Pricing and Hedging the Smile with SABR: Evidence from the Interest Rate Caps Market

Tao L. Wu*

December 27, 2011

Abstract

This is the first comprehensive study of the SABR (Stochastic Alpha-Beta-Rho) model (Hagan et. al (2002)) on the pricing and hedging of interest rate caps. We implement several versions of the SABR interest rate model and analyze their respective pricing and hedging performance using two years of daily data with seven different strikes and ten different tenors on each trading day. In-sample and out-of-sample tests show that in addition to having stochastic volatility for the forward rate, it is essential to recalibrate daily either the "vol of vol" or the correlation between forward rate and its volatility, although recalibrating both further improves pricing performance. The fully stochastic version of the SABR model exhibits excellent pricing accuracy and more importantly, captures the dynamics of the volatility smile over time very well. This is further demonstrated through examining delta hedging performance based on the SABR model. Our hedging result indicates that the SABR model produces accurate hedge ratios that outperform those implied by the Black model.

---

* Illinois Institute of Technology, 565 W. Adams Street, Chicago, IL 60661, USA. Tel: (312) 906-6553, Fax: (312) 906-6549, email: tw33_99@yahoo.com. I would like to thank Bob Webb and Gyoocheol Shim for very helpful comments that significantly improved this paper.

1. Introduction

The empirical performance of interest rate models in pricing and hedging derivatives has attracted the attention of academic researchers. Its importance has been foreshadowed in Dai and Singleton (2003). Jaganathan, Kaplin and Sun (2003) show that three-factor CIR models produce large pricing errors for caps and swaptions. Collin-Dufresne and Goldstein (2002), as well as Heidari and Wu (2003) present evidence for the "unspanned stochastic volatility" puzzle by noting that risk factors driving interest rate derivatives prices are unspanned by factors underlying LIBOR and swap rates. They suspect stochastic volatility factors unspanned by bond yields are needed to explain derivatives prices. Li and Zhao (2006) report that although quadratic term structure models fit LIBOR and swap rates well, they lead to poor hedging performance of cap straddles. Longstaff, Santa-Clara, Schwartz (2001), Driessen, Klaassen and Melenberg (2003), and Duarte (2008), among others, investigate the pricing of at-the-money interest rate options. Jarrow, Li, and Zhao (2007) introduce jumps to LIBOR market models to capture the implied volatility smile in the interest caps market. Alternatively, Trolle and Schwartz propose a stochastic volatility model. However, neither Trolle and Schwartz (2008) nor Jarrow, Li, and Zhao (2007) conduct out-of-sample hedging analysis. Gupta and Subrahmanyam (2005) study pricing and hedging of caps/floors using models with only yield-based factors (without stochastic volatility). But both the length and moneyness/tenor span of their data sample are more limited than those of our data. Surprisingly, the most popular but simpler SABR model used extensively in practice has eluded the examination by academics. We fill this void in the literature and study the empirical performance of the SABR model in the pricing and hedging of interest rate caps.

The SABR model is a stochastic volatility model that attempts to capture the volatility smile in derivatives markets. The name SABR stands for "Stochastic Alpha, Beta, Rho", referring to the variables of the model. The SABR model is widely used by practitioners in the financial industry, especially in the interest rate derivatives markets. We believe this is the first systematic empirical study of the SABR model in pricing interest rate derivatives using a comprehensive dataset.

The SABR model describes the dynamics of the forward F, such as a LIBOR forward rate, a forward swap rate, or a forward stock price. The volatility of the forward F is described by $\alpha > 0$. SABR is a dynamic model in which both F and α are represented by stochastic state variables whose time evolution is given by the following system of stochastic differential equations under the forward risk neutral measure:

$$dF = \alpha F^{\beta} dW$$

$$d\alpha = \upsilon \alpha dZ$$

$$dWdZ = \rho dt$$

where ν is the volatility of α, and ρ is the instantaneous correlation between the two Brownian Motions W and Z. The SABR model can be viewed as a stochastic extension of the CEV model (contant α) with skewness parameter β. β is between 0 and 1.

Caplets are European options on the LIBOR. Under the SABR model, the price of

European options is given by Black's formula:

$$V_{call} = D(t_{set})\{ fN(d_1) - KN(d_2)\}$$
$$V_{put} = V_{call} + D(t_{set})[K - f]$$

with $d_{1,2} = \dfrac{\log f / K \pm \frac{1}{2}\sigma_B^2 \tau_{ex}}{\sigma_B \sqrt{\tau_{ex}}}$

where σ_B is the Black implied volatility and $D(t_{set})$ is the discount factor.

f is the forward price and K is the strike price. Define log moneyness $x = \ln(f/K)$. Then the implied volatility in the Hagan et. al. (2002) SABR paper is given by:

$I(x,\tau) = I^0(x)(1 + I^1(x)\tau) + O(\tau^2)$, with

$$I^0(0) = \alpha K^{\beta-1}$$

$$I^0(x)|_{\nu=0} = \frac{x\alpha(1-\beta)}{f^{1-\beta} - K^{1-\beta}}$$

$$I^0(x)|_{\beta=1} = \nu x / \ln\left(\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho}\right),\ z = \frac{\nu x}{\alpha}$$

$$I^0(x)|_{\beta<1} = \nu x \frac{\zeta}{z} / \ln\left(\frac{\sqrt{1-2\rho\zeta + \zeta^2} + \zeta - \rho}{1-\rho}\right),\ z = \frac{\nu}{\alpha}\frac{f^{1-\beta} - K^{1-\beta}}{1-\beta},\ \zeta = \frac{\nu}{a}\frac{f - K}{(fK)^{\beta/2}}$$

where β is the coefficient in the local volatility model.

However, Obloj made the correction to $I^0(x)|_{\beta<1}$ such that

$$I^0(x)|_{\beta<1} = \nu x / \ln\left(\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho}\right),\ z = \frac{\nu}{\alpha}\frac{f^{1-\beta} - K^{1-\beta}}{1-\beta}$$

We adopt Obloj (2008)'s correction in our implementation. Finally,

$$I^1(x) = \frac{(\beta-1)^2}{24}\frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4}\frac{\rho\nu\alpha\beta}{(fK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\nu^2.$$

Although other stochastic volatility models have been proposed[1], SABR is the simplest stochastic volatility model that is homogeneous in F and α. Although the formula appears complicated, it can be easily programmed due to its accurate closed-form approximations. We will show that the SABR model can fit the market volatility curve well by looking at the pricing and more importantly it captures the

[1] See for example, Hull and White (1987), Heston (1993), Lewis (2000), Fouque et. al. (2000).

time-series dynamics of the volatility curve by looking at hedging. As a result, it becomes an attractive tool to manage the smile risk.

We first bootstrap the cap flat volatilities in conjunction with linear interpolation[2] to obtain caplet spot volatilities. We implement several versions of the interest rate SABR model and analyze their respective pricing performance. Both in-sample and out-of-sample tests show that the SABR model exhibits excellent pricing accuracy. Moreover, delta hedging indicates that the SABR model generates accurate hedge ratio that out-performs the hedge ratio given by the Black's model.

