North American Actuarial Journal

ISSN: 1092-0277 (Print) 2325-0453 (Online) Journal homepage: http://www.tandfonline.com/loi/uaaj20

A Long-Term Model of the Dynamics of the S&P500 Implied Volatility Surface

Martin le Roux FSA, FCIA

A Long-Term Model of the Dynamics of the S&P500 Implied Volatility Surface

Martin le Roux*

Abstract

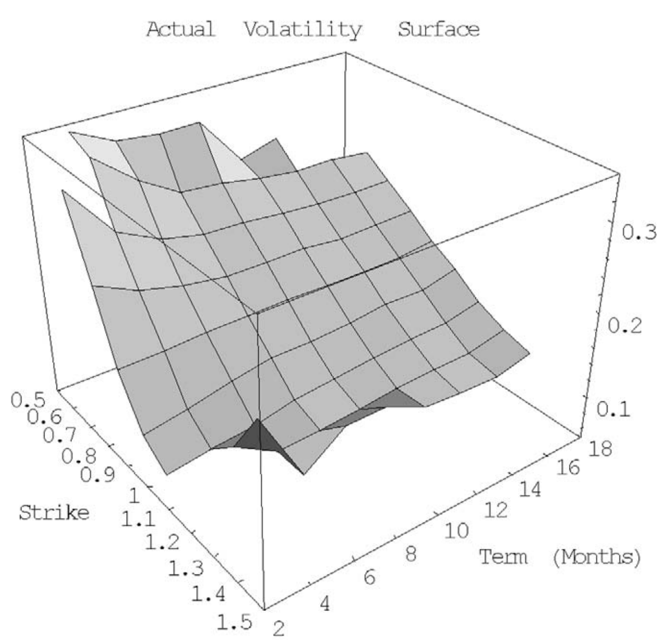
In this paper we present an econometric model of implied volatilities of S&P500 index options. First, we model the dynamics the CBOE VIX index as a proxy for the general level of implied volatilities. We then describe a parametric model of the implied volatility surface for options with a term of up to two years. We show that almost all of the variation in the implied volatility surface can be explained by the VIX index and one or two other uncorrelated factors. Finally, we present a model of the dynamics of these factors.

1. Introduction

Over the past decade North American life insurance companies have underwritten large volumes of annuity contracts with embedded guarantees tied to the performance of stock portfolios or equity indices, and actuaries have devoted considerable attention to modeling the risks inherent in these contracts. For example, see Hardy (2001) or the Life Capital Adequacy Subcommittee (2005). More recently some insurers have implemented derivative-based hedging strategies for managing these exposures. Modeling such a hedging strategy requires that one capture not only the dynamics of the underlying equity portfolio or index but also the dynamics of factors affecting the value of the insurer's hedge portfolio.

In this paper we focus on econometric modeling of one such set of factors, namely, the implied volatility surface for options on the S&P500 index. By "implied volatility surface" we mean implied volatility expressed as a function of strike price and term to maturity. According to the original Black-Scholes option-pricing model, all options on the same underlying should have the same implied volatility, regardless of strike or term, but it is well known that this is not the case in reality. Nevertheless, it is common practice to express option prices in terms of Black-Scholes implied volatility, or as it is sometimes described, "the wrong number to plug into the wrong formula to get the right price." Figure 1 shows a recent example of the implied volatility surface. The variation of volatility with respect to strike is known as the "smile," although recently the term "sneer" or "smirk" would be more descriptive. The gradient of the smile is known as the "skew."

Figure 1
S&P500 Implied Volatility Surface as of September 30, 2005



* Martin le Roux, FSA, FCIA, ING Investment Management, 1290 Broadway, Denver, CO 80203, martin.le.roux@inginvestment.com. The views expressed in this paper are the author's and not necessarily ING's.

The models in this paper are econometric models and are intended for risk management purposes. This is in contrast to models whose purpose is to describe the implied volatility surface on a given date in terms of a risk-neutral process for the underlying equity index, so as to be able to infer the latter's parameters by reference to observed values of the former. For a summary of recent work in this area see Gatheral (2006).

To analyze the dynamics of the implied volatility surface we used three sources of historical data. First, we used the Chicago Board Options Exchange (CBOE) VIX index, which dates back to January 1986. The VIX index can be thought of as a weighted average of implied volatilities for options with an average term to maturity of 30 days, across a range of strikes. The index does not provide information about the shape of the volatility surface, but it does provide information about the overall level of implied volatilities. The methodology for the VIX index was revised in 2003, and the index was recalculated retrospectively to January 1990. For 1986–89 we used the "old" VIX, or VXO, which is based on a narrower range of strikes and references options on the S&P100 index, not the S&P500. The VXO and VIX have tracked each other fairly closely since the latter's inception. We used VXO data mainly because we wanted to ensure that the 1987 crash was included in the model's calibration data. For details of the construction of the VIX index, see the Chicago Board Options Exchange (2003).

Second and third, we used implied volatilities of individual S&P500 option contracts as reported by OptionMetrics and iVolatility.com. OptionMetrics was used from January 1996 to November 2000, and iVolatility.com from November 2000 to November 2005. Both databases report end-of-day implied volatilities for all listed S&P500 option contracts. Each day's observations typically included about 200 usable data points, covering options with terms of up to about two years. We used these data to model the shape of the implied volatility surface, that is, smile and term structure effects. To be able to compare specific points on the volatility surface on successive dates, we followed the lead of other researchers such as Dumas, Fleming, and Whaley (1998) and Alentorn (2004), and fitted a smooth parametric surface to each set of observations before investigating how this surface fluctuates over time.

All the analyses described in this paper are based on weekly data. We chose weekly intervals to balance accuracy and practicality. On the one hand, a longer time interval (e.g., monthly) may miss important features of the data and may not be very accurate for modeling hedging strategies that involve dynamic rebalancing. On the other hand, actuarial applications typically focus on time horizons of a year or more, and a model based on a shorter time interval (e.g., daily) would be impractical in this context.

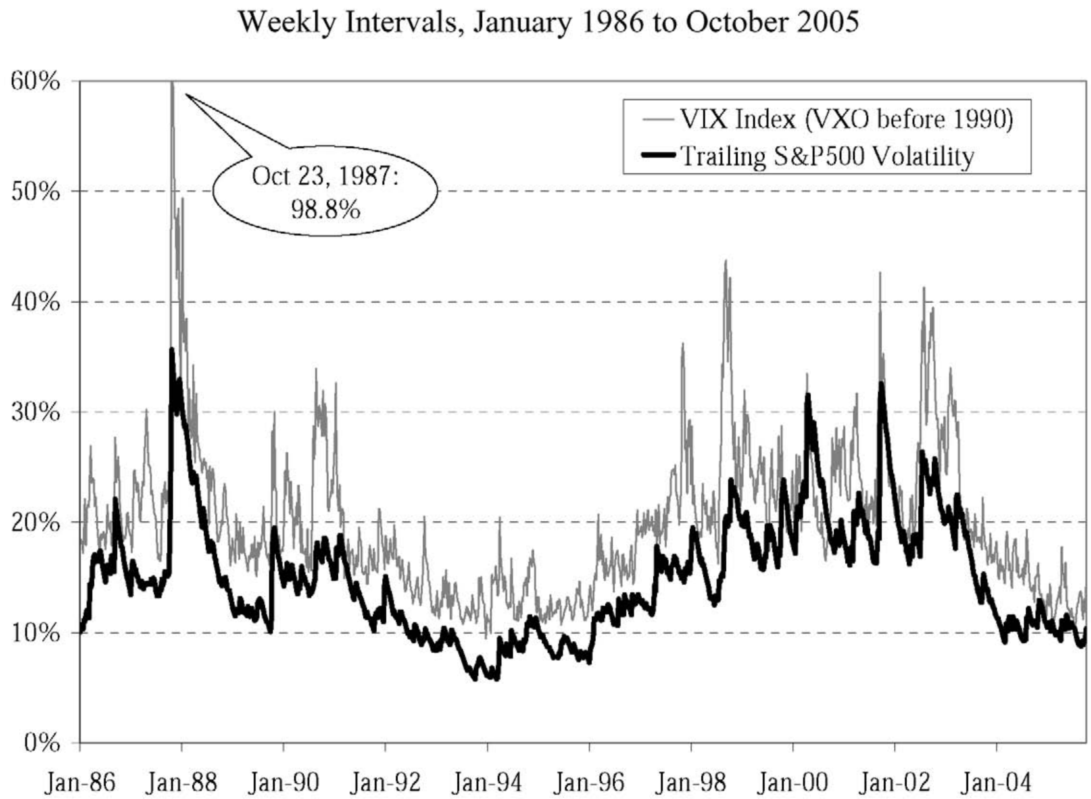
In principle, the methods described in this paper could be applied to other indices for which historical option pricing data are available, but the S&P500 is distinguished by the length of time for which options have been traded and the depth and liquidity of its options market. Volatility indices such as the VIX are available only for a few other markets, so some changes in methodology would be necessary for markets that lack such an index. The same methods could also be applied to long-dated over-the-counter options, but it is not clear how reliable the results would be. Some historical pricing data are available for over-the-counter-dated options, but typically these are brokers' "indicative" prices and not necessarily prices at which actual transactions took place.

