Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Banking & Finance 29 (2005) 701–733
Journal of BANKING & FINANCE
www.elsevier.com/locate/econbase

Pricing and hedging interest rate options: Evidence from cap–floor markets

Anurag Gupta [a,1], Marti G. Subrahmanyam [b,*]

[a] *Department of Banking and Finance, Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7235, USA*
[b] *Department of Finance, Leonard N. Stern School of Business, New York University, 44 West Fourth Street 9-15, New York, NY 10012-1126, USA*

Received 14 February 2003; accepted 10 February 2004
Available online 25 June 2004

Abstract

We examine the pricing and hedging performance of interest rate option pricing models using daily data on US dollar cap and floor prices across both strike rates and maturities. Our results show that fitting the skew of the underlying interest rate probability distribution provides accurate *pricing* results within a one-factor framework. However, for *hedging performance*, introducing a second stochastic factor is more important than fitting the skew of the underlying distribution. This constitutes evidence against claims in the literature that correctly specified and calibrated one-factor models could replace multi-factor models for consistent pricing and hedging of interest rate contingent claims.

JEL classification: G12; G13; G19
Keywords: Interest rate options; Caps/floors; Term structure of interest rates; Model performance; Hedging

* Corresponding author. Tel.: +1-212-998-0348; fax: +1-212-995-4233.
E-mail addresses: anurag.gupta@case.edu (A. Gupta), msubrahm@stern.nyu.edu (M.G. Subrahmanyam).
[1] Tel.: +1-216-368-2938; fax: +1-216-368-6249.

0378-4266/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2004.05.025

1. Introduction

Interest rate option markets are amongst the largest and most liquid option markets in the world today, with daily trading volumes of trillions of US dollars, especially for caps/floors and swaptions. [2] These options are widely used both for hedging as well as speculation against changes in interest rates. Theoretical work in the area of interest rate derivatives has produced a variety of models and techniques to value these options, some of which are widely used by practitioners. [3] The development of many of these models was mainly motivated by their analytical tractability. Therefore, while these models have provided important theoretical insights, their empirical validity and performance remain to be tested. Empirical research in this area has lagged behind theoretical advances partly due to the difficulty in obtaining data, as most of these interest rate options are traded in over-the-counter markets, where data are often not recorded in a systematic fashion. This gap is being slowly filled by recent research in this area.

This paper provides empirical evidence on the validity of alternative interest rate models. We examine the pricing and hedging performance of interest rate option pricing models in the US dollar interest rate cap and floor markets. For the first time in this literature, a time series of actual cap and floor prices across strike rates and maturities is used to study the systematic patterns in the pricing and hedging performance of competing models, on a daily basis. The one-factor models analyzed consist of two spot rate specifications (Hull and White, 1990 [HW]; Black and Karasinski, 1991 [BK]), five forward rate specifications (within the general Heath et al., 1990 [HJM] class), and one LIBOR market model (Brace et al., 1997 [BGM]). For two-factor models, two alternative forward rate specifications are implemented within the HJM framework. The analysis in this paper, therefore, sheds light on the empirical validity of a broad range of models for pricing and hedging interest rate caps and floors, especially across different strikes, and suggests directions for future research.

The interest rate derivatives market consists of instruments that are based on different market interest rates. Interest rate swaps and FRAs are priced based on the *level* of different segments of the yield curve; caps and floors are priced based on the level and the *volatility* of the different forward rates (i.e., the diagonal elements of the covariance matrix). Since caps and floors do not price the correlations among forward rates, it appears, at first glance, that one-factor models might be sufficiently

[2] The total notional principal amount of over-the-counter interest rate options such as caps/floors and swaptions outstanding at the end of December 2002 was about \$13.7 trillion, as per the BIS Quarterly Review, Bank for International Settlements, September 2003.

[3] The early models, many of which are still widely used, include those by Black (1976), Vasicek (1977), Cox et al. (1985), Ho and Lee (1986), Heath et al. (1990), Hull and White (1990), Black et al. (1990), and Black and Karasinski (1991). Several variations and extensions of these models have been proposed in the literature in the past decade.

accurate in pricing and hedging them, and the added numerical complexity of multi-factor models (in particular, two-factor models) may not be justified. [4] This is one of the key issues that this paper seeks to investigate.

We evaluate the empirical performance of analytical models along two dimensions – their pricing and hedging accuracy. Pricing performance refers to the ability of a model to price options accurately, conditional on the term structure. Hedging performance refers to the ability of the model to capture the underlying movements in the term structure in the future, after being initially calibrated to fit current market observables. The pricing accuracy of a model is useful in picking out deviations from arbitrage-free pricing. The tests for hedging accuracy examine whether the interest rate dynamics embedded in the model is similar to that driving the actual economic environment that the model is intended to represent.

Our results show that, for plain-vanilla interest rate caps and floors, a one-factor lognormal forward rate model outperforms other competing one-factor models, in terms of out-of-sample pricing accuracy. In addition, the estimated parameters of this model are stable. In particular, the one-factor BGM model outperforms other models in pricing tests where the models are calibrated using option pricing data for the same day on which they are used to estimate prices of other options. We also find that the assumption of lognormally distributed interest rates results in a smaller "skew" in pricing errors across strike rates, as compared to other distributions assumed in alternative interest rate models. Two-factor models improve pricing accuracy only marginally. Thus, for accurate pricing of caps and floors, especially away-from-the-money, it is more important for the term structure model to fit the skew in the underlying interest rate distribution, than to have a second stochastic factor driving the term structure. However, the hedging performance improves significantly with the introduction of a second stochastic factor in term structure models, while fitting of the skew in the distribution improves hedging performance only marginally. This occurs because two-factor models allow a better representation of the dynamic evolution of the yield curve, which is more important for hedging performance, as compared to pricing accuracy. Thus, even for simple interest rate options such as caps and floors, there is a significant advantage to using two-factor models, over and above fitting the skew in the underlying (risk-neutral) interest rate distribution, for consistent pricing and hedging within a book. This refutes claims in the literature that correctly specified and calibrated one-factor models could eliminate the need to have multi-factor models for pricing and hedging interest rate derivatives. [5] We also find that simple two-factor models of the term structure are able to hedge caps and floors across strikes quite well as far as 1 month out-of-sample,

[4] One-factor term structure models imply perfectly correlated spot/forward rates, while two-factor (and multi-factor) models allow for imperfect correlation between spot/forward rates of different maturities.

[5] For instance, Hull and White (1995) state that "the most significant difference between models is a strike price bias ... the number of factors in a term structure model does not seem to be important except when pricing spread options ... one-factor Markov models when used properly do a good job of pricing and hedging interest rate sensitive securities".

indicating that there may not be a strong need to incorporate stochastic volatility into the model explicitly, if the objective is to price and hedge caps and floors.

We examine two alternative calibrations of the spot rate models. In the first implementation, the volatility and mean-reversion parameters are held constant. As a result, while the models are calibrated to fit the current term structure exactly, the model prices match the current cap/floor prices only with an error, albeit by minimizing its impact. In the alternative implementation, an additional element of flexibility is introduced by making the parameters time-varying. This enables us to fit both the current term structure and the cap/floor prices exactly, although this renders the parameter estimates unstable.

The paper is organized as follows. Section 2 presents an overview of the different term structure models used for pricing and hedging interest rate contracts, and the empirical studies in this area so far. In Section 3, details of estimation and implementation of these term structure models are discussed, along with the experimental design and the different methodologies used in evaluating the alternative models. Section 4 describes the data used in this study, along with the method used for constructing the yield curve. The results of the study are reported and discussed in Section 5. Section 6 concludes.

2. Literature review

2.1. Term structure models

There are numerous models for valuing interest rate derivatives, which, broadly speaking, can be divided into two categories: spot rate models and forward rate models. In the case of spot rate models, the entire term structure is inferred from the evolution of the spot short-term interest rate (and, in case of two-factor models, by another factor such as the long-term interest rate, the spread, the volatility factor, or the futures premium). [6] A generalized one-factor spot rate specification, that explicitly includes mean reversion, has the form

$$df(r) = [\theta(t) - af(r)]\,dt + \sigma\,dz, \tag{1}$$

where

$f(r)$ = some function f of the short rate r,
$\theta(t)$ = a function of time chosen so that the model provides an exact fit to the initial term structure, usually interpreted as a time-varying mean,
a = mean-reversion parameter,
σ = volatility parameter.

[6] This includes the traditional models by Vasicek (1977), Brennan and Schwartz (1979), Cox et al. (1985), Longstaff and Schwartz (1992), Stapleton and Subrahmanyam (1999), Peterson et al. (2003) and others.

Two special cases of the above model are in widespread use. When $f(r) = r$, the resultant model is the HW model (also referred to as the extended-Vasicek model):

$$dr = [\theta(t) - ar]\,dt + \sigma\,dz. \tag{2}$$

$f(r) = \ln r$ leads to the BK model:

$$d\ln r = [\theta(t) - a\ln r]\,dt + \sigma\,dz. \tag{3}$$

The probability distribution of the short rate is Gaussian in the HW model and lognormal in the BK model. These models can be modified to match the term structure *exactly* (i.e., taking the current term structure as an input rather than as an output), in an arbitrage-free framework by making one or more of the parameters time-varying, so that, at least, there is no mispricing in the underlying bonds. [7] However, this may result in unstable parameter estimates and implausible future evolutions of the term structure. [8] This would be reflected in poor out-of-sample performance of these models. Hence, there is a trade-off between a perfect fit of the current term structure and the stationarity of the model parameters. [9]

In forward rate models (starting with Ho and Lee (1986) and Heath et al. (1990)), the instantaneous forward rate curve is modeled with a fixed number of unspecified factors that drive the dynamics of these forward rates. The form of the forward rate changes can be specified in a fairly general manner. In fact, some of the processes specified in the literature for the evolution of the spot interest rate can be treated as special cases of HJM models by appropriately specifying the volatility function of the forward interest rates. [10]

Let $f(t,T)$ be the forward interest rate at date t for instantaneous riskless borrowing or lending at date T. In the HJM approach, forward interest rates of every maturity T evolve simultaneously according to the stochastic differential equation

$$df(t,T) = \mu(t,T,.)\,dt + \sum_{i=1}^{n} \sigma_i(t,T,f(t,T))\,dW_i(t), \tag{4}$$

where $W_i(t)$ are n independent one-dimensional Brownian motions and $\mu(t,T,.)$ and $\sigma_i(t,T,f(t,T))$ are the drift and volatility coefficients for the forward interest rate of maturity T. The volatility coefficient represents the instantaneous standard deviation (at date t) of the forward interest rate of maturity T, and can be chosen arbitrarily. For each choice of volatility functions $\sigma_i(t,T,f(t,T))$, the drift of the forward rates under the risk-neutral measure is uniquely determined by the no-arbitrage condition.

[7] This is implemented in the models by Hull and White (1990), Black et al. (1990), Black and Karasinski (1991), Peterson et al. (2003) and others.

[8] This non-stationarity would be more problematic for derivative instruments whose prices depend on future volatility term structures (like American/Bermudan options, spread options, captions, etc.). For standard caps and floors, as in this paper, this is likely to be less important.

[9] See Hull and White (1996) for a discussion on this issue.

[10] For example, an exponential volatility function gives rise to the Ornstein–Uhlenbeck process as in Vasicek (1977). A constant volatility results in the continuous-time version of the Ho and Lee model. In these two cases, closed form solutions are available for discount bonds and option prices.

