On Pricing and Hedging in the Swaption Market: How Many Factors, Really?*

Rong Fan† Anurag Gupta‡ Peter Ritchken§

October 1, 2001

*The authors would like to thank seminar participants at presentations made at the Eleventh Annual Derivatives Securities Conference in New York, The 2001 European Financial Management Association Meetings in Paris, and the Seventh Annual US Derivatives and Risk Management Congress in Boston.

†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

On Pricing and Hedging in the Swaption Market: How Many Factors, Really?

Abstract

This article examines how the number of stochastic drivers and their associated volatility structures affect pricing accuracy and hedging performance in the swaption market. In spite of the fact that low dimensional one and two-factor models do not reflect historical correlations that exist among forward rates, we show that they are capable of accurately pricing swaptions as well as higher order multifactor models, across all expiry dates and over all underlying swap maturities. Effective out-of-sample pricing is necessary but not sufficient for good hedging. Indeed, regarding hedging, we show there are significant benefits in using multifactor models. This is true even if one accounts for the fact that fewer hedging instruments are required when single factor models are used to hedge swaptions. Our empirical findings have strong implications for the modeling and risk management of an array of actively traded derivatives that closely relate to swaptions.

JEL Classification: G12; G13; G19.

Keywords: Interest Rate Derivatives; Swaptions; Volatility Structures; Pricing; Hedging.

According to the Option Clearing Corporation, the notional amount of derivatives held by US commercial banks is about $40 trillion, with interest rate derivatives contracts accounting for 80% of this total. Over-the-counter contracts account for over 90% of the total notional amount, with exchange-traded contracts making up the rest. In this large market, the primary option contracts are caps and floors on interest rates and swaptions, which are options to enter or cancel swaps. The notional amount of these over the counter contracts exceeds 5 trillion dollars, making them amongst the most important interest rate claims that trade.

Based on the size of this market, it is not surprising that significant effort has been placed on developing pricing models for these claims.[1] Despite the importance of caps and swaptions, there is still wide divergence of opinion on how to best value these claims. It is widely believed that since the term structure of interest rates is driven by multiple factors, interest rate claims should be valued using multifactor models. Standard arbitrage arguments of Heath, Jarrow and Morton (1992) have shown that for pricing purposes, models differ according to the assumptions imposed on the volatility and correlation structures of forward rates. The exact specification of the volatility structures for forward rates, and the appropriate number of factors to be used, are considered to depend on the particular application. For example, Rebonato (1999) argues that while a one-factor model, in which forward rates are instantaneously perfectly correlated, might suffice for the pricing of caps, it is very unlikely to be useful for pricing swaptions, since they depend heavily on the correlation among forward rates.[2]

Recent advances in modeling methodology have made it possible to use multi-factor models to price even complex interest rate claims, like American swaptions.[3] This has led to a deeper discussion on how many factors are really necessary to model interest rate claims such as caps and swaptions, as well as more complex claims like Bermudan and American swaptions. Several studies have attempted to address these issues and, as we discuss in the next section, the results are mixed and somewhat confusing. We attribute much of the confusion to the fact that some studies focus on pricing issues, while others focus on hedging. Actually, the answer depends on whether the model is to be used for pricing alone, or whether it is to be used for hedging.

[1] Indeed, given the liquidity of caps and swaptions, traders often demand that models for interest rate exotics, such as Bermudan swaptions and resettable caps, have the property that they price these claims at, or at least very close to, their market values.

[2] A cap consists of a portfolio of caplets, each caplet representing an option on an individual forward rate. In contrast, a swaption can be viewed as an option on a portfolio of forward rates. As a result, the relationship between caps and swaptions is determined largely by the correlation structure among the forward rates. Rebonato claims that since in a one-factor model forward rates are perfectly correlated, such models will tend to overprice swaptions.

[3] Recent studies by Andersen (2000), Carr and Yang (1999), Longstaff, Santa-Clara and Schwartz (2001b), Longstaff and Schwartz (2000) and Pedersen (1999), for example, contribute to the literature by developing methodology that allows contracts such as Bermudan swaptions to be numerically priced, relative to a given array of European swaptions, and consistent with the current term structure.

As we discuss later on, if models are evaluated solely on pricing performance, then lower order models might be acceptable. However, if models are evaluated based on hedging performance, then a more demanding standard is established, and lower order models might be unacceptable. *Good pricing performance is a necessary condition for a useful model; good hedging performance, however, is sufficient!*

In this article we investigate a broad range of one, two, three and four-factor models with different volatility structures, that incorporate varying degrees of level and maturity dependence to pick up skewness and volatility hump effects. We closely examine how these models perform in pricing and hedging swaptions. Typical data sets in this market consist of at-the-money contracts, with an array of expiry dates and maturity dates of the underlying swaps. Characterizing the biases in prices produced by the models along these two dimensions is important, and has not been well documented by empiricists to date.[4] A good model should have the property that all expiry and maturity effects are well explained. Of course, if we had data on the prices of away-from-the-money swaptions, then skew effects could also be examined. While such data is not available to us, we are fortunate to have price data on caps with multiple strikes, which we use to explore skew effects.

In the first part of the paper we investigate the impact of adding additional factors on pricing swaptions. To measure pricing effectiveness, we calibrate a model using known swaption and term structure data. Once the parameters are estimated we can price claims in the future *contingent* on the future term structure. If the "out-of-sample" residuals are "small" and have no biases, then the model is viewed positively. In our analysis, the "out-of-sample" tests are conducted one, two, three, and four weeks after calibration. We replicate some of the results of Longstaff, Santa-Clara and Schwartz (2001a), hereafter LSS. They consider models with up to eight factors, and show that a four-factor model is necessary to price swaptions accurately. LSS argue that the large improvements obtained by adding factors is due to the fact that low dimensional models are unable to produce realistic correlations among forward rates and getting these correlation effects right is crucial for pricing swaptions. We provide an alternative explanation to their results, and identify specific one and two-factor models that can price swaptions as effectively as their four-factor model. Indeed, for the purpose of pricing swaptions, one-factor models may suffice.

In the second part of the paper we examine hedging effectiveness. Given a model, we can establish a hedge position in bonds. Then, *if the model is correct*, holding the hedge position for a short time increment should lead to small price changes relative to the unhedged position, regardless of the future term structure. The models can then be evaluated based on their hedging errors. We show that models that price well do not necessarily hedge well. In particular, multiple

[4] An exception is Jaganathan, Kaplin and Sun (2001) who characterize the pricing performance of their models according to the expiry dates of the swaption.

factors play a bigger role here. Our one-factor model, which was competitive with a four-factor model for pricing, is much less precise when viewed from the hedging perspective. For hedging swaptions, multi-factor models are desirable! Given these results for vanilla swaptions, it is clear that hedging products such as Bermudan swaptions, and other exotics, which typically are priced relative to a core set of swaptions, will be more effective with multi-factor models.

The paper proceeds as follows. In the first section we review the literature, sort through the current set of confusing empirical results, and highlight the contributions that this article makes to the literature. In the second section we discuss the set of 18 different models that we evaluate. In the third and fourth sections we discuss our data and model implementation. In section five we discuss our experimental design for examining pricing accuracy and hedging precision. In section six we closely examine a nested set of principal component based models. In section seven we compare the performance of our best principal component based model with alternative lower order models. In section eight we examine the hedging effectiveness of our models. Section nine concludes.

1 Literature Review

Amin and Morton (1994) present one of the early tests of alternative forward rate volatility structures. They find that the single factor generalized Vasicek model provides the best out-of-sample pricing performance. Using caplet data, where maturities ranged from 3 months to 10 years, Ritchken and Chuang (1999) show that a generalization of this model, that captures the hump in the volatility of forward rates would lead to significant improvements. Gupta and Subrahmanyam (2001) examine many one and two-factor models for pricing and hedging interest rate caps and floors. Their data set was unique since it contained cap and floor prices with multiple strike prices. Unlike Amin and Morton, they conclude that a one-factor lognormal forward rate model outperforms other competing one-factor models in pricing accuracy, with two-factor models improving pricing performance only marginally. However, for hedging, they find a significant advantage in moving from one to two-factor models. Bühler, Uhrig, Walter and Weber (1999) test different one and two-factor models in the German fixed-income warrants market, where an array of claims trade with maturities up to 3 years. They reject the deterministic volatility structure for forward rates in favor of a model where volatility is proportional to the level of rates. However, unlike Gupta and Subrahmanyam (2001), they find no advantage in moving beyond a one-factor model.

Few empirical studies have been conducted on swaptions. LSS use a string model framework to test the relative valuation of caps and swaptions using at-the-money cap and swaptions data, and find evidence for using at least a four-factor model for swaptions. Their criterion for evaluating models is based on the sum of squared percentage pricing errors. In other words, their criterion is based on pricing accuracy, not on hedging precision.[5] Peterson, Stapleton and Subrahmanyam (2001) develop an extension of the lognormal model of Black and Karasinski (1991) to multiple factors and provide evidence that the addition of a third factor is helpful in pricing swaptions. In this regard, these studies provide support for Rebonato's claim of the importance of the correlation structure.

Not all studies, however, indicate that multiple factors are necessary for improving pricing performance for swaptions. For example, Driessen, Klaassen, and Melenberg (2001), (hereafter DKM) investigate the performance of several Gaussian models, where volatility structures are deterministic functions of their maturities. They show that the out-of-sample pricing performance of swaption pricing models does not necessarily improve as the number of factors increases. Indeed, one of their one-factor models prices swaptions no worse than their multi-factor models and to the same degree of accuracy as LSS's multi-factor model. Jagannathan, Kaplin and Sun (2001) investigate the pricing of swaptions using multifactor Cox, Ingersoll and Ross models. Their preliminary conclusions suggest that increasing the number of factors does not necessarily improve pricing performance. Indeed, adding factors makes the pricing of short term contracts worse.

Very few studies have compared the abilities of different models for hedging swaptions. LSS briefly consider hedging, in the context of their four-factor model, relative to the Black model, but they do not evaluate the benefits of hedging using an increasing number of factors.[6] Perhaps the most comprehensive study is by DKM, who use their Gaussian models to demonstrate that if the number of hedge instruments is equal to the number of factors, multi factor models outperform one-factor models in hedging caps and swaptions. However, they claim that by using a large set of hedge instruments, their one-factor models perform as well as multi-factor models. This last finding is the opposite of what Gupta and Subrahmanyam (2001) find in the cap market. We are unaware of any other studies specific to the swaption market.

Recently, significant research has 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.[7] 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

[5] Hull and White (1999) develop very similar models to the LSS models using the LIBOR market based model of Brace, Gatarek and Musiela (1997) and the extended LIBOR market model of Andersen and Andreasen (2000). A big motivation for extending the existing models was to permit an analysis of volatility structures that were not proportional to their levels. Hull and White test their extended model using data for a single day, and provide preliminary support for multi-factor models where volatilities are not linear in forward rates.

[6] In addition, their hedging tests are not based on the construction of portfolios of traded instruments.

[7] Radhakrishnan (1998) also shows that one-factor models underprice Bermudan swaptions relative to two-factor models.

Andreasen (2001) conclude that the standard market practice of recalibrating one-factor models does not necessarily understate the price of Bermudan swaptions. While their study is useful since it suggests that practitioners are not making systematic errors in marking their Bermudan swaptions to market, it does not fully resolve the issue of how many factors are necessary to model Bermudan swaptions, since they do not investigate any hedging issues.

In summary, some studies find that four factors are necessary for pricing swaptions, while others find that one and two-factor models are satisfactory. Some studies find that multifactor models are necessary for pricing Bermudan swaptions, while others find the current market practice of recalibrating one-factor models to be satisfactory. The few studies on hedging swaptions have also produced mixed results, with one study finding that hedging with a one-factor model, but using multiple hedging instruments, is not worse than hedging swaptions in a higher order model, and another study drawing the exact opposite conclusion.