2. Data

Our dataset are obtained from ICAP, a major inter-dealer broker in fixed income and derivatives. It consists of daily cap flat Black implied volatilities on 3 month LIBOR from 9/30/2004 to 9/29/2006. These caps have tenors of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 years. Their strikes are 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08. We use the data from the first half of our sample period (from 9/30/2004 to 9/29/2005) to calibrate the SABR model and in-sample pricing. The data from the second half of the sample is used for out-of-sample pricing and hedging performance evaluation. We also obtained daily 3, 6, 9 month LIBOR rates and swap rates with tenors of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 years for the same time period.

3. Calibration

We follow common practice among interest rate derivatives traders to fix β at 0.5. This convention is analogous to the CIR model that guarantees non-negativity for the

[2] We also tried other interpolation method and found that linear interpolation produces the most reasonable results.

forward interest rate, level-dependent volatility and tractability. The quantities we then need to estimate are α, ν and ρ. To estimate these parameters, we need spot Black implied volatilities for each caplet. A cap is a basket of options with different maturities but with the same strike rate. For simplicity, the market quotes cap prices in terms of a single number, the flat volatility. This is the single volatility which, when substituted into the Black's formula (for all caplets), reproduces the market price of the instrument. Clearly, flat volatility is a dubious concept: since a single caplet may be part of different caps it gets assigned different flat volatilities. The process of constructing actual implied caplet volatility from market quotes is called stripping cap volatility. The result of stripping is a sequence of caplet volatilities for all maturities ranging from 3 month to, say, 10 years. Convenient benchmarks are 3-month, 6-month, 9-month… every 3 months up to 10 years. Following standard market practice, we use bootstrap in combination with linear spline. One starts at the short end and moves further trying to match the prices of caps.

Once we have the spot volatilities, we can calculate the caplet price using the SABR model. The objective function is then constructed as the sum of the vega-weighted squared differences between SABR implied spot volatilities and Black implied spot volatilities, as the importance of the volatility fit is proportional to the vega:

$$ObjectiveFunction(\alpha,\nu,\rho)=\sum_{i=1}^{N} Vega_i(BlackSpotVol_i - SABRSpotVol_i)^2$$

where N is the total number of caplets. We then find the parameters by minimizing this objective function over the sample period used for estimation.

4. Pricing Performance Analysis

In this section, we examine the pricing performance of different versions of the SABR model. The goal is to examine the economic significance of allowing various variables to be free variables. Among the specifications we analyze, v and ρ are fixed to be both constant, only one of them to be constant, or both of them to be free variables. For example, for the constant v and ρ model, we calibrate α, v and ρ on the first day. On the following days, we fix v and ρ, only calibrate the model with the market cap prices to get a different α. On the other hand, we recalibrate the model daily to get different α, v and ρ for the fully stochastic model. In total, we obtain four versions of the SABR model: stochastic model, constant v and ρ model, constant v model, and constant ρ model. We examine their respective pricing performance by various criteria as follows.

4.1 RMSE

4.1.1 In-Sample Pricing Analysis

We calculate and plot the time series of the root mean square errors (RMSEs) of the four models over our first half of the sample period. The RMSE at time t is calculated as the square root of the average of square of percentage pricing errors over all maturities and all strikes, or mathematically, $\mathrm{RMSE}_t = \sqrt{\frac{\sum_{i=tenor} \sum_{j=strike} \varepsilon_{ij}^2(t)}{MN}}$, where ε_{ij} is the percentage difference of the SABR implied Black volatility and the market implied Black's volatility, and M is the total number of tenors, N is the total number of strikes, MN is the total number of caplets.

We plot the time series of the root mean square errors (RMSEs) of the four models, namely the stochastic model and the constant ν and/or ρ model over the out-of-sample period. RMSEs are used instead of sum of squared errors (SSEs) because the former provides a more direct measure of average percentage pricing errors of caplets with different maturities. Figure 1 shows the RMSEs for in-sample pricing of caplets:

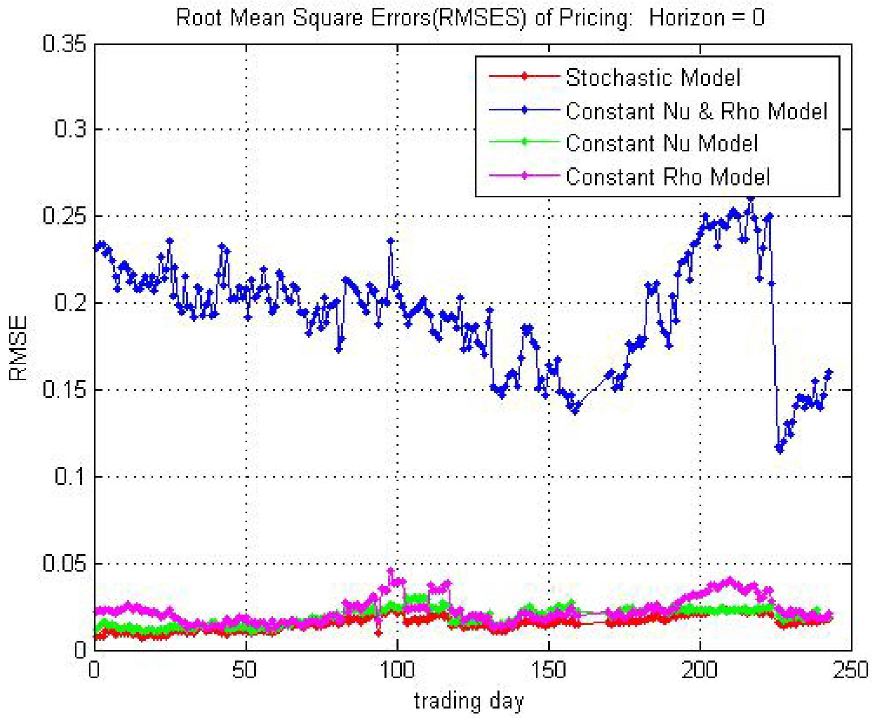

Figure 1. In-sample RMSEs for caplets.

We see from the graph that the constant ν and ρ model has significantly larger pricing errors, while the other three models have similar performances. We shall explore their pricing accuracies in more detail in the out-of-sample period and by examining in-sample average percentage pricing errors.

The in-sample pricing results for caps and diffcaps (defined as the difference between two adjacent maturity caps) are also shown in Figure 2 and Figure 3.