The choice of the volatility function $\sigma_i(t, T, f(t, T))$ determines the interest rate process that describes the stochastic evolution of the entire term structure. If the volatility function is stochastic, it *may* make the interest rate process non-Markovian, in which case no closed-form solutions are possible for discount bonds or options. [11] Hence, it is preferable to restrict the nature of the volatility functions in order to obtain manageable solutions.

The volatility functions analyzed in this paper, $\sigma_i(t, T, f(t, T))$, are time invariant functions. In these functions, the volatility depends on t and T only though $T - t$. Therefore, given a term structure at time t, the form of its subsequent evolution through time depends only on the term structure, not on the specific calendar date t. Even with this restriction, a rich class of volatility structures can be analyzed. We focus on the following volatility functions, and models that they imply:

One-factor models: [12]

1.	Absolute:	$\sigma(\cdot) = \sigma_0$
2.	Linear absolute:	$\sigma(\cdot) = [\sigma_0 + \sigma_1(T - t)]$
3.	Square root:	$\sigma(\cdot) = \sigma_0 f(t, T)^{1/2}$
4.	Proportional:	$\sigma(\cdot) = \sigma_0 f(t, T)$
5.	Linear proportional:	$\sigma(\cdot) = [\sigma_0 + \sigma_1(T - t)]f(t, T)$

Two-factor models: [13]

1.	Absolute exponential:	$\sigma_1(\cdot) = \sigma_1 \exp[-\kappa_1(T - t)]$
		$\sigma_2(\cdot) = \sigma_2 \exp[-\kappa_2(T - t)]$
2.	Proportional exponential:	$\sigma_1(\cdot) = \sigma_1 \exp[-\kappa_1(T - t)]f(t, T)$
		$\sigma_2(\cdot) = \sigma_2 \exp[-\kappa_2(T - t)]f(t, T)$

where the σ represents the volatility, and κ, the mean-reversion coefficient.

In recent years, the so-called "market models" have become very popular amongst practitioners. These models recover market-pricing formulae by directly modeling market quoted rates. This approach overcomes one of the drawbacks of the traditional HJM models: that they involve instantaneous forward rates that are not directly observable (and are hence difficult to calibrate). A model that is popular among practitioners is the one proposed by Brace et al. (1997) [BGM]. [14] They derive the processes followed by market quoted rates within the HJM framework,

[11] Ritchken and Sankarasubramanian (1995) identify restrictions on volatility structures that are necessary and sufficient to make the process Markovian with respect to two state variables.

[12] The absolute volatility specification leads to the continuous-time version of the Ho–Lee model, with Gaussian interest rates. The HJM framework requires that the volatility functions be bounded. Hence the proportional volatility function is capped at a sufficiently high level of f^*, such that there is no effect on prices.

[13] The use of exponential volatility functions with different decay parameters makes the two-factor models identifiable, since the two-factor model cannot be reduced to a single-factor equivalent. We thank an anonymous referee for pointing this out.

[14] A similar model has also been proposed by Miltersen et al. (1997).

and deduce the restrictions necessary to ensure that the distribution of market quoted rates of a *given* tenor under the risk-neutral forward measure is lognormal. With these restrictions, caplets of that tenor satisfy the Black (1976) formula for options on forward/futures contracts.

For a particular tenor, τ, market quoted forward rates are required to be lognormal. The tenor is fixed once and for all, since the requirement is that rates of only that tenor are lognormal. If $L(t,x)$ is the market quoted forward rate at time t for time $t+x$ of tenor τ, then the process for the market quoted rate is required to be lognormal as follows:

$$\mathrm{d}L(t,x) = \mu(t,x)\,\mathrm{d}t + \gamma(t,x)L(t,x)\,\mathrm{d}z_t, \tag{5}$$

where $\gamma(t,x)$ is a d-dimensional vector. BGM show that for this restriction to hold, the drift $\mu(t,x)$ must have the form

$$\frac{\partial}{\partial x}L(t,x) + L(t,x)\gamma(t,x)\sigma(t,x) + \frac{\tau L^2(t,x)}{1+\tau L(t,x)}|\gamma(t,x)|^2, \tag{6}$$

where $\sigma(t,x)$ is related to $\gamma(t,x)$ by

$$\sigma(t,x) = \begin{cases} 0, & 0 \leqslant x \leqslant \tau, \\ \sum_{k=1}^{\left|\frac{x}{\tau}\right|} \frac{\tau L(t,x-k\tau)}{1+\tau L(t,x-k\tau)}\gamma(t,x-k\tau), & \tau \leqslant x. \end{cases} \tag{7}$$

The BGM functions $\gamma(t,x)$ are calibrated to the observed Black implied volatilities using the following relation:

$$\sigma_i^2 = \frac{1}{t_{i-1}-t}\int_t^{t_{i-1}} |\gamma(s,t_{i-1}-s)|^2\,\mathrm{d}s. \tag{8}$$

Since the BGM models focus on market quoted instruments, there is no need for instantaneous rates, which are required in the other models.

2.2. Empirical studies

There are very few papers that study the empirical performance of these models in valuing interest rate derivatives. [15] Bühler et al. (1999) test one- and two-factor models in the German fixed-income warrants market, and report that the one-factor forward rate model with linear proportional volatility outperforms all other models. However, their study is limited to options with maturities of less than 3 years; the underlying asset for these options is not homogenous; the estimation of model parameters is based on historical interest rate data rather than on current derivative prices; and most importantly, the paper does not analyze strike-rate biases. The last

[15] Some of the early studies include Flesaker (1993), Amin and Morton (1994), and Canabarro (1995). Amin and Morton analyze only short-term Eurodollar futures options, in a one-factor world. The Canabarro study does not use market data, and tests some models based on simulated Treasury yield curves. Hence the inferences drawn in these studies are not convincing.

point is particularly significant, since in practice, the calibration of the volatility "skew" or "smile" is an important step in the pricing and hedging of options.

Ritchken and Chuang (1999) test a three-state variable Markovian HJM model using a humped volatility structure of forward rates, but only using price data for at-the-money (ATM) caplets. They find that with three-state variables, the model captures the full dynamics of the term structure without using any time-varying parameters. Hull and White (2000) test the LIBOR market model for swaptions and caps across a range of strike rates, but with data for only 1 day. They find that the absolute percentage pricing error for caps is greater than for swaptions. In a similar vein, Peterson et al. (2003) test alternative calibrations in the context of their multi-factor model. Longstaff et al. (2001) [LSS] use a string model framework to test the relative valuation of caps and swaptions using ATM cap and swaptions data. Their results indicate that swaption prices are generated by a four-factor model, and that cap prices periodically deviate from the no-arbitrage values implied by the swaption market. Moraleda and Pelsser (2000) test three spot rate and two forward rate models on cap and floor data from 1993–1994, and find that spot rate models outperform the forward rate models. However, as they acknowledge, their empirical tests are not very formal.

Jagannathan et al. (2003) evaluate the empirical performance of one, two, and three factor CIR models and show that as the number of factors increases, the fit of the models to LIBOR and swap rates improves. However, none of these models is able to price swaptions accurately, leading them to conclude that there may be need for non-affine models to price interest rate derivatives. In fact, Collin-Dufresne and Goldstein (2002) argue that there is a missing stochastic volatility factor that affects the prices of interest rate options, but does not affect the underlying LIBOR or swap rates. They propose models with explicit factors driving volatility, and suggest that cap prices may not be explained well by term structure models that only include yield curve factors. In a similar vein, Heidari and Wu (2002) claim that at least three additional volatility factors are needed to explain movements in the swaption volatility surface. [16] In contrast, Fan et al. (2003) show that swaptions can be well hedged using LIBOR bonds alone.

Two prior papers examine the hedging performance of the alternative models. One is by LSS, where they test their four-factor model against the Black model, and show that the performance of the two models is statistically indistinguishable. The other is by Driessen et al. (2003) [DKM] whose analysis runs parallel to the direction of our paper. [17] DKM test one-factor and multi-factor HJM models with respect to their pricing and hedging performance using ATM cap and swaption volatilities. They find that a one-factor model produces satisfactory pricing results for

[16] Other related papers include De Jong et al. (2002), and Han (2001). De Jong et al. show that historical correlations are significantly higher than those implied by cap and swaption data, hence a volatility risk premium may be present. Han explicitly models the covariances of bond yields as a linear function of a set of state variables, and finds some empirical support for the model.

[17] Fan et al. (2001) also examine the hedging performance of alternative term structure models, but only in the swaption market, not for caps/floors.

caps and swaptions. In terms of hedging performance, for both caps and swaptions, they find that the choice of hedge instruments affects the hedging accuracy more than the particular term structure model chosen. However, as with all other studies cited above, their data set is restricted to ATM options. As noted earlier, the strike rate effect may be extremely important since many of the model imperfections are more evident when one analyzes options away-from-the-money. While it is interesting that they find satisfactory pricing and hedging performance using a one-factor model, even for swaptions, their results are not surprising. The question is whether this conclusion holds up for options that are away-from-the-money. In our paper, we specifically focus on cap and floor prices across *different* strike rates and maturities, to examine how alternative term structure models are affected by strike biases.

The previous studies have important implications regarding the structure of interest rate models appropriate for the interest rate derivative markets. If there is need to explicitly incorporate stochastic volatility factors in the model, then it should be difficult to hedge interest rate options accurately using models consisting of just yield curve factors. In our paper, all the models assume that the term structure is driven by yield curve factors alone. Therefore, our pricing and hedging results for caps and floors, across strikes, have important implications regarding the need for stochastic volatility factors in term structure models, when they are applied to derivative markets.

3. Model implementation and experimental design

The spot rate models (HW and BK) are implemented by constructing a recombining trinomial lattice for the short-term interest rate (as in Hull and White, 1994). The current term structure is estimated from spot LIBOR rates and Eurodollar futures prices. [18] The volatility parameter σ and the mean-reversion parameter a are chosen so as to provide a "best fit" to the market prices of caps and floors, by minimizing the sum of squared residuals. The delta hedge ratios are computed using the quadratic approximation to the first derivative of the option price with respect to the short rate.

Forward rate models are implemented under the HJM framework, with the specified volatility functions, using discrete-time, non-recombining binomial trees (which are computationally efficient). The forward rate process described above is arbitrage-free only in continuous time and, therefore, cannot be directly used to construct a discrete-time tree for the evolution of the forward curve. Therefore, the drift term

[18] Market swap rates can also be used to estimate the LIBOR term structure. However, Eurodollar futures prices are available for maturities upto 10 years in increments of 3 months, and they are very liquid contracts, hence they are likely to reflect the best available information about the term structure. The futures yields are corrected for convexity using standard methods (see Gupta and Subrahmanyam (2000) for a detailed discussion on convexity adjustments, and the methods that can be used to estimate them).

in the forward rate process needs to be reformulated in discrete time. [19] The delta hedge ratios are again computed as before, using the quadratic approximation to the first derivative.

The BGM model is implemented using Monte Carlo simulation, in the interest of computational efficiency. We simulate 5000 different paths, using the initial term structure, and use antithetic variance reduction techniques, to price all our options. Extensive robustness checks were done to ensure that the results were not sensitive to the number of simulated paths. The discretization of the forward rate process and its drift are taken from Hull and White (2000). The delta hedge ratios are computed using a central difference approximation.

3.1. Hedging interest rate caps and floors

Since caplets and floorlets are essentially options on the forward interest rate, they can be hedged with appropriate positions in the LIBOR forward market. In practice, they are most commonly hedged using Eurodollar futures contracts, due to the liquidity of the futures market, as well as the availability of contracts up to a maturity of 10 years, in increments of 3 months. A short position in a caplet (floorlet) can be hedged by going short (long) an appropriate number of futures contracts. The hedge position of the cap (floor) is the sum of the hedge positions for the individual caplets (floorlets) in the cap (floor), i.e., a series of futures contracts of the appropriate maturities, known as the futures strip.