The organization of this paper is as follows. In Section 2 we describe the qualitative features of the VIX index and its relationship to the S&P500 index, and we evaluate a variety of models designed to capture these features. In Section 3 we describe a parametric model of the implied volatility surface. In Section 4 we examine changes over time in the volatility surface. We show how these changes can be explained first by the VIX index and second by a small number of uncorrelated random factors, similar to the conclusions reached by Cont and Fonseca (2002). In Section 5 we present a model of the dynamics of these factors.

2. DYNAMICS OF THE VIX INDEX

Figure 2 is a chart of the VIX index at weekly intervals from January 1986 to October 2005. This also shows the trailing volatility of weekly S&P500 returns, measured as an exponentially weighted moving average (see below for details of

Figure 2
VIX Index and Trailing S&P500 Volatility



the formula). Several features are apparent from this figure:

- The VIX generally moves in line with the trailing volatility of the S&P500, but the VIX is usually higher. On average the difference is 5.18 percentage points, and the difference is positive about 95% of the time.
- The VIX appears to follow a mean-reverting process. About 75% of the time it has fluctuated between 12.5% and 27.5%, and deviations outside this range have been generally short-lived.
- Index values have a skewed distribution. On the downside, the VIX has almost never fallen below 10%, but on the upside the VIX has occasionally spiked to extreme levels (almost 100% on October 23, 1987).

Expanding on the last point, Figure 3 shows the unconditional distribution of the VIX and compares this to a lognormal distribution. As can be seen, a lognormal distribution provides a reasonable fit to the data. For this reason all the models presented below focus on the process for $\log(VIX_t)$. This also ensures that the modeled process for VIX_t is strictly positive.

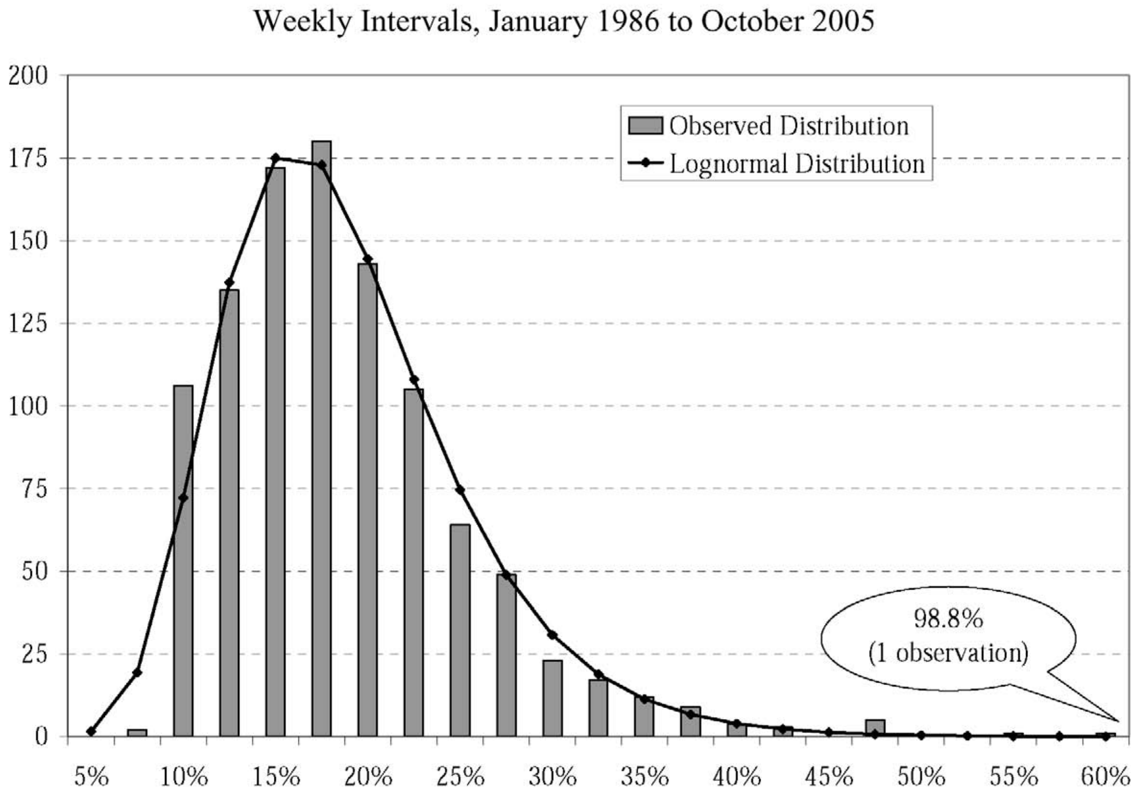
Another well-known phenomenon is that changes in the VIX are inversely related to S&P500 index returns. When the S&P500 falls the VIX generally rises, and vice versa. This is illustrated in Figure 4, which indicates an approximately linear relationship between weekly log returns on the S&P500 and weekly changes in $\log(VIX_t)$. It is also possible that this relationship may be asymmetric, that is, the VIX's response to negative S&P500 returns may be different from its response to positive returns, although this is not clear from the figure.

A general model that encompasses all of these observations is as follows:

$$\log(VIX_t) = \nu_0 + \nu_1 \log(VIX_{t-1}) + \nu_2 \log(\sigma_{SPX,t}) + \nu_3 y_t^- + \nu_4 y_t^+ + \sigma_{VIX}\varepsilon_{VIX,t}. \quad (2.1)$$

Here $\nu_0 \ldots \nu_4$ and σ_{VIX} are fitted parameters, VIX_t is the VIX index at time t, y_t is the S&P500 log return for the period $t-1$ to t, $y_t^- = \min[y_t, 0]$,

Figure 3
Observed Distribution of VIX versus Lognormal Distribution



$y_t^+ = \max[y_t, 0]$, $\varepsilon_{VIX,t}$ is a standard normal deviate, and $\sigma_{SPX,t}$ is an exponentially weighted moving average (EWMA) measure of trailing S&P500 volatility, determined using the recursive relationship

$$\sigma^2_{SPX,t} = \lambda\sigma^2_{SPX,t-1} + (1 - \lambda)52y_t^2,$$

where λ is another fitted parameter. The constant 52 arises because the returns y_t are measured in weekly time steps whereas VIX_t is measured annualized terms. The variable $\sigma_{SPX,t}$ is the same EWMA measure of trailing volatility as is shown in Figure 2.

We fitted this model to 1,031 weekly observations of the VIX index from January 10, 1986, to October 7, 2005, inclusive. (The initial value of σ_{SPX} was based on 20 prior weeks of S&P500 returns, from August 16, 1985 to January 10, 1986.) Using these data, maximum likelihood parameter estimates are as follows:

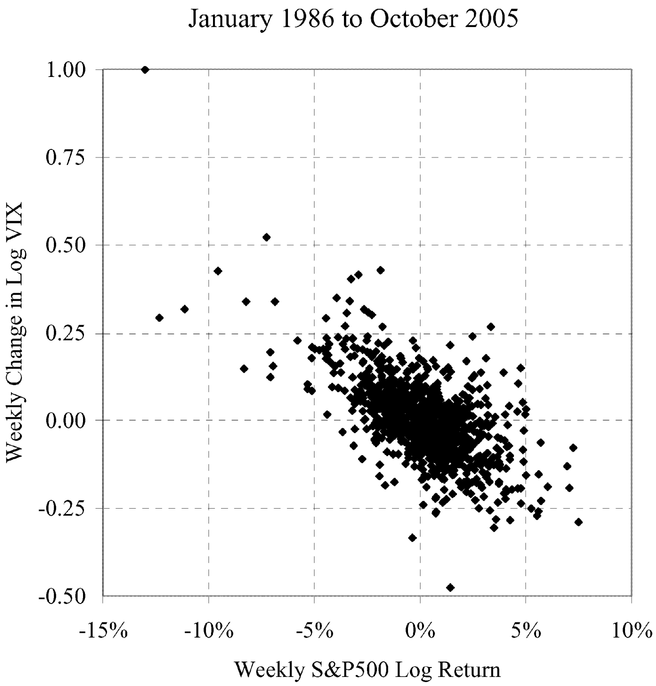
$$\begin{aligned} \nu_0 &= -0.1219 & \nu_1 &= 0.8208 \\ \nu_2 &= 0.0992 & \nu_3 &= -4.5968 \\ \nu_4 &= -1.5666 & \lambda &= 0.9249 \\ \sigma_{VIX} &= 0.0795. \end{aligned}$$

The fitted value of $\lambda = 0.9249$ means that the EWMA volatility measure $\sigma_{SPX,t}$ has a half-life of 8.9 weeks, since $0.9249^{8.9} = 0.50$.

