Two Essays on Options Market

by

Yang-Ho Park

M.S., Florida State University, 2006

M.S., Seoul National University, 2000

B.S., Seoul National University, 1998

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Leeds School of Business,

Department of Finance

2011

This thesis entitled:
Two Essays on Options Market
written by Yang-Ho Park
has been approved for the Leeds School of Business,
Department of Finance

_____

Dr. Garland B. Durham

_____

Dr. Michael J. Stutzer

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Park, Yang-Ho (Ph.D., Finance)

Two Essays on Options Market

Thesis directed by Dr. Garland B. Durham

Essay I: The Roles of Short-Run and Long-Run Volatility Factors in Options Market: A Term Structure Perspective

This paper examines the option pricing implications of short-run and long-run volatility factors, which are assumed to be driven by short-run and long-run news events, respectively. Using a comprehensive dataset of S&P 500 index options over 1993-2008, I find that the proposed two-factor volatility models have two desirable properties that help capture the term structures of option-implied volatility and skewness. First, the options data show evidence of time-variation in the long-run expectation of volatility, which may be caused by long-run news events. While this feature is inconsistent with a single-factor volatility assumption, the two-factor volatility models do a good job of matching the entire term structure of implied volatility. Second, the options data reveal that the term structure of implied skewness is nearly flat on average. This feature is hard to reconcile with single-factor volatility models and jumps in returns. In contrast, I find that the two-factor volatility models can generate flat term structures much like those seen in the data. In particular, the short-run volatility factor is dominant in generating short-term skewness, while the long-run volatility factor plays a pivotal role in generating long-term skewness.

Essay II: Beyond Stochastic Volatility and Jumps in Returns and Volatility

While a great deal of attention has been focused on stochastic volatility in stock returns, there is strong evidence suggesting that return distributions have time-varying skewness and kurtosis as well. This can be seen, for example, from variation across time in the shape of Black-Scholes implied volatility smiles. This paper investigates model characteristics that are consistent with variation in the shape of return distributions using a standard stochastic volatility model with a regime-switching feature to allow for random changes in the parameters governing volatility of volatility and leverage effect. The analysis consists of two steps. First, the models are estimated using only information from observed returns and option-implied volatility. Standard model assessment tools indicate a strong preference in favor of the proposed models. Since the information from option-implied skewness and kurtosis is not used in fitting the models, it is available for diagnostic purposes. In the second step of the analysis, regressions of option-implied skewness and kurtosis on the filtered state variables (and some controls) suggest that the models have strong explanatory power. This is important because it suggests that variation in the shape of risk-neutral return distributions (and of the Black-Scholes implied volatility smile) is not just due, for example, to changes in risk premia, but is associated with changes in related characteristics of the physical dynamics.

Acknowledgements

I feel so lucky to have Garland Durham as my dissertation chair. I owe him an enormous amount of intellectual and moral debt during my doctoral program.

I would also like to express my thanks to my dissertation committee: Sanjai Bhagat, Bjorn Jorgensen, Iulian Obreja, and Mike Stutzer. Sanjai Bhagat has given me a great deal of advice for my academic and non-academic growth. I have learned a lot from Bjorn Jorgensen about professional attitude including presentation and interview skills; his advice has always been very delicate. Iulian Obreja has always kindly discussed with me about interesting research topics. I would not forget about the support of Mike Stutzer, who has given me academic advice and, more importantly, had me employed upon my graduation.

Finally, I thank my doctoral program friends, with whom I have spent enjoyable and productive times: Brian Burnett, Jeffrey Merrell, Nhan Le, Jun Lu, and Ying Xiao. I also thank my friends, with whom I have so much fun: Woong Jae Baek, Seoung Chung, Jamie Frade, Ahmet Goncu, Ki Chul Koo, Seung Bo Shim, and Songbo Sim.