In performing empirical tests on pricing and hedging effectiveness, there are some important features that we have to consider. First, all the models that we study have time stationary volatility structures.[8] As discussed earlier, our volatility structures are also chosen to accommodate a large number of different skew patterns in implied volatilities. This is accomplished by introducing varying degrees of level dependence into the volatility structures. Both level dependence and time-stationarity assumptions come at some cost. Level dependence results in increasing the number of state variables that are necessary to characterize the dynamics of the term structure, thereby increasing the computational complexity. Without time dependence in the volatility structures, it is not possible to construct models that can match an array of swaption prices exactly.

Introducing time dependence in models can serve to illustrate interesting properties of models and is a common practice at investment banks. Andersen and Andreasen (2001), for example, show that the popular market practice of using continuously recalibrated one-factor models, with time varying parameters, may be a good proxy, even if true prices are generated by higher factor models where the correlation among forward rates is lower. Using cleverly introduced time varying parameters in one-factor models may help explain whether exercise decisions are being well proxied, and whether such models serve a role for correctly marking exotic products to market. However, even if a clever choice of adjustment factors is made in a one-factor model so as to produce the same price as a properly calibrated higher factor model, the hedges produced by the two models will be distinct, and the hedging effectiveness of the one-factor model will be suspect, especially if the time varying parameters used in the model are not stable over time. Since our goal is not only to assess whether one-factor models are appropriate for pricing accuracy, but also to assess their hedging effectiveness, we restrict our analysis to

[8] If market participants had strong beliefs that volatilities would change over time in a specific manner, then we could accommodate this. However, in practice, this assumption may be unreasonable.

stationary volatility structures, where the number of free parameters that need to be estimated are limited, and hence exact matching of European swaptions to their market prices is unlikely. In addition to not using time varying parameters, all the models that we examine have at most four free parameters in the volatility structures, so at best we could only match four swaptions at any point in time. The models considered by LSS have the property that the number of free parameters equals the number of stochastic drivers. In contrast, some of our low order one and two-factor models contain as many free parameters as our four-factor models.

Our first goal is to untangle some of the conflicting *pricing* results in the swaption market. In particular, we compare the pricing performance of several single and multi-factor models with different volatility structures and identify those models that eliminate most of the pricing biases in the swaption market. In this regard, our paper is closely related to DKM and LSS. However, the former only study Gaussian models, where the volatility structures are independent of their levels, while the latter only investigate a specific family of nested models that have proportional volatility structures.

For swaptions, we first confirm the results of LSS, that from a pricing perspective, increasing the number of factors up to four (within principal component based models) provides improved precision in estimating out-of-sample prices. LSS argue that the reason for such large improvements is due to correlation effects. We show that similar results are obtained for caps which may be less sensitive to correlation effects. We then test whether the improvement in these types of models is due to increasing the number of factors, or increasing the number of parameters. We do this by considering several one and two-factor models that have the same number of free parameters as our four-factor models, and compare their pricing performance. In this case, we find that for pricing swaptions, the benefits of increasing the number of factors beyond one is minor. Indeed, we conclude that from a pricing perspective only, there appears to be little advantage in moving from a one to a four-factor model. Our pricing results hold true even when swaption prices are generated upto four weeks after the parameters were estimated. For pricing purposes, the importance of models that better reflect correlation structures among forward rates is minor.

We also address the importance of level dependence in the volatility structures. Since volatility skews are hard to assess in the swaption market, because the prices of contracts that are available are restricted to at-the-money contracts, we turn to the cap market to get evidence for the pricing performance over the strike price domain.[9] Our results show that incorporating level dependence in the volatility structure is extremely important for away-from-the-money caps, and that proportional dependent structures are better than both square root or level independent structures. For at-the-money swaptions, the level dependence issue is minor. However, the

[9]Incorporating level dependence in the volatility structure results in distributions of forward rates that are no longer Gaussian. This has large implications for pricing claims that derive value from the tails of the distribution.

evidence from cap prices suggests that away-from-the-money swaptions would be better priced using proportional models rather than level independent structures.

The removal of systematic expiry date and underlying swap maturity biases in out-of-sample pricing is a necessary condition for a model to be useful for hedging, but it is not sufficient. The second goal of our research is to investigate the ability of alternative models in *hedging* swaptions. In particular, we want to carefully quantify the benefits, if any, of using higher order factor models over lower order models. We investigate the effectiveness of alternative hedging strategies using different models and differing numbers of hedging instruments and produce convincing evidence that multifactor models are essential for reducing the risk in hedged positions. We also demonstrate that allowing additional hedging instruments in a one and two factor model does not improve the results. Our main conclusion is that while accurate swaption prices can be obtained from a one-factor model, one and even two-factor models cannot hedge swaptions well, and the benefits of multifactor models are significant.

The results of our research have obvious implications for practitioners who are uncertain as to which models to implement for the pricing and hedging of swaptions. While one-factor models can be as effective as multifactor models for marking to market purposes, for hedging purposes, multifactor models are preferable. Typically, interest rate derivatives desks trade many exotic contracts, including resettable caps and Bermudan swaptions. Since traders often demand that models for these products have the property that they price liquid swaptions and caps at or at least close to their observed market prices, modelers often begin the process by specifying the number of stochastic drivers and the nature of the volatility structures. While these choices may depend on the particular application, our research casts light on the benefits of beginning with particular structures. Our research also casts light on value at risk systems. In such systems many interest rate claims have to be priced under different scenarios, and the question of how many factors to incorporate needs to be addressed. In such cases, given a future possible forward rate curve, an accurate set of prices with one-factor models may be satisfactory. On the other hand, if the desk is setting up a hedge, then the requirements of a model are more demanding and multifactor models are required.

2 The Basic Models

Caps and 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 maturities and strikes trading at different implied volatilities. Since the volatilities used in the Black model are for forward rates, for the case of caps, and swap rates, for the case of swaptions, direct comparisons of cap and swaption volatilities are not meaningful. Indeed, 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.

As discussed, all the volatility structures that we consider are time homogeneous. Let

$$\sigma_{f_1}(t,T) = \left[[a + b(T-t)]e^{-\kappa(T-t)} + c\right] f(t,T)^{\gamma}. \tag{2}$$

This parametric volatility structure nests many well known models. First, when $\gamma = 0$, the volatility structure is a deterministic function of maturity. With $c = b = 0$ the model reduces to the generalized Vasicek model, commonly referred to as the single factor Hull and White (1993) model, in which forward rate volatilities dampen with their maturity. With b and c released from 0, the model can accommodate hump shapes in forward rate volatilities. Such models have been considered by Moraleda and Vorst (1997) and Ritchken and Chuang (1999). They have the attractive feature that analytical solutions can be set up for pricing many derivative contracts. However, they have also been criticized since interest rate volatilities do not depend on their levels and can therefore become negative. When γ is positive, the volatility structure does depend on the level of rates. In this case, however, computational problems emerge. In particular, the term structure is no longer Markovian in a finite number of state variables.[10] Fortunately, Monte Carlo simulation provides a powerful tool for computing European claims, and recently, Longstaff and Schwartz (2001) and Andersen and Broadie (2001), among others, have shown that American claims can also be computed using rather simple and efficient methods.[11]

[10]For a discussion of Markovian Heath, Jarrow, Morton models see Ritchken and Sankarasubramanian (1995a) and Bhar and Chiarella (1995)

[11]For a review of simulation approaches for pricing claims see Boyle, Broadie and Glasserman (1997).

We consider three parametric one-factor models, where the volatility structure is of the form in equation (2). The models differ according to the level dependence parameter, γ. In particular, we consider models where γ is zero, one half, and one. All these models have four parameters.

In our two-factor models, the first volatility structure is of the form in equation (2), with $c = 0$. The second volatility structure is of the form:

$$\sigma_{f_2}(t, T) = d[r(t)]^{\gamma}. \tag{3}$$

For this two-factor model, the shocks to forward rates consist of two types. The first consist of a shock that depends on the level of the forward rate and on the maturity. Over the short end the structure permits an increasing volatility, but eventually the shock dampens with maturity. The second shock has a "parallel" effect over maturity. The absolute magnitude of this effect is driven by the level of the short rate. A structure similar to this model has been empirically examined by Inui and Kijima (1998). For $\gamma = 0$ this model nests the two-factor generalized Vasicek model of Hull and White (1993). For $\gamma = 0.5$ and 1 the model is similar to a generalized Cox, Ingersoll and Ross (1985) model and a proportional model respectively. Note that all three of our two-factor models also have four free parameters.

The final set of models are based on modifying 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:

$$\sigma_{f_j}(t, t + \tau_j) = g(\tau_j) f(t, t + \tau_j)^{\gamma} \tag{4}$$

where $g(.)$ is a deterministic function of the maturity of the forward rate, that is estimated primarily using principal component analysis on historical data. In particular, take the case where $\gamma = 0$. Using historical data on weekly forward rates, a correlation matrix of forward rate changes, separated by three months for maturities less than a year, and six months thereafter (up to ten years maturity), is obtained. In particular, 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

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

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:

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

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.

When γ differs from 0, the same analysis is done, except the correlation structure for the principal component analysis is estimated over transformed values of forward rates. For example, when $\gamma = 1$, the correlation is estimated over percentage changes in forward rates.

The above method, which we term an adapted Principal Component Analysis (PCA) method has been used by LSS for proportional models ($\gamma = 1$) and by DKM for absolute models ($\gamma = 0$). In addition to these, we also permit $\gamma = 0.5$. Notice that, like all our other models, our four-factor model has four free parameters that can be implied out using option data. Notice too, that if a three factor model is used, then only the first three principal components are retained and the number of free parameters drops by one.

In summary, we consider twelve PCA models (3 one-factor, 3 two-factor, 3 three factor and 3 four-factor models) and six parametric models (3 one-factor and 3 two-factor models). All the parametric models and the four-factor PCA models have four free parameters. The other models have as many free parameters as stochastic drivers.

3 Data

The data for this study consists of USD swaption and cap prices. The swaptions data set comprises volatilities of swaptions of maturities 6 months, 1-, 2-, 3-, 4-, and 5-years, with the underlying swap maturities of 1-, 2-, 3-, 4, and 5-years each (in all, there are 30 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 cap prices are for a ten-month period (March 1 - December 31, 1998), across four different strikes (6.5%, 7%, 7.5%, and 8%) and four maturities (2-, 3-, 4-, and 5-year), obtained from Bloomberg Financial Markets. For swaptions, the data consists of at-the-money volatilities for a 32 month period (March 1, 1998 - October 31, 2000), obtained from DataStream.

For constructing the yield curve, we use futures and swap data. For the short end of the curve (upto 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 at 6 month intervals from market swap rates. The futures and swap data is obtained from DataStream. Eventually, we obtain weekly forward rate curves that start one year before our cap and swaption data begins, and extends to the end of our swaption data period.

For the principal component analysis we follow the procedure used in LSS. Specifically, we use the one year history of forward rates that exist prior to the beginning of our swaption data, to estimate the correlation structure of forward rates. For example, for $\gamma = 1$, we estimate the percentage changes in forward rates from the historical time series 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. A similar analysis is done for the models with $\gamma = 0$ and $\gamma = 0.5$.

4 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 and on 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:

$$\Delta f^{\Delta t}(t, j) = \mu_f^{\Delta t}(t, j)\Delta t + \sum_{n=1}^{N} \sigma_{f_n}^{\Delta t}(t, j)\sqrt{\Delta t} Z_{t+1}^{(n)}. \qquad (5)$$

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

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

and

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

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

Prices of European interest rate claims can be computed using Monte Carlo simulation. Specifically, we simulate $K = 2000$ different paths, 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 (5). 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 $t\Delta$ 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 $K = 2000$ 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.