It can be seen that S&P500 returns enter into the model both directly through the terms y_t^- and y_t^+ as well as indirectly through the term $\sigma_{SPX,t}$. To ensure that these terms are not redundant and that the model is not overfitted, we compared it to the following simpler alternatives:

$$\log(VIX_t) = \nu_0 + \nu_1 \log(VIX_{t-1}) + \nu_3 y_t + \sigma_{VIX}\varepsilon_{VIX,t}, \qquad (2.2)$$

Figure 4
Weekly Changes in Log VIX versus Weekly S&P500 Log Returns



$$\log(VIX_t) = \nu_0 + \nu_1 \log(VIX_{t-1}) + \nu_2 \log(\sigma_{SPX,t}) + \sigma_{VIX}\varepsilon_{VIX,t}, \quad (2.3)$$

$$\log(VIX_t) = \nu_0 + \nu_1 \log(VIX_{t-1}) + \nu_3 y_t^- + \nu_4 y_t^+ + \sigma_{VIX}\varepsilon_{VIX,t}, \quad (2.4)$$

$$\log(VIX_t) = \nu_0 + \nu_1 \log(VIX_{t-1}) + \nu_2 \log(\sigma_{SPX,t}) + \nu_3 y_t + \sigma_{VIX}\varepsilon_{VIX,t}. \quad (2.5)$$

To assess the relative performance of model (2.1) versus these alternatives we use the following criteria to evaluate the trade-off between the number of model parameters and goodness of fit. First, since all the alternatives are special cases of model (2.1), we used the likelihood ratio test. For this test one computes the statistic $2(l_1 - l_2)$, where l_1 is the log-likelihood of the more complex model and l_2 the log-likelihood of the simpler alternative. Under the null hypothesis that the more complex model is not a significant improvement over the simpler version, the test statistic has a χ^2 distribution with degrees of freedom equal to the difference in the number of parameters between the two models.

Second, we used the Schwartz Bayes criterion. This selects the model that maximizes $l_j - \frac{1}{2} k_j \log n$, where l_j is the log-likelihood of model j, k_j is the number of model parameters, and n is the number of observations. For a discussion of model selection criteria and further details of these tests see Hardy (2001).

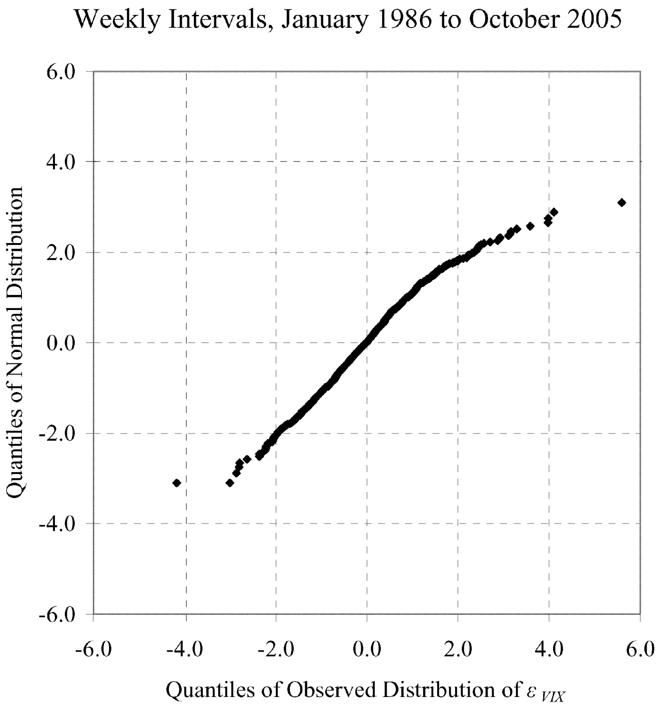
The results of this comparison are shown in Table 1. As can be seen, model (2.1) is a significant improvement over the simpler alternatives under both these criteria. Figure 5 shows a quantile-quantile plot of the residuals $\varepsilon_{VIX,t}$ under model (2.1) versus a standard normal distribution. This figure suggests that a normal distribution provides a good fit to the residuals.

Note that the VIX model depends on S&P500 index returns, which means that it is necessary to specify a model for the latter to simulate the former. (However, calibration of the VIX model is based solely on historical data.) If the model for the S&P500 is unrealistic—for instance, if it does not capture observed features such as occasional extreme price movements, skewed returns, or vol-

Table 1
Comparison of Alternative VIX Models

Model	Number of Parameters (k_j)	Log-Likelihood (l_j)	Likelihood Ratio Test	Schwartz Bayes Criterion
2.1	7	1,148.05		**1,123.76**
2.2	4	1,075.51	$< 10^{-10}$	1,061.64
2.3	5	1,026.58	$< 10^{-10}$	1,009.24
2.4	5	1,122.49	$< 10^{-5}$	1,105.14
2.5	6	1,125.10	$< 10^{-5}$	1,104.29

Figure 5
Quantile-Quantile Plot: Observed Distribution of ε_{VIX} versus Normal Distribution



atility clustering—then the VIX model will not produce realistic results either.

3. A PARAMETRIC MODEL OF THE VOLATILITY SURFACE

Having modeled the process for the VIX, our next objective is to model the dynamics of the volatility surface. However, we first need to address the fact that implied volatilities for specific options as reported by OptionMetrics and iVolatility.com on successive dates are not directly comparable with each other, because of the passage of time and changes in the value of the underlying S&P500 index. For example, an option that has a 60-day term to maturity on a given date will have only a 53-day term to maturity one week later. Likewise, if its strike was originally 100% of spot, this is unlikely to be exactly the case one week later. To be able to compare specific points on successive dates, we therefore follow the lead of other researchers such as Dumas, Fleming, and Whaley (1998), Cont and Fonseca (2002), and Alentorn (2004) and fit a smoothed surface to each weekly set of observations. (This also addresses the fact that the historical data have many omissions and irregularities, especially for options that are deeply in- or out-of-the-money.) Mainly for ease of computation we use a parametric model, as opposed to the nonparametric kernel estimator used by Cont and Fonseca. The model we use is as follows:

$$\sigma_t(x, T) = a_{0,t} + a_{1,t}\frac{\log x}{T^{1/2}} + a_{2,t}\frac{1}{T^{1/2}} + a_{3,t}\frac{(\log x)^2}{T} + a_{4,t}\frac{\log x}{T} + a_{5,t}\frac{(\log x)^2}{T^{3/2}}.$$

Here $\sigma_t(x, T)$ is the modeled value of implied volatility at time t, x is the strike (as a fraction of the current spot price), T is the unexpired term (in years), and $a_{0,t} \ldots a_{5,t}$ are six time-varying parameters. We denote the vector $\{a_{0,t} \ldots a_{5,t}\}$ as $\mathbf{a}_t$. The expression $(\log x)T^{-1/2}$ is sometimes referred to as the "moneyness" of the option.

