the box and whiskers plots for the one week hedging errors, for swaptions across all expirations and underlying swap maturities. The corresponding plots for the unhedged swaptions are presented for comparison purposes. In each figure, the first box corresponds to the unhedged swaption, the second box to the one-factor model, the third box to the two-factor model, the fourth box to the three-factor model, and the fifth box to the four-factor model.

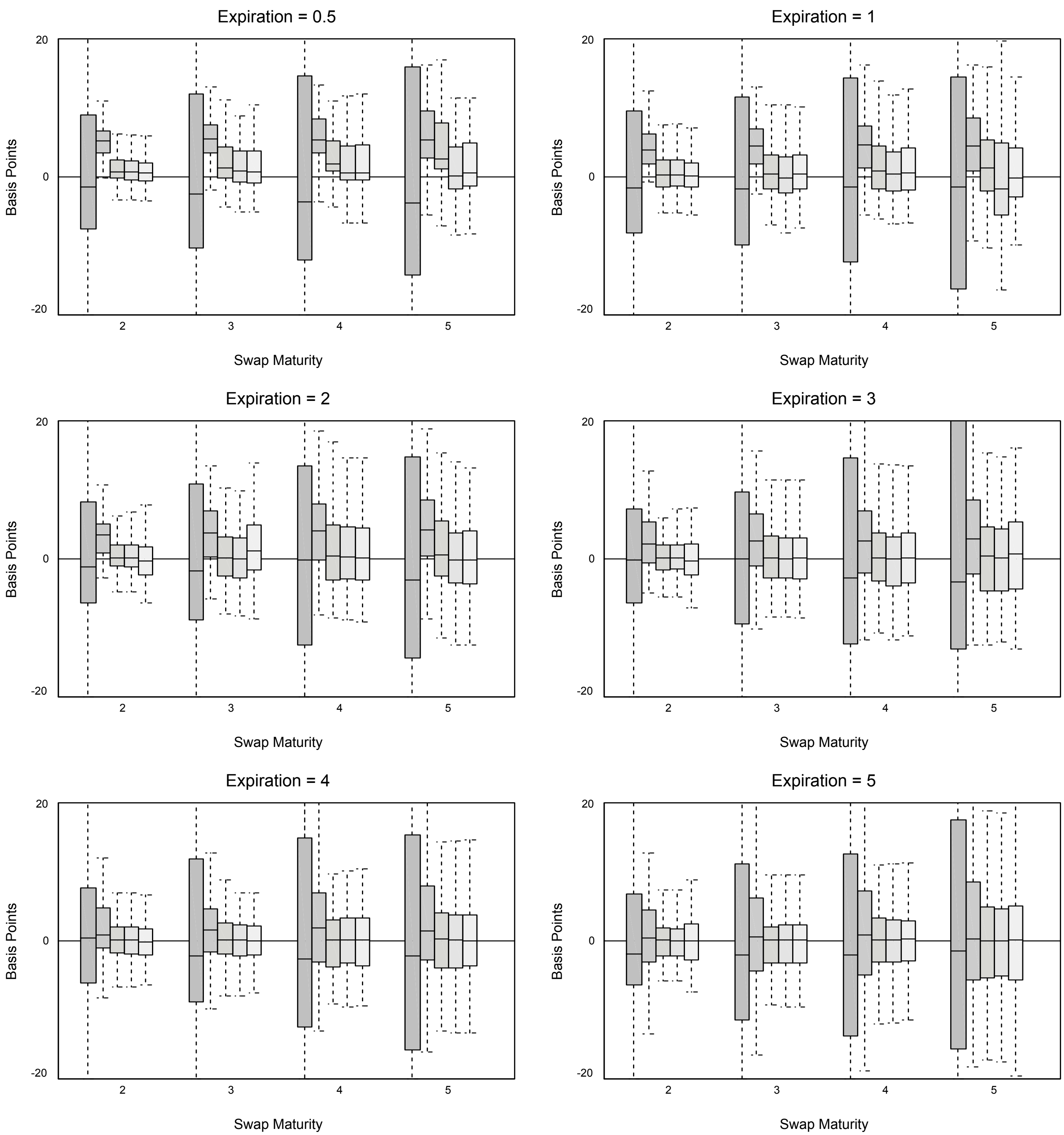

Figure 4

Box and Whiskers Plots of Swaption Straddle Gamma Hedging Errors

This figure presents the box and whiskers plots for the one week hedging errors, for swaption straddles across all expirations and underlying swap maturities. The corresponding plots for the unhedged straddles are presented for comparison purposes. In each figure, the first box corresponds to the unhedged swaption, the second box to the one-factor model, the third box to the two-factor model, the fourth box to the three-factor model, and the fifth box to the four-factor model.