The hedge position is constructed by computing the change in the price of the caplets for a unit (say 1 basis point) change in the forward rate, relative to the number of futures contracts of *appropriate* maturity that give the same change in value for the same unit change in the forward rate. This is the delta hedge ratio for the caplet. In the context of a particular term structure model, the delta can sometimes be defined in closed form. In this paper, the hedge ratios are calculated numerically as explained above. Various robustness checks are done to ensure that the discretization of the continuous-time process does not materially affect the accuracy of the computed delta.

A portfolio of short positions in a cap and an appropriate number of futures contracts is *locally* insensitive to changes in the forward rate, thus making it "delta-neutral". In theory, this delta-neutral hedge requires continuous rebalancing to reflect the changing market conditions. In practice, however, only discrete rebalancing is possible. The accuracy of a delta hedge depends on how well the model's assumptions match the actual movements in interest rates.

A caplet/floorlet can also be gamma-hedged in addition to being delta-hedged, by taking positions in a variety of LIBOR options. Gamma hedging refers to hedging against changes in the hedge ratio. Setting up a gamma-neutral hedge results in a lower hedge slippage over time. However, in principle, the accuracy of the gamma hedge in the context of a particular model could be different from the accuracy of

[19] The discrete time no-arbitrage conditions for the drift term have been adapted from Jarrow (1996) and Radhakrishnan (1998).

the delta hedge within the same model. Therefore, the hedging performance of the models could be different if they were evaluated using both delta and gamma hedging, instead of just delta hedging. In this paper, term structure models are tested based only on their delta hedging effectiveness.

There is a conceptual issue relating to hedging that needs to be defined explicitly. The hedging for any interest rate derivative contract can be done either "within the model" or "outside the model." The "within the model" hedge neutralizes the exposure only to the model driving factor(s), which, in the case of a one-factor model, is the spot or the forward rate. The "outside the model" hedge is determined by calculating price changes with respect to exogenous shocks, which, per se, would have a virtually zero probability of occurrence within the model itself. [20] This "outside the model" procedure is, hence, conceptually internally inconsistent and inappropriate when testing one model against another. [21] The "within the model" hedging tests give very useful indications about the realism of the model itself. The discussion about "delta-hedging" in the previous paragraphs of this section deals only with "within the model" hedging. This is the type of hedging that is empirically examined in this paper.

3.2. Empirical design for testing pricing accuracy

Pricing performance shows how capable a model is of predicting future option prices conditional on term structure information. It is important for valuation models to capture information from current observable market date, and translate them into accurate option prices. [22] Therefore, in this study, the models are calibrated based on the market data on term structure parameters as well as option prices at the current date. Then, at a future date, the same model is used along with current term structure to estimate option prices, which results in a "static" test of the models.

We measure the comparative pricing performance of the models for pricing caps/floors by analyzing the magnitude of the out-of-sample cross-sectional pricing errors. The spot rate models are first estimated using constant parameters so that the models fit the current term structure exactly, but the volatility structure only approximately (in a least squares sense). In the second estimation, the parameters in the spot rate models are made time-varying so that the models fit the volatility term structure exactly as well, by calibration to the observed prices of caps/floors. To examine the out-of-sample pricing performance of each model, the prices of interest rate caps and floors at date t_i are used to calibrate the term structure model and back out the

[20] Examples of such exogenous shocks include jumps in the yield curve or in individual forward rates, changes in the volatilities of interest rates, etc. These are ruled out within the structure of all of the models examined in this paper.

[21] From a practitioner's viewpoint, this inconsistency may be less important than the actual hedge accuracy.

[22] This is especially true for Value-at-Risk systems, where the objective is to be able to accurately estimate option prices in the future, conditional on term structure information.

implied parameters. Using these implied parameter values and the current term structure at date t_{i+1}, the prices of caps and floors are computed at date t_{i+1}. The observed market price is then subtracted from the model-based price, to compute both the absolute pricing error and the percentage pricing error. This procedure is repeated for each cap and floor in the sample, to compute the average absolute and the average percentage pricing errors as well as their standard deviations. These steps are followed separately for each of the models being evaluated. Then, the absolute as well as percentage pricing errors are segmented by type of option (cap or floor), "moneyness" (in-the-money, at-the-money, and out-of-the-money) and maturity to test for systematic biases and patterns in the pricing errors. To analyze the impact of increasing the out-of-sample period on the comparative model performance, we also estimate the pricing errors for each model 1 week out-of-sample. The coefficients of correlation between the pricing errors across the various models are also computed to examine how the models perform with respect to each other.

The cross-sectional pricing performance of the models is further examined using two different calibration methods. The objective of estimating pricing errors using alternative calibration methods is to test the robustness of the pricing results to estimation methodology. In the first one, the prices of ATM caps (of all maturities) are used to calibrate the term structure model. [23] This model is then used to price the away-from-the-money caps of all maturities on the *same* day. The same procedure is repeated for the floors. The model prices are compared with market prices, and the errors are analyzed in a manner similar to the one before. In the second method, the cap prices (of all strike rates and maturities) are generated using the models calibrated to floor prices (of all strike rates and maturities), and floor prices generated by calibrating the models to cap prices. These two tests are strictly cross-sectional in nature, as the prices of options on 1 day are used to price other options on the *same* day, while in the earlier procedure the prices of options on the previous day were used to estimate current option prices.

To study the possible systematic biases in the pricing performance of the models in more detail, the pricing errors for these models are analyzed. The pricing error (in Black vol. terms) is regressed on a series of variables such as moneyness, maturity, etc. to analyze the biases in the pricing errors and identify the model that is most consistent with the data.

3.3. Empirical design for testing hedging accuracy

The hedging tests of these models examine the fundamental assumption underlying the construction of arbitrage-free option pricing models, which is the model's ability to replicate the option by a portfolio of other securities that are sensitive to the same source(s) of uncertainty. [24] This test is conducted by first constructing a

[23] The ATM cap is taken to be the one with the strike that is closest to ATM, since, in general, no fixed strike cap (or floor) will be exactly ATM.

[24] With continuous trading and continuous state variable sample paths, the only sensitivities that matter for hedging are the deltas, hence the higher order sensitivities need not be explicitly considered.

hedge based on a given model, and then examining how the hedge performs over a small time interval subsequently. An accurate model to hedge interest rate exposures must produce price changes similar to those observed in the market, conditional on the changes in its state variables. Hence, the hedging tests are indicative of the extent to which the term structure models capture the future movements in the yield curve, i.e., the dynamics of the term structure. In principle, it is possible for a model to perform well in pricing tests and yet fail in hedging tests, since the two types of tests are measuring different attributes of the model.

These tests are implemented by analyzing the magnitude of the out-of-sample cross-sectional hedging errors. To examine the hedging performance of the models, the term structure models are calibrated at date t_i using the current prices of interest rate caps and floors, and the requisite parameters are backed out. Using the current term structure of interest rates as well as spot cap/floor prices, the delta-hedge portfolio is constructed. The hedge portfolio is constructed separately for caps and floors. Each of these hedge portfolios consists of individual caps (or floors) of the four maturities (2-, 3-, 4- and 5-year), across the four strike prices, and the appropriate number of Eurodollar futures contracts required for hedging it.

In constructing the delta hedge for a caplet/floorlet with interest rate futures contracts, the hedge position must account for an institutional factor. Caps/floors are negotiated each trading date for various maturities; hence, the expiration dates of caplets could be any date in the month. In contrast, exchange-traded futures contracts expire on a particular date. The expiration dates of the futures contracts generally do not coincide with the expiration dates of the individual caplets (floorlets) in the cap (floor). Therefore, it is necessary to create a "synthetic" (hypothetical) futures contract whose expiration date coincides with that of a particular caplet/floorlet, by combining (via interpolation) two adjacent futures contracts with maturity dates on either side of the expiration date of the caplet/floorlet being hedged.

Using this hedge portfolio, the hedging error is computed at date t_{i+k}, to reflect a k-day rebalancing interval. The hedging error corresponds to the change in the value of the hedge portfolio over these k days. In order to test for the effect of the rebalancing interval, the hedging errors are computed using a 5-day and a 20-day rebalancing interval. [25] In both cases, the procedure is repeated for each model, and the hedging errors are analyzed. [26]

[25] A 5-day rebalancing interval corresponds to weekly portfolio rebalancing, while a 20-day rebalancing interval approximates monthly rebalancing. The results using daily rebalancing are not reported in the paper as there was very little hedge slippage over one trading day, thereby leading to almost perfect hedging using any model. Generally speaking, longer term rebalancing intervals provide a more stringent test of the extent to which the dynamics of the underlying interest rate are embedded in the model. The longer rebalancing intervals are in line with the spirit of capital adequacy regulations based on the guidelines of the Bank for International Settlements.

[26] The results reported in this paper are robust to the specific number of time steps in the discrete interest rate trees. Tests were done to study the differences in results by using a larger number of time steps, and the differences were insignificant.

4. Data

The data for this study consist of daily prices of US dollar (USD) caps and floors, for a 10-month period (March 1–December 31, 1998), i.e., 219 trading days, across four different strike rates (6.5%, 7%, 7.5%, 8% for caps, and 5%, 5.5%, 6%, 6.5% for floors) and four maturities (2-, 3-, 4-, and 5-year). [27] These data were obtained from Bloomberg Financial Markets.

Table 1 presents descriptive statistics of the data set. The prices of the contracts are expressed in basis points, i.e., a price of 1 bp implies that the price of the contract for a notional principal of $10,000 is $1. The average, minimum and maximum price of the respective contracts over the sample period are reported in this table. The table indicates that the prices of both caps and floors increase with maturity. The prices of caps (floors) decrease (increase) with the strike rate.

It should be noted that our sample period witnessed considerable volatility in the global fixed-income markets. Several major events triggered by the Russian default and the long term capital management (LTCM) crisis jolted the fixed-income cash and derivatives markets. Hence, the dollar cap and floor markets experienced greater variation in prices than usual. This is fortuitous since it implies that the empirical tests of the various models are that much more stringent and, as a result, our conclusions are likely to be robust.

Since interest rate caps and floors are contracts with specific maturity *periods* rather than specific maturity *dates*, a complication arises while doing the hedging tests. For these tests, we need the market prices of the original cap/floor contract that was hedged using futures. However, each day, the reported prices of caps and floors refer to prices of *new* contracts of corresponding maturities, and not to the prices of the contracts quoted before. Hence, there is no *market price* series for any *individual* cap/floor contract. For example, consider a 5-year cap quoted at date t_i, which is also hedged at date t_i. To evaluate the performance of this hedge at date t_{i+1}, we need the price of the *same* cap at date t_{i+1}, i.e., at date t_{i+1}, we need the price of a cap expiring in 5 years *less 1 day*. However, the cap price that is observed at date t_{i+1} is the price of a *new* cap expiring in 5 years, not 5 years less 1 day. This data problem is not specific to just caps and floors – it is present for all OTC contracts that are fixed maturity rather than fixed maturity date contracts.

To overcome this problem, we construct a price series for each cap/floor contract, each day, until the end of the hedging rebalancing interval. The current term structure and the current term structure of volatilities (from the current prices of caps/floors) are used to price the original cap/floor contract each day. This price is used as a surrogate for the market price of the cap/floor contract on that particular day. This price is a model price, and not a real market price. However, the hedging performance tests are still useful in identifying models that can set up more accurate hedges for the cap/floor contracts. At the very least, the tests will evaluate models in terms of their internal consistency in terms of hedging performance.