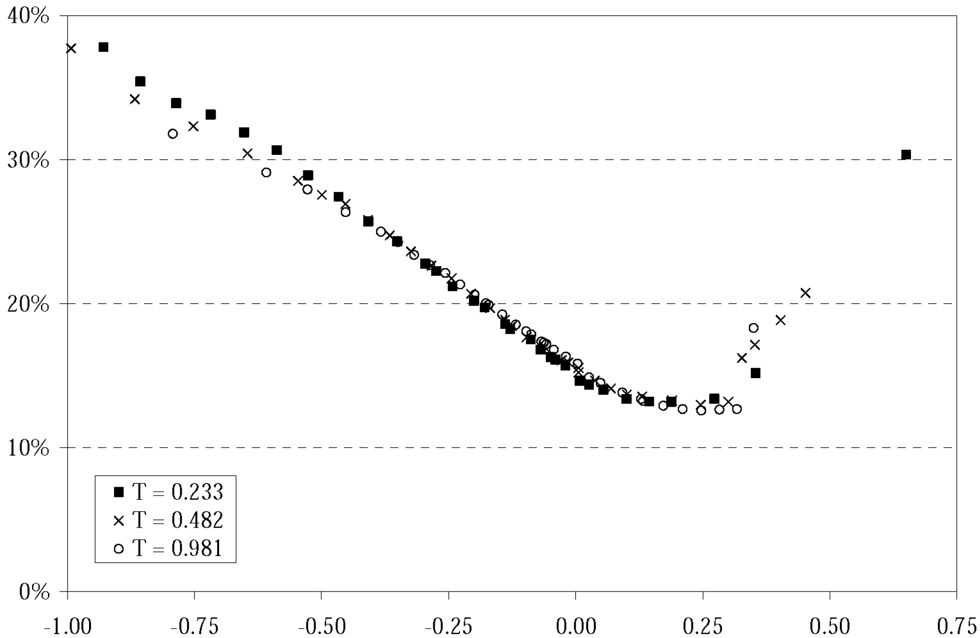
This model is motivated by the following empirical observations. First, if implied volatility for a given term is plotted as a function of moneyness, then the function typically has a roughly parabolic shape (Fig. 6). (The parabolic shape breaks down for deep out-of-the-money options, but these options tend to be thinly traded, and it is not clear how meaningful their quoted prices are.) Second, if implied volatilities for a given moneyness are plotted as a function of $T^{-1/2}$, then this function tends to be roughly linear. (This is sometimes known as the "square root of time" rule.) This is illustrated in Figure 7.

We used a least-squares procedure to estimate the model parameters. However, instead of minimizing the sum of the squared absolute differences $(\sigma_{act} - \sigma_{est})^2$, where σ_{act} is the observed value of implied volatility and σ_{est} is the modeled value, we minimized the sum of the squared relative differences $(1 - \sigma_{est}/\sigma_{act})^2$. In effect this weights each observation by $1/\sigma_{act}$ and hence gives least weight to observations with the highest implied volatility, which are usually thinly traded options that are deeply in- or out-of-the-money.

We fitted this model to OptionMetrics data at weekly intervals from January 5, 1996, to

Figure 6
Implied Volatility as a Function of Moneyness, (log *x*) · $T^{-1/2}$

(1) As of December 26, 2003



(2) As of September 30, 2005

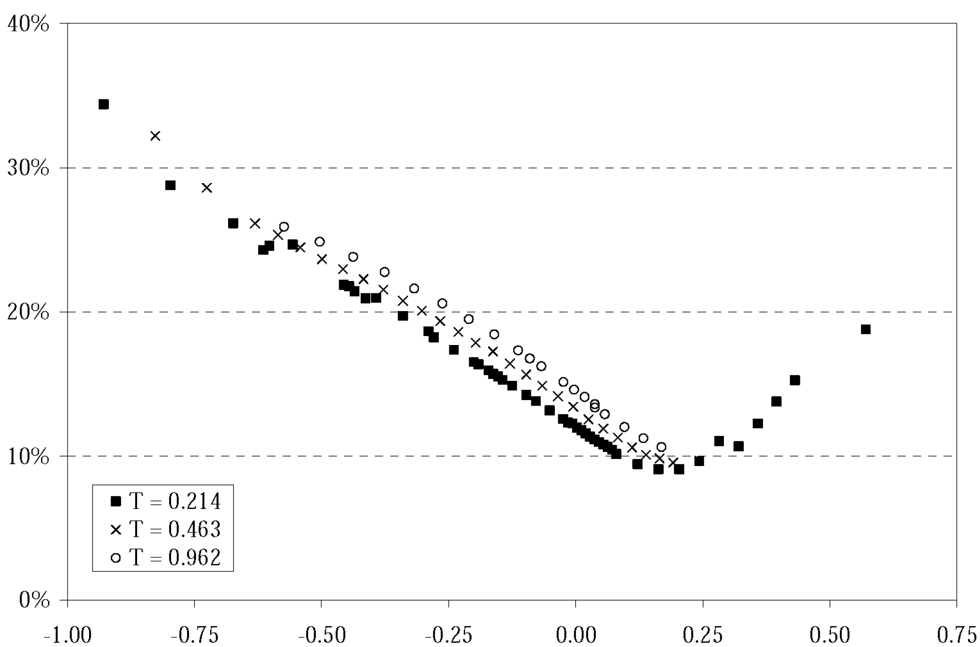
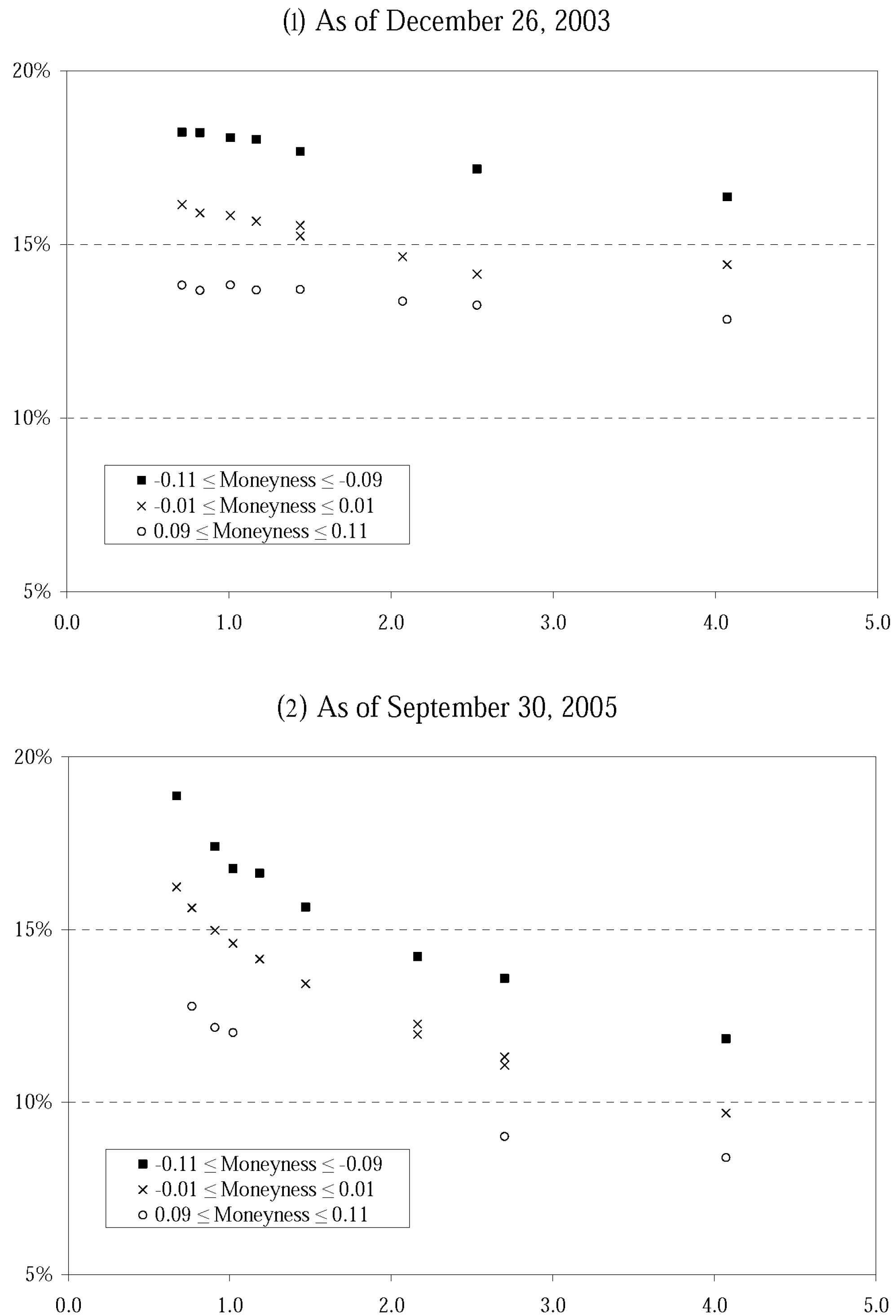


Figure 7
Implied Volatility as a Function of $T^{-1/2}$



November 24, 2000, and iVolatility.com data from December 1, 2000, to November 25, 2005. The reason for switching from OptionMetrics to iVolatility.com half-way through is that OptionMetrics shows occasional significant differences in implied volatilities for call and put options with identical terms and strikes, which we suspect may be a problem with OptionMetrics's methodology for calculating implied volatilities as opposed to actual arbitrage opportunities. However, iVolatility.com data were not available prior to November 2000, whereas OptionMetrics data were available back to January 1996.