5 Experimental Design

Like Amin and Morton (1994), DKM, LSS, and Moraleda and Pelsser (2000), 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 four orders of magnitude apart, such a minimization would place more weight on the expensive contracts.

The first set of experiments examine the pricing issues for at-the-money swaptions. Using the swaption data, for each odd week, using mid-week data, we establish the best fit for the prices of all swaptions. Given these parameter estimates, we forecast one, two, three and four week out-of-sample residuals, which are all labeled according to their in-sample time period, model and contract.

A similar set of experiments are conducted on cap prices. For each of our models we establish the best fit for the prices of caps for the 4 strikes and 4 maturities. We do this using mid-week data, in separate optimizations, for every odd week. We then use the parameter estimates, together with the term structure the following week, to generate one week "out-of-sample" residuals. In addition, we also compute two, three and four week out-of-sample residuals. These residuals are stored for each model, for each contract and for each date.

Our hedging experiments are 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. The analysis is repeated for holding periods of two, three and four weeks as well. Finally, this entire procedure is repeated every second week, over all 70 weeks, for which data was available. Further, as we will discuss, alternative criteria are applied to select the hedging instruments, and in some cases, more hedging instruments were used than factors.[12]

[12]For a discussion of how best to use simulation models for establishing hedge ratios see Jäckel (2001).

6 Pricing Performance of PCA Models

Table 1 presents the average absolute percentage errors for one week out-of-sample swaption and cap prices produced by the PCA models.

Table 1 here

The results show that for any given value of γ, increasing the number of factors improves the pricing performance for swaptions. This confirms the results obtained by LSS, who conclude that four factors are necessary for pricing swaptions. Table 1 also shows that for a fixed number of factors, increasing γ has no significant effect on reducing the average absolute percentage errors for swaptions. Since the level dependence parameter, γ, controls the skewness of the risk neutral distribution, and since all the swaptions are at-the-money contracts, these results are not surprising.[13]

Table 1 also presents the results for caps. For any given number of factors, the average absolute percentage errors for models with $\gamma = 1$ are consistently lower than those for models with $\gamma = 0$, with the errors for models with $\gamma = 0.5$ being in between the two. In contrast to swaptions, the advantage of increasing the number of factors appears to be small.

In order to further investigate the effects of level dependence, Table 2 presents the proportion of times that a PCA model with a particular γ value produces one week out-of-sample cap prices closer to actual market prices, than the same model with a different γ value. These proportion tests are based on a total of 352 residuals each, since there are 16 cap contracts on each of the 22 dates. For example, for the one-factor PCA model, 66% of the residuals from a $\gamma = 0.5$ model are closer to zero than the residuals produced by the $\gamma = 0$ model. For each of the PCA models, the precision increases with an increase in the γ value.

Table 2 Here

These results provide conclusive evidence for the use of proportional models over square root and absolute models for explaining the skew effects in cap prices, since they price away-from-the-money contracts better. This suggests that away-from-the-money swaptions would also be better priced by proportional models.

In order to further examine the effects of increasing the number of factors, Table 3 reports the proportion of times a higher order PCA model outperforms a lower order PCA model in pricing caps. Specifically, for each of the four PCA models and the three γ values, we estimate the one

[13] Prices of away-from-the-money contracts convey more information about the tails of the conditional distribution of forward rates and are more sensitive to γ. For example, see Ritchken and Sankarasubramanian (1995b).

week out-of-sample prices, and compute the residuals. Again, since there are 16 contracts at each date and 22 dates, we have 352 residuals for each model-gamma combination.

Table 3 here

For each pair of models, Table 3 reports the proportion of times (out of 352) that one model produces smaller residuals than the other. For example, for γ =1, the two-factor PCA model produces prices closer to the actual market prices 57% of the time, when compared to the one-factor PCA model. The results show that, at the 1% level of significance, for all levels of γ, the two-factor, three-factor and four-factor PCA models outperform the one-factor PCA model. The three-factor model outperforms the two-factor model. However, the four-factor model does not produce results significantly different from the three-factor model.

Table 4 presents similar tests for swaptions. In each of the 70 time periods there are 30 contracts, for a total of 2100 residuals. The table presents the percentage of times the higher order model produces smaller absolute residuals than the lower order model. The results are presented for the in-sample residuals, as well as for one, two and four week out-of-sample residuals, only for the proportional model.

Table 4 Here

These results show that within a PCA framework, the two-factor model outperforms the one-factor model, the three-factor model outperforms the two-factor model, and the four-factor model outperforms the three-factor model. These results are significant at the 1% level of significance, and are consistent with those reported in other studies. When the results are broken down and analyzed by expirations (or underlying swap maturities), the same trend is observed. Overall, the marginal improvement from adding an additional factor decreases as the number of factors increases, and the contribution of the fourth factor, while statistically significant, is small. The relative performance of PCA models for swaptions remains the same, even as the out-of-sample period is increased to four weeks. Although not reported in the paper, we also observe that the inference drawn from caps remains unchanged as the out-of-sample period increases from one to four weeks.

In summary, we show that for at-the-money swaptions, level dependence is not important, but increasing the number of factors (within PCA models) to four improves pricing performance. Since level dependence is important for away-from-the-money caps, we postulate that it is important for away-from-the-money swaptions as well. Finally, for caps, increasing the number of factors to three or four significantly improves pricing performance.

LSS attribute the better performance of the multi-factor PCA models for swaptions to the fact that they are capable of producing more realistic correlation structures for the forward rates. On the other hand, since caps are less sensitive to correlation effects, the improvement in their pricing performance, as the number of factors increases, might be surprising, and raises the possibility that there are other explanations. For example, the improvements in performance might be due to the fact that higher order models have more free parameters than lower order models, and the improvements arise because of these extra degrees of freedom.

7 Pricing Performance of Parametric and PCA Models

In this section, we establish whether the improvement in pricing swaptions is due to increasing the number of free parameters in the model, or due to increasing the number of stochastic drivers. In order to do this, we focus on the two parametric models with one and two stochastic drivers, and on the four-factor PCA model (which is the best model among the PCA models), all of which have four free parameters. The four-factor PCA model clearly comes closest to matching the historical correlations among forward rates. If this, indeed, is an important feature of the swaption market, then this model should outperform one and two-factor models with the same degrees of freedom. On the other hand, if correlations are less important, than the one and two-factor models should be comparable. In addition, the previous results for PCA models were based on aggregate statistics. In this section, we examine the potential biases in the models when the contracts are broken down by expiry and underlying swap maturity dates.

Table 5 compares the average absolute swaption errors in the out-of-sample prices by each model.

Table 5 Here

The one week out-of-sample results are similar to the results obtained for the PCA models, where the importance of γ was found to be minor. In light of the importance of γ in pricing away-from-the-money caps, we therefore focus on models with $\gamma = 1$. Table 5 shows the two, three and four week out-of-sample performance of the models with $\gamma = 1$.

What immediately stands out is the performance of the one-factor parametric model. In particular, one week out-of-sample, the model prices swaptions with average absolute errors less than three percent, which is typically within the bid-ask spreads in the swaptions market. Indeed, for the one-factor parametric model, of the $2100 (70 \times 30)$ one week out-of-sample residuals, 84% (56%) were within one (one-half) Black vol, and almost one-third were within one-quarter of a Black vol.

Table 5 also shows that, in aggregate, the two-factor parametric model outperforms the one-factor parametric and the four-factor PCA models. Therefore, of all the models, there is a slight preference for the two-factor model. The important point here, however, is that the four-factor model, at the aggregate level, does not dominate the lower order models. However, since the above conclusions are based on aggregate results, they may mask any expiration date and swap maturity biases.

7.1 Effects of Level Dependence on Pricing

To more closely examine the effects of level dependence on the pricing of swaptions, we compute the average percentage errors, one week out-of-sample, for each swaption expiration and underlying swap maturity, for each model. Figure 1 presents the average percentage errors for each swaption expiration, plotted against the underlying swap maturity, for all three models and for all three γ values.

Figure 1 Here

In general, the one-factor model underprices swaptions with lower underlying swap maturities, and overprices swaptions with higher underlying swap maturities. However, the figures clearly show that the inclusion of level dependence does not eliminate this bias. Indeed, the pricing patterns are strikingly similar for all three values of γ, across all swaption expirations and underlying swap maturities. There appears to be no benefit of incorporating level dependence in the one-factor model, which confirms the results that were obtained at the aggregate level in Table 5.

The last two columns of Figure 1 show that the two-factor parametric model and the four-factor PCA model produce similar results. Interestingly, in the PCA models, it is not just the lower underlying swap maturity swaptions that are underpriced. For some swaption expirations, even the longer underlying swap maturity swaptions are underpriced, with the medium swap maturity swaptions being overpriced. Figure 1 indicates that incorporating level dependence has little effect on pricing at-the-money swaptions.

As discussed earlier for the PCA models, we turn to our cap data to assess the effects of level dependence in forward rate volatility structures. Table 6a presents the average absolute pricing errors for caps, one week out-of-sample.

Tables 6a and 6b Here

This table shows that models with $\gamma = 1$ are significantly better for pricing sway-from-the-money caps. In Table 6b we provide pairwise comparisons across models with different γ values. For all cases, the models with $\gamma = 1$ are the best, and the models with $\gamma = 0$ are the worst. Figure 2 presents the breakdown of the biases by moneyness for differing maturity dates and models.

Figure 2 Here

The figure clearly identifies skew effects, especially for the models with $\gamma = 0$ and $\gamma = 0.5$. The models with $\gamma = 1$ explain more of the skew effects than the other models. In light of the evidence on skews in the cap market, and the indifference of level dependence for at-the-money swaptions, we only consider models with $\gamma = 1$.

7.2 Effects of the Number of Factors on Pricing

So far, we have presented results for swaptions at the aggregate level. We now examine individual contract pricing errors in more detail, across the models, to determine if there is any significant benefit to moving beyond a one-factor model for pricing at-the-money swaptions. Figure 3 presents the box and whiskers plots of one week out-of-sample errors for the one- and two-factor models, and the four-factor PCA model. The plots are presented separately for each swaption expiration, with the errors for the three models plotted across underlying swap maturities.

Figure 3 Here

For the six month expiry contracts, the one-factor model produces fairly unbiased prices with small variances for all 5 underlying swap maturities. Indeed, with few exceptions, over all 30 expiry-swap maturity combinations, the bias and the interquartile range for the one-factor model appear to be no worse than those for the four-factor model. Further, for any expiry date, the trends of the biases across swap maturities appear to be random. The plots suggest that the largest differences between the models arise for short term expiration dates.

The percentage errors for almost all of the 30 expiration - swap maturity buckets are within 3 percent. Table 7 presents the proportion of one week out-of-sample residuals that are within one, one-half, and one-quarter of a Black vol. of the actual price.

Table 7 Here

As can be seen, the precision of the three models is similar. Over 80% of all contracts are within one vol, with almost 60% of the contracts being within half a vol. This represents a reasonable bid-ask spread during the time period analyzed in this study.

Table 8 provides pairwise comparisons among the three models for each expiry-underlying swap maturity combination. While there are some expiry dates-swap maturity combinations where the four-factor model performs the best, overall, there appears to be a preference for the lower order parametric models. Indeed, Table 7 suggests that the differences in the performance of the models is small, with all models producing a high percentage of prices within bid-ask spreads.

Table 8 confirms this inference, with more formal pairwise proportion tests across the models. The table presents the proportion of times that one model produces more precise one week out-of-sample prices than another, for contracts broken down by expiration and swap maturity. We first examine the summary statistics. In comparing the residuals produced by the different models, contract by contract, the two-factor parametric model produces smaller residuals than the one-factor model 55% of the time, and smaller residuals than the four-factor PCA model 54% of the time. The four-factor PCA model does not outperform the one-factor parametric model. Therefore, we conclude that on an aggregate basis at the 1% level of significance, the two-factor parametric model is the best, with no significant differences between the one-factor parametric and the four-factor PCA models. However, the table allows us to identify where some models fail. Specifically, the reason for the poor performance of the four-factor PCA model is its inability to price both very short and very long dated swaptions.