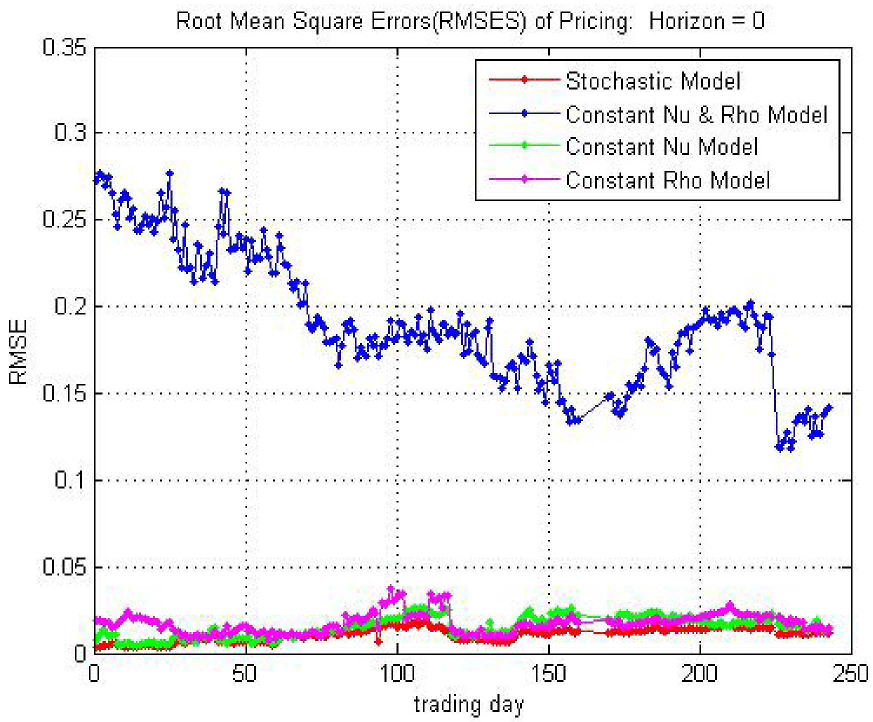

Figure 2. In-sample RMSEs for caps

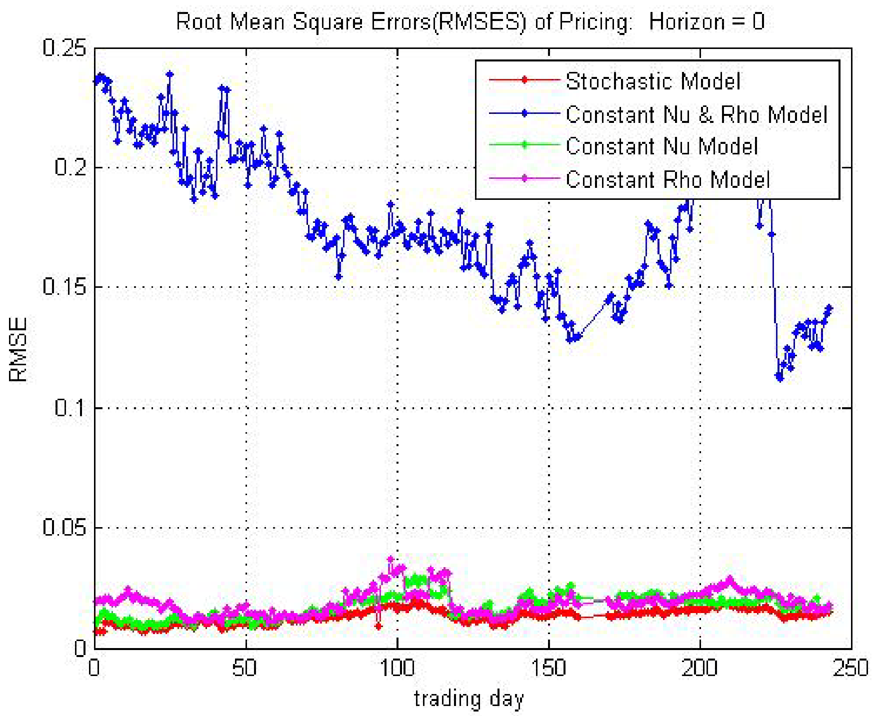

Figure 3. In-sample RMSEs for diffcaps.

4.1.2 Out-of-Sample Pricing Analysis

For out-of-sample pricing analysis, we plot the time series of the root mean square errors (RMSEs) of the four models over our second half of the sample period in Figure 4a and 4b. (This is to be consistent with the constant v and ρ, constant v or constant ρ models, which used the first half sample for calibration). Horizon represents the time interval between model calibration and the time of pricing. For example, when horizon is 1 day, we use yesterday's calibration result (α, v, ρ) and today's term structure to price today's caplet, then compare it to the price observed at market, and the error is the difference.

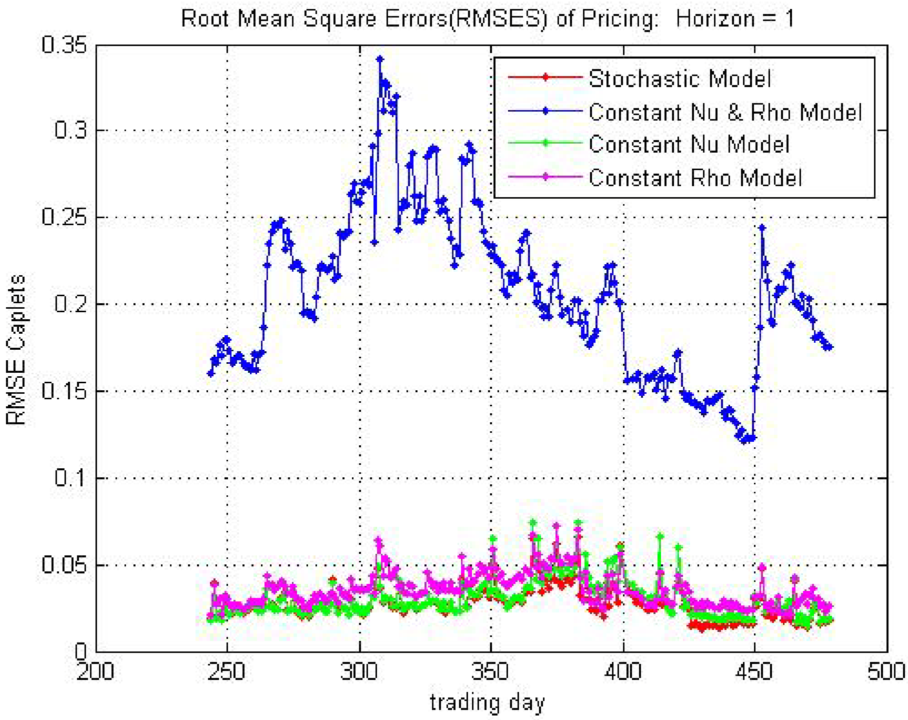

Figure 4a. Out-of-sample RMSEs for caplets (horizon=1 day).

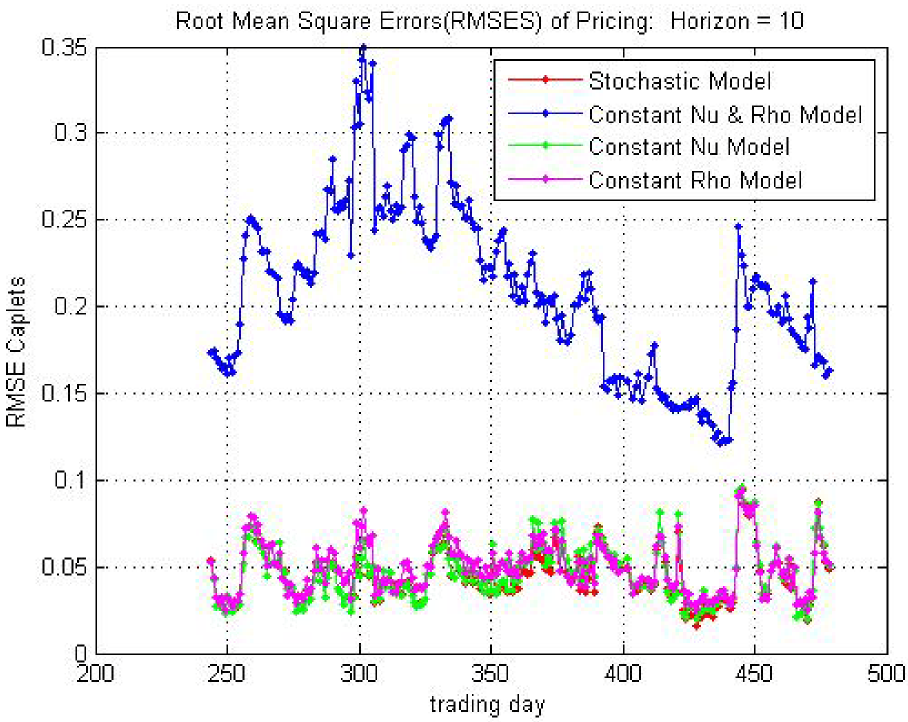

Figure 4b. Out-of-sample RMSEs for caplets (horizon=10 days).