Our analysis was based on implied volatilities of call options with delta between 0% and 50%, and put options with delta between −50% and 0%. We disregarded all options with an unexpired term of less than 14 days because of the general difficulty of calculating implied volatilities for such short terms. In some cases implied volatilities as reported by iVolatility.com for deep out-of-the-money options were identical to implied volatilities for options that were closer to the money. We disregarded such data points. Generally this was a problem only for calls with a delta less than 1% or puts with a delta greater than −1%. This typically left us with about 200 observations on each date, covering options with terms of up to about two years.

Table 2 shows fitted values of $\mathbf{a}_t$ on selected dates as well as the corresponding r^2 statistic. Here r^2 is the square of the Pearson correlation coefficient between σ_{act} and σ_{est} on each selected date (note that the above fitting procedure does not necessarily maximize r^2). The aggregate r^2 across the entire period is 92.2%. Figure 8 shows a comparison of the raw data and the modeled volatility surface on two selected dates.

4. Principal Component Analysis of the Volatility Surface

In the previous section we showed how the volatility surface could be represented using a six-factor model, as opposed to working directly several hundred underlying data points. In this section we show that the model can be further simplified and that most of the time-series variation in the volatility surface can be explained by just the VIX index and one or two other uncorrelated factors, using the technique of principal component analysis (PCA). An analogy that will be familiar to some actuaries is the use of PCA for modeling the dynamics of the yield curve: instead of directly modeling each of the dozen or so (highly correlated) points that make up the yield curve, changes in the yield curve can be represented as a linear combination of just two or three uncorrelated factors, usually known as a "shift" factor, a "twist" factor, and (sometimes) a "butterfly" factor. For an introduction to PCA, see Frye (1997).

Because we have used a six-factor model, we do not lose any information if we arbitrarily select six points $\{x_i, T_i\}$ and focus on the dynamics of the fitted values $\sigma_{t,i} = \sigma_t(x_i, T_i)$ at just these points. Provided the choice of points meets certain conditions, we can always reconstruct the parameter vector $\mathbf{a}_t$ from the vector of fitted values $\boldsymbol{\sigma}_t = \{\sigma_{t,1} \ldots \sigma_{t,6}\}$ at the six chosen points, and hence we can reconstruct the rest of the fitted surface.

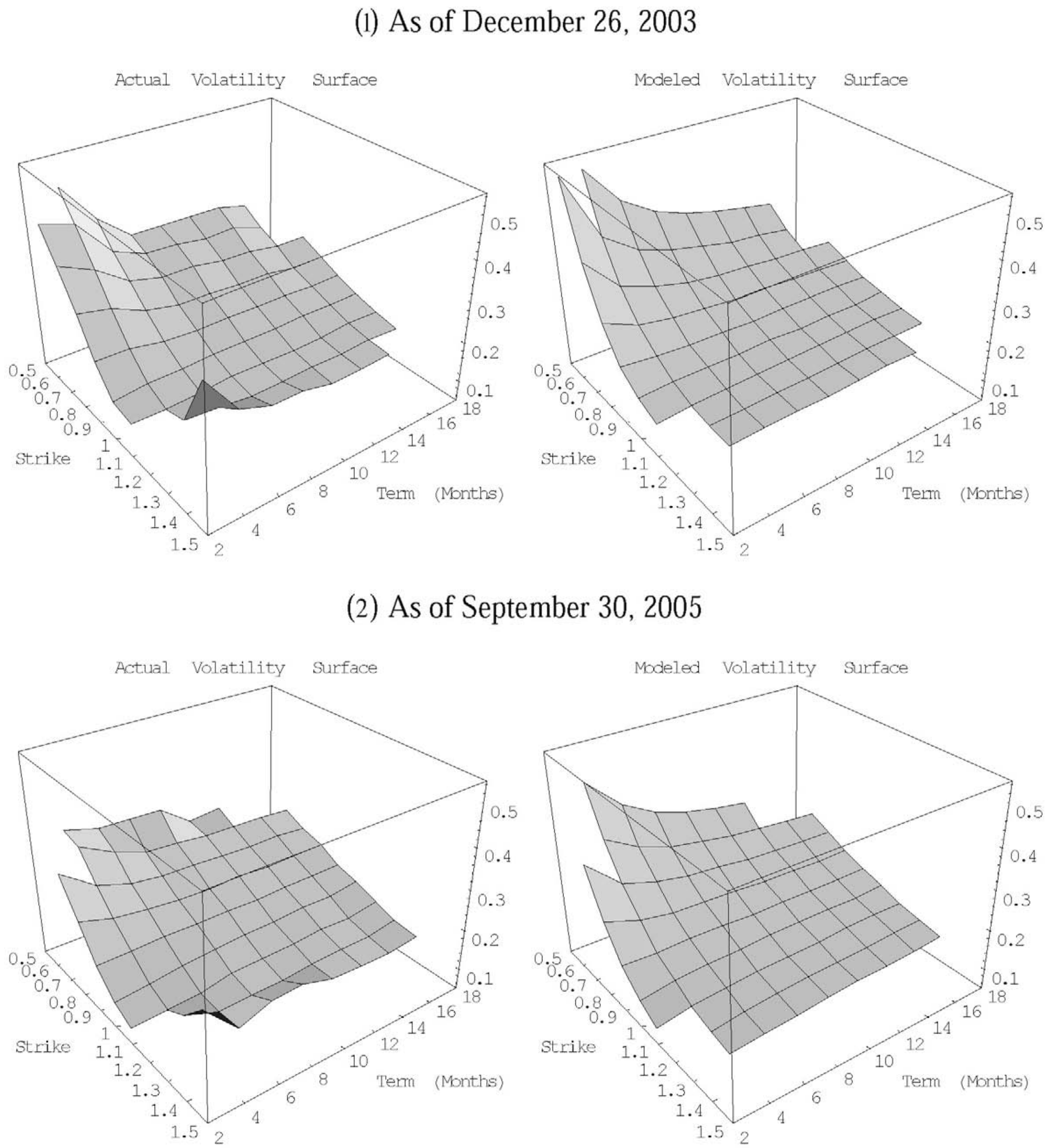
More precisely, we can write $\boldsymbol{\sigma}_t = \mathbf{a}_t \cdot \mathbf{A}$ and hence $\mathbf{a}_t = \boldsymbol{\sigma}_t \cdot \mathbf{A}^{-1}$ where the matrix $\mathbf{A}$ is defined as

Table 2

Parameters and r^2 for Volatility Surface Model on Selected Dates

Date	$a_{0,t}$	$a_{1,t}$	$a_{2,t}$	$a_{3,t}$	$a_{4,t}$	$a_{5,t}$	r^2
Dec. 27, 1996	0.1753	−0.2171	0.0008	−0.0337	0.0130	0.0293	98.64%
Dec. 26, 1997	0.2197	−0.2543	0.0129	−0.0288	0.0135	0.0153	98.28
Dec. 31, 1998	0.2803	−0.3284	−0.0167	−0.0394	0.0489	0.0342	97.36
Dec. 31, 1999	0.2332	−0.2830	−0.0030	−0.0635	0.0369	0.0405	97.29
Dec. 29, 2000	0.2321	−0.2603	−0.0033	0.1852	0.0555	0.0059	70.06
Dec. 28, 2001	0.2168	−0.2428	−0.0084	0.0822	0.0594	0.0441	78.46
Dec. 27, 2002	0.2401	−0.2099	0.0030	0.3496	0.0554	−0.0359	82.17
Dec. 26, 2003	0.1648	−0.1517	−0.0034	0.1117	0.0128	0.0170	92.18
Dec. 31, 2004	0.1625	−0.2245	−0.0121	−0.0150	0.0388	0.0680	92.18
Sept. 30, 2005	0.1590	−0.1970	−0.0152	0.0184	0.0166	0.0392	96.99

Figure 8
Actual and Modeled Volatility Surface



$$\mathbf{A} = \begin{bmatrix} 1 & \cdots & 1 \\ (\log x_1)T_1^{-1/2} & \cdots & (\log x_6)T_6^{-1/2} \\ T_1^{-1/2} & \cdots & T_6^{-1/2} \\ (\log x_1)^2 T_1^{-1} & \cdots & (\log x_6)^2 T_6^{-1} \\ (\log x_1)T_1^{-1} & \cdots & (\log x_6)T_6^{-1} \\ (\log x_1)^2 T_1^{-3/2} & \cdots & (\log x_6)^2 T_6^{-3/2} \end{bmatrix}.$$

The six points $\{x_i, T_i\}$ need to be chosen so that the matrix **A** is invertible, meaning that none of its six columns can be expressed as a linear combination of the others.