[27] Therefore, there are 218 days for which the model forecasts are compared with market prices.

Table 1
Descriptive statistics for cap/floor prices

	2 yr	3 yr	4 yr	5 yr	2 yr	3 yr	4 yr	5 yr
	6.5% Caps				7% Caps			
Mean	16	37	72	117	8	22	47	82
Min	4	13	32	57	2	8	21	42
Max	33	64	109	164	18	38	74	120
	7.5% Caps				8% Caps			
Mean	4	13	31	57	3	8	20	40
Min	2	3	12	29	1	2	8	21
Max	10	24	55	94	5	17	41	75
	5% Floors				5.5% Floors			
Mean	37	132	163	197	67	186	234	284
Min	7	80	98	115	20	112	143	169
Max	129	267	328	385	190	359	445	523
	6% Floors				6.5% Floors			
Mean	116	262	332	401	182	363	461	557
Min	51	166	213	254	106	251	322	385
Max	262	465	580	682	341	583	731	864

This table presents descriptive statistics of the data set used in this paper. The data consists of cap and floor prices across four different maturities (2-, 3-, 4-, and 5-year) and across four different strike rates, for each maturity (6.5%, 7%, 7.5%, and 8% for caps and 5%, 5.5%, 6%, and 6.5% for floors). The sample period consists of 219 trading days of daily data, from March 1 to December 31, 1998. The prices of the contracts are expressed in basis points, i.e., a price of 1 bp implies that the price of the contract for a notional principal of $10,000 is $1. The average, minimum and maximum price of the respective contracts over the sample period are reported in this table.

5. Results

5.1. Parameter stability

To examine the stability of the parameters of the estimated models, summary statistics for the estimated parameters are reported in Table 2. The parameter estimates across models are not directly comparable for several reasons. First, the models use different factors (spot rates and forward rates), with some of them being two-factor models. Second, the drift and volatility functions differ in functional form. Third, the number of parameters estimated varies across models. However, the stability of these parameters can be inferred from the estimate of the coefficient of variation for each parameter. In our two-factor specification, with an exponential structure for the volatility, there are two parameters for each factor, the volatility, σ, and the mean-reversion coefficient, κ.

Our results show that there is some variation in the parameter estimates across time. By definition, the models posit that the drift and volatility parameters are constant. One explanation for this divergence from theory is that there is a second

Table 2
Model parameter estimates

Model	Parameter	Mean	Min	Max	std. dev.	coef. of var.
Spot rate models						
Hull and White	a	0.045	0	0.088	0.027	0.61
	σ	0.0109	0.0051	0.0172	0.0035	0.32
Black and Karasinski	a	0.055	0	0.097	0.025	0.45
	σ	0.194	0.131	0.284	0.056	0.29
Forward rate models – one factor						
Absolute	σ_0	0.0113	0.0075	0.0214	0.0035	0.31
Linear absolute	σ_0	0.0098	0.0031	0.018	0.0043	0.44
	σ_1	0.0007	−0.0029	0.053	0.0018	2.60
Square root	σ_0	0.0456	0.0273	0.0874	0.0105	0.23
Proportional	σ_0	0.1851	0.1169	0.2741	0.0407	0.22
Linear proportional	σ_0	0.1759	0.0799	0.2632	0.0721	0.41
	σ_1	0.0053	−0.0005	0.0138	0.0037	0.70
Forward rate models – two factor						
Absolute exponential	σ_1	0.0062	0.0019	0.0113	0.0032	0.52
	κ_1	0.041	0.005	0.087	0.019	0.46
	σ_2	0.0115	0.0037	0.0221	0.0059	0.51
	κ_2	0.059	0.011	0.098	0.023	0.39
Proportional exponential	σ_1	0.1043	0.0615	0.1592	0.0285	0.27
	κ_1	0.052	0.012	0.095	0.016	0.31
	σ_2	0.1719	0.1077	0.2841	0.0412	0.24
	κ_2	0.035	0.004	0.069	0.013	0.37

This table presents summary statistics for the parameter estimates for the one-factor and two-factor spot rate, forward rate, and market models tested in this paper. The summary statistics for each parameter are computed using daily parameter estimates over the sample period, March 1–December 31, 1998. The models are estimated each day over the 219 day sample period, by calibrating them to the market prices of caps and floors across four different maturities (2-, 3-, 4-, and 5-year) and across four different strike rates for each maturity (6.5%, 7%, 7.5%, 8% for caps, and 5%, 5.5%, 6%, 6.5% for floors).

or third factor driving the evolution of rates, which is manifesting itself in the form of time-varying parameters. Possible candidates for the additional factor could be stochastic volatility, or a curvature factor. However, though the parameters vary over time, they are not unstable. The mean, standard deviation, coefficient of variation, minimum value and the maximum value of the parameters are reported in Table 2. The coefficient of variation for most parameters is below 0.5, and for many parameters it is below 0.33. Comparable statistics are difficult to provide for the BGM model, since model estimation involves calibration of many volatility functions, not specific parameters, each day.

For the one-factor and the two-factor models, the parameter values are more stable for one-parameter models, while the coefficients of variation are significantly higher for the two- and four-parameter models. In the case of spot rate models, the mean-reversion rate has a small absolute value and high standard error relative to the mean estimate, indicating that it is observed with significant error. In the forward rate models, the slope parameters for the linear absolute and linear proportional models have high coefficients of variation and small absolute values, making their estimates less reliable. The parameters of the proportional exponential two-factor model are more stable than those for the absolute exponential model, since the variation in the forward rates absorbs some of the time-series fluctuations. These results indicate that adding more parameters to the model improves the ability of the model to fit prices, but hampers the stability of the estimated model. Therefore, from a practical perspective, the one-parameter one-factor models provide accurate, stable results as far as the model parameters are concerned.

5.2. Pricing performance

The tests for the comparative pricing performance of the models are implemented using the methodology described in Section 3. The results for these tests are reported in Tables 3–7. These results are for out-of-sample fits of model-based prices to the observed market prices. [28]

We present four sets of statistics for model performance, based on the average absolute percentage error, the average percentage error, the average absolute error and the average error. Our main criterion for the fit of the various models is the average absolute percentage error, since it measures the error relative to the price, and hence is not biased heavily towards the more expensive options, which tend to be long-dated and in-the-money. However, we also look at the average percentage error, to check whether there is a bias in a model's forecasts. The average absolute and the average errors are useful to get an order of magnitude estimate that can be compared to the bid–ask spreads.

The summary statistics of the forecast errors, based on out-of-sample estimates 1-day ahead, are presented in Table 3. The table provides a first impression about the empirical quality of the models. The average percentage error is less than 2% in most of the cases, indicating a very small bias in the predictions of the different models. In terms of the error in basis points, the average is below 1 bp for caps, indicating a very small bias in terms of prices across models. For floors, the error is close to 3 bp for the HW – time-varying, and the absolute and linear absolute forward rate models, while it is less than 1 bp for most of the other models. Since the bid–ask spread in these markets is of the order of 2 bp, the fit of the models is generally good.

The average absolute errors and the average absolute percentage errors display a clear pattern. The average absolute percentage errors are roughly similar for caps and floors. Within the class of one-factor models, the average absolute percentage errors are highest for the absolute and linear absolute forward rate models (10.1% and 6.9% for caps and 6.0% and 6.4% for floors) and lowest for the BK – time-varying model (3.3% for caps and 2.4% for floors). All the other models fall in between these models,

[28] Note that these models use 1–4 parameters estimated out-of-sample to simultaneously generate 16 cap and 16 floor prices each day. In terms of the number of options, the models price 304 caplets (19 caplets for four maturities and four strikes each) and 304 floorlets (19 floorlets for four maturities and four strikes each) every day.

Table 3
Pricing performance (1-day ahead)

Model	Caps				Floors			
	Avg error (bp)	Avg abs error (bp)	Avg % error	Avg % abs error	Avg error (bp)	Avg abs error (bp)	Avg % error	Avg % abs error
Spot rate models								
Hull and White	−1.0	2.3	−1.8%	6.9%	1.0	4.6	−1.3%	5.3%
HW – time-varying	−0.2	1.3	−0.9%	4.5%	2.5	3.9	0.5%	3.8%
Black and Karasinski	0.1	1.4	0%	4.3%	−0.1	3.0	−1.3%	3.1%
BK – time-varying	0.4	1.1	0.7%	3.3%	0.2	2.5	−0.7%	2.4%
Forward rate models – one factor								
Absolute	0.8	3.5	1.4%	10.1%	2.9	6.8	−0.3%	6.0%
Linear absolute	0.1	2.3	−0.2%	6.9%	2.6	6.2	−0.7%	6.4%
Square root	−1.2	1.7	−2.3%	4.9%	0.5	3.8	−1.0%	3.9%
Proportional	0.1	1.2	0%	4.0%	−0.1	2.7	−1.3%	2.9%
Linear proportional	0.2	1.2	0.6%	3.9%	−0.1	2.5	−1.1%	2.7%
Forward rate models – two factor								
Absolute exp.	0.9	2.2	1.8%	6.6%	1.8	4.0	0%	3.9%
Proportional exp.	0.1	0.9	0%	3.2%	0.2	2.0	−0.8%	2.1%
Market model – one factor								
BGM	0.5	1.2	0.7%	3.9%	0.1	2.6	−1.2%	2.8%

This table presents summary statistics for the forecast errors (in basis points and percentage terms), 1-day ahead, for the one-factor and two-factor spot rate, forward rate, and market models analyzed in the paper. The average error is defined as the predicted model price minus the observed market price, averaged for the 32 caps and floors (four strike rates each for caps and floors, for each of the four maturities) over the 219 days (March–December, 1998) for which the study was done. The average percentage error is defined as the (model price–market price)/market price, averaged in a similar way.

in terms of prediction errors. [29] The two-factor models have marginally lower pricing errors as compared to the one-factor models that they nest. For example, the two-factor lognormal model has an average absolute percentage error of 3.2% for caps and 2.1% for floors, as compared to 4.0% and 2.9% respectively for the one-factor lognormal model. Also, the spot rate models with time-varying parameters have lower pricing errors for caps as well as floors, as compared to those for the models with constant parameters. Making the parameters time-varying brings down the errors to almost the level of two-factor models. In this case, the time-varying parameters appear to be acting as "pseudo-factors". The one-factor BGM model works as well as the one-factor proportional volatility model. Perhaps, the one-factor lognormal structure that is common to both models is more important than other aspects of the two models.

Table 4 presents the pricing errors similar to the previous table, but based on out-of-sample estimates 1-week ahead. As expected, these errors are more than twice as

[29] The average absolute errors follow a similar pattern.