In the out-of-sample test, all models except the constant v and ρ model exhibit RMSEs less than 0.1, which are significantly smaller than the results (up to 0.3) of the stochastic model reported in Jarrow, Li and Zhao (2007). These figures show that the stochastic model produces the minimum pricing errors among all four models. Moreover, the pricing errors of all models have a similar trend. This result is more obvious when the horizon is set to 1 day. Restricting both v and ρ to be constant seem to produce larger pricing errors than the rest of the models. On the other hand, the performance of the constant v or constant ρ model (only one of v and ρ is restricted to be constant) are quite close to that of the stochastic model. The out-of-sample RMSE analysis for diffcaps as plotted in Figure 5a and 5b show similar results of what we obtained for caplets.

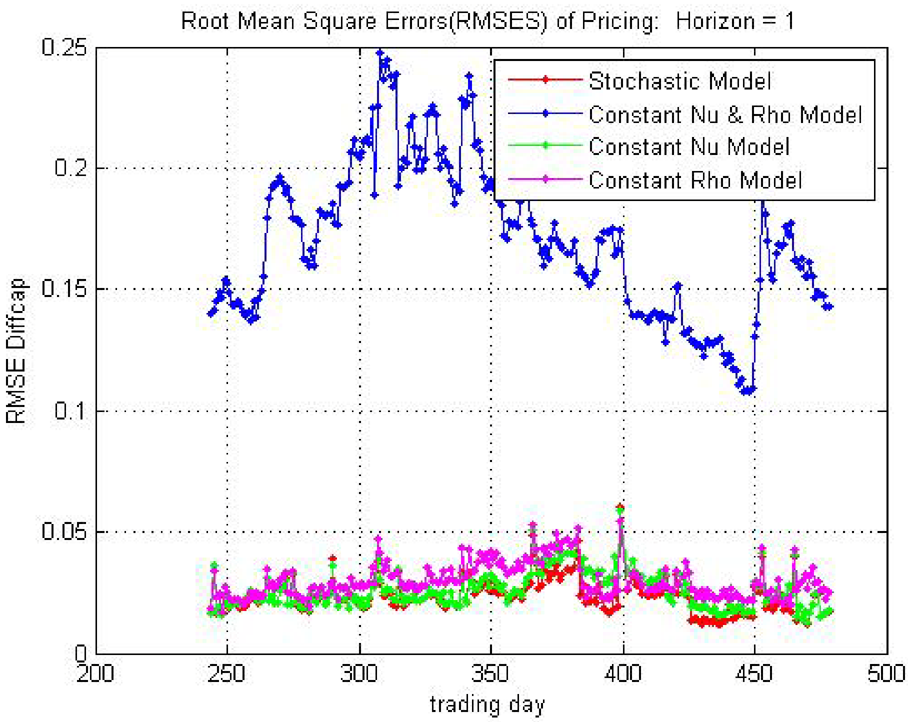

Figure 5a. Out-of-sample RMSEs for diffcaps (horizon=1 day).

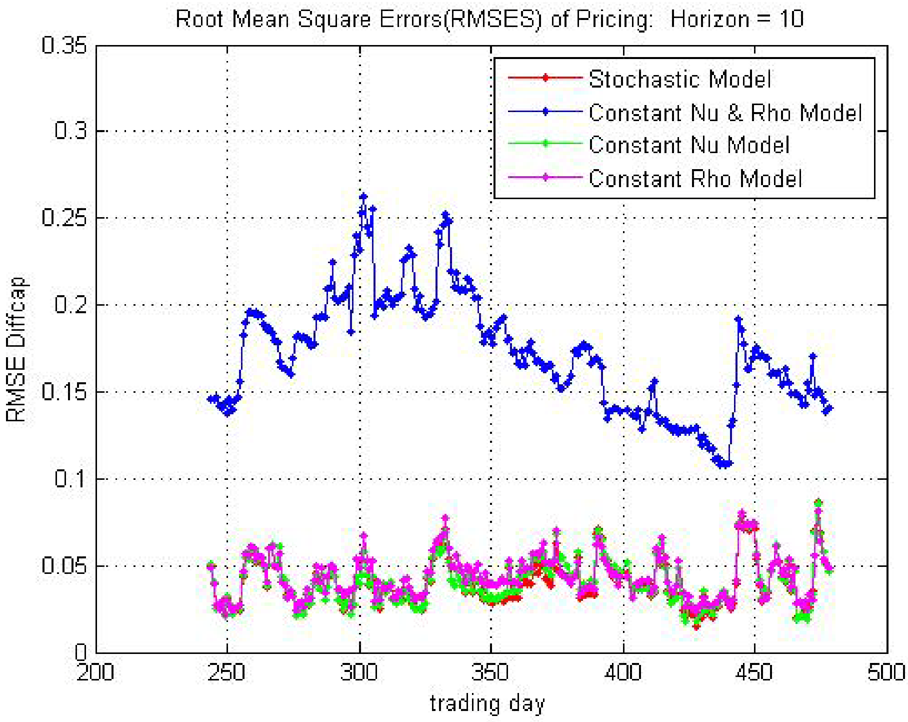

Figure 5b. Out-of-sample RMSEs for diffcaps (horizon=10 days).

4.2 Average Percentage Pricing Error Analysis

4.2.1 In-sample Average Percentage Pricing Error Analysis

In this section, we calculate in-sample average percentage pricing errors (in absolute value) over the first half period for caplets of various strikes and tenors. We show the results in the Table 1. Consistent with the RMSE analysis, the stochastic SABR model produce slightly smaller average percentage pricing errors than the constant ν model or the constant ρ model, while the constant ν and constant ρ model produces much larger average percentage pricing errors. For all models, the average percentage pricing errors are relatively larger for short-dated caps than those for longer-dated ones possibly due to the smaller price premiums for the short-dated caps. Within a particular maturity, there is no consistent pattern for the pricing errors across caps with different strikes. (Note that the stochastic model may have slightly higher pricing errors at a particular strike than the more constrained model as the optimization is over all the strikes.)

4.2.2 Out-of-Sample Average Percentage Pricing Error Analysis

In this section, we calculate out-of-sample average percentage pricing errors over the second half of the sample period for various strikes and tenors. We also perform the analysis for different horizons in the out-of-sample analysis just as what we did in the RMSE analysis.

Tables 2a and 2b show the average percentage pricing errors over the second half of our sample period for a given maturity and given strike for the four models, with horizon respectively set to 1 day and 10 days respectively.

Table 1. In-sample average percentage pricing errors.