Having selected six such points, we use a two-step process to model the dynamics of the surface. First, we use linear regression to model the relationship between $\sigma_{t,i}$ and VIX_t at each point, as follows:

Table 3
Parameters of Volatility Surface Model: Regression Coefficients of σ_t against VIX_t, and First and Second Principal Components of ξ_t

	$\sigma_{t,1}$	$\sigma_{t,2}$	$\sigma_{t,3}$	$\sigma_{t,4}$	$\sigma_{t,5}$	$\sigma_{t,6}$	$\sigma_{k,i}$
$\boldsymbol{\alpha_0}$	0.0836	0.0526	0.0294	0.1230	0.0857	0.0525	
$\boldsymbol{\alpha_1}$	0.7201	0.6851	0.6688	0.6453	0.5581	0.5416	
$\mathbf{b_1}$	−0.3492	−0.3184	−0.2216	−0.5596	−0.5203	−0.3792	0.0456
$\mathbf{b_2}$	−0.3485	0.0581	0.5578	−0.4636	0.0294	0.5901	0.0159

$$\sigma_{t,i} = \alpha_{0,i} + VIX_t\alpha_{1,i} + \boldsymbol{\xi}_{t,i}.$$

Here $\alpha_{0,i}$ and $\alpha_{1,i}$ are chosen so as to minimize the sum across time of the squared residuals $\Sigma_t\xi_{t,i}^2$. We can rewrite this in vector notation as

$$\boldsymbol{\sigma}_t = \boldsymbol{\alpha}_0 + VIX_t\boldsymbol{\alpha}_1 + \xi_t,$$

where $\boldsymbol{\alpha}_0 = \{\alpha_{0,1} \ldots \alpha_{0,6}\}$, $\boldsymbol{\alpha}_1 = \{\alpha_{1,1} \ldots \alpha_{1,6}\}$, and $\boldsymbol{\xi}_t = \{\xi_{t,1} \ldots \xi_{t,6}\}$.

Second, we use PCA to approximate the residuals $\xi_{t,i}$ using the model

$$\boldsymbol{\xi}_t \approx k_{1,t}\mathbf{b}_1 + k_{2,t}\mathbf{b}_2 + \cdots + k_{n,t}\mathbf{b}_n,$$

where only the scalars $k_{1,t}, k_{2,t} \ldots k_{n,t}$ change over time, and n is less the number of elements of $\boldsymbol{\xi}_t$. The n vectors $\mathbf{b}_i = \{b_{i,1} \ldots b_{i,6}\}$ are known as the principal components of $\boldsymbol{\xi}_t$ and are chosen so as to satisfy the following properties: (1) $\mathbf{b}_i \cdot \mathbf{b}_j = 0$ for all $i \neq j$, that is, the vectors $\mathbf{b}_i$ are orthogonal to each other; (2) $\|\mathbf{b}_i\| = 1$ for all i; and (3) the $k_{i,t}$ on any given date are independent of each other, that is, the covariance between $k_{i,t}$ and $k_{j,t}$ is zero for $i \neq j$. In formal terms the vectors $\mathbf{b}_i$ are the principal eigenvectors of the covariance matrix of $\boldsymbol{\xi}_t$, and the corresponding eigenvalues are the variances of $k_{i,t}$.

Given $\boldsymbol{\xi}_t$ and $\mathbf{b}_i$, each $k_{i,t}$ is calculated so as to minimize $\|\boldsymbol{\xi}_t - k_{i,t}\mathbf{b}_i\|$. Making use of the fact that $\|\mathbf{b}_i\| = 1$ it can be shown that the solution to this equation is $k_{i,t} = \boldsymbol{\xi}_t \cdot \mathbf{b}_i$.

Using the OptionMetrics and iVolatility.com data described above we have analyzed the fitted volatility surface over the period January 1996 to November 2005 at the following six points $\{x_i, T_i\}$: $\{90\%, \frac{1}{4}\}$, $\{100\%, \frac{1}{4}\}$, $\{110\%, \frac{1}{4}\}$, $\{75\%, 1\frac{1}{4}\}$, $\{100\%, 1\frac{1}{4}\}$, and $\{125\%, 1\frac{1}{4}\}$. We find that 80.2% of the variation in $\boldsymbol{\sigma}_t$ is explained by VIX_t. Of the variation of the residuals $\boldsymbol{\xi}_t$, 86.0% is explained by the first principal component $\mathbf{b}_1$, and another 10.4% is explained by the second principal component $\mathbf{b}_2$. This suggests that for most practical purposes it would be sufficient to model just the VIX index and process for $k_{1,t}$ (i.e., the magnitude of $\mathbf{b}_1$), but for completeness we will also describe the process for $k_{2,t}$ (i.e., the magnitude of $\mathbf{b}_2$).

With two factors $\mathbf{b}_1$ and $\mathbf{b}_2$ the above model can be written as

$$\boldsymbol{\sigma}_t \approx \boldsymbol{\alpha}_0 + VIX_t\boldsymbol{\alpha}_1 + k_{1,t}\mathbf{b}_1 + k_{2,t}\mathbf{b}_2.$$

Table 3 shows numerical values of $\boldsymbol{\alpha}_0$, $\boldsymbol{\alpha}_1$, $\mathbf{b}_1$, and $\mathbf{b}_2$ as well as the weekly standard deviations of $k_{1,t}$ and $k_{2,t}$, which we denote by $\sigma_{k,1}$ and $\sigma_{k,2}$, respectively. Figures 9–11 show a graphical representation of these factors. Specifically, Figure 9 shows the average profile of the volatility surface, calculated by setting $\boldsymbol{\sigma}_t$ equal to $\boldsymbol{\alpha}_0 + \mu_{VIX}\boldsymbol{\alpha}_1$, where μ_{VIX} is the mean value of VIX_t over the above period. Figure 10 shows the sensitivity of the volatility surface to a 1% change in the VIX index, calculated by setting $\boldsymbol{\sigma}_t$ equal to $0.01\boldsymbol{\alpha}_1$. Figure 11 shows the first and second principal

Figure 9
Average Profile of Volatility Surface

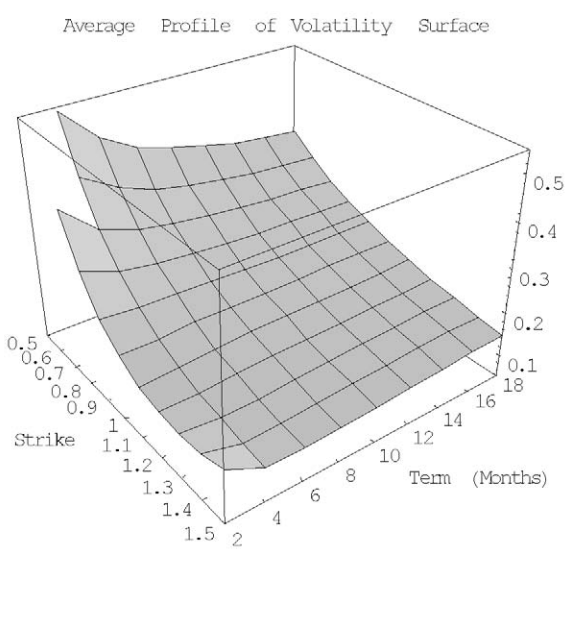
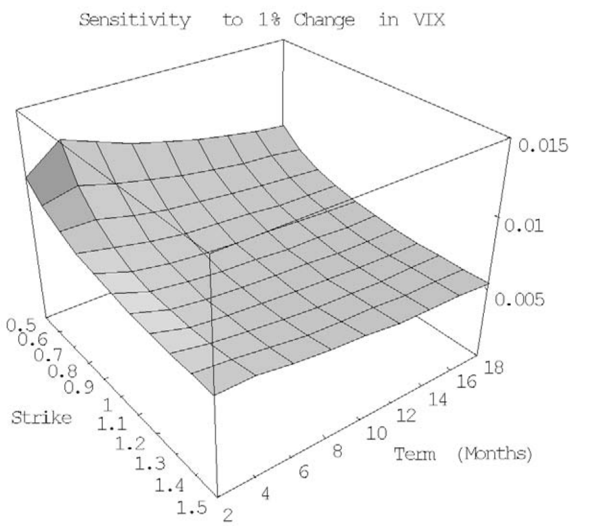


Figure 10
Sensitivity of Volatility Surface to 1% Change in VIX



components of the residuals $\boldsymbol{\xi}_t$, calculated by setting $\boldsymbol{\sigma}_t$ equal to $\sigma_{k,1}\mathbf{b}_1$ and $\sigma_{k,2}\mathbf{b}_2$ respectively.