Table 8 Here

The results in Tables 7 and 8 reaffirm the conclusion that was derived from the aggregate results, that all three models produce results that are satisfactory, with the two-factor parametric model being marginally better. Unlike LSS, we do not conclude that higher order models are necessary for effectively pricing swaptions, one-week out-of-sample and even beyond, up to four weeks.

7.3 Implied Volatility Structures from Swaptions and Caps

Figure 4 presents the time series of forward rate volatilities, estimated using our three models with $\gamma = 1$. The left panel shows the estimated volatility structures implied out using swaption data for each second week over the entire period of 140 weeks, while the right panel shows the same volatility structures implied from the cap market for each second week over the data period of 44 weeks (which falls almost in the middle of the swaption data).

The volatility surface for the swaption data is clearly humped, with the maximum volatility occuring in the second year. The two-factor model also displays the hump, although there are a few periods, in the latter part of the time series, where the peaks are very high and occur in the first year. Finally, the volatility structures for the four-factor model are more spiked, but similar.

The volatility surfaces for forward rates implied out from the cap market are also, for the most part, humped. For the first 20 weeks, the one-factor model produces a volatility hump of forward rates that is fairly stable over time. However, in a few of the latter periods, the implied structure of the volatility of forward rates expands over the five year horizon. The estimated volatility structures for the two-factor model and the four-factor PCA model also reflect the expanded volatilities over the latter part of the data set.[14]

Figure 4 Here

In comparing these two panels, we observe that the volatility structures implied out by swaptions are more stable over time, for all the models. However, the overall patterns are fairly similar.

Unlike the volatility structures of forward rates, the implied correlations produced by the three models are very distinct. The average correlation between forward rates and the spot rate produced by the two parametric models is much higher than the actual correlations over the time period, while the four-factor model, as expected, produces values fairly close to the actual correlations over the time period.

Our analysis shows that for pricing swaptions it may not be necessary to require accurate calibration of correlations among forward rates, and high dimensional models, such as three or four-factor models, may not be necessary.

8 Hedging Performance of Parametric and PCA Models

Our analysis shows that *conditional* on a future term structure, all three models are capable of producing fairly precise estimates of swaption prices one to four weeks after the parameters for the volatility structure have been estimated. The removal of systematic expiry date and swap maturity biases from a pricing model is a necessary condition for any viable model for swaptions. However, a good model should also be able to hedge effectively. If the 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 known with certainty and continuous revisions were not accomplished over the time increment. 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. Evaluating models based on how accurately they price in the future *conditional* on the future set of bond prices, is much less demanding than evaluating models based on whether changes in swaption prices can be replicated by changes in particular portfolios of bonds, where the hedge ratios are determined by the model. The pricing accuracy

[14] The latter part of this data set corresponds to the period immediately after the solvency threat of Long term Capital Management, when interest rate volatilities did spike up. For a discussion on this point see LSS (2001a).

tests are in some sense conditional tests, where the future term structure is taken as given. In contrast, the hedging tests are more demanding in that they are unconditional tests that require more precision in estimating how swaption price changes are related to the underlying dynamics of bond price changes. A model that prices well and removes systematic biases may be a good model. A model that can be consistently used to construct efficient hedges *is* a good model. Pricing accuracy is necessary for a good model, but hedging precision is sufficient!

We first compare the hedging effectiveness of the one-factor model using one discount bond, with the two-factor model using two discount bonds and the four-factor model that hedges using four discount bonds. For the one-factor model, we take the discount bond corresponding to the maturity date of the underlying swap. For the two-factor model, the two hedging instruments are taken as the discount bonds corresponding to the expiration date of the swaption and the maturity date of the underlying swap. 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 choice of maturity, the hedges are uniquely determined. Since no analytical solutions are available for the hedge ratios, we use Monte Carlo simulation to set them up. Specifically, the initial term structure is perturbed by a small shock to a specific factor, and then the price of each swaption and hedging instrument is recomputed. The sensitivity of the swaption price to the underlying hedge instruments is then established. For the one-factor model, only one-factor is perturbed, and the hedge ratio is readily computed. For the two-factor model, two separate perturbations are involved that lead to two equations in the hedging instruments with two unknowns, and a unique hedge position that immunizes shocks to the two model factors is then computed. Finally, for our four-factor model, each shock corresponds to a shock to the principal component. Given four hedging instruments, the hedge portfolio immunizes shocks to the first four factors.

8.1 Effects of the Number of Factors on Hedging

The benchmark hedging period is one week, i.e., the hedge is set up and then evaluated after one week. For each swaption contract, a time series of weekly unhedged and hedged residuals are obtained. The ratio of the standard deviation of the hedged position versus the standard deviation of the unhedged swaption over the 70 weeks provides one measure of the effectiveness of the hedge for a particular swaption contract. This is the criterion used by DKM and is equivalent to investigating the R^2 values or percentage of variance explained by the hedging variables in a regression analysis.[15]

Table 9 presents the ratio of the standard deviations of the hedged versus unhedged positions for the three models. The analysis is limited to contracts with at least two years between the

[15] This latter methodology is a popular method for evaluating hedging effectiveness.

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.

Table 9 Here

The magnitude of the standard deviation ratios is impressive. For example, reported at the bottom of the table is the average hedge effectiveness over all contracts. For the one-factor model, the ratio is 0.34. Squaring this number leads to 0.1156, which implies that the hedge reduces the variance by over 88%. The average performance if the hedge is maintained unchanged for two, three and four weeks is also reported in the table. Over the four week period, the average ratio is 0.37, which translates into a variance reduction of about 86%.[16] In contrast, the four-factor model accounts for about 91% of the variance of the unhedged position.

In comparing the ratios of standard deviations across models, contract by contract, there appears to be very minor improvement as the number of factors increases. At first glance, this indicates that there is little benefit in increasing the number of stochastic factors in a model beyond one. Indeed, this criterion is used by DKM in comparing alternative models. However, comparing standard deviations 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. Figure 5 compares the box and whisker plots of hedging errors for each contract type across the three models. It also presents the unhedged pricing errors.

Figure 5 Here

The figure immediately shows that all hedges are doing their job! However, it also shows that the one-factor model has larger errors. While the spread of the errors, as indicated by the width of the inner box, are of similar sizes for the three models, the bias in the results are greatest for the one-factor model. 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 and two-factor models. In contrast, the interquartile ranges are somewhat similar. This phenomenon holds true for almost all expiry dates, with the exception

[16] DKM keep their hedge in place for two weeks and their best one-factor Gaussian model reduced variance by 60%. Of course, their data period is different, and they included caps in addition to swaptions, so a direct comparison may not be fair.

of the long term contracts. Here, the bias is small, but the variance of the residuals produced by the one-factor model is larger.

Figure 5 clearly shows that the analysis in Table 9 cannot be used to infer hedging effectiveness, and that a more appropriate hedging test is that based on the root mean squared error of the hedging errors. Table 10 presents the root mean squared error (multiplied by 10000) for each contract type. As a result, each entry can be interpreted in basis points.

Table 10 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 PCA 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 four-factor model become apparent.

As a more formal test of hedging effectiveness, we conduct pairwise comparisons of the hedging residuals produced by each model. Specifically, 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 presented in Table 11. For example, for the six month contract, with swap maturity of two years, out of 70 hedging experiments, the two and four-factor models produce smaller absolute hedging residuals than the one-factor model on 87% of occasions. In 66% of occasions, the four-factor model outperforms the two-factor model. If the hedges are maintained unchanged for two weeks, the performance of the multifactor models improves even further. Indeed, over the four week period, the four-factor model outperforms the one-factor model in all but one week.

Table 11 Here

The table clearly shows that over all contracts, and for all weeks, the multifactor model is dominant. For hedging purposes, multifactor models are necessary!

8.2 The Effects of Increasing the Number of Hedging Instruments

DKM show that when bucket strategies are used for hedging, the performance of the one-factor models improves significantly. In light of their result, we set up experiments where the benefits of using additional hedging instruments in a one and two-factor model could be assessed.

The hedging instruments for these tests are selected as follows. Let the swaption expiry date be labeled date 1, and the underlying swap maturity date be labeled date 4. So far, for the one-factor model, we only considered hedging using the date 4 maturity discount bond. In contrast, the four-factor model uses the date 1 and 4 bonds, together with two equally spaced bonds in between. Let these two corresponding dates be 2 and 3, respectively. For the one-factor model, we now consider hedging the swaption with two discount bonds, with three bonds, and with four bonds. In particular, with two bonds, we choose bonds 3 and 4; with three bonds, we choose bonds 2,3 and 4; and with 4 bonds we use the same set of bonds as the four-factor model.

If more than one hedging instrument is used in a one-factor model, the hedge ratios are not unique, and a somewhat arbitrary rule must be made to construct the specific bond hedges. We use equal number of bonds in each hedge. Of course, other allocations can be considered, as well as other rules for obtaining unique hedge ratios, but our goal was only to assess if there were any strategies using lower order models that could lead to significant improvements.

For the two-factor model, we use two hedging instruments corresponding to dates 1 and 4. When using three hedging instruments, we choose the third bond to be the middle maturity between 1 and 4, and with four hedging instruments, we again use bonds 1,2,3, and 4. With three hedging instruments, we assume that the numbers in the short bond and the middle bond are equal; for the four hedging instruments, unique hedges are constructed by assuming the number of bonds in the first two maturities to be equal, as well as the number of bonds in the last two maturities to be equal.

The one-factor model with one hedging instrument, corresponding to the swap maturity (labeled bond 4), is used as the benchmark. The number of times the benchmark model produces smaller absolute residuals than each of the challenging models is recorded, across all weeks for each contract. Table 12 summarizes the results aggregated over all contracts.

Table 12 Here

For the one-factor model, none of the potential enhancements produce better results. Similarly, no improvements are obtained for the two-factor model. Other bond allocations were tested, but we could not identify systematic ways of improving the performance of lower order models using multiple hedging instruments.[17]

Since we conclude that the four-factor PCA model is the best model from the hedging perspective, it is worthwhile to look at the composition of the hedge positions over time. Since

[17] For example, in the one-factor model, it is true that choosing a different maturity bond as the single bond to hedge with can make a difference. For example, using the expiry date bond as the only hedging instrument gave very poor results relative to using the swap maturity dated bond. Since the choice of hedging instruments should not really matter, if the model is correct, this only provides more evidence that the one-factor model is a poor model for hedging.

the hedge ratios depend on the term structure and since absolute volatilities of forward rates depend on the level of forward rates, one cannot expect the hedges to be identical over time. Moreover, it could be argued that since the hedge is constructed based on solving a system of four equations in four unknowns, the hedge portfolio, being unique, might be unstable over time periods and subject to problems from measurement errors. However, the hedge compositions are fairly stable over time. To illustrate this, Figure 6 presents the time series of all four hedge ratios for a typical swaption contract, namely a contract with expiry date in 3 years and a swap maturity of 3 years. The underlying hedging instruments are the bonds with maturities of 3, 4, 5 and 6 years. Similar trends were observed for all other contracts - the hedge compositions using the four-factor model were remarkably stable. Since the hedge ratios are all based off cross sectional estimations, the stability of the time series results suggest that our models are indeed capturing a somewhat stable underlying volatility structure.

Figure 6 Here

9 Conclusion

This article carefully examines the role of volatility structures and factors in the pricing of swaptions. Among the PCA models, we show that increasing the number of factors up to four improves pricing performance. Similar results are obtained for caps. Since away-from-the-money caps provide information about skew effects, we use that data to establish that level dependence is important and that proportional models are better than absolute or square root models. Our results reconfirm the results of LSS and suggest that the Gaussian models of DKM could be improved upon by incorporating level dependence in the volatility structures.