Panel A: Average Percentage Pricing Errors of the Stochastic Model

STRIKE	1Yr	2Yr	3Yr	4Yr	5Yr	6Yr	7Yr	8Yr	9Yr	10Yr
0.02	0.014	0.008	0.005	0.005	0.004	0.005	0.005	0.004	0.004	0.004
0.03	0.039	0.015	0.010	0.011	0.011	0.011	0.011	0.011	0.010	0.010
0.04	0.015	0.007	0.007	0.007	0.006	0.005	0.004	0.004	0.005	0.005
0.05	0.012	0.008	0.006	0.007	0.008	0.010	0.010	0.009	0.009	0.009
0.06	0.022	0.007	0.007	0.006	0.005	0.004	0.003	0.004	0.004	0.004
0.07	0.009	0.004	0.006	0.006	0.006	0.006	0.005	0.005	0.006	0.006
0.08	0.024	0.006	0.007	0.006	0.004	0.003	0.003	0.004	0.004	0.005

Panel B: Average Percentage Pricing Errors of the Constant ν and ρ Model

STRIKE	1Yr	2Yr	3Yr	4Yr	5Yr	6Yr	7Yr	8Yr	9Yr	10Yr
0.02	0.159	0.117	0.152	0.157	0.151	0.142	0.130	0.114	0.099	0.082
0.03	0.286	0.082	0.150	0.178	0.181	0.177	0.167	0.154	0.140	0.123
0.04	0.294	0.127	0.149	0.176	0.183	0.181	0.174	0.161	0.148	0.131
0.05	0.247	0.151	0.194	0.209	0.208	0.200	0.188	0.174	0.158	0.141
0.06	0.210	0.178	0.223	0.230	0.223	0.213	0.199	0.183	0.165	0.146
0.07	0.212	0.193	0.220	0.222	0.213	0.202	0.187	0.170	0.152	0.131
0.08	0.228	0.195	0.199	0.200	0.190	0.179	0.164	0.147	0.128	0.106

Panel C: Average Percentage Pricing Errors of the Constant ν Model

STRIKE	1Yr	2Yr	3Yr	4Yr	5Yr	6Yr	7Yr	8Yr	9Yr	10Yr
0.02	0.019	0.020	0.017	0.017	0.015	0.014	0.014	0.013	0.013	0.012
0.03	0.039	0.017	0.013	0.014	0.014	0.013	0.013	0.013	0.012	0.012
0.04	0.023	0.017	0.016	0.014	0.012	0.010	0.009	0.009	0.009	0.009
0.05	0.010	0.012	0.009	0.008	0.006	0.005	0.005	0.005	0.005	0.005
0.06	0.022	0.006	0.009	0.008	0.008	0.007	0.007	0.007	0.007	0.008
0.07	0.011	0.008	0.006	0.006	0.005	0.004	0.003	0.004	0.004	0.004
0.08	0.027	0.018	0.019	0.018	0.016	0.015	0.014	0.014	0.014	0.015

Panel D: Average Percentage Pricing Errors of the Constant ρ Model

STRIKE	1Yr	2Yr	3Yr	4Yr	5Yr	6Yr	7Yr	8Yr	9Yr	10Yr
0.02	0.031	0.012	0.009	0.009	0.008	0.007	0.007	0.006	0.005	0.005
0.03	0.030	0.013	0.010	0.010	0.009	0.008	0.008	0.008	0.008	0.007
0.04	0.024	0.009	0.009	0.010	0.010	0.009	0.008	0.008	0.007	0.008
0.05	0.017	0.011	0.009	0.010	0.011	0.012	0.012	0.012	0.012	0.011
0.06	0.019	0.007	0.008	0.007	0.006	0.005	0.004	0.004	0.004	0.004
0.07	0.028	0.011	0.014	0.015	0.014	0.013	0.011	0.011	0.010	0.010
0.08	0.052	0.021	0.024	0.023	0.020	0.018	0.016	0.015	0.014	0.013

Table 2a. Out-of-sample average percentage pricing errors (horizon = 1 day).

STRIKE	1Yr	2Yr	3Yr	4Yr	5Yr	6Yr	7Yr	8Yr	9Yr	10Yr
Panel A: Average Percentage Pricing Errors of the Stochastic Model										
0.02	0.031	0.023	0.016	0.014	0.013	0.012	0.011	0.011	0.011	0.012
0.03	0.049	0.038	0.024	0.021	0.019	0.017	0.016	0.015	0.015	0.016
0.04	0.044	0.015	0.015	0.014	0.013	0.012	0.011	0.011	0.010	0.010
0.05	0.038	0.019	0.014	0.013	0.012	0.011	0.011	0.010	0.010	0.009
0.06	0.031	0.018	0.014	0.013	0.013	0.013	0.012	0.012	0.012	0.012
0.07	0.021	0.018	0.014	0.012	0.011	0.010	0.010	0.010	0.010	0.009
0.08	0.032	0.026	0.019	0.017	0.015	0.014	0.013	0.013	0.013	0.013
Panel B: Average Percentage Pricing Errors of the Constant ν and ρ Model										
0.02	0.082	0.071	0.056	0.063	0.061	0.059	0.055	0.057	0.056	0.056
0.03	0.061	0.042	0.034	0.037	0.037	0.035	0.032	0.033	0.031	0.030
0.04	0.035	0.031	0.035	0.037	0.036	0.035	0.033	0.033	0.032	0.032
0.05	0.037	0.018	0.013	0.012	0.012	0.011	0.011	0.012	0.012	0.013
0.06	0.043	0.015	0.015	0.013	0.013	0.012	0.012	0.011	0.011	0.011
0.07	0.081	0.041	0.032	0.025	0.020	0.018	0.017	0.017	0.016	0.015
0.08	0.113	0.046	0.034	0.025	0.021	0.019	0.018	0.017	0.016	0.015
Panel C: Average Percentage Pricing Errors of the Constant ν Model										
0.02	0.050	0.029	0.020	0.020	0.019	0.018	0.018	0.018	0.019	0.019
0.03	0.047	0.038	0.024	0.021	0.019	0.017	0.016	0.016	0.015	0.016
0.04	0.048	0.018	0.015	0.014	0.014	0.014	0.013	0.013	0.012	0.012
0.05	0.040	0.022	0.017	0.019	0.020	0.020	0.020	0.020	0.019	0.019
0.06	0.021	0.018	0.013	0.012	0.011	0.011	0.011	0.011	0.011	0.011
0.07	0.022	0.018	0.015	0.013	0.012	0.011	0.011	0.011	0.011	0.010
0.08	0.041	0.028	0.019	0.020	0.018	0.017	0.016	0.016	0.016	0.016
Panel D: Average Percentage Pricing Errors of the Constant ρ Model										
0.02	0.034	0.024	0.017	0.015	0.014	0.012	0.012	0.012	0.012	0.012
0.03	0.056	0.033	0.019	0.017	0.016	0.014	0.014	0.013	0.013	0.013
0.04	0.033	0.029	0.032	0.030	0.027	0.025	0.023	0.021	0.020	0.019
0.05	0.044	0.024	0.015	0.013	0.011	0.011	0.011	0.011	0.012	0.012
0.06	0.025	0.017	0.012	0.011	0.011	0.011	0.011	0.011	0.011	0.011
0.07	0.063	0.031	0.033	0.030	0.025	0.022	0.020	0.020	0.019	0.018
0.08	0.104	0.052	0.053	0.050	0.045	0.041	0.037	0.037	0.035	0.033

Table 2b. Out-of-sample average percentage pricing errors (horizon = 10 days).