Thank you all for everything.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

The Roles of Short-Run and Long-Run Volatility Factors in Options Market: A Term Structure Perspective

1.1 Introduction

The discounted cash flow model predicts that the volatility of a financial asset should be linked to uncertainty over the future cash flows and discount rate of that asset. In this framework, volatility changes are driven by arrivals of new information pertaining to uncertainty over the cash flows and discount rate, including macroeconomic news events, market microstructure shocks, and firm-specific shocks. Schwert (1989) studies the relation of stock market volatility with macroeconomic volatility, stock market activity, and financial leverage. Every news event may have a distinct term structure effect on volatility, depending on the market's perception of the persistence of the news event along a horizon. Macroeconomic news events tend to have more persistent effects on volatility than market microstructure ones. Adrian and Rosenberg (2008) link short-run volatility to the tightness of financial constraints and long-run volatility to business cycle risk. Therefore, volatility is comprised of multiple components, each of which has its distinct term structure effect on volatility according to its economic nature.

The fact that volatility consists of multiple components has important implications particularly for the options market, as opposed to the stock market. Since there are options with different investment horizons on any given day, the term structure of volatility is observable in the options market. The option-implied volatility reflects the day's market expectation of future volatility. The dynamics of the term structure of implied volatility are then determined by the market's perception

of the term structure effect of each news event on volatility. If a news event is considered as short-run (or temporary), long-term implied volatility will not be affected as much as short-term implied volatility. In contrast, if a news event is interpreted as long-run (or permanent), short-term and long-term implied volatilities are likely to move in lockstep. Therefore, the options market will show various movements in the term structure of volatility, depending on the size and persistence of an arriving news event.

The objective of this article is to investigate the option pricing implications of short-run and long-run volatility factors, which are assumed to be directed by short-run and long-run news events, respectively. There is a related literature. Bates (2000), Christoffersen, Heston, and Jacobs (2009), and Christoffersen, Jacobs, Ornthanalai, and Wang (2008), among others, study option pricing by using two-factor volatility models. However, all of these papers are based on a narrow class of affine models. Although the affine models are econometrically convenient due to the existence of closed-form option pricing formulas, they have trouble fitting the returns and options data in practice. First of all, the affine models do a poor job of capturing the time-series dynamics of returns relative to non-affine models (see, e.g., Ait-Sahalia and Kimmel 2007, Jones 2003, and Durham 2008). More importantly, Christoffersen, Jacobs, and Mimouni (2010) compare several single-factor stochastic volatility models in terms of option pricing performance and find that the CEV model with elasticity of variance equal to 1 (hereafter, the ONE model) performs better than the affine model. To the best of my knowledge, a study of non-affine option pricing models with multiple stochastic volatility factors is still lacking.

In response, this paper develops a family of log volatility models with short-run and long-run volatility factors. The short-run (long-run) factor controls short-term (long-term) trends and fluctuations in volatility. In addition, I augment the two-factor log volatility model by including jumps in both returns and the short-run volatility factor. The log stochastic volatility framework can be translated into the ONE model with a nonlinear drift using Ito's Lemma. To the extent that nonlinearity in the drift is insignificant, the models in this paper can be regarded as two-factor extensions of the ONE model.

Although the log stochastic volatility models under consideration offer more realistic return dynamics than the affine models, a great deal of econometric challenge arises. It does not seem feasible to estimate the log volatility models by minimizing the sum of option pricing errors for a long panel dataset of options, because there is no closed-form option pricing formula. To address this issue, I develop a computationally efficient two-step procedure. In the first step, I pin down two latent volatility factors from a pair of model-free implied volatilities on each day; and in the second step, I apply an implied-state maximum likelihood estimation by treating those volatility factors as observable. Joint analysis of this kind using information on both returns and options has been used by Chernov and Ghysels (2000), Jones (2003), Ait-Sahalia and Kimmel (2007), and Pan (2002), among others.

The focus of my analysis is on why two-factor volatility models perform better in explaining option prices than one-factor volatility models. In particular, I find it very informative to compare the term structures of option-implied volatility and skewness to their model-implied counterparts. Comparing the term structures of volatility and skewness makes it easy to identify the sources of misspecification in the models, because each stochastic factor has its own term structure signature that allows it to be distinguished from other factors. For example, a misfitting in skewness implied by options with short times to maturity is more likely to be associated with the short-run volatility factor (and possibly a jump process), while a misfitting in skewness implied by options with longer times to maturity is most likely to be associated with the long-run volatility factor. Similarly, a misfitting in short-term (long-term) volatility is likely to be related to the short-run (long-run) volatility factor.