Table 4
Pricing performance (1-week ahead)

Model	Caps				Floors			
	Avg error (bp)	Avg abs error (bp)	Avg % error	Avg % abs error	Avg error (bp)	Avg abs error (bp)	Avg % error	Avg % abs error
Spot rate models								
Hull and White	−1.5	6.4	−2.6	16.2%	1.3	8.2	−0.7	14.1%
HW – time-varying	0.2	4.3	0.3%	11.5%	2.9	6.1	0.7%	10.2%
Black and Karasinski	0.3	3.2	0.2%	7.9%	0.2	4.9	−0.5%	7.1%
BK – time-varying	0.5	2.8	0.9%	6.8%	0.3	3.4	−0.4%	6.5%
Forward rate models – one factor								
Absolute	1.1	8.3	1.8%	21.4%	3.4	10.4	0.1%	15.5%
Linear absolute	0.7	5.8	0.4%	13.7%	2.7	9.1	−0.5%	14.4%
Square root	−1.0	4.1	−1.9%	10.5%	0.8	7.9	−0.7%	9.8%
Proportional	0.3	3.3	0.1%	7.1%	0	4.2	−0.4%	6.8%
Linear proportional	0.4	3.1	0.8%	6.9%	0.1	4.0	−0.5%	6.6%
Forward rate models – two factor								
Absolute exp.	1.1	7.4	1.7%	17.1%	2.2	8.7	0.1%	13.1%
Proportional exp.	0.2	2.7	0.1%	6.3%	0.1	3.9	−0.3%	6.5%
Market model – one factor								
BGM	0.6	3.2	0.8%	7.2%	0.3	4.5	−0.3%	7.0%

This table presents summary statistics for the forecast errors (in basis points and percentage terms), 1-week ahead, for the one-factor and two-factor spot rate, forward rate, and market models analyzed in the paper. The average error is defined as the predicted model price minus the observed market price, averaged for the 32 caps and floors (four strike rates each for caps and floors, for each of the four maturities) over the 219 days (March–December 1998) for which the study was done. The average percentage error is defined as the (model price – market price)/market price, averaged in a similar way.

large as the errors in Table 3. However, the comparative performance of the models is very similar. For example, the average absolute percentage error for the one-factor proportional model is 7.1% for caps, while it is 6.3% for the two-factor proportional model. Similarly, for floors, the one-factor proportional model has an average absolute percentage error of 6.8%, while it is 6.5% for the two-factor proportional model. Again, the introduction of a second stochastic factor does not improve the pricing performance of the models significantly.

Tables 5 and 6 present the absolute and percentage errors for the caps/floors for all the models, for the cross-sectional tests using different calibration methods. For results in Table 5, the models are first calibrated using ATM cap/floor prices, and then the ITM and OTM cap/floor prices are estimated. The absolute and percentage errors in this case are lower than those in Table 3, where the models are calibrated using cap/floor prices from the previous day. For example, the two- factor proportional exponential model has an average absolute percentage error of 1.7% and 1.3% for caps and floors respectively, compared with 3.2% and 2.1% in the previous calibration. For this calibration, the one-factor BGM model has the lowest percentage pricing errors, while the constant volatility Gaussian model has the highest error.

The proportional volatility models have low pricing error, but they are outperformed by the BGM model. Again, the spot rate models with time-varying parameters have much lower pricing errors. The two-factor models have marginally lower pricing errors than the one-factor models that they nest.

The pricing errors are lowered further in Table 6, where the models are calibrated using caps to estimate floor prices, and using floors to estimate cap prices. [30] Across models, the pattern of errors is similar to the previous table. The one-factor BGM model (along with the two-factor proportional exponential model) outperforms all other models. The lognormal forward rate model provides fairly accurate pricing performance, but not the most accurate in these two calibrations. However, it should be noted that the BGM model is designed to fit contemporaneous cap prices exactly. Hence, in these two alternative calibrations, it performs better that the other models, since these tests, strictly speaking, are not out-of-sample – some of the option prices are being used to price the rest of the options the *same* day. The magnitudes of the pricing errors from the cross-sectional tests reinforce the conclusion that two-factor models are only marginally better than one-factor models for pricing these options (0.9% and 0.8% compared to 1.4% and 1.1%, for caps and floors respectively). The results from the two alternative calibration methods for the models reaffirm that the pricing results reported in Table 3 are robust to changes in model calibration methods. They also show that calibrating models to current option prices (as done in Tables 5 and 6) and to a full range of strike rates (done only in Table 6) results in more accurate pricing performance. Overall, a comparison of all the results for the one-factor models with those for the two-factor models shows that fitting the skew in the distribution of the underlying interest rate improves the static performance of the model more than by introducing another stochastic factor in the model.

Figs. 1 and 2 plot the percentage errors for the models, as a function of their strikes, based on the errors 1-day ahead (corresponding to Table 3). All the models tend to overprice short-dated caps/floors and under-price long-dated ones. However, the over- and under-pricing patterns are different for one-parameter and two-parameter models. The one-parameter models tend to compensate the over-pricing of short-dated options by under-pricing long-dated options. The two-parameter models display a slight hump at the 3-year maturity stage. They overprice medium-term

[30] We did direct put–call parity (cap–floor = swap) tests for the 6.5% strike caps and floors (that is the only strike for which we have both cap and floor data). Considering that we have 219 days of data for four maturities each, it amounted to put–call parity tests on 876 sets of cap/floor prices (219×4). For these options, the put call parity relationship states that it should not be possible to create a long cap, short floor, short swap (with a swap rate of 6.5%) position at negative cost (reflecting an arbitrage). In reality, there will be bid–ask spreads that the arbitrageur will face, especially for the swap which will, in general, be an off-market swap subject to much higher bid–ask spreads than the average 2 bp bid–ask spreads for vanilla swaps. Out of the 876 tests, we find that the cost of creating such an arbitrage portfolio is negative in 42 cases. However, the cost of creating this arbitrage portfolio is less than −5 bp in only 7 cases (we believe that across these 3 trades, the arbitrageur will at least face cumulative trading costs of 5 bp). Therefore, at most, the put call parity relationship is violated in 7 out of 876 cases or about 0.8% of the cases. Since 99.2 of the option prices are definitely not in violation of the put call parity relationship, the errors presented in Table 6 are primarily attributable to model differences, not to data errors.

Table 5
Pricing performance – models calibrated to ATM options

Model	Caps				Floors			
	Avg error (bp)	Avg abs error (bp)	Avg % error	Avg % abs error	Avg error (bp)	Avg abs error (bp)	Avg % error	Avg % abs error
Spot rate models								
Hull and White	−0.7	1.8	−1.3%	4.9%	0.8	3.3	−0.8%	3.7%
HW – time-varying	0.1	1.1	0%	3.3%	1.9	3.0	0.1%	3.0%
Black and Karasinski	0.1	0.9	0.1%	2.6%	0	1.9	−0.8%	1.9%
BK – time-varying	0.4	0.7	0.5%	1.9%	1.2	1.7	0.1%	1.5%
Forward rate models – one factor								
Absolute	0.6	2.7	1.2%	7.8%	2.2	5.2	−0.2%	4.5%
Linear absolute	0.1	1.8	0.1%	5.1%	2.0	4.5	−0.5%	4.7%
Square root	−0.9	1.2	−1.7%	3.4%	0.4	3.0	−0.7%	2.9%
Proportional	0.1	0.9	0%	2.4%	−0.04	1.9	−0.8%	1.9%
Linear proportional	−0.1	0.8	0.1%	2.4%	0.05	1.8	−0.6%	1.8%
Forward rate models – two factor								
Absolute exp.	0.4	1.3	0.5%	3.7%	1.7	4.3	0%	3.3%
Proportional exp.	0	0.4	0%	1.7%	0	1.4	−0.5%	1.3%
Market model – one factor								
BGM	0	0.5	0%	1.6%	0	0.9	0%	1.2%

This table presents summary statistics for the cross-sectional out-of-sample forecast errors (in basis points and percentage terms) for the one-factor and two-factor spot rate, forward rate, and market models. The models are calibrated using the prices of ATM options (out of the four strike rates, the one that is closest to ATM). Then, the prices of the away-from-the-money (ITM and OTM) caps and floors are estimated using the models (for the three remaining strike rates). This is done for all maturities, and for caps and floors separately. The average error is defined as the predicted model price minus the observed market price, averaged for the 12 caps/floors (the three remaining strike rates for each of the four maturities) over the 219 days (March–December 1998) for which the study was done. The average percentage error is defined as the (model price–market price)/market price, averaged in a similar way.

caps/floors more than the short-term ones, and then compensate by under-pricing the long-dated caps/floors. In terms of fitting errors, the two-parameter models are a marginally better fit than the one-parameter models that they nest.

The patterns of mispricing display a clear skew across strike rates, for all maturities. All the models tend to over-price in-the-money (low strike) caps and underprice out-of-the-money (high strike) caps. In the case of floors, the models underprice out-of-the-money (low strike) and overprice in-the-money (high strike). These patterns are consistent across all maturities. The asymmetry or "skew" is the greatest for the constant volatility (Ho–Lee Gaussian model) and the least for the proportional volatility models (one-factor and two-factor lognormal models). For the square root volatility model, in which the distribution of the underlying rate is non-central chi-square (which is less skewed than lognormal), the extent of skew in the pricing errors is also in between the Gaussian and the lognormal models. These patterns are similar

Table 6
Pricing performance – calibrating to caps (floors) for pricing floors (caps)

Model	Caps				Floors			
	Avg error (bp)	Avg abs error (bp)	Avg % error	Avg % abs error	Avg error (bp)	Avg abs error (bp)	Avg % error	Avg % abs error
Spot rate models								
Hull and White	−0.5	1.2	−0.9%	3.5%	0.6	2.3	−0.6%	2.5%
HW – time-varying	0.2	0.8	0.1%	2.5%	1.3	2.0	0.1%	2.0%
Black and Karasinski	0.1	0.6	0.1%	1.8%	0	1.2	−0.5%	1.2%
BK – time-varying	0.3	0.5	0.5%	1.3%	0.3	0.9	−0.2%	0.9%
Forward rate models – one factor								
Absolute	0.5	2.0	0.9%	5.7%	1.6	3.7	−0.2%	3.2%
Linear absolute	0.1	1.3	0.1%	3.8%	1.5	3.3	−0.4%	3.3%
Square root	−0.6	0.8	−1.1%	2.3%	0.3	2.0	−0.5%	1.9%
Proportional	0	0.5	0%	1.5%	0	1.2	−0.5%	1.2%
Linear proportional	−0.1	0.5	0%	1.4%	0	1.1	−0.4%	1.1%
Forward rate models – two factor								
Absolute exp.	0.3	1.2	0.5%	3.3%	0.8	1.9	−0.1%	1.7%
Proportional exp.	0	0.3	0%	0.9%	0	0.7	−0.2%	0.8%
Market model – one factor								
BGM	0	0.4	0%	0.9%	0	1.0	0%	0.8%

This table presents summary statistics for the cross-sectional out-of-sample forecast errors (in basis points and percentage terms) for the one-factor and two-factor spot rate, forward rate, and market models. For pricing caps, the models are calibrated using the current prices of floors, and vice versa. The average error is defined as the predicted model price minus the observed market price, averaged for the 16 caps or floors (four strike rates for each of the four maturities) over the 219 days (March–December 1998) for which the study was done. The average percentage error is defined as the (model price – market price)/market price, averaged in a similar way.

for caps and floors, and are consistent across spot rate and forward rate models, as well as one-factor and two-factor models.

This negative skew in the pricing errors is consistent with the hypothesis that fatter right tails in the distribution of the underlying interest rate would lead to under-pricing in out-of-the-money caps and floors. The results indicate that the risk-neutral distribution of the underlying interest rate has a thinner left tail and a fatter right tail than the assumed distribution for any of these models. The partial correction of the skew by the lognormal model suggests that a skewness greater than that in the lognormal distribution may help to predict away-from-the-money cap and floor prices better.