STRIKE	1Yr	2Yr	3Yr	4Yr	5Yr	6Yr	7Yr	8Yr	9Yr	10Yr
Panel A: Average Percentage Pricing Errors of the Stochastic Model										
0.02	0.060	0.048	0.035	0.033	0.032	0.030	0.029	0.028	0.029	0.028
0.03	0.065	0.050	0.036	0.033	0.030	0.028	0.026	0.025	0.025	0.026
0.04	0.068	0.032	0.027	0.025	0.024	0.022	0.021	0.020	0.020	0.020
0.05	0.059	0.032	0.026	0.025	0.024	0.024	0.023	0.023	0.023	0.022
0.06	0.065	0.030	0.023	0.022	0.021	0.020	0.020	0.019	0.019	0.019
0.07	0.061	0.033	0.025	0.024	0.023	0.022	0.021	0.021	0.021	0.021
0.08	0.071	0.044	0.031	0.030	0.027	0.026	0.025	0.025	0.025	0.024
Panel B: Average Percentage Pricing Errors of the Constant ν and ρ Model										
0.02	0.086	0.075	0.058	0.064	0.062	0.060	0.057	0.058	0.057	0.057
0.03	0.074	0.052	0.042	0.042	0.040	0.037	0.035	0.034	0.033	0.031
0.04	0.062	0.040	0.040	0.040	0.039	0.037	0.035	0.034	0.033	0.033
0.05	0.060	0.032	0.026	0.025	0.024	0.024	0.024	0.024	0.024	0.024
0.06	0.074	0.030	0.024	0.023	0.022	0.021	0.020	0.020	0.020	0.019
0.07	0.098	0.047	0.039	0.034	0.030	0.027	0.026	0.026	0.026	0.025
0.08	0.129	0.052	0.040	0.034	0.031	0.028	0.027	0.027	0.026	0.025
Panel C: Average Percentage Pricing Errors of the Constant ν Model										
0.02	0.074	0.049	0.034	0.033	0.032	0.031	0.029	0.029	0.030	0.030
0.03	0.064	0.050	0.036	0.033	0.030	0.028	0.026	0.025	0.025	0.026
0.04	0.070	0.033	0.029	0.028	0.027	0.026	0.024	0.024	0.023	0.023
0.05	0.061	0.031	0.028	0.030	0.031	0.031	0.031	0.031	0.030	0.030
0.06	0.057	0.029	0.022	0.022	0.021	0.021	0.021	0.021	0.021	0.021
0.07	0.061	0.033	0.025	0.024	0.023	0.022	0.021	0.021	0.021	0.021
0.08	0.075	0.045	0.032	0.031	0.028	0.026	0.026	0.025	0.025	0.025
Panel D: Average Percentage Pricing Errors of the Constant ρ Model										
0.02	0.061	0.049	0.035	0.033	0.032	0.030	0.029	0.028	0.029	0.028
0.03	0.070	0.046	0.032	0.030	0.028	0.025	0.024	0.023	0.023	0.023
0.04	0.062	0.038	0.036	0.034	0.031	0.029	0.027	0.025	0.024	0.024
0.05	0.061	0.035	0.026	0.024	0.023	0.022	0.022	0.022	0.022	0.023
0.06	0.062	0.029	0.022	0.021	0.020	0.020	0.019	0.019	0.019	0.019
0.07	0.084	0.039	0.040	0.037	0.033	0.031	0.029	0.028	0.028	0.027
0.08	0.122	0.056	0.057	0.055	0.050	0.046	0.042	0.042	0.040	0.038

The result of the out-of-sample average absolute percentage pricing errors is consistent with the result of out-of-sample RMSE analysis as well as in-sample percentage pricing analysis. The stochastic SABR model produces slightly smaller average percentage pricing errors than the constant v model or the constant ρ model, while the constant v and constant ρ model produces much larger average percentage pricing errors. For all models, the average percentage pricing errors are relatively larger for short-dated caps than those for longer-dated ones due to smaller price premium for the short-dated caps. Within a particular maturity, there is no consistent pattern for the pricing errors across caps with different strikes.

4.3 Test Using Diebold-Mariano Statistics

To provide more rigorous statistical tests comparing two models, we compute the Diebold-Mariano Statistics, defined as the following:

$$S = \frac{\overline{d}}{\sqrt{f_d(0)/N}}$$

where

$$\overline{d} = \frac{1}{N}\sum_{t=1}^{N}[\varepsilon_1(t) - \varepsilon_2(t)]$$

$$f_d(0) = \sum_{q=-\infty}^{q=\infty} \gamma_d(q)$$

$$\gamma_d(q) = Cov(d_t, d_{t-q})$$

where $d_t = \varepsilon_1(t) - \varepsilon_2(t)$.

$\varepsilon_1(t)$ and $\varepsilon_2(t)$ are the sums of the squares of percentage pricing errors calculated based on Model 1 and Model 2 over all maturities and all strikes at time *t*. *N* is the number of observation points in the time-series dimension (such as the number of days). It allows us to test if the outperformance of Model 1 over the Model 2 is statistically significant. The results are shown in Table 3:

Table 3. Diebold-Mariano Statistic using caplets.

Horizon	Constant ν vs Stochastic		Constant ρ vs Stochastic		Constant ν and ρ vs Stochastic	
	Mean Difference	D-M statistic	Mean Difference	D-M statistic	Mean Difference	D-M statistic
1 day	.12172	14.380	.05408	4.6609	.13969	9.8494
5 day	.12181	14.214	.04936	4.3180	.13024	7.4632
10 day	.12158	13.841	.04349	3.7418	.11677	5.5095

Horizon	Constant ν and ρ vs Constant ν		Constant ρ vs Constant ν		Constant ν and ρ vs Constant ρ	
	Mean Difference	D-M statistic	Mean Difference	D-M statistic	Mean Difference	D-M statistic
1 day	.12118	14.253	.08562	4.5524	.12032	14.371
5 day	.12131	14.094	.08088	3.8370	.12051	14.252
10 day	.12115	13.739	.07328	3.0844	.12042	13.944

We calculate the Diebold-Mariano statistics of squared percentage pricing errors of caplets for all six pairs of the four models. All the D-M statistics are significant. The pricing accuracy rank from high to low is the stochastic model, the constant ν model, the constant ρ model, and the constant ν and constant ρ model.

Table 4. Diebold-Mariano Statistic using diffcaps.

Horizon	Constant ν vs Stochastic		Constant ρ vs Stochastic		Constant ν and ρ vs Stochastic	
	Mean Difference	D-M statistic	Mean Difference	D-M statistic	Mean Difference	D-M statistic
1 day	.020579	17.433	.00914	4.7040	.02439	9.4503
5 day	.020503	16.937	.00812	4.4169	.02231	7.3261
10 day	.020396	16.241	.00678	3.5163	.01928	5.4402

Horizon	Constant ν and ρ vs Constant ν		Constant ρ vs Constant ν		Constant ν and ρ vs Constant ρ	
	Mean Difference	D-M statistic	Mean Difference	D-M statistic	Mean Difference	D-M statistic
1 day	.020488	17.268	.01525	4.9846	.020336	17.353
5 day	.020422	16.795	.01419	4.2748	.020280	16.925
10 day	.020329	16.130	.01250	3.4996	.020204	16.324

We also present the D-M statistics for diffcaps in Table 4 that shows similar results.