As can be seen from Figure 11, the first principal component corresponds to an increase in the curvature of the smile for short-dated options, but with relatively little effect on longer-dated options. The second principal component corresponds to a decrease in skew for all terms, and an increase in the curvature of the smile for short-dated options. Note that these factors can be multiplied by negative values of $k_{i,t}$ as well as positive values.

As noted above, we do not lose any information by focusing on the dynamics of the volatility surface at six arbitrarily chosen points $\{x_i, T_i\}$, because we can always reconstruct the entire fitted surface given values of $\sigma_{t,i}$ at any six points (subject to certain technical conditions). However, the results of the PCA decomposition will differ depending on which particular points we select. The points should be chosen so as to be representative of the types of options we are interested in simulating, otherwise the PCA decomposition will produce misleading results.

5. Dynamics of Changes in the Shape of the Implied Volatility Surface

In the previous section we showed how the vector of residual volatilities $\boldsymbol{\xi}_t$ could be approximated as $k_{1,t}\,\mathbf{b}_1 + k_{2,t}\,\mathbf{b}_2$, where only $k_{1,t}$ and $k_{2,t}$ vary over time. In this section we describe a model for the processes followed by $k_{1,t}$ and $k_{2,t}$, historical values of which are shown in Figure 12.

These figures suggest that successive values of $k_{1,t}$ and $k_{2,t}$ are not independent of each other. In fact, we find that the correlation between $k_{1,t}$ and $k_{1,t-1}$ is 91.0%, and between $k_{2,t}$ and $k_{2,t-1}$ the correlation is 84.6%. This leads us to model them as autoregressive processes of the form

Figure 11
Volatility Surface Model: Principal Components of ξ_t

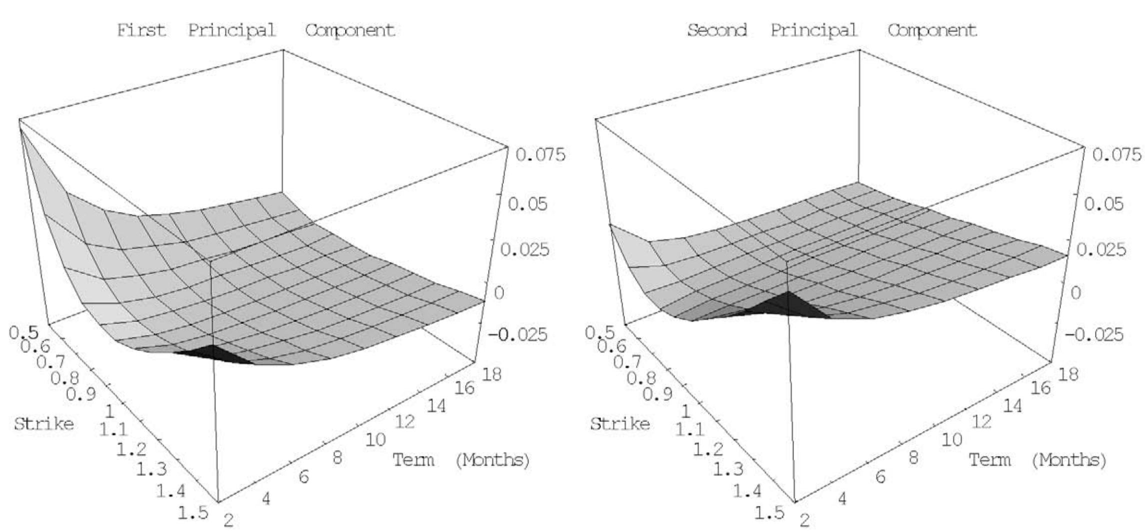
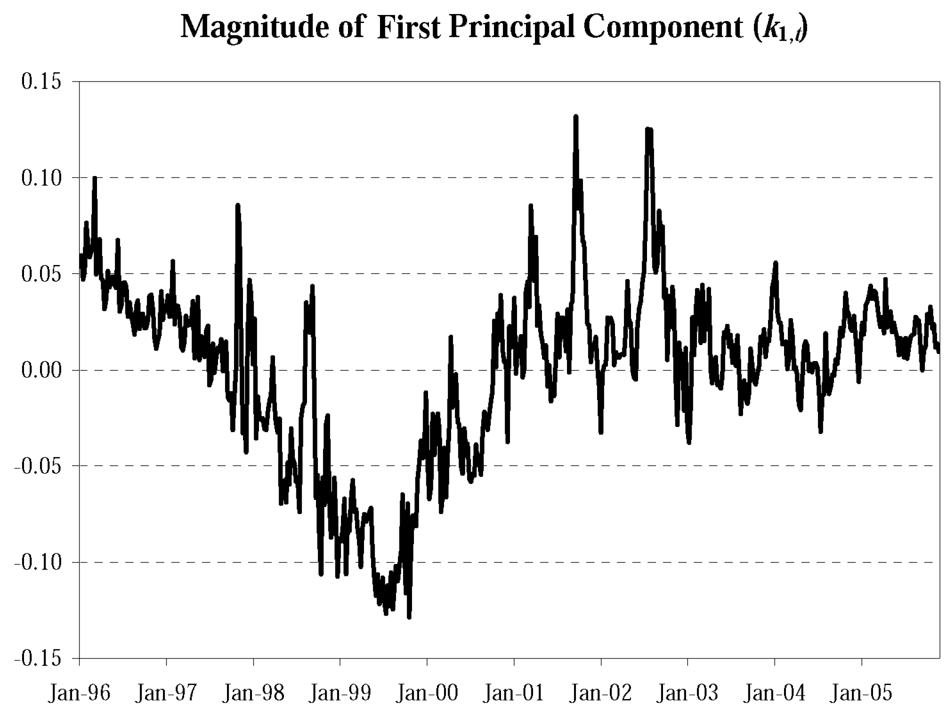
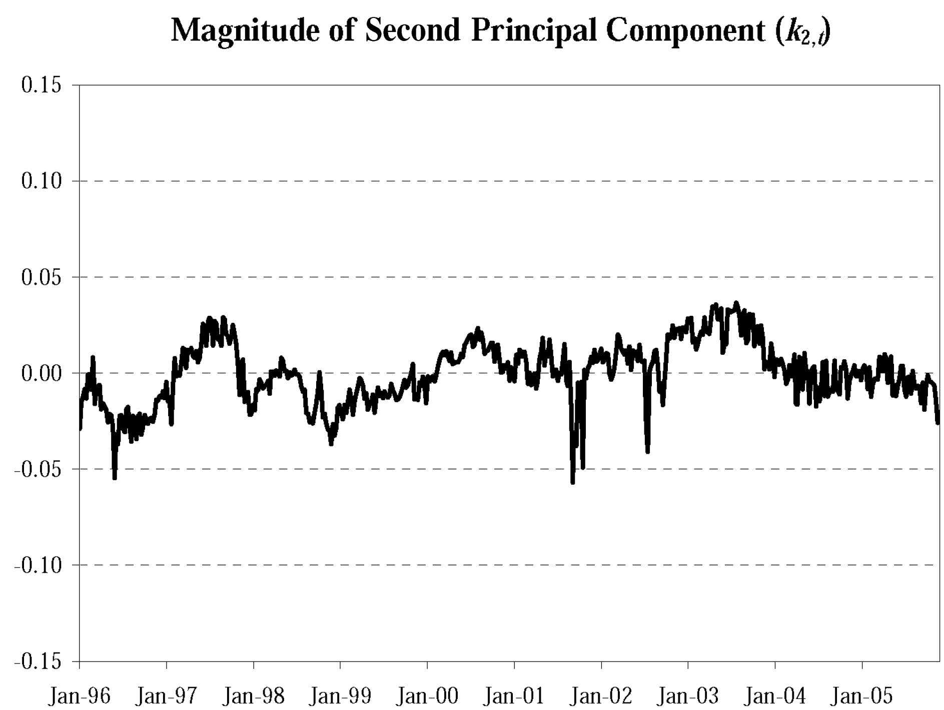