First, I investigate whether the two-factor volatility models are capable of capturing the entire term structure of model-free implied volatility. The model-free implied volatilities are computed using the approach of Carr and Wu (2009) during the sample period 1993-2008. For an intuitive exposition, I provide evidence that it is often possible that (i) when short-term implied volatility is low (high), the term structure is downward (upward) sloping, and (ii) short- and long-term volatilities do not always move in lockstep. Such empirical features are suggestive of a time-

varying long-run expectation of volatility, which is probably caused by long-run news events. Single-factor volatility models, which assume constancy in the long-run volatility level, are clearly deficient for the purpose of option pricing. My results show that the two-factor volatility models do a good job of capturing the entire term structure of implied volatility, as opposed to the one-factor volatility models. In particular, the inclusion of a long-run volatility factor is critical in capturing levels of and changes in long-term implied volatility.

Second, I look at the effect of short-run and long-run volatility factors on the term structure of skewness. Since the 1987 market crash, options markets have exhibited an empirical regularity commonly referred to as the "volatility smirk": out-of-the-money puts (in-the-money calls) have higher Black-Scholes implied volatilities than in-the-money puts (out-of-the-money calls). Although there has been abundant research on negative skewness (or volatility smirk), little is known about the *term structure* of negative skewness. Do short-term options have more or less left-skewed return distributions than long-term options? To see this, I calculate implied skewness via the model-free approach of Bakshi, Kapadia, and Madan (2003) and find that the term structure of implied skewness is nearly *flat* on average. Simply put, long-dated skewness is as low (negative) as short-dated skewness. This empirical regularity is associated with the finding of Carr and Wu (2003) that volatility smirks are parallel across different terms to maturity when a standard measure of moneyness is used. To address this issue, they introduce a log stable Levy process with maximum negative skewness that deliberately violates the central limit theorem.

Alternatively, I find that the two-factor volatility models are capable of imitating the flat term structure of implied skewness, at least qualitatively. The reason is as follows. While it is well-known that a leverage effect can generate negative skewness, it is not the only variable to do this. There are two more parameters that are relevant for skewness: the persistence of volatility and the volatility of volatility. The persistence of volatility determines how long the leverage effect will hold up along an increasing horizon. For example, if a volatility factor is weakly persistent, the leverage effect (and thus the generated skewness) disappears relatively quickly as a horizon increases. The volatility of volatility is a key variable in determining the level of skewness as well.

With the leverage effect equal, a high level of volatility of volatility will generate more left-skewed return distributions than a low one.

According to my estimation results, the short-run and long-run factors involve very different levels of the skewness-related parameters. The short-run factor is characterized by weak persistence and high volatility of volatility, while the long-run factor is characterized by strong persistence and low volatility of volatility. Because of these differences, the two factors have opposing impacts on the term structure of skewness. The short-run factor can generate a large amount of negative skewness at short horizons thanks to the high volatility of volatility, but the skewness dissipates quickly along an increasing horizon due to the weak persistence. In contrast, the long-run factor can generate only a small amount of negative skewness at short horizons due to the low volatility of volatility, but the leverage effect holds up until a longer horizon due to the strong persistence. That is, while the short-run factor is dominant in generating short-term skewness, the long-run factor plays a pivotal role in generating long-term skewness. The different yet complementary roles played by the two volatility factors can produce flat term structures qualitatively.

However, the two-factor volatility models fall short of matching observed option-implied skewness precisely; specifically, model-implied skewness is higher (less negative) than option-implied skewness. This undesirable consequence will yield underpricings of out-of-the-money (OTM) puts and overpricings of OTM calls. This problem is not unique to this paper but has been widely reported in the existing option pricing literature (e.g., Bakshi, Cao, and Chen 1997; Christoffersen, Jacobs, and Mimouni 2010; Christoffersen, Jacobs, Ornthanalai, and Wang 2008). It is also in line with the empirical finding of Bakshi, Cao, and Chen (1997) and Bates (2000) that the volatility of volatility and the absolute value of leverage effect implied by option prices are implausibly large relative to their time-series-based counterparts. My estimates for the volatility of volatility and the absolute value of the leverage effect may not be sufficiently large to rationalize the levels of skewness implicit in option prices.[1]

[1] Since the maximum likelihood estimation used in this paper optimizes the models' one-step forecasting ability, my estimates are time-series estimates.