One limitation of the PCA models is that the number of free parameters implied out from derivatives data equals the number of stochastic drivers. Therefore, it is unclear whether the four-factor PCA model outperforms lower order models due to the number of factors or due to the number of free parameters. If it is due to more factors, it supports the importance of forward rate correlations for modeling swaption products. However, if it is due to the fact that higher order models have more degrees of freedom which permit better fits to swaptions, then the need to model correlations well is questionable. By analyzing lower order models with the same degrees of freedom, we address this issue, and conclude that, from a pricing perspective, one-factor models are sufficient.

Regarding hedging effectiveness, however, the one-factor models are significantly outperformed by multifactor models. We show that the higher order models perform consistently better for all contracts, regardless of their expiration date and underlying swap maturity. In the analysis, we show 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. This is due to the fact that some models have systematic biases in their hedges, since the model upon which the hedge is based is flawed. Using *root mean squared error* as the criterion, which accounts for variance and bias, leads to more precise conclusions. In the hedging tests, the performance of multifactor models is not only superior for hedges maintained over one week, but also for hedges maintained unchanged for as long as four weeks. In addition, the four-factor model produces remarkably stable hedges.

While at a superficial level our results echo the final conclusions reached by LSS, our conclusions are arrived at for different reasons. In particular, for pricing purposes they find it necessary to use four-factor models. In contrast, we find a one-factor model that is satisfactory for pricing, and only recommend multifactor models for hedging. Our results are consistent with Andersen and Andreasen (2001), who identify useful one-factor models for pricing Bermudan swaptions. With regard to both pricing and hedging, our results also align with those obtained by Gupta and Subrahmanyam (2001), who show that in the cap market, one-factor models are good enough for pricing, but two-factor models are required for hedging. Our hedging results for swaptions also show that using multiple instruments within a lower order model does not improve hedging performance. These results differ from DKM, possibly due to their incorrect use of variance reduction as a criterion of hedging effectiveness. Bias effects are important when a misspecified model is used.

It would be interesting to examine how well the models we examine perform on swaptions with strikes away-from-the-money. Unfortunately, we do not have prices of such contracts. For at-the-money swaptions, the effects of level dependence in the forward rate volatility structure are not apparent. Indeed, absolute and square root models work as well as proportional models. For the Gaussian models, extremely accurate analytical approximations are available for swaption prices. However, based on our results for caps, it is likely that away-from-the-money swaptions might be more sensitive to the level dependence parameter, γ, hence models with $\gamma = 1$ are preferable.

While our study indicates that there is a need for multi factor models for hedging exotics like Bermudan swaptions, much work remains here. For example, Andersen and Andreasen (2001) explore the common market practice of pricing swaption based products by continuously refitting a model with time varying parameters to swaption prices. The common market practice for hedging, as far as we can assess, is to perturb selected forward rates, one at a time, and to reprice under each shocked curve. The resulting price changes of claims can then be hedged using Eurodollar futures of the appropriate maturities. It remains for future research to assess if the hedges set up based on this popular market approach are competitive with hedges set up using our higher order multi-factor models.

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. A deeper understanding of these structures is crucial for interest rate risk management and value at risk.

References

Amin, K., and A. Morton, 1994, "Implied volatility functions in arbitrage-free term structure models," *Journal of Financial Economics*, 35, 141-180.

Andersen, L., 2000, "A Simple Approach to the Pricing of Bermudan Swaptions in the Multifactor Libor Market Model," *Journal of Computational Finance*, 3, 5-32.

Andersen, L. and J. Andreasen, 2000, "Volatility Skews and Extensions to the LIBOR Market Model," *Applied Mathematical Finance*, 7, 1-32.

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

Andersen, L. and M. Broadie, 2001, "A Primal-Dual Simulation Algorithm for Pricing Multi-Dimensional American Options," working paper, Graduate School of Business, Columbia University.

Bhar, R., and C. Chiarella (1995),"Transformation of Heath-Jarrow-Morton Models to Markovian Systems," *The European Journal of Finance*, 3 , 1-26.

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

Black, F., and P. Karasinski, 1991, "Bond and Option Pricing when Short Rates are Lognormal," *Financial Analysts Journal*, 47, 52-59.

Boyle, P., M. Broadie, and P. Glasserman, 1997, "Monte Carlo Methods for Security pricing," Journal of Economic Dynamics and Control, 21, 1267-1322.

Broadie, M. and P. Glasserman, 1997, "A Stochastic Mesh Method for Pricing High Dimensional American Options," working paper, Columbia University.

Bühler, W., M. Unrig, U. Walter, and T. Weber, 1999, "An Empirical Comparison of Forward and Spot-rate Models for Valuing Interest-rate Options", *Journal of Finance*, 54, 269-305.

Carr, P. and G. Yang, 1999, "Simulating Bermudan Interest Rate Derivatives,", working paper, Bank of America Securities.

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

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.

Gupta, A., and M. Subrahmanyam, 2001, "An Examination of the Static and Dynamic Performance of Interest Rate Option Pricing Models in the Dollar Cap-Floor Markets", Working Paper, New York University.

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

Hull, J. and A. White, 1993, "one-factor Interest Rate Models and the Valuation of Interest Rate Derivative securities", *Journal of Financial and Quantitative Analysis*, 28, 235-254.

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.

Inui, K. and M. Kijima, 1998, "A Markovian Framework in Multi-Factor Heath-Jarrow-Morton Models," *Journal of Financial and Quantitative Analysis,* 33, 423-440.

Jäckel, P. 2001,"Greeks With Monte Carlo," *Wilmott* July, 55-63.

Jagannathan, R., A. Kaplin, and S. Sun, 2001, "An Evaluation of Multi-Factor CIR Models Using LIBOR, Swap Rates, and Cap and Swaption Prices", working paper, Northwestern University.

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

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

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

Longstaff, F., and E. Schwartz, 2001, "Valuing American Options by Simulation: A Simple Least-Squares Approach", *The Review of Financial Studies*, 14, 113-147

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.

Moraleda J. and T. Vorst, 1997, "Pricing American Interest Rate Claims with Humped Volatility Models," *Journal of Banking and Finance*, 21, 1131-1137.

Pedersen, M., 1999, "Bermudan Swaptions in the LIBOR Market Model," working paper, Simcorp.

Peterson, S., R. Stapleton, and M. Subrahmanyam, 2001, "The Valuation of Caps, Floors and Swaptions in a Multi-Factor Spot-Rate Model," working paper, Stern School of Business, New York University.

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

York University.

Ritchken, P. and I. Chuang, 1999, "Interest Rate Option Pricing with Volatility Humps", *Review of Derivatives Research* 3, 237-262.

Ritchken, P. and L. Sankarasubramanian, 1995a, "Volatility Structures of Forward Rates and the Dynamics of the Term Structure", *Mathematical Finance*, 5, 55-72.

Ritchken, P and L. Sankarasubramanian, 1995b, "The Importance of Forward Rate Volatility Structures in Pricing Interest Rate Sensitive Claims", *Journal of Derivatives*, 25-41.

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.

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, 5, 177-188.

Table 1

Absolute Average Pricing Errors - PCA Models

This table presents the average absolute percentage errors, one week out-of-sample, for swaptions and caps, for the PCA models tested in this paper. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. Hence each error reported for swaptions is an average across 30 contracts over 70 dates (hence an average of 2100 individual errors). The cap data corresponds to biweekly data from March 1 – December 31, 1998, consisting of 22 data sets. So each error reported for caps is an average across 16 contracts over 22 dates (hence an average of 352 individual errors). The standard error of the mean is reported in parenthesis. The options are priced using Monte Carlo simulation with 4000 paths for the evolution of the term structure. In generating the paths, the same seeds for the random number generator were used to ensure consistency across the models.

	Swaptions			Caps		
	γ=0	γ=0.5	γ=1	γ=0	γ=0.5	γ=1
One-factor	4.22	4.22	4.52	18.9	16.4	15.8
	(0.08)	(0.08)	(0.09)	(0.79)	(0.75)	(0.76)
Two-factor	3.81	3.57	3.64	18.1	16.0	14.9
	(0.07)	(0.07)	(0.08)	(0.75)	(0.72)	(0.80)
Three-factor	3.39	3.15	3.18	18.1	16.1	14.6
	(0.07)	(0.07)	(0.08)	(0.75)	(0.75)	(0.79)
Four-factor	2.99	3.03	3.01	18.1	16.5	14.7
	(0.06)	(0.07)	(0.07)	(0.75)	(0.78)	(0.79)

Table 2

Level Dependence Comparisons for Caps using PCA Models

This table compares the one week out-of-sample predictions of each PCA model for each of the 16 contracts over all 22 dates, for different γ values. For example, for the one-factor model, based on all 352 pairwise comparisons of residuals, the model with γ=0.5 beat the model with γ=0 66% of the time. A starred cell indicates that the proportion is significantly different from 50% at the 5% level of significance.

No. of Factors	γ=0.5 vs γ=0	γ=1 vs γ=0	γ=1 vs γ=0.5
1	0.66*	0.59*	0.53*
2	0.72*	0.69*	0.73*
3	0.78*	0.76*	0.70*
4	0.78*	0.79*	0.75*

Table 3

Comparison of Out-of-Sample Performance of PCA Models for Caps

This table compares the one week out-of-sample predictions for the PCA models for each of the 16 contracts over all 22 dates. For example, the first row of the table shows the proportion of the times that the two-factor model produces a residual smaller than the one-factor model. All four PCA models are compared pairwise. A starred cell indicates that the proportion is statistically significantly different from 50% at the 1% level of significance. The data corresponds to biweekly data from March 1 – December 31, 1998.

Models	$\gamma=0$	$\gamma=0.5$	$\gamma=1$
2 vs 1	0.56*	0.57*	0.57*
3 vs 1	0.57*	0.56*	0.62*
4 vs 1	0.54*	0.57*	0.59*
3 vs 2	0.63*	0.53*	0.58*
4 vs 2	0.55*	0.50	0.54*
4 vs 3	0.51	0.46	0.51

Table 4

Comparison of Out-of-Sample Performance of PCA Models for Swaptions

This table compares the one, two, and four week out-of-sample predictions of each PCA model (for gamma = 1), for each of the 30 contracts over all 70 dates. The results are aggregated over all expirations and underlying maturities for the swaptions. Results for in-sample tests are also reported for comparison. Each entry in the table is based on the number of times a particular model, represented by the first number in the row, produces a residual closer to zero than the model it is competing with. Each entry is based on 2100 residuals. All four PCA models are compared pairwise. A starred cell indicates that the proportion is statistically significantly different from 50% at the 1% level of significance. The data corresponds to biweekly data from March 1, 1998 – October 31, 2000.

	Weeks Out-Of-Sample			
Models	0	1	2	4
2 vs 1	0.66*	0.66*	0.64*	0.63*
3 vs 1	0.75*	0.72*	0.69*	0.67*
4 vs 1	0.71*	0.69*	0.66*	0.65*
3 vs 2	0.68*	0.66*	0.65*	0.62*
4 vs 2	0.63*	0.63*	0.59*	0.58*
4 vs 3	0.53*	0.52*	0.53*	0.53*

Table 5

Average Absolute Pricing Errors for Swaptions - Parametric and Four-Factor PCA Models

This table presents the average absolute percentage errors, one, two, three and four weeks out-of-sample, for swaptions across expirations and maturities (of underlying swaps), for the one-factor and two-factor parametric models and the four-factor PCA model. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. Hence each error reported in this table is an average across 30 contracts over 70 dates (hence an average of 2100 individual errors). The standard error of the mean is reported in parenthesis. The swaptions are priced using Monte Carlo simulation with 4000 paths for the evolution of the term structure. In generating the paths, the same seeds for the random number generator were used to ensure consistency across the models.