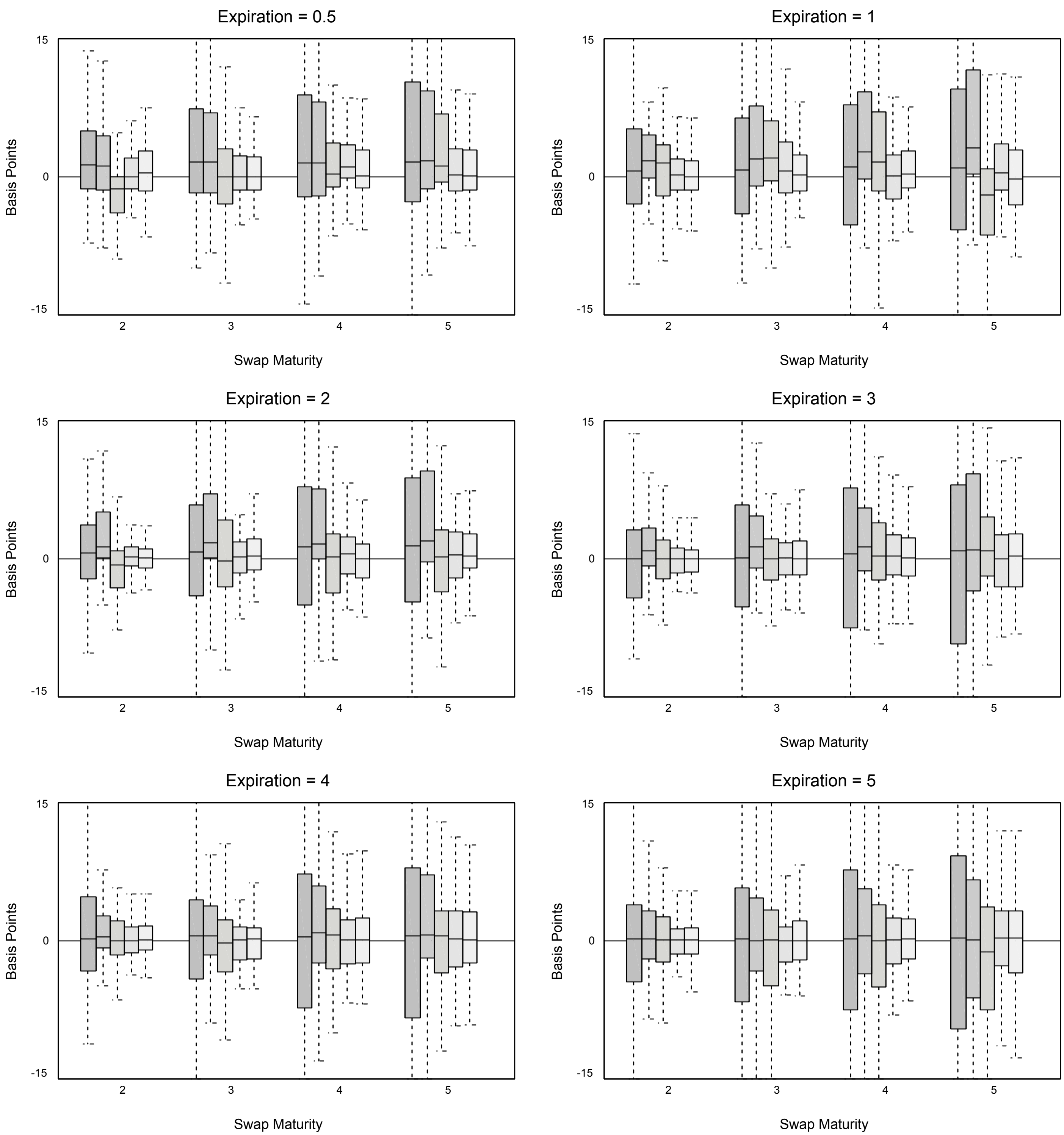

Figure 5

Time Series of Swaption Straddle Hedging Errors Explained by the Four-Factor Model

The figure compares the time series of the percentage of straddle hedging errors explained by the four-factor model, across all contracts by quarter, when the eigenvalues of the four-factor model are updated every second week, with the percentage from the same model with no parameter updating. In the figure, the lower line represents the percentage of the unhedged straddle error, as measured by the total sum of squared error for all contracts over the quarter, explained by the four-factor model with no recalibration. The upper line represents the corresponding percentage for the four-factor model when it is periodically recalibrate