Having examined the term structures of volatility and skewness, I have done most of the work actually. In line with the preceding results, the two-factor volatility models enhance option pricing performance over the one-factor volatility models, mainly because they are able to fit the term structure of volatility much better. I also compare the proposed two-factor volatility models to the two-factor GARCH model of Christoffersen, Jacobs, Ornthanalai, and Wang (2008) and find that the former outperform the latter by 66.2% in terms of the root mean squared errors of the Black-Scholes implied volatilities. This improvement occurs largely because the former generates a more realistic term structure of skewness than the latter.

At last, I discuss a possible misspecification in (multifactor) stochastic volatility models. While stochastic volatility models have been successful in capturing time-variation in volatility, they can explain little of the variation in skewness that is a prominent feature of the options data. My investigation shows that the level of volatility is weakly correlated with the level of skewness, and that changes in volatility have little correlation with changes in skewness. This evidence suggests that the dynamics of skewness are very different from those of volatility. It may be necessary to introduce a novel *skewness factor* that is associated with the time-varying skewness, in addition to multiple volatility factors.[2]

The estimation strategy used in this paper makes it possible to identify skewness factors by analyzing option pricing errors. While I force the models to fit the observed term structure of volatility, the term structure of skewness remains unfit. As a result, a large fraction of option pricing errors occur along the moneyness dimension (for example, when the option market implies more negative skewness, pricing errors are typically larger). For this reason, I run regressions of option pricing errors on three potential skewness factors: stochastic volatility of volatility, stochastic leverage effect, and stochastic jumps. The regression results show that a large fraction of the option pricing errors can be explained by the skewness factors (with other control variables) and that the stochastic volatility of volatility is the most promising skewness factor among the variables

[2]A multifactor affine or CEV model is capable of generating some of time-variation in skewness. However, this model would be imperfect in fitting the stochastic skewness because the dynamics of volatility are loosely related to those of skewness.

considered in this paper. The results indicate that one potential extension of the two-factor volatility models would be to randomize the volatility of volatility in order to capture time-variation in skewness.

The rest of the paper is organized as follows. Section 1.2 clarifies the motivation of the paper, focusing on the empirical regularities of the options market. In Section 1.3, I describe the one- and two-factor volatility models under consideration. Section 1.4 develops the estimation strategy, focusing on how to handle unobservable volatility factors. The data are described in Section 1.5. The empirical results are provided in Section 1.6. Finally, Section 1.7 concludes and discusses possible directions for future work.

1.2 Empirical features of options data

I start my analysis by exploring the empirical features of option prices that are inconsistent with the implications of one-factor volatility models. In particular, I look at the term structures of volatility and skewness, as implied by the option market. This work is facilitated by the recently developed model-free approaches to implied volatility and skewness.

1.2.1 Term structure of implied volatility

The issue of whether a long-run volatility factor is necessary in describing the behavior of implied volatility dates back to as early as the overreaction puzzle introduced by Stein (1989). Stein finds that long-dated implied volatility changes more in response to changes in short-dated implied volatility than in response to those implied by the one-factor volatility model. The subsequent papers such as those of Diz and Finucane (1993) and Xu and Taylor (1994) argue that the overreaction puzzle is suggestive of time-variation in the long-run expectation of volatility.

The term structure of volatility will react in various ways according to the size and persistence of an arriving news event. Single-factor volatility models may be too restrictive to capture the dynamics of the term structure of implied volatility. To see this, assume the one-factor CEV

model (encompassing the affine model) under a risk-neutral $\mathbb{Q}$-measure:

$$dh_t = \kappa(\overline{h} - h_t)dt + \sigma h_t^{\gamma} dW_t^{\mathbb{Q}}, \tag{1.1}$$

where h_t is a volatility state; κ is the persistence of volatility; $\overline{h}$ is the long-run mean; σ is volatility of volatility; γ is the elasticity of volatility; and $W_t^{\mathbb{Q}}$ refers to the standard Brownian motion under the $\mathbb{Q}$-measure.

Under the one-factor CEV specification, the model-implied volatility $\mathrm{IV}_{t,\tau}^2$ can be derived as a linear function of h_t:[3]

$$\begin{aligned} \mathrm{IV}_{t,\tau}^2 &\equiv \frac{1}{\tau} E_t^{\mathbb{Q}} \left[\int_t^{t+\tau} h_s ds \right] \\ &= \beta(\tau) h_t + (1 - \beta(\tau))\overline{h}, \end{aligned} \tag{1.2}$$

where $\beta(\tau) = \frac{1-e^{-\kappa\tau}}{\kappa\tau}$.