Figure 12
Volatility Surface Model: Dynamics of Principal Components of ξ_t
Weekly Intervals, January 1996 to November 2005





$$k_{i,t} = \beta_i k_{i,j-1} + \gamma_i \varepsilon_{i,t},$$

where the residuals $\varepsilon_{i,t}$ have zero mean and unit variance (note that both $k_{1,t}$ and $k_{2,t}$ have zero means, by construction). The parameters β_i and γ_i are estimated by linear regression of $k_{i,t}$ against $k_{i,t-1}$. Their values are as follows:

$$\beta_1 = 0.9084 \qquad \gamma_1 = 0.0189$$
$$\beta_2 = 0.8445 \qquad \gamma_2 = 0.0085.$$

Figure 13 shows quantile-quantile plots of the residuals $\varepsilon_{i,t}$ versus a standard normal distribution. These figures suggest that a normal distribution provides a good fit to the residuals $\varepsilon_{1,t}$. The residuals $\varepsilon_{2,t}$ appear to have fatter tails than a normal distribution, but this component has relatively little effect on the overall volatility surface.

In the simplified case where the model for $\boldsymbol{\xi}_t$ is based solely on the first principal component $\mathbf{b}_1$, we can drop references to $k_{i,t}$ and write the process for $\boldsymbol{\xi}_t$ as follows:

$$\boldsymbol{\xi}_t = k_{1,t}\mathbf{b}_1 = \beta_1 \boldsymbol{\xi}_{t-1} + \gamma_1 \varepsilon_{1,t} \mathbf{b}_1.$$

Finally, we note that for simulation purposes it is more convenient to work with the process for the parameter vector $\mathbf{a}_t$ rather than $\boldsymbol{\sigma}_t$. As noted in Section 3, we can write $\mathbf{a}_t = \sigma_t \cdot \mathbf{A}^{-1}$. Consequently, the process for $\mathbf{a}_t$ can be written as

$$\mathbf{a}_t = \boldsymbol{\alpha}_0' + VIX_t \boldsymbol{\alpha}_1' + \boldsymbol{\xi}_t',$$

$$\boldsymbol{\xi}_t' = \beta_1 \boldsymbol{\xi}_{t-1}' + \gamma_1 \varepsilon_{1,t} \mathbf{b}_1',$$

where $\boldsymbol{\alpha}_0' = \boldsymbol{\alpha}_0 \cdot \mathbf{A}^{-1}$, $\boldsymbol{\alpha}_1' = \boldsymbol{\alpha}_1 \cdot \mathbf{A}^{-1}$, and $\mathbf{b}_1' = \mathbf{b}_1 \cdot \mathbf{A}^{-1}$.

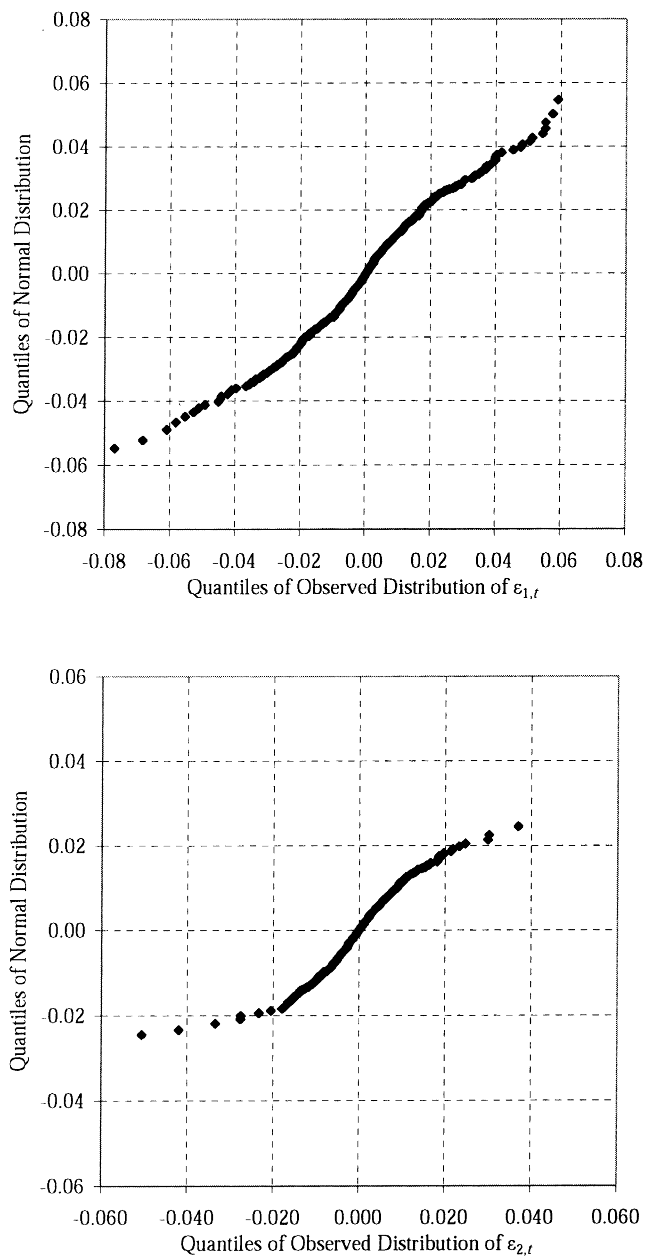
6. SUMMARY AND CONCLUSIONS

We have presented a model of the S&P500 implied volatility surface based on the following components: first, the general level of implied volatilities, as measured by the VIX index; and second, an additional factor (or pair of factors) describing changes in the shape of the volatility surface that are not explained by the VIX index.

To utilize this model for simulation purposes it is also necessary to simulate the underlying S&P500 index, since this is an input to the modeled process for the VIX index. However, the latter model is not dependent on any specific model for the S&P500.

Figure 13

Quantile-Quantile Plots: Observed Distribution of $\varepsilon_{i,t}$ versus Normal Distribution



Our model is intended for projection periods of a year or more, and as such it has been calibrated using as much historical data as were available to us—since January 1986 for the VIX index, and since January 1996 for the OptionMetrics and iVolatility.com databases describing the volatility surface.

7. ACKNOWLEDGMENTS

The author thanks the following for helpful comments: colleagues at ING, participants at the Canadian Institute of Actuaries' 2006 Stochastic Modeling Symposium, and an anonymous referee.

REFERENCES

ALENTORN, A. 2004. Modelling the Implied Volatility Surface: An Empirical Study for FTSE Options. Working paper. privatewww.essex.ac.uk/~aalent/.

CHICAGO BOARD OPTIONS EXCHANGE. 2003. The New CBOE Volatility Index—VIX. www.cboe.com/micro/vix/vixwhite.pdf.

CONT, R., AND J. DA FONSECA. 2002. Dynamics of Implied Volatility Surfaces. *Quantitative Finance* 2: 45–60.

DUMAS, B., J. FLEMING, AND R. E. WHALEY. 1998. Implied Volatility Functions: Empirical Tests. *Journal of Finance* 53: 2059–2106.

FRYE, J. 1997. Principals of Risk: Finding Value-at-Risk through Factor-Based Interest Rate Scenarios. In *VAR: Understanding and Applying Value at Risk,* pp. 275–88. London: Risk Publications.

GATHERAL, J. 2006. *The Volatility Surface: A Practitioner's Guide*. Hoboken, NJ: John Wiley & Sons.

HARDY, M. R. 2001. A Regime-Switching Model of Long-Term Stock Returns. *North American Actuarial Journal* 5: 41–53.

LIFE CAPITAL ADEQUACY SUBCOMMITTEE. 2005. *Recommended Approach for Setting Regulatory Risk-Based Capital Requirements for Variable Annuities and Similar Products*. American Academy of Actuaries. www.actuary.org/pdf/life/c3_june05.pdf.

Discussions on this paper can be submitted until April 1, 2008. The author reserves the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.