	Out-of-Sample Period					
	1 week			2 weeks	3 weeks	4 weeks
	γ=0	γ=0.5	γ=1	γ=1	γ=1	γ=1
One-factor Parametric	2.43	2.57	2.96	3.23	3.48	3.70
	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)
Two-factor Parametric	2.57	2.36	2.48	2.74	3.03	3.16
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
Four-factor PCA	2.99	3.03	3.01	3.34	3.56	3.80
	(0.06)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)

Table 6a

Absolute Average Pricing Errors for Caps - Parametric and Four-Factor PCA Models

This table presents the average absolute percentage errors, one week out-of-sample, for caps across strikes and maturities, for the one- and two-factor parametric models and the four-factor PCA model. The cap data corresponds to biweekly data from March 1 – December 31, 1998, consisting of 22 data sets. So each error reported in this table is an average across 16 contracts over 22 dates (hence an average of 352 individual errors). The standard error of the mean is reported in parenthesis. The caps are priced using Monte Carlo simulation with 4000 paths for the evolution of the term structure. In generating the paths, the same seeds for the random number generator were used to ensure consistency across the models.

	γ=0	γ=0.5	γ=1
One-factor Parametric	17.9	15.0	14.4
	(0.73)	(0.70)	(0.76)
Two-factor Parametric	18.4	15.6	15.1
	(0.75)	(0.68)	(0.72)
Four-factor PCA	18.1	16.5	14.7
	(0.75)	(0.78)	(0.79)

Table 6b

Level Dependence Comparisons for Caps - Parametric and Four-Factor PCA Models

This table presents the results of the proportion tests for the one-and two-factor parametric models and the four-factor PCA model, for different values of γ. For example, the first column of this table shows the proportion of times a model with γ=0.5 outperforms a model with γ=0. The total number of contracts used for each proportion was 352 (22 x 16). The second and third columns show the proportion of times a γ=1 model beats a γ=0.5 and a γ=0 model respectively. The starred cells indicate the cases where the null hypothesis that the proportion of wins is 50% is rejected at the 5% level of significance.

	γ=0.5 vs γ=0	γ=1 vs γ=0.5	γ=1 vs γ=0
One-factor Parametric	0.61*	0.51	0.67*
Two-factor Parametric	0.70*	0.59*	0.63*
Four-factor PCA	0.72*	0.65*	0.71*

Table 7

Proportion of One Week Out-of-Sample Residuals for Swaptions Within Bounds

This table presents the proportion of swaption contracts, one week out-of-sample, that are within 1, 0.5, and 0.25 Black vols, for the one-factor and two-factor parametric and four-factor PCA models, with γ=1. The contracts are aggregated over their maturities. Since there are 70 dates and 5 maturities each, the proportions are each based on 350 residuals.

Expiration	Within 1 Black Vol.			Within 0.5 Black Vol.			Within 0.25 Black Vol.		
	1-factor	2-factor	4 PCA	1-factor	2-factor	4 PCA	1-factor	2-factor	4 PCA
0.5	0.78	0.78	0.71	0.58	0.53	0.39	0.33	0.24	0.19
1	0.81	0.80	0.84	0.54	0.55	0.59	0.34	0.36	0.32
2	0.83	0.84	0.90	0.49	0.57	0.62	0.27	0.38	0.34
3	0.89	0.92	0.87	0.64	0.61	0.59	0.36	0.31	0.36
4	0.83	0.94	0.91	0.56	0.73	0.63	0.32	0.45	0.38
5	0.84	0.90	0.86	0.57	0.63	0.56	0.31	0.36	0.29
Total	0.83	0.86	0.85	0.56	0.60	0.56	0.32	0.35	0.31

Table 8

Comparison of Parametric and Four-Factor PCA Models for Swaptions

This table compares the one week out-of-sample predictions for the one-factor and two-factor parametric models and the four-factor PCA model, with γ=1, for each of the 30 contracts over all 70 dates. For example, the first entry in the top table shows that in 39% of the 70 cases, the two-factor model outperforms the one-factor model for one week ahead pricing of the 0.5x1 swaptions. The starred cells indicate that the proportion is significantly different from 50%. All tests are conducted at the 1% level of significance.

Expiration	Swap Maturity	2P vs 1P	4PCA vs 2P	4PCA vs 1P
0.5	1	.39	.22*	.28*
0.5	2	.53	.37	.50
0.5	3	.32*	.25*	.24*
0.5	4	.31*	.49	.28*
0.5	5	.57	.34*	.28*
1	1	.53	.70†	.56
1	2	.44	.59	.37
1	3	.46	.60	.50
1	4	.60	.50	.51
1	5	.61	.46	.57
2	1	.31	.67†	.46
2	2	.49	.59	.57
2	3	.50	.61	.63
2	4	.78†	.44	.72†
2	5	.82†	.31*	.84†
3	1	.74†	.22*	.36
3	2	.41	.53	.37
3	3	.31*	.70†	.44
3	4	.36	.65†	.57
3	5	.67†	.57	.70
4	1	.88†	.31*	.90†
4	2	.71†	.51	.75†
4	3	.65†	.53	.71†
4	4	.63	.37	.40
4	5	.53	.27*	.37
5	1	.81†	.53	.61
5	2	.64	.27*	.36
5	3	.44	.40	.37
5	4	.50	.31*	.46
5	5	.64	.46	.61
	Average	.55†	.46*	.51

Table 9

Standard Deviation Ratios for Hedging Swaptions

This table presents the ratio of standard deviations of the hedged and unhedged portfolios for the one-factor and two-factor parametric and the four-factor PCA models, for swaptions, one week out-of-sample. 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. Therefore, the standard deviations are computed using the values of the hedged and unhedged swaption portfolios over 70 weeks. The averages of the standard deviation ratios across all contracts are also reported, one to four weeks out-of-sample, for an aggregate analysis.

Expiration	Swap Maturity	One-Factor	Two-factor	Four-factor
0.5	2	0.23	0.26	0.23
0.5	3	0.20	0.40	0.21
0.5	4	0.19	0.44	0.20
0.5	5	0.19	0.46	0.20
1	2	0.25	0.26	0.24
1	3	0.23	0.26	0.23
1	4	0.23	0.35	0.23
1	5	0.22	0.33	0.21
2	2	0.33	0.26	0.33
2	3	0.33	0.26	0.42
2	4	0.31	0.28	0.25
2	5	0.30	0.30	0.25
3	2	0.41	0.30	0.43
3	3	0.39	0.49	0.29
3	4	0.37	0.32	0.29
3	5	0.37	0.31	0.31
4	2	0.47	0.38	0.33
4	3	0.41	0.34	0.33
4	4	0.39	0.30	0.30
4	5	0.38	0.29	0.29
5	2	0.49	0.38	0.42
5	3	0.47	0.34	0.35
5	4	0.46	0.34	0.33
5	5	0.44	0.33	0.36
	Average (1 week out)	0.34	0.33	0.29
	Average (2 weeks out)	0.35	0.35	0.30
	Average (3 weeks out)	0.36	0.39	0.33
	Average (4 weeks out)	0.37	0.40	0.33

Table 10

Absolute Hedging Errors for Swaptions

This table presents the root mean squared errors (in basis points) of the hedged and unhedged portfolios for the one-factor and two-factor parametric and the four-factor PCA models, one week out-of-sample. 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.

Expiration	Swap Maturity	Unhedged Swaption	One-Factor	Two-factor	Four-Factor
0.5	2	12.1	6.2	3.6	3.0
0.5	3	18.1	7.0	5.9	4.2
0.5	4	23.3	7.7	6.3	5.2
0.5	5	28.2	8.2	6.5	5.9
1	2	13.1	5.5	3.6	3.2
1	3	18.9	6.6	5.0	4.6
1	4	24.1	7.7	8.8	5.7
1	5	28.9	8.2	8.7	6.2
2	2	13.1	5.6	3.6	4.3
2	3	18.7	7.4	5.0	8.4
2	4	23.0	8.4	6.6	5.8
2	5	27.5	9.5	8.4	7.0
3	2	12.6	5.8	3.8	5.4
3	3	17.3	7.4	8.5	5.0
3	4	21.9	8.7	7.2	6.3
3	5	26.9	10.6	8.1	8.4
4	2	11.4	5.7	4.3	3.7
4	3	15.8	7.0	5.5	5.3
4	4	20.9	8.5	6.3	6.3
4	5	25.9	10.2	7.7	7.6
5	2	10.8	5.4	4.1	4.5
5	3	16.4	7.8	5.6	5.7
5	4	21.5	9.9	7.2	7.1
5	5	26.3	11.7	8.8	9.5

Table 11

Comparison of Parametric and PCA Models in Hedging Swaptions

This table presents the fraction of times one model outperforms the other model in hedging forecasts, for the one-factor and two-factor parametric and the four-factor PCA models, one week out-of-sample. 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. Therefore, for each contract, the proportions are computed from a comparison of 70 hedging errors.

Expiration	Swap Maturity	1 week out-of-sample			2 weeks out-of-sample			4 weeks out-of-sample		
		2 vs 1	4 vs 1	4 vs 2	2 vs 1	4 vs 1	4 vs 2	2 vs 1	4 vs 1	4 vs 2
0.5	2	0.87	0.87	0.66	0.93	0.99	0.68	0.94	0.99	0.75
0.5	3	0.81	0.86	0.67	0.93	0.90	0.71	0.93	0.96	0.84
0.5	4	0.67	0.87	0.76	0.77	0.88	0.75	0.82	0.94	0.79
0.5	5	0.59	0.81	0.73	0.70	0.84	0.72	0.78	0.91	0.78
1	2	0.81	0.77	0.61	0.91	0.90	0.64	0.97	0.97	0.65
1	3	0.71	0.73	0.54	0.86	0.88	0.58	0.94	0.94	0.65
1	4	0.63	0.70	0.57	0.74	0.86	0.57	0.79	0.93	0.62
1	5	0.56	0.69	0.61	0.62	0.78	0.74	0.71	0.84	0.72
2	2	0.73	0.66	0.50	0.78	0.75	0.57	0.90	0.85	0.53
2	3	0.67	0.53	0.31	0.72	0.65	0.30	0.94	0.74	0.31
2	4	0.59	0.69	0.59	0.65	0.68	0.51	0.82	0.84	0.62
2	5	0.61	0.64	0.63	0.62	0.67	0.57	0.68	0.78	0.53
3	2	0.63	0.61	0.46	0.70	0.67	0.52	0.76	0.68	0.49
3	3	0.66	0.67	0.60	0.64	0.70	0.55	0.74	0.74	0.53
3	4	0.66	0.69	0.57	0.68	0.67	0.57	0.74	0.72	0.56
3	5	0.69	0.66	0.56	0.65	0.65	0.67	0.75	0.74	0.62
4	2	0.64	0.63	0.53	0.71	0.72	0.54	0.79	0.74	0.51
4	3	0.70	0.69	0.60	0.65	0.67	0.61	0.76	0.68	0.51
4	4	0.77	0.73	0.61	0.70	0.65	0.62	0.79	0.75	0.59
4	5	0.74	0.71	0.60	0.70	0.67	0.59	0.72	0.66	0.53
5	2	0.73	0.66	0.50	0.75	0.62	0.45	0.74	0.68	0.49
5	3	0.81	0.76	0.54	0.72	0.74	0.62	0.76	0.71	0.60
5	4	0.80	0.77	0.66	0.72	0.68	0.62	0.71	0.69	0.62
5	5	0.73	0.66	0.47	0.72	0.70	0.58	0.69	0.66	0.50
Average (1 week out)		0.70	0.71	0.58	0.73	0.75	0.59	0.80	0.80	0.60

Table 12

Comparison of Models Using Different Hedging Instruments

This table presents the fraction of times one model specification outperforms the same model using different hedging instruments, one week out-of-sample, for the one-factor and two-factor parametric models. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. Underlying discount bonds are used as hedging instruments, and are labeled as follows. The bond expiring on the swaption expiration date is labeled "1", while the bond expiring on the underlying swap maturity date is labeled "4". Bonds "2" and "3" correspond to bonds with maturities equally spaced between the swaption expiration date (bond 1) and the underlying swap maturity date (bond 4). For the one-factor model, the hedge using bond 4 is used as the benchmark hedge. For the two-factor model, the hedge using bonds 1 and 4 is the benchmark. When using three hedging instruments within the two-factor model, the third bond is chosen to be the middle maturity between bonds 1 and 4. The four hedging instruments within the two-factor model are bonds 1, 2, 3, and 4.