Under the single-factor volatility assumption, the state h_t will determine all implied volatilities at time t with different terms to maturity. In particular, equation (1.2) implies the following linear relation between short-term implied volatility and the term premium of implied volatility:

$$\mathrm{IV}_{t,\tau_2}^2 - \mathrm{IV}_{t,\tau_1}^2 = \frac{\beta(\tau_1) - \beta(\tau_2)}{\beta(\tau_1)} \left(\overline{h} - \mathrm{IV}_{t,\tau_1}^2 \right), \tag{1.3}$$

where $\mathrm{IV}_{t,\tau_2}^2 - \mathrm{IV}_{t,\tau_1}^2$ measures the term premium of implied volatility between a long horizon τ_2 and a short horizon τ_1. $\beta(\tau_1) > \beta(\tau_2)$ since $\tau_2 > \tau_1$. This equation implies that the term premium is positive (negative) when the short-term implied volatility IV_{t,τ_1}^2 is lower (higher) than the long-run mean $\overline{h}$. That is, the term structure of implied volatility should be upward (downward) sloping when the short-term implied volatility is historically low (high). Furthermore, the relation should be deterministically linear up to measurement errors.

The strict relation above is often in contrast to the actual relation between implied volatilities. To see this, I extract model-free implied volatilities, following the approach of Carr and Wu

[3] With a slight abuse of terminology, $\mathrm{IV}_{t,\tau}$ and $\mathrm{IV}_{t,\tau}^2$ are both referred to as implied volatility in this paper.

(2009).[4] I then compute constant-maturity implied volatilities by interpolating/extrapolating two surrounding implied volatilities for each of 1, 2, 3, 6, 9, and 12 months to maturity. The top panel of Figure 1.1 presents a scatter plot between the term premium and the one-month implied volatility. The plot shows that the actual relation is far from a linear relation. In particular, the bottom-left (top-right) quadrant of the plot corresponds to the cases in which the term structures are downward (upward) sloping when short-term implied volatilities are historically low (high).

Another examination of the options data concerns changes in implied volatilities. Equation (1.2) implies the linear relation between changes in long- and short-term implied volatilities up to measurement errors:

$$d\mathrm{IV}^2_{t,\tau_2} = \frac{\beta(\tau_2)}{\beta(\tau_1)} d\mathrm{IV}^2_{t,\tau_1}. \tag{1.4}$$

The bottom panel of Figure 1.1 presents a scatter plot between changes in one-year and one-month implied volatilities. There is considerable deviation from a linear relation. Besides, short- and long-term volatilities do not always move in lockstep; the top-left and bottom-right quadrants of the plot correspond to the cases in which short- and long-term volatilities move in the opposite directions.

As is seen in the above, the single-factor volatility assumption is too restrictive to describe the term structure of implied volatility. An approach to addressing this issue is to recalibrate structural parameters on a daily basis (e.g., Bakshi, Cao, and Chen 1997; Duffie, Pan, and Singleton 2000; Barone-Adesi, Engle, and Mancini 2008). The recalibration, however, makes it difficult to understand the true data generating process (or the stochastic factors implicit in option prices) because it does not exploit any time series property of returns and implied volatility. In fact, when recalibrated, a dynamic model for option pricing loses its dynamic appeal, becoming close to the spirit of the local volatility model of Derman and Kani (1994) and Dupire (1994) and the implicit volatility function of Dumas, Fleming, and Whaley (1998).

Alternatively, the above empirical features can be reconciled with the existence of multiple

[4] Alternative approaches to model-free implied volatility have been developed by Neuberger (1994), Carr and Madan (1998), Demeterfi, Derman, Kamal, and Zou (1999), and Britten-Jones and Neuberger (2000), among others.

volatility factors, each of which has a distinct term structure effect according to its economic nature.[5] For this reason, I adopt a class of stochastic volatility models with short-run and long-run volatility factors, which are assumed to be directed by short-run and long-run news events. I then test their option pricing performance with a single set of parameters without recalibration.