To study the systematic biases in more detail, the following cross-sectional regression model is estimated for caps and floors separately:

$$(\mathrm{IV}_{\mathrm{mkt}} - \mathrm{IV}_{\mathrm{model}})_t = \beta_0 + \beta_1 \mathrm{LMR}_t + \beta_2 \mathrm{MAT}_t + \beta_3 \mathrm{ATMVol}_t + \beta_4 r_t + \beta_5 \mathrm{Slope}_t + \varepsilon_t. \tag{9}$$

We regress the model error (in terms of Black volatilities) on the logarithm of the moneyness ratio (ratio of the par swap rate to the strike rate of the cap/floor)

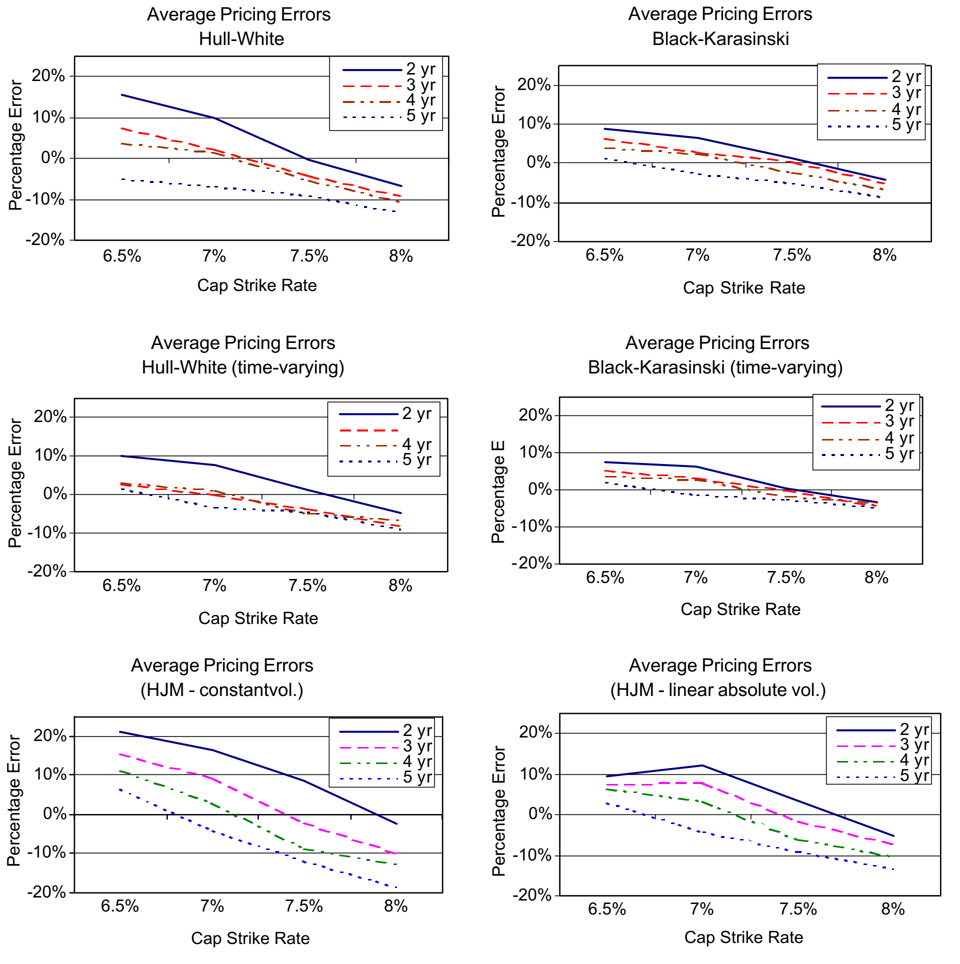

Fig. 1. Pricing performance for caps. These figures present the average percentage pricing errors in predicting the prices of caps, 1-day ahead, using the one-factor and two-factor spot rate, forward rate, and market models. The errors presented pertain to caps of 2-, 3-, 4- and 5-year maturity for strike rates of 6.5%, 7%, 7.5% and 8%. These errors are averaged over the 219 trading day sample period, March 1–December 31, 1998.

(LMR), the maturity of the cap/floor (MAT), the ATM volatility of a similar maturity option (ATMVol), and the 3-month LIBOR (r_t), and the slope of the term structure (Slope, defined as the difference between the 5-year rate and the 3-month rate). The objective of this analysis is to identify any biases in our pricing errors, in order to identify the model that is most consistent with data. The model error is represented in terms of Black volatilities, since that removes the term structure effects from the option price, thereby eliminating any effects of interest rate drifts, strike rates, etc.

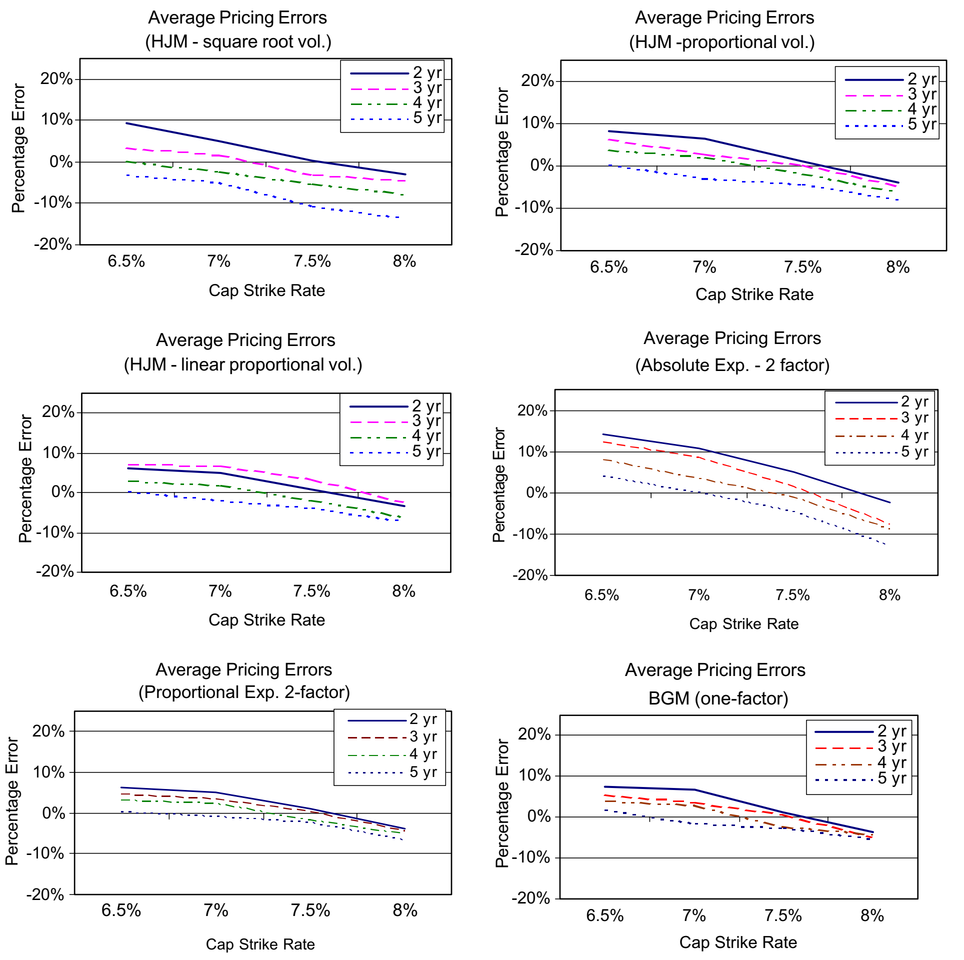

Fig. 1 (*continued*)

First, the dependence on the depth in-the-money, the "smile" or "skew" effect, is captured by the LMR. Second, the maturity or term structure effect is measured by the MAT variable. Third, the volatility variable is added to examine whether the patterns of smile vary significantly with the level of uncertainty in the market. During periods of greater uncertainty, there is likely to be more information asymmetry than during periods of lower volatility. If there is significantly greater information asymmetry, market makers may charge higher than normal prices for away-from-the-money options, since they may be more averse to taking short positions at these strikes. This will lead to a steeper smile, especially on the ask side of the smile curve. Fourth, the absolute level of interest rates is also indicative of general economic conditions, as well as the direction of interest rate changes in the future – for example, if

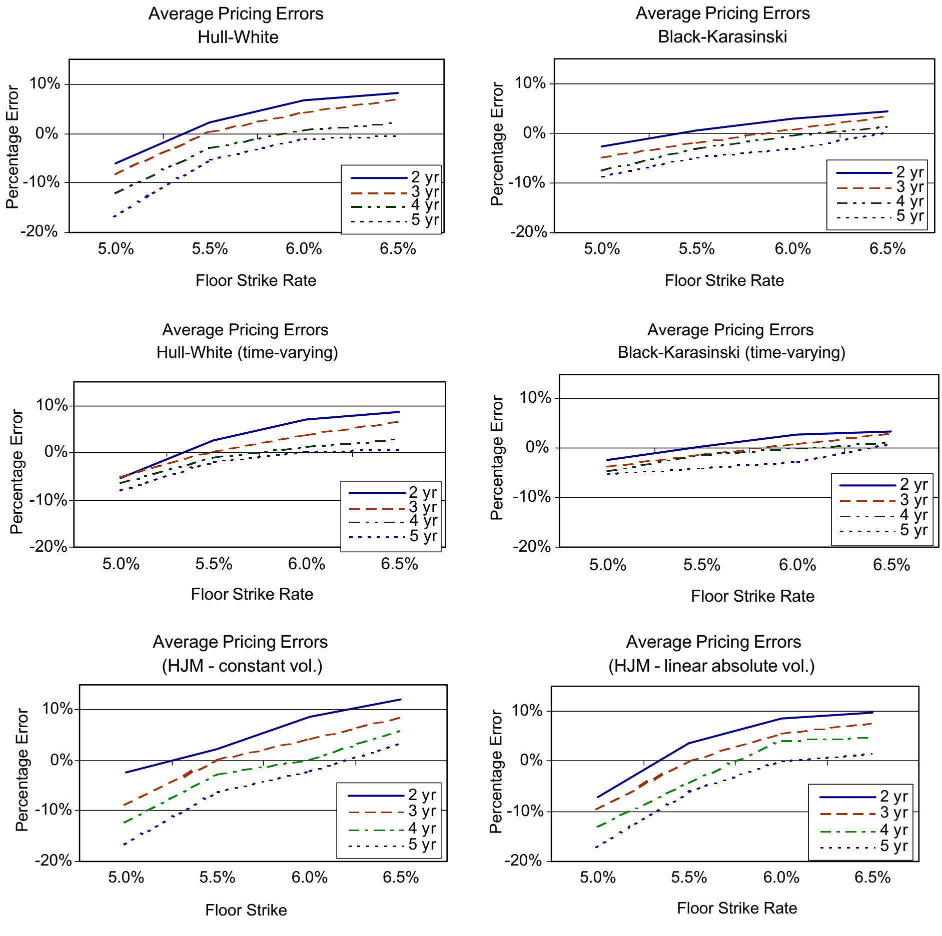

Fig. 2. Pricing performance for floors. These figures present the average percentage pricing errors in predicting the prices of floors, 1-day ahead, using the spot rate, forward rate, and market models. The errors presented pertain to floors of 2-, 3-, 4- and 5-year maturity for strike rates of 5%, 5.5%, 6% and 6.5%. These errors are averaged over the 219 trading day sample period, March 1–December 31, 1998.

interest rates are mean-reverting, very low interest rates are likely to be followed by rate increases. This would manifest itself in a higher demand for out-of-the-money caps in the market, thus affecting the prices of these options, and possibly the shape of the implied volatility smile itself. Last, the slope of the yield curve is added as an explanatory variable, as it is widely believed to proxy for general economic conditions, in particular the stage of the business cycle. The slope of the yield curve is also an indicator of future interest rates, which affects the demand for away-from-the-money options: if interest rates are expected to increase steeply, there will be a high demand for out-of-the-money caps, resulting in a steepening of the smile curve.