Model Comparison	Fraction of wins
One-Factor Parametric Model	
One vs two instruments (4 vs 4,3)	0.52
One vs three instruments (4 vs 4,3,2)	0.54
One vs four instruments (4 vs 4,3,2,1)	0.55
One instrument (4 vs 3)	0.53
One instrument (4 vs 2)	0.57
One instrument (4 vs 1)	0.66
Two-Factor Parametric Model	
Two vs three instruments	0.50
Two vs four instruments	0.53

Figure 1

Average Percentage Errors for Swaptions

This figure presents plots of the average percentage errors, one week out-of-sample, for swaptions across underlying swap maturities, for each expiration, for the one-factor and two-factor parametric and the four-factor PCA models. The plots are presented for three values of γ; γ=0 (dotted line), γ=0.5 (hashed line), and γ=1 (solid line). The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. The swaptions are priced using Monte Carlo simulation with 4000 paths for the evolution of the term structure. In generating the paths, the same seeds for the random number generator were used to ensure consistency across the models.

One-Factor Model　　　　Two-Factor Model　　　　Four-Factor PCA Model

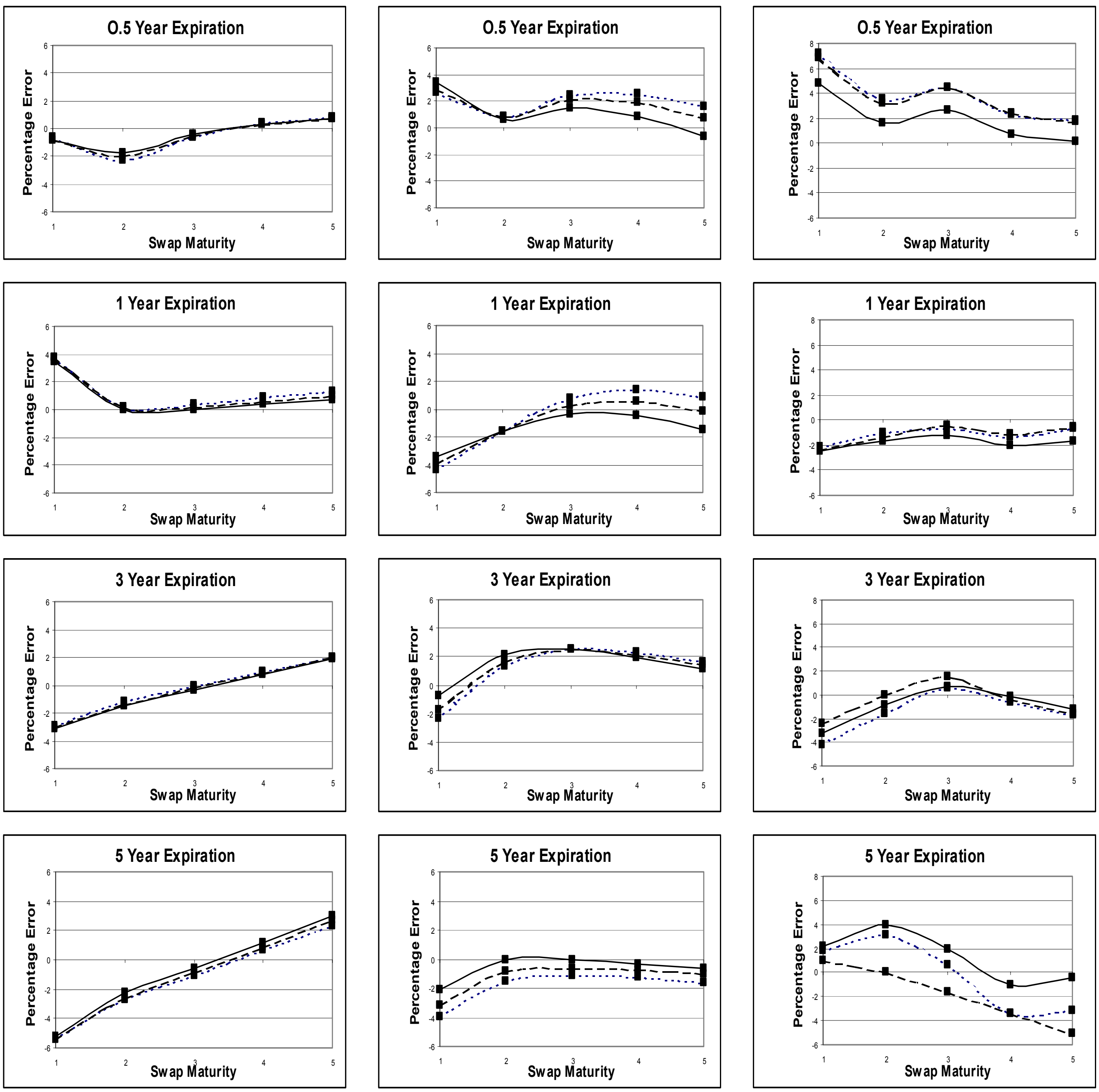

Figure 2

Average Percentage Errors for Caps

This figure presents plots of the average percentage errors, one week out-of-sample, for caps across strikes for each maturity, for the one-factor and two-factor parametric and the four-factor PCA models. The plots are presented for three values of γ; γ=0 (dotted line), γ=0.5 (hashed line), and γ=1 (solid line). The cap data corresponds to biweekly data from March 1 – December 31, 1998, consisting of 22 data sets. The in sample optimizations for the models were conducted over four unknown parameters, with caps being priced using Monte Carlo simulation with 4000 paths for the evolution of the term structure. In generating the paths, the same seeds for the random number generator were used to ensure consistency across the models.

One-Factor Model Two-Factor Model Four-Factor PCA Model

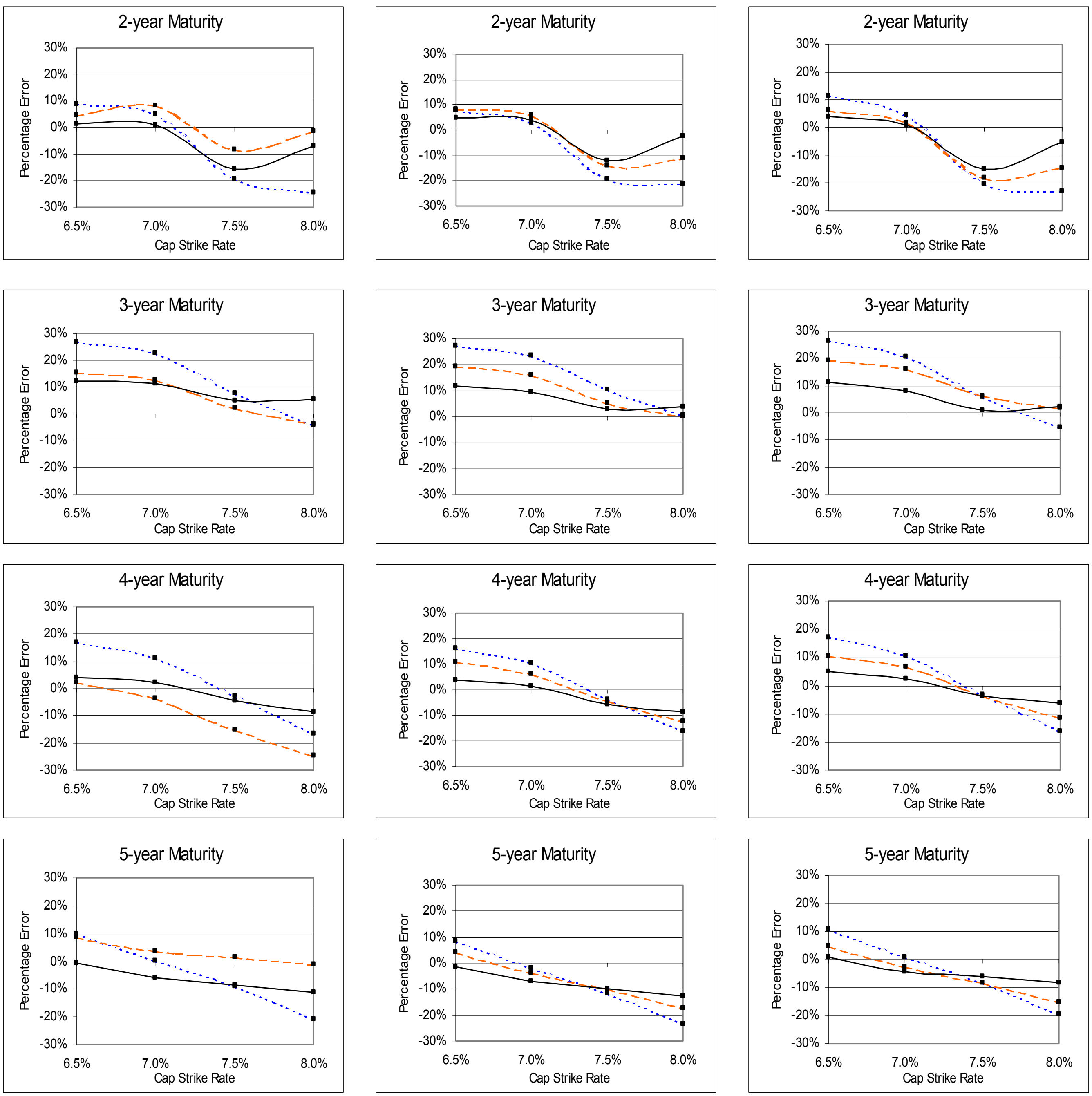

Figure 3

Box and Whiskers Plots of Swaption Pricing Errors

This figure presents the box and whiskers plots of the one week out-of-sample pricing errors for the one-factor and two-factor parametric and the four-factor PCA models with γ=1, for swaptions across expirations and underlying swap maturities. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. The swaptions are priced using Monte Carlo simulation with 4000 paths for the evolution of the term structure. In generating the paths, the same seeds for the random number generator were used to ensure consistency across the models. In each figure, the first box corresponds to the one-factor parametric model, the second box to the two-factor parametric model, and the third box to the four-factor PCA model.