For now, a question arises: is a third volatility factor necessary? To see this, I apply a principal component analysis of constant-maturity implied volatilities. Table 1.1 shows that the first two principal components combined explain 99.8% of variations in implied volitilities. The first and second components are associated with the level and the slope effects, respectively. The curvature effect of the third component appears to be trivial (at least up to the one-year horizon).[6] This result is consistent with the arguement of Li and Zhang (2010) that the two-factor structure proxied by short- and long-term implied volatilities is good enough to fit options data.

1.2.2 Term structure of implied skewness

Since the 1987 market crash, options markets have exhibited an empirical regularity commonly referred to as the “volatility smirk": out-of-the-money puts (in-the-money calls) have higher Black-Scholes implied volatilities than in-the-money puts (out-of-the-money calls). A commonly used approach to addressing this pricing bias is to introduce a stochastic volatility factor with a leverage effect or superimpose a jump process on top of a diffusive process.

Although a great deal of attention has been focused on understanding negative skewness, few studies have investigated the *term structure* of negative skewness. In this subsection, I fill this void by looking at the term structure of implied skewness. To do this, I compute the model-free skewness of Bakshi, Kapadia, and Madan (2003) and then run a kernel regression of the model-free skewness against term to maturity. The top panel of Figure 1.2 shows the regression result with a Gaussian kernel and a manually chosen bandwidth. Interestingly, the term structure of skewness

[5]The multifactor volatility models have been used to describe physical return dynamics (e.g., Gallant, Hsu, and Tauchen 1999; Chacko and Viceira 2003; Durham 2006).

[6]The curvature effect might be important if option-implied volatilities with more than one year to maturity were included.

is nearly flat on average. Simply put, long-dated skewness is as low (negative) as short-dated skewness.

The above empirical pattern is rather challenging from a modeling perspective. It cannot be explained by jumps in returns because the effect of the jumps on skewness dissipates quickly due to the central limit theorem; recall that jumps contribute to multihorizon return distributions in an independent fashion. A short-run volatility factor is also hard to reconcile with this feature because its weak persistence dissipates the leverage effect quickly as a horizon increases.[7] I find, on the other hand, that a long-run volatility factor can generate a significant amount of negative skewness at long horizons because its strong persistence delays the dissipation of the leverage effect along an increasing horizon.

In fact, the flat term structure of skewness is related to the finding of Carr and Wu (2003) that slopes of volatility smirks are almost the same across different times to maturity when a standard measure of moneyness is used. Similarly, I also look at slopes of volatility smirks with different times to maturity but with a new measure of moneyness:

$$d = \frac{\log(K/F_{t,\tau})}{\mathrm{IV}_{t,\tau}\sqrt{\tau}}, \tag{1.5}$$

where $\mathrm{IV}_{t,\tau}$ and $F_{t,\tau}$ refer to the model-free implied volatility and the implied future price at time t with time to maturity τ, respectively.[8] Notice that the moneyness is divided by the corresponding implied volatility in order to account for time-variation in volatility during a long sample period; unlike this moneyness measure, the moneyness measure of Carr and Wu (2003) does not account for time-varying volatility by using the all-time average of Black-Scholes implied volatilities in the denominator. An OTM put (call) option with a moneyness d will become exactly-at-the-money if the log future price decreases (increases) by d times the corresponding implied volatility.

Now, I obtain the average volatility surface over d and τ by using a kernel regression with

[7] For example, in the affine stochastic volatility model, the absolute value of skewness is proportional to the reciprocal of $\kappa\sqrt{\tau}$ as τ goes to infinity, where κ is the persistence of volatility.

[8] $F_{t,\tau}$ is calculated from the put-call parity relation using at-the-money call and put options for a given time to maturity (see Ait-Sahalia and Lo 1998; Ait-Sahalia and Lo 2000).

a Gaussian kernel and a manually chosen bandwidth. The bottom panel of Figure 1.2 shows an array of volatility smirks that are generated by slicing the volatility surface along 1, 3, 6, 9, and 12 months to maturity. Interestingly, the volatility smirks with different times to maturity not only are parallel, but also collapse into a single line. As is mentioned previously, this similarity of volatility smirks is related to the fact that the term structure of skewness is nearly flat on average.