5. Hedging Performance Analysis

In this section, we calculate the Hedging Variance Ratios (HVR) to compare the delta hedging performance between the Black model and the SABR model. The definition of HVR is defined by the following (assuming the hedging horizon is *m* days): Assuming on day 0 (t_0), we delta hedge the derivative (caplet in our analysis) using the hedging instrument. Following Li and Zhao (2006), we use forward rate as hedging instrument, which has the same tenor as the caplet. For example, if the reset and payment date for the caplet are in 6 months and 9 months respectively, we use the forward rate between the 6th month and the 9th month as the hedging instrument:

$$Caplet(t_0) = \alpha + \Delta_0 FwdRate(t_0)$$

On day 2, we have the following:

$$Caplet(t_1) = \alpha + \Delta_0 FwdRate(t_1) + \varepsilon_{t_1}$$

where ε_{t_1} is the hedging error after one day, and can be calculated as:

$$\varepsilon_{t_1} = Caplet(t_1) - Caplet(t_0) - \Delta_0[FwdRate(t_1) - FwdRate(t_0)]$$

We rebalance the delta daily to comply with common practice, and accumulate the hedging error up to the hedging horizon. Thus, the accumulated hedging error after *m* days is

$$Hedging\_Error = \sum_{k=1}^{m} Caplet(t_k) - Caplet(t_{k-1}) - \Delta_{k-1}[FwdRate(t_k) - FwdRate(t_{k-1})]$$

We will have time series of accumulated hedging error and the corresponding un-hedged derivative premium (caplet price). The HVR is then calculated as the following:

$$HVR = 1 - \frac{Var(HedgingError)}{Var(Caplet\,\mathrm{Pr}\,ice)}$$

The higher the HVR, the better the hedging performance.

5.1 SABR Delta vs Black's Delta

Delta hedging by the SABR model is different from that by the Black model in the way delta is calculated. For Black's model, delta is calculated by shifting the underlying (forward rate) while keeping other parameters constant:

$\Delta_{Black} = \frac{\partial B}{\partial f}$; B is the market price of caplet, and f is the forward rate.

In the SABR model, since the implied volatility (σ) is a function of forward rate and the volatility of the forward rate α, which is correlated with forward rate. Therefore the SABR delta is given by[3]:

$$\Delta_{newSABR} = \frac{\partial B}{\partial f} + \frac{\partial B}{\partial \sigma}\left(\frac{\partial \sigma}{\partial f} + \frac{\partial \sigma}{\partial \alpha}\frac{\rho \upsilon}{f^{\beta}}\right)$$

In our analysis, we also compute $\Delta_{oldSABR}$, which was first given in Hagan et al (2002)'s original paper and ignored the correlation between α and forward rate:

$$\Delta_{oldSABR} = \frac{\partial B}{\partial f} + \frac{\partial B}{\partial \sigma}\frac{\partial \sigma}{\partial f}$$

Our purpose is to assess the importance for account for the terms related to the correlation between the forward rate and its volatility when calculating delta.

[3] See Bartlett (2006).

Once we calculate the delta, we follow the approach described in the previous section to calculate the hedging variance ratio and to compare the hedging performance of the Black model and the SABR model.

Black's delta for caplet is calculated by shifting the corresponding forward rate, and using the central difference method to approximate $\frac{\partial B}{\partial f}$. The SABR Delta for caplet involves three more partial derivatives:

$\frac{\partial B}{\partial \sigma}$: This is just the Black's Vega, and we calculated it by shifting the corresponding spot volatilities for the caplet and using the central difference approximation.

$\frac{\partial \sigma}{\partial f}$ *and* $\frac{\partial \sigma}{\partial \alpha}$: Both can be computed analytically as SABR implied volatility (σ) is an analytical function of forward rate and SABR parameter α.

5.2 Hedging Performance

Figure 6a and 6b plot the Hedging Variance Ratios for caplets of various maturities, with strikes from 0.02 to 0.08, and for hedging horizons of 1 day and 10 days respectively.

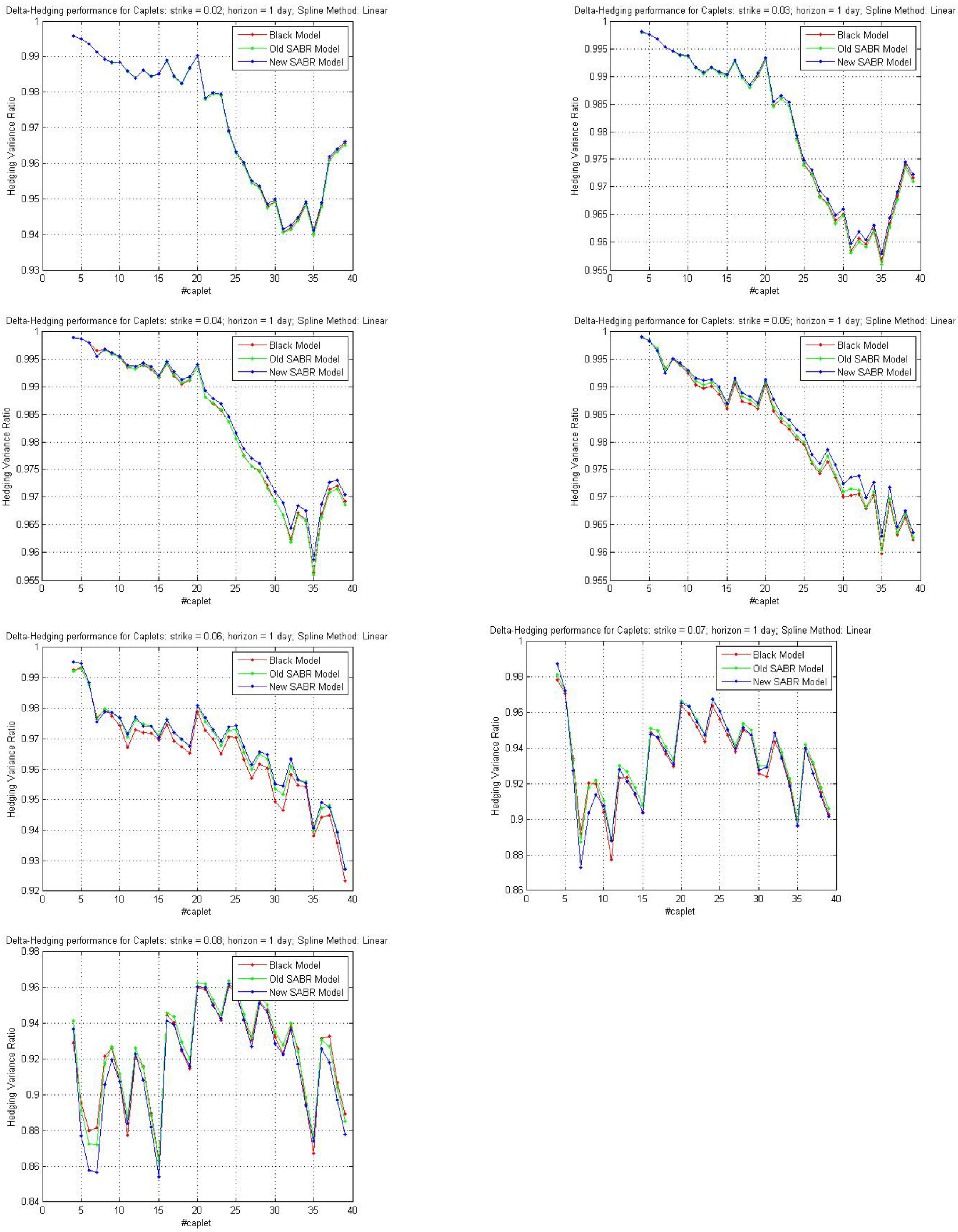

Figure 6a. Hedging Variance Ratio for caplets (horizon = 1 day).