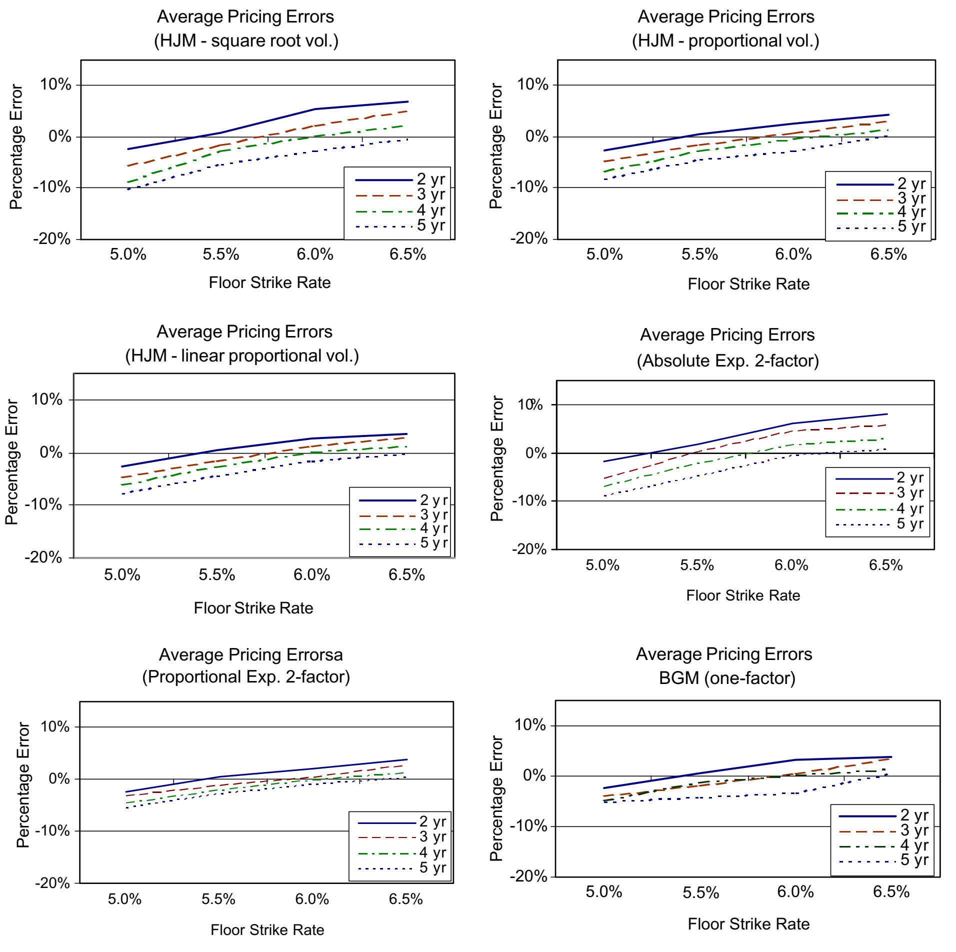

Fig. 2 (*continued*)

The results of this estimation are presented in Table 7. All the models exhibit some biases, though the nature and severity of the biases vary across the models. We reject the null hypothesis that all coefficients are jointly 0 for all the models. For caps, for most of the models, as the strike rate increases, the pricing errors change from positive to negative, which is reflected in the generally positive (and significant) coefficient on LMR. The opposite trend is observed for floors. All options show an inverse relationship between maturity and pricing errors, i.e., shorter term options are generally overpriced, while the longer term options are underpriced. Most of the other coefficients are not significant. The adjusted R^2 coefficients also indicate the severity of the bias for each model. From a pricing perspective, the log-normal models appear to have the least amount of bias in the pricing errors, generally speaking.

5.3. Hedging performance

The tests for the comparative dynamic accuracy of the models are conducted using the methodology described in Section 3.3. The results for this analysis are presented in Table 8. The accuracy of hedging, and hence the accuracy of replication of the interest rate options, differs significantly across term structure models. The average absolute percentage hedging errors reported in Table 8 show that two-factor models perform significantly better than one-factor models in hedging interest rate risk in caps and floors. The difference is more significant for longer rebalancing intervals. With a 5-day rebalancing interval, most one-factor model hedges typically result in an average absolute percentage error of less than 0.5% of the hedge portfolio value for caps and 0.75% for floors. In the case of two-factor models, the 5-day average percentage error is reduced to less than 0.2%. With a 20-day rebalancing interval, the average absolute percentage hedging error reduces from 1.6–3% for various one-factor models to 0.5–0.7% for the two-factor models. Interestingly, the hedging results for the time-varying implementation of the spot rate models are very different from the pricing results – making the parameters time-varying actually leads to consistently *larger* hedging errors, indicating that the stability of model parameter estimation is important for accurate hedging performance. The hedging errors are evidence of the overall effectiveness of the interest rate hedges created by the models over time. Hence, the hedging performance reflects the dynamic accuracy of the various term structure models.

Within the class of one-factor and two-factor models, the hedging errors do depict the trend observed in the pricing errors, of a higher skew in the underlying distribution leading to smaller errors. For example, for the 5-day rebalancing interval, the average absolute percentage error for caps goes down from 0.68% for the Gaussian one-factor forward rate model to 0.33% for the lognormal one-factor forward rate model. Similarly, for the 20-day rebalancing interval, the error goes down from 2.44% to 1.62%, respectively. However, adding a second stochastic factor leads to a much larger reduction in the hedging errors. This result is different from the pricing results where fitting the skew correctly dominated the introduction of a second stochastic factor. The Gaussian two-factor forward rate model has an average percentage absolute error of 0.18% for 5-day rebalancing and 0.51% for 20-day rebalancing, which is significantly lower than those for the one-factor lognormal forward rate model.

A principal component analysis of changes in the zero coupon yields constructed from swap rates (USD) over our sample period (March–December 1998) reveals the various factors that drive the evolution of the term structure. [31] The first factor, interpreted as the "level" factor capturing parallel shifts in the term structure, contributes 92.47% of the overall explained variance of interest rate changes. The second factor, interpreted as a "twist" factor in the yield curve, incorporating changes in the slope of the term structure, contributes another 5.57% of the overall explained

[31] See, for example, Brown and Schaefer (1994) and Rebonato (1998).

Table 7
Model error regressions

Model	β_0	β_1	β_2	β_3	β_4	β_5	Adj. R^2
Panel A: Caps							
Spot rate models							
Hull and White	−0.29	0.28*	−0.09*	0.23	0.11	−0.02	21%
HW – time-varying	−0.14	0.19*	−0.04*	0.11	−0.24	0.07	14%
Black and Karasinski	−0.09	0.15*	−0.05*	0.17*	0.09	0.05	11%
BK – time-varying	−0.11	0.12*	−0.02	−0.05	0.15	−0.17	10%
Forward rate models – one-factor							
Absolute	−0.37	0.31*	−0.14*	−0.41	0.22	0.05	23%
Linear absolute	−0.20	0.04	−0.09*	0.44*	0.21	0.07	12%
Square root	−0.07	0.05*	−0.12*	0.32	0.04	0.15	13%
Proportional	0.13	0.07*	−0.03	0.21	0.17	−0.15	7%
Linear prop.	−0.15	0.03	−0.04	0.09	0.41	0.33	6%
Forward rate models – two-factor							
Absolute exp.	0.14	0.15*	−0.11*	−0.41	0.35	0.22	13%
Proportional exp.	0.03	−0.02	−0.03	0.21	−0.19	0.08	4%
Market model – one-factor							
BGM	0.10	0.03	0.02	−0.25	0.13	0.19	6%
Panel B: Floors							
Spot rate models							
Hull and White	−0.07	−0.09*	−0.05*	0.41	0.33	−0.04	16%
HW – time-varying	0.01	−0.07*	−0.04*	0.31	−0.18	−0.07	11%
Black and Karasinski	0.12	−0.03	−0.04*	−0.28*	−0.11	0.05	7%
BK – time-varying	−0.10	0.01	−0.02	0.17	0.09	0.13	5%
Forward rate models – one-factor							
Absolute	0.15	−0.14*	−0.11*	0.22	0.19	−0.07	21%
Linear absolute	0.24	−0.06*	−0.05*	−0.31	0.44	0.09	14%
Square root	0.25	−0.04	−0.07*	−0.49*	0.52	−0.19	12%
Proportional	−0.05	−0.01	−0.03*	0.22	0.15	−0.04	5%
Linear prop.	0.16	−0.02	−0.03	0.17	−0.12	−0.07	4%
Forward rate models – two-factor							
Absolute exp.	0.33	−0.04*	−0.05*	0.19	−0.27	−0.22	9%
Proportional exp.	0.19	−0.02	0.01	0.15	0.03	−0.21	3%
Market model – one-factor							
BGM	0.24	0.01	−0.03	−0.21	0.34	−0.18	5%

This table presents results for model performance by estimating the following regression model for each of the one-factor and two-factor models examined in the paper:

$$(\mathrm{IV}_{\mathrm{mkt}} - \mathrm{IV}_{\mathrm{model}})_t = \beta_0 + \beta_1 \mathrm{LMR}_t + \beta_2 \mathrm{MAT}_t + \beta_3 \mathrm{ATMVol}_t + \beta_4 r_t + \beta_5 \mathrm{Slope}_t + \varepsilon_t.$$

The model and market prices of the caps and floors are expressed in basis points, for the 219 daily observations during the sample period March–December 1998. All the caps (6.5%, 7%, 7.5%, and 8% strike) and floors (5%, 5.5%, 6%, 6.5%) for each of the four maturities (2-, 3-, 4-, and 5-year) are used in the regression model to test for biases in model performance. * denotes statistical significance at the 5% level.

variance of interest rate changes. The third factor, interpreted as the "curvature" factor, incorporating changes in the curvature of the term structure, explains 0.71% of

Table 8
Hedging performance

Model	Caps				Floors			
	5-day rebal.		20-day rebal.		5-day rebal.		20-day rebal.	
	Avg. % error	Avg. % abs error	Avg. % error	Avg. % abs error	Avg. % error	Avg. % abs error	Avg. % error	Avg. % abs error
Spot rate models								
Hull and White	0.05%	0.56%	0.17%	2.67%	0.03%	0.76%	0.22%	3.04%
HW – time-varying	0.04%	0.51%	0.29%	3.22%	0.05%	0.59%	0.35%	4.22%
Black and Karasinski	−0.03%	0.41%	−0.09%	2.05%	0.06%	0.58%	0.19%	2.41%
BK – time-varying	−0.12%	0.32%	0.03%	2.11%	0.11%	0.53%	0.25%	2.78%
Forward rate models – one factor								
Absolute	0.08%	0.68%	0.07%	2.44%	0.12%	0.81%	0.04%	3.15%
Linear absolute	0.11%	0.52%	0.13%	2.23%	0.09%	0.75%	0.14%	2.57%
Square root	0.10%	0.46%	0.21%	1.98%	−0.13%	0.44%	−0.08%	2.16%
Proportional	0.04%	0.33%	0.07%	1.62%	0.07%	0.31%	0.11%	1.55%
Linear proportional	0.04%	0.37%	0.08%	1.67%	0.05%	0.29%	0.09%	1.69%
Forward rate models – two factor								
Absolute exp.	0.02%	0.18%	0.04%	0.51%	0.01%	0.13%	0.01%	0.69%
Proportional exp.	0.02%	0.10%	0.04%	0.45%	−0.01%	0.14%	−0.01%	0.53%
Market model – one factor								
BGM	0.05%	0.38%	0.07%	1.65%	0.06%	0.33%	0.12%	1.59%

This table presents summary statistics for the hedging errors for the one-factor and two-factor spot rate, forward rate, and market models. The hedging error is defined as the percentage change in the value of the hedge portfolio over a 5-day and a 20-day rebalancing interval. This error is averaged over the 219 days (March–December 1998) for which the study was done. The hedge portfolio consists of one each of all the caps (floors) in the sample, across the four strike rates and the four maturities, and the appropriate

the variance of interest rate changes. [32] The results in this paper show that, for accurate hedging of even simple interest rate options like caps and floors, it is not enough to correctly model just the first factor. Modeling the second factor allows the incorporation of expected twists in the yield curve while determining state variable sensitivities, thereby leading to more accurate hedging. This also constitutes evidence against claims in the literature, that correctly specified and calibrated one-factor models can replace multi-factor models for hedging purposes. [33]

[32] This third factor may be important for pricing swaptions and bond options, but not for pricing interest rate caps and floors.