1.2.3 Stochastic skewness

While the preceding subsection 1.2.2 concerns the average term structure of skewness, this subsection discusses variation over time in skewness. The options data show that implied skewness is time-varying like implied volatility. The variation in skewness is potentially affected by a combination of multiple stochastic factors. Durham and Park (2010) find that stochastic volatility of volatility and a stochastic leverage effect are both associated with time-variation in implied skewness. Moreover, jumps are probably related to variation in implied skewness (at least at short horizons). Carr and Wu (2007) propose a stochastic skew model for foreign exchange rates with positive and negative jumps driven by independent Levy processes.

While most of the literature on option pricing focuses on modeling stochastic volatility, the stochastic volatility is unlikely to address stochastic skewness. The top panel of Figure 1.3 presents a scatter plot of implied skewness against implied volatility. The plot suggests that implied volatility is weakly related to implied skewness. The bottom panel of Figure 1.3 displays a scatter plot of changes in implied skewness against changes in implied volatility. The plot shows that there is little correlation between changes in volatility and changes in skewness. In sum, the dynamics of skewness are different from those of volatility, so that modeling stochastic volatility appears to have little chance of success in fully capturing stochastic skewness.

What causes skewness to change? The answer to this question will guide a further extension of the existing two-factor volatility models. Recall that while I force the models to fit the term structure of implied volatility, I do not use any information on skewness in estimation. As a result, option pricing errors should be associated with stochastic skewness. By analyzing option pricing

errors, I can identify skewness factors that are responsible for variation in implied skewness.[9]

How many skewness factors are necessary to describe the term structure of implied skewness? To see this, I calculate constant-maturity implied skewness for each of 1, 2, 3, 6, 9, and 12 months to maturity and run a principal component analysis of the implied skewness. Table 1.2 shows that the level effect (first principal component) and the slope effect (second principal component) can explain 77% and 13% of variation in skewness, respectively. Even the curvature effect (third principal component) is not trivial, explaining 4.6% of variation in skewness. Clearly, there are multiple skewness factors that drive the evolution of the term structure of skewness.

1.3 Models

The modeling framework in this paper is based on a log stochastic volatility model with short-run and long-run volatility factors. Each of the two volatility factors is assumed to follow the Ornstein-Uhlenbeck process. The short-run (long-run) factor determines short-term (long-term) trends and fluctuations in volatility. This form is sometimes referred to as the volatility component model (e.g., Lu and Zhu 2010). An alternative approach to constructing a multifactor volatility model is the stochastic long-run mean model in which the long-run mean of volatility is randomized (e.g., Duffie, Pan, and Singleton 2000; Egloff, Leippold, and Wu 2010; Christoffersen, Jacobs, Ornthanalai, and Wang 2008). These two forms of multifactor volatility models are equivalent within the log volatility framework.[10]

A large body of literature documents the significance of jumps in returns and volatilities (e.g., Eraker, Johannes, and Polson 2003; Bates 2000; Pan 2002; Andersen, Benzoni, and Lund 2002). In particular, the addition of jumps helps to generate a significant amount of negative skewness at a very short horizon by making conditional distributions non-normal. Accordingly, I augment the underlying two-factor log volatility model by incorporating jumps in returns and the short-run factor.

[9] I use the term skewness factor to refer to the stochastic factor associated with time-variation in option-implied skewness.

[10] This is not the case for the affine and CEV frameworks.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and information filtration $\{\mathcal{F}_t\}$, the ex-dividend log stock price, y_t, is assumed to evolve as

$$\begin{aligned} dy_t &= \left[\mu - \frac{1}{2}h_t - \lambda\overline{\mu}_{1J}\right]dt + \sqrt{h_t}\left[dW_{1t} + J_{1t}dN_t\right] \\ dv_t &= -\kappa_v v_t dt + \sigma_v\left[dW_{2t} + J_{2t}dN_t - \lambda\mu_{2J}dt\right] \\ du_t &= -\kappa_u u_t dt + \sigma_u dW_{3t} \\ h_t &= \exp(\alpha + v_t + u_t), \end{aligned} \tag{1.6}$$

where v_t and u_t indicate the short-run and long-run volatility factors, respectively; h_t is a volatility state; μ means the drift rate of the stock price; W_{1t}, W_{2t}, and W_{3t} are the standard Brownian motions under the $\mathbb{P}$-measure, with $\text{corr}(W_{1t}, W_{2t}) = \rho_v$, $\text{corr}(W_{1t}, W_{3t}) = \rho_u$, and $\text{corr}(W_{2t}, W_{3t}) = 0$.