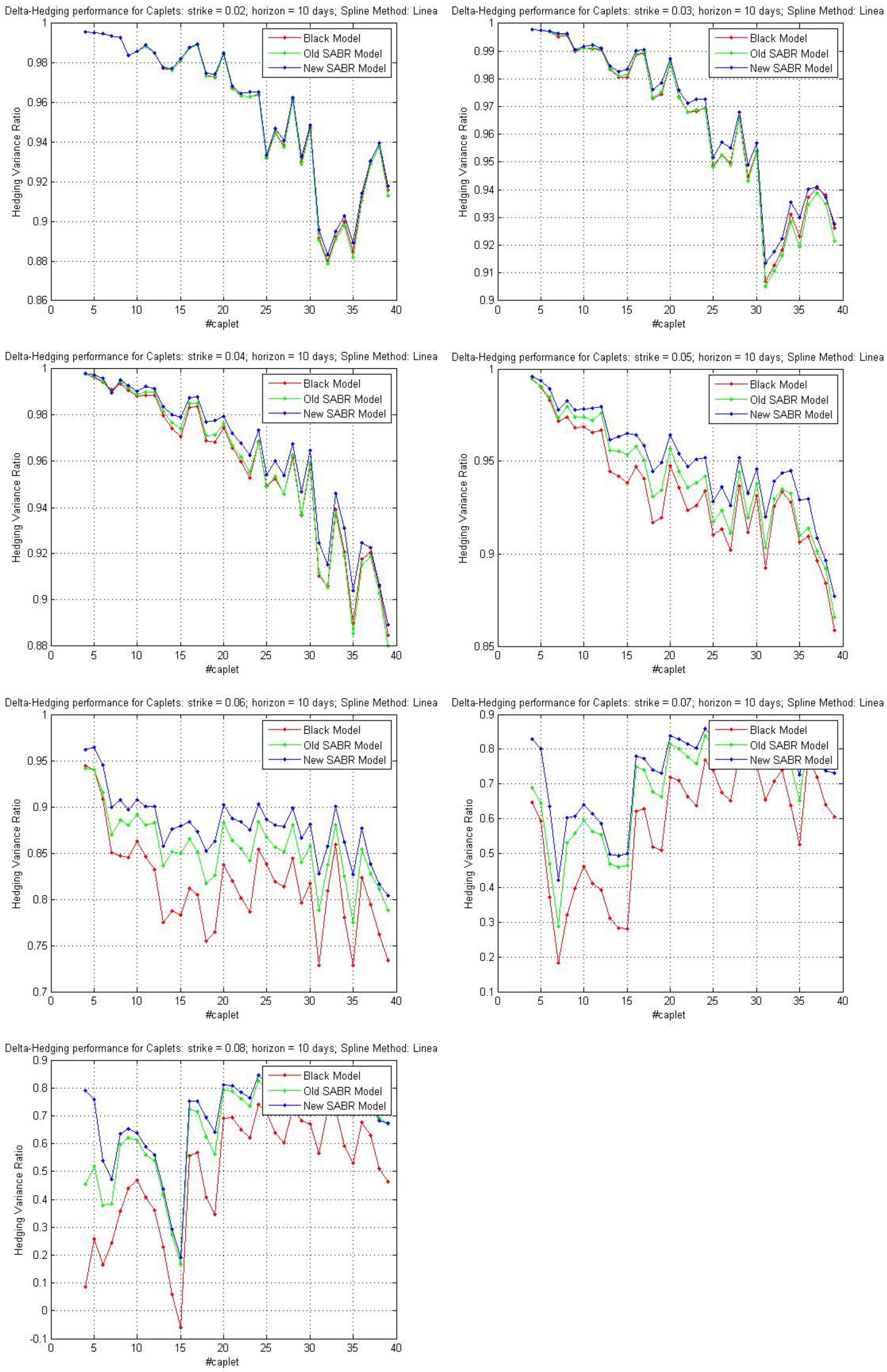

Figure 6b. Hedging Variance Ratio for caplets (horizon = 10 days).

Overall, we observe that delta hedging performances rank as follows: new SABR delta is better than old SABR delta, which in turn is better than Black's delta based on variance ratio analysis. In term of absolute hedging performance, all models are better at hedging close-to-the-money options than deep-in/out-of-the-money options. When options move deep in- or out-of-the-money, the hedging performances of all delta calculation methods are weakened.

6. Conclusion

We empirically investigate the pricing and hedging performance of a popular stochastic volatility model, the SABR model using 2 years of daily prices of interest rate caps of 7 different strikes and 10 different tenors. Success in both in-sample and out-of-sample pricing, as well as in its hedging performance indicates that the SABR model not only fits the implied volatility smile very well (accurate pricing), it also does a good job in capturing the dynamics of the smile over time (hedging). Moreover, to achieve accurate pricing both in-sample and out-of-sample, it is essential to recalibrate daily the volatility of the forward rate, and to recalibrate daily either the "vol of vol" or the correlation between forward rate and its volatility, although recalibrating all three further improves pricing performance.

REFERENCES

Bartlett, B., Hedging under SABR Model, 2006, *Wilmott Magazine*, July/August, 2-4.

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

Cox, J., J. Ingersoll and S. Ross, 1985, A Theory of the Term Structure of Interest rates, *Econometrica* 53, 385-403.

Dai, Q., and K. Singleton, 2003, Term Structure Dynamics in Theory and Reality, *Review of Financial Studies* 16, 631-678.

Diebold, F.X., and R.S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 253–263.

Driessen, J., P. Klaassen and B. Melenberg, 2003, The Performance of Multi-Factor Term Structure Models for Pricing and Hedging Caps and Swaptions, *Journal of Financial and Quantitative Analysis* 38, 635-672.

Duarte, J., 2008, The Causal Effect of Mortgage Refinancing on Interest Rate Volatility: Empirical Evidence and Theoretical Implications, *Review of Financial Studies* 21, 1689-1731.

Fouque, J., G. Papanicolaou and K. Sirclair, 2000, Derivatives in Financial Markets with Stochastic Volatility, Cambridge University Press.

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

Hagan, P., D. Kumar, A. Lesniewski, and D. Woodward, 2002, Managing Smile Risk, *Wilmott Magazine*.

Heidari, M., and L. Wu, 2003, Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates? *Journal of Fixed Income* 13, 75-86.

Heston, S., 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327–343.

Hull, J., and A. White, 1987, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42, 281–300.

Jagananthan, R., A. Kaplin, S. Sun, 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., Li, H., and F. Zhao, 2007, Interest Rate Caps Smile Too! But Can the LIBOR Market Models Capture It? *Journal of Finance* 62, 345-382.

Li, H., and F. Zhao, 2006, Unspanned Stochastic Volatility: Evidence from Hedging Interest Rate Derivatives, *Journal of Finance* 61, 341-378.

Lewis, A., 2000, Option Valuation Under Stochastic Volatility, Financial Press.

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

Obloj, J., 2008, Fine-tune Your Smile, Correction to Hagan et al. Working Paper.