[33] See, for example, Hull and White (1990), and Buser et al. (1990).

6. Conclusions

A variety of models of interest rate dynamics have been proposed in the literature to value interest rate contingent claims. While there has been substantial theoretical research on models to value these claims, their empirical validity has not been tested with equal rigor. This paper presents extensive empirical tests of the pricing and hedging accuracy of term structure models in the interest rate cap and floor markets. The paper also examines, for the first time in the literature, actual price data for caps and floors across strike rates, with maturities extending out to 5 years.

Alternative one-factor and two-factor models are examined based on the accuracy of their out-of-sample price prediction, and their ability to hedge caps and floors. Within the class of one-factor models, two spot rate, five forward rate, and one market model specifications are analyzed. For two-factor models, two forward rate specifications are examined. Overall, in terms of the out-of-sample static tests, the one-factor lognormal (proportional volatility) forward rate model is found to outperform the other competing one-factor models in pricing accuracy. The estimated parameters of this model are more stable than those for corresponding two-parameter models, indicating that one-parameter models result in more robust estimation. In contrast, the pricing errors allowing for time-varying implementation of the one-factor models are at the level of those for the two-factor models: the time-varying parameters appear to be acting as "pseudo-factors." However, making the parameters time-varying actually leads to consistently *larger* hedging errors, indicating that the stability of model parameter estimation is important for accurate hedging performance. The one-factor BGM model also provides accurate pricing results, but outperforms the lognormal model only in tests which are not strictly out-of-sample.

More significantly, the lognormal assumption in the distribution of the underlying forward rate leads to a smaller "skew" in pricing errors across strike rates, as compared to the errors obtained by using a Gaussian interest rate process. The pricing accuracy of two-factor models is found to be only marginally better than the corresponding one-factor models that they nest. Therefore, the results show that a positive skewness in the distribution of the underlying rate helps to explain away-from-the-money cap and floor prices more accurately, while the introduction of a second stochastic factor has only a marginal impact on pricing caps and floor.

On the other hand, the tests for the hedging performance of these models show that two-factor models are more effective in hedging the interest rate risk in caps and floors. While fitting the skew improves hedging performance marginally, introducing a second stochastic factor in the term structure model leads to significantly more accurate hedging. The one-factor BGM model provides hedging accuracy similar to the one-factor lognormal forward rate model, perhaps due to the common lognormal structure, but is outperformed by two-factor models. The two-factor models allow a better representation of the dynamic evolution of the yield curve, by incorporating expected changes in the slope of the term structure. Since the interest rate dynamics embedded in two-factor models is closer to the one driving the actual economic environment, as compared to one-factor models, they are more accurate in hedging interest rate caps and floors. This result is also evidence against claims in the literature that correctly specified and calibrated one-factor models could replace multi-factor models for hedging.

So what are the implications of these results for the pricing and hedging of caps and floors in particular, and interest rate contingent claims in general? For interest rate caps and floors, one-factor lognormal and BGM models have been found to be sufficiently accurate in pricing performance. However, even for these plain-vanilla options, there is a need to use two-factor models for accurate hedging. Therefore, for consistent pricing and hedging within a book, even for plain-vanilla options like caps and floors, there is evidence that strongly suggests using two-factor models, over and above fitting the skew in the underlying interest rate distribution. Whether there is need for a third factor driving the term structure is still an open question for research. [34] Introducing more stochastic factors in the model makes computations more time consuming, so there is a trade-off between the cost of implementing a model and the stability of the model parameters, on the one hand, and its accuracy, on the other. However, for consistent pricing and hedging of the interest rate exposures of more complicated interest rate contingent claims like swaptions and yield spread options, there may be significant benefits to using term structure models with three or more factors. In addition, since our models are able to hedge caps and floors accurately even 1 month out-of-sample, there seems to be little need to explicitly incorporate stochastic volatility factors in the model, if the objective is just to hedge these options. We defer these issues to be explored in future research.

Acknowledgements

We thank Viral Acharya, Andrew Carverhill, Stephen Figlewski, Kenneth Garbade, A.R. Radhakrishnan, Matthew Richardson, Peter Ritchken, and Richard Stapleton for valuable comments on earlier drafts. We acknowledge with thanks the insightful comments provided by two anonymous referees on a previous draft of the paper, which led to a substantial improvement in our model specifications. We are also thankful to the participants in the seminars at Baruch College, Case Western Reserve University, McGill University, New York University, Rutgers University, University of South Carolina, University of Strathclyde, University of Toronto, and conference participants at the Derivatives Securities Conference at Boston University, the Western Finance Association meetings, the European Financial Management Association meetings, the European Finance Association meetings and the Financial Management Association meetings for comments and suggestions. The usual disclaimer applies.

[34] Litterman and Scheinkman (1991) report that the third factor, modeling changes in the curvature of the term structure, is important in explaining price changes.

References

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

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

Black, F., Derman, E., Toy, W., 1990. A one-factor model of interest rates and its application to Treasury bond options. Financial Analysts Journal 46, 33–39.

Black, F., Karasinski, P., 1991. Bond and option pricing when short rates are lognormal. Financial Analysts Journal 47, 52–59.

Brace, A., Gatarek, D., Musiela, M., 1997. The market model of interest rate dynamics. Mathematical Finance 7, 127–155.

Brennan, M., Schwartz, E., 1979. A continuous time approach to the pricing of bonds. Journal of Banking and Finance 3, 133–155.

Brown, R.H., Schaefer, S.M., 1994. Why do long-term forward rates (almost always) slope downwards? Working Paper, London Business School.

Bühler, W., Uhrig, M., Walter, U., Weber, T., 1999. An empirical comparison of forward- and spot-rate models for valuing interest-rate options. Journal of Finance 54, 269–305.

Buser, S.A., Hendershott, P., Sanders, A., 1990. Determinants of the value of the call options on default-free bonds. Journal of Business 63, 33–50.

Canabarro, E., 1995. Comparing the dynamic accuracy of yield-curve-based interest rate contingent claim pricing models. Journal of Financial Engineering 2, 365–401.

Collin-Dufresne, P., Goldstein, R.S., 2002. Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility. Journal of Finance 57, 1685–1730.

Cox, J.C., Ingersoll, J.E., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–408.

De Jong, F., Driessen, J., Pelsser, A., 2002. On the information in the interest rate term structure and option prices. Working paper, University of Amsterdam.

Driessen, J., Klaassen, P., Melenberg, B., 2003. The performance of multi-factor term structure models for pricing and hedging caps and swaptions. Journal of Financial and Quantitative Analysis 38 (3), 635–672.

Fan, R., Gupta, A., Ritchken, P., 2001. On pricing and hedging in the swaption market: How many factors, really? Working paper, Case Western Reserve University.

Fan, R., Gupta, A., Ritchken, P., 2003. Hedging in the possible presence of unspanned stochastic volatility: Evidence from swaption markets. Journal of Finance 58, 2219–2248.

Flesaker, B., 1993. Testing the Heath–Jarrow–Morton/Ho–Lee model of interest-rate contingent claims pricing. Journal of Financial and Quantitative Analysis 28, 483–495.

Gupta, A., Subrahmanyam, M.G., 2000. An empirical examination of the convexity bias in the pricing of interest rate swaps. Journal of Financial Economics 55, 239–279.

Han, B., 2001. Stochastic volatilities and correlations of bond yields. Working paper, UCLA.

Heath, D., Jarrow, R.A., Morton, A., 1990. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. Econometrica 60, 77–105.

Heidari, M., Wu, L., 2002. Are interest rate derivatives spanned by the term structure of interest rates? Working paper, Fordham University.

Ho, T.S.Y., Lee, S.B., 1986. Term structure movements and pricing interest rate contingent claims. Journal of Finance 41, 1011–1029.

Hull, J., White, A., 1990. Pricing interest rate derivative securities. Review of Financial Studies 3, 573–592.

Hull, J., White, A., 1994. Numerical procedures for implementing term structure models: Single-factor models. Journal of Derivatives 2, 7–16.

Hull, J., White, A., 1995. A note on the models of Hull and White for pricing options on the term structure: Response. Journal of Fixed Income 5, 97–102.

Hull, J., White, A., 1996. Using Hull–White interest rate trees. Journal of Derivatives (Spring), 26–36.

Hull, J., White, A., 2000. Forward rate volatilities, swap rate volatilities, and the implementation of the LIBOR market model. Journal of Fixed Income 10, 46–62.

Jagannathan, R., Kaplin, A., Sun, S., 2003. An evaluation of multi-factor CIR models using LIBOR, swap rates, and cap and swaption prices. Journal of Econometrics 116, 113–146.
Jarrow, R., 1996. Modeling Fixed Income Securities and Interest Rate Options. McGraw-Hill.
Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. Journal of Fixed Income 1, 54–61.
Longstaff, F.A., Schwartz, E., 1992. Interest-rate volatility and the term structure: A two-factor general equilibrium model. Journal of Finance 47, 1259–1282.
Longstaff, F.A., Santa-Clara, P., Schwartz, E., 2001. The relative valuation of caps and swaptions: Theory and empirical evidence. Journal of Finance 56, 2067–2109.
Miltersen, K.R., Sandmann, K., Sondermann, D., 1997. Closed form solutions for term structure derivatives with log-normal interest rates. Journal of Finance 52, 409–430.
Moraleda, J.M., Pelsser, A., 2000. Forward versus spot interest rate models of the term structure: An empirical comparison. The Journal of Derivatives 7, 9–21.
Peterson, S., Stapleton, R.C., Subrahmanyam, M.G., 2003. A multi-factor spot rate model for the pricing of interest rate derivatives. Journal of Financial and Quantitative Analysis 38, 847–880.
Radhakrishnan, A.R., 1998. An empirical study of the convergence properties of non-recombining HJM forward rate tree in pricing interest rate derivatives. Working Paper, NYU.
Rebonato, R., 1998. Interest-Rate Option Models, second ed. John Wiley & Sons, England.
Ritchken, P., Chuang, I., 1999. Interest rate option pricing with volatility humps. Review of Derivatives Research 3, 237–262.
Ritchken, P., Sankarasubramanian, L., 1995. Volatility structures of forward rates and the dynamics of the term structure. Mathematical Finance 5, 55–72.
Stapleton, R.C., Subrahmanyam, M.G., 1999. A two-factor stochastic model of interest rate futures prices. Working Paper, New York University.
Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177–188.