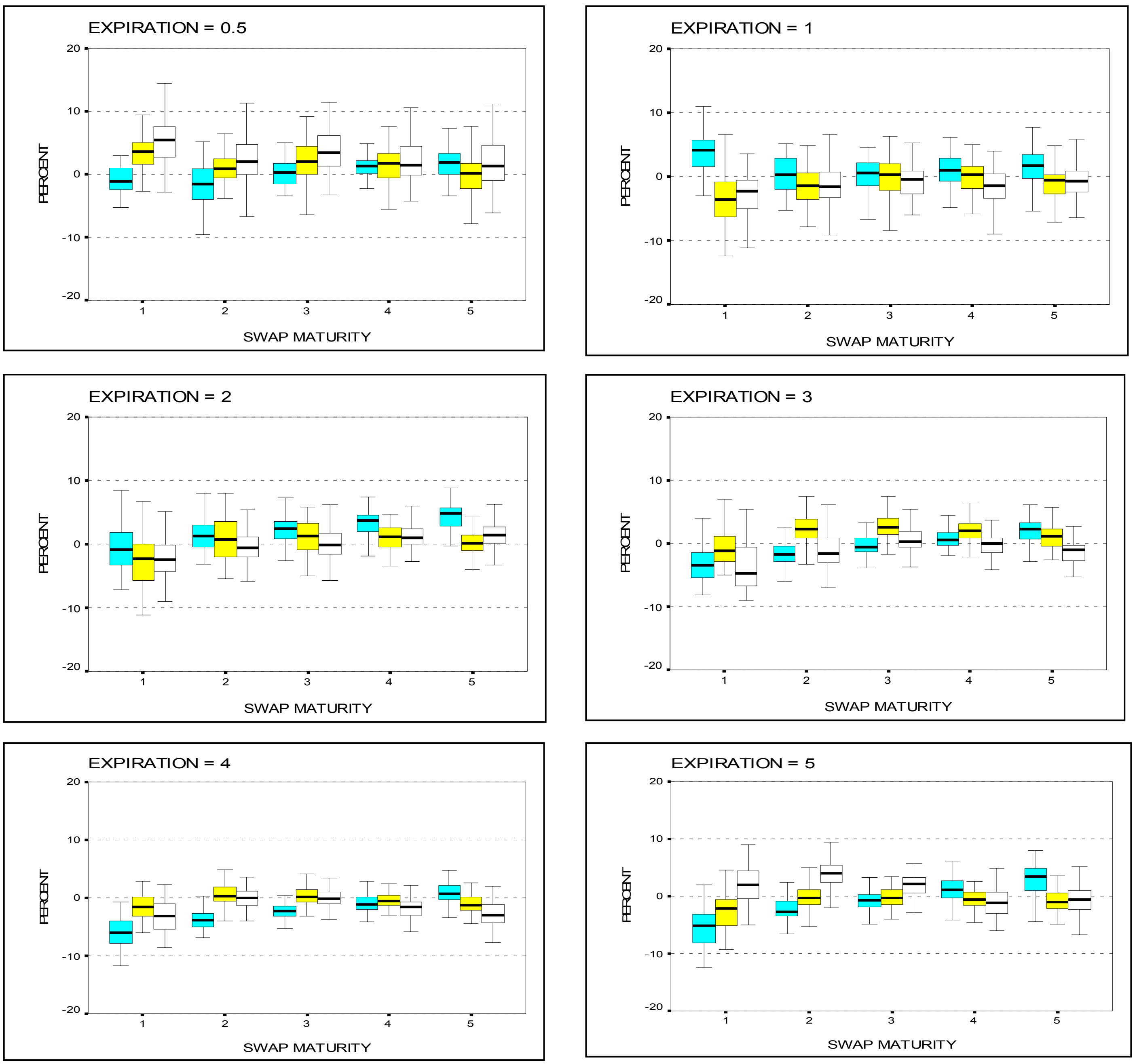

Figure 4

Estimated Volatilities of Forward Rates Implied out from Swaption and Cap Prices

The figures presents the time series of estimated forward rate volatilities for the one-factor and two-factor parametric and the four-factor PCA models. There are 70 volatility curves (22 for caps) on each figure, each curve separated by two weeks. The data for these curves are derived from the 70 optimization problems (22 optimizations for caps).

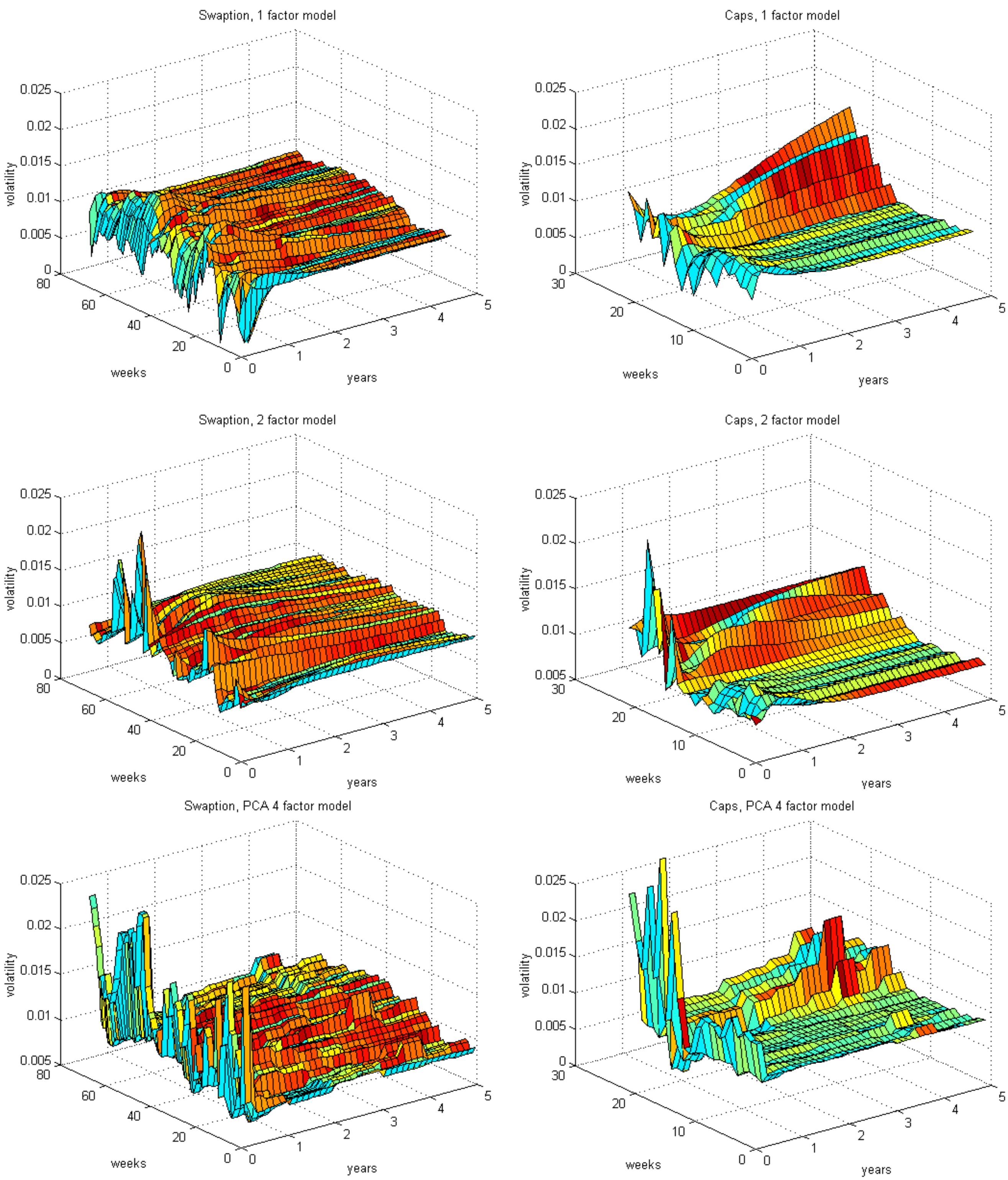

Figure 5

Box and Whiskers Plots of Swaption Hedging Errors

This figure presents the box and whiskers plots for the one week out-of-sample hedging errors for the one-factor and two-factor parametric and the four-factor PCA models (with γ=1), 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 parametric model, the third box to the two-factor parametric model, and the fourth box to the four-factor PCA model.

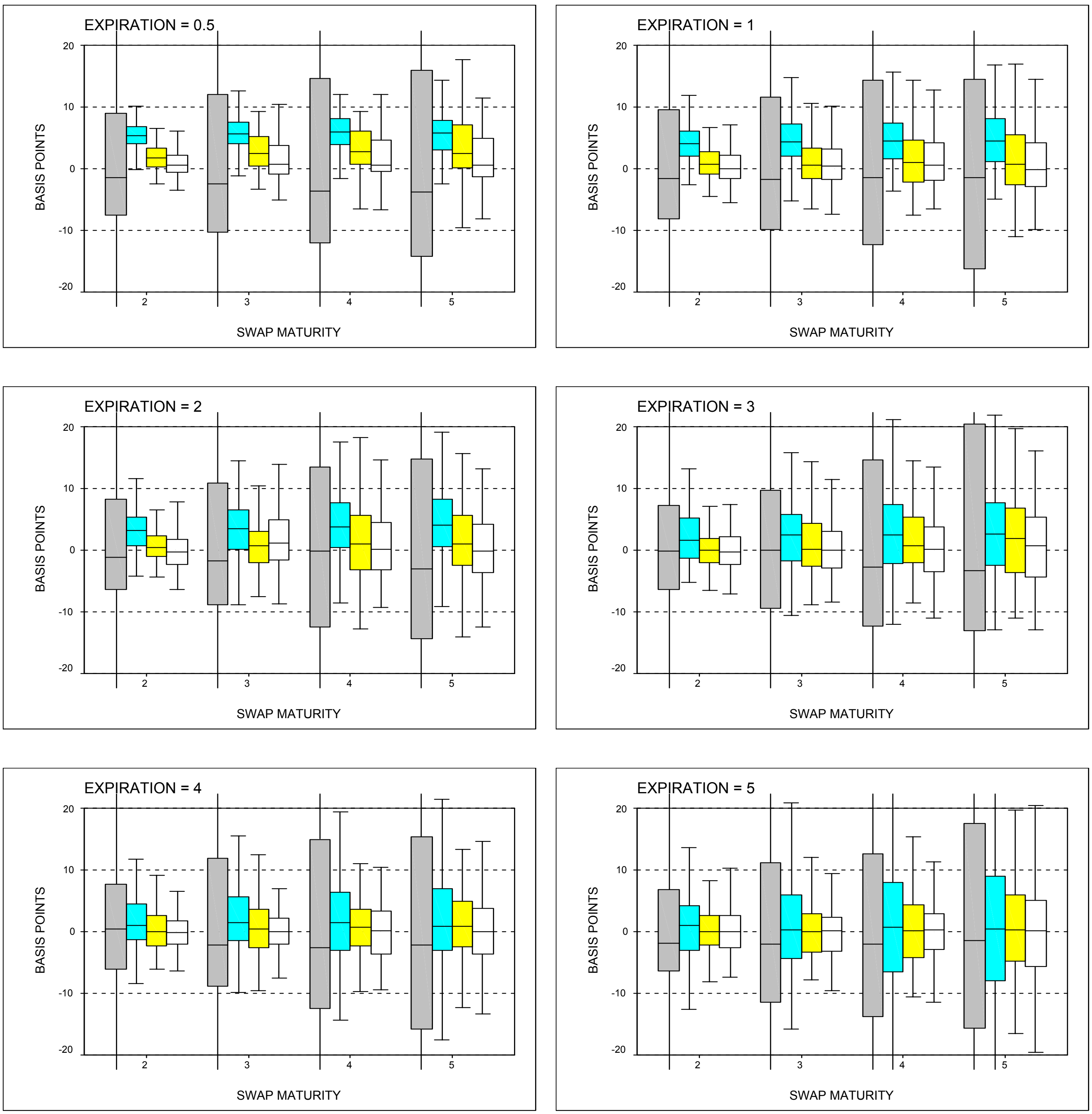

Figure 6

Swaption Hedge Portfolio for Four-factor PCA Model

This figure presents the time series of the positions in each of the four hedging instruments in the hedge portfolio for a sample swaption (3 x 3), using the four-factor PCA model. The swaption data corresponds to biweekly data from March 1, 1998 – October 31, 2000, consisting of 70 data sets. Therefore, the time series of bond positions is for each of the 70 weeks. The bond expiring on the swaption expiration date is labeled "1", while the bond expiring on the underlying swap maturity date is labeled "4". Bonds "2" and "3" correspond to bonds with maturities equally spaced between the swaption expiration date (bond 1) and the underlying swap maturity date (bond 4).