Notice that the two volatility factors have their own persistence κ_v and κ_u, their own volatility of volatility σ_v and σ_u, and their own leverage effects ρ_v and ρ_u. While both v_t and u_t are set to have the long-run mean of zero under the $\mathbb{P}$-measure, the long-run mean of h_t is captured through α under the $\mathbb{P}$-measure.

The jump structures are modeled as low-frequency compound Poisson processes. I posit that jumps in y_t and v_t are contemporaneously driven by a common Poisson process N_t with intensity λ. (J_{1t}, J_{2t}) are bivariate normal with $J_{1t} \sim N\left(\mu_{1J}, \sigma_{1J}^2\right)$, $J_{2t} \sim N\left(\mu_{2J}, \sigma_{2J}^2\right)$, and $\text{corr}(J_{1t}, J_{2t}) = \rho_J$. Notice that the final jump distribution in y_t is $\sqrt{h_t}J_{1t}$. Durham and Park (2010) find that this form of jumps fits the data better than the identically distributed jump structure used by Bakshi, Cao, and Chen (1997), Bates (2000), and Broadie, Chernov, and Johannes (2007). The volatility-scaled jump form offers more realistic dynamics by allowing for larger (smaller) jumps in higher (lower) volatility states. The terms $\lambda\overline{\mu}_{1J}$ and $\lambda\mu_{2J}$ are jump compensators for y_t and v_t, respectively, where $\overline{\mu}_{1Jt} = E_t(e^{\sqrt{h_t}J_{1t}} - 1)$ is time-varying.

There are several differences between two-factor affine and log volatility models from a modeling perspective. The two-factor affine model is able to produce a stochastic volatility of volatility and a stochastic leverage effect, determined by relative levels of the two volatility factors. In contrast, the two-factor log volatility model has a constant volatility of volatility and a constant

leverage effect. To see this, let me define volatility of volatility as the variance of a rate of changes in volatility per a unit time, $\text{var}(dh_t/h_t)/dt$, and the leverage effect as the correlation between changes in y_t and h_t, $\text{corr}(dy_t, dh_t)$. Using Ito's lemma, these variables are readily derived in the case without jumps

$$\begin{aligned} \frac{\text{var}(dh_t/h_t)}{dt} &= \left(\sigma_v^2 + \sigma_u^2\right) \\ \text{corr}(dy_t, dh_t) &= \frac{\rho_v \sigma_v + \rho_u \sigma_u}{\sigma_v^2 + \sigma_u^2}. \end{aligned} \tag{1.7}$$

While the observed time series data are driven by the above model under the $\mathbb{P}$-measure, options are priced under a risk-neutral $\mathbb{Q}$-measure. Under the $\mathbb{Q}$-measure, the stock price dynamics take the following form:

$$\begin{aligned} dy_t &= \left[r_t - q_t - \frac{1}{2}h_t - \lambda\bar{\mu}_{1J}^{\mathbb{Q}}\right]dt + \sqrt{h_t}\left[dW_{1t}^{\mathbb{Q}} + J_{1t}^{\mathbb{Q}}dN_t\right] \\ dv_t &= (-\kappa_v v_t - \eta_v)dt + \sigma_v\left[dW_{2t}^{\mathbb{Q}} + J_{2t}dN_t - \lambda\mu_{2J}dt\right] \\ du_t &= (-\kappa_u u_t - \eta_u)dt + \sigma_u dW_{3t}^{\mathbb{Q}} \\ h_t &= \exp(\alpha + v_t + u_t), \end{aligned} \tag{1.8}$$

where r_t and q_t denote the risk-free rate and the dividend rate, respectively. $W_{1t}^{\mathbb{Q}}$, $W_{2t}^{\mathbb{Q}}$, and $W_{3t}^{\mathbb{Q}}$ are the standard Brownian motions under the $\mathbb{Q}$-measure, with the same correlations as those under the $\mathbb{P}$-measure.

Any discrepancy in the stock price dynamics between the $\mathbb{P}$- and $\mathbb{Q}$-measures is attributed to risk premiums. I account for two types of risk premiums: variance and jump risk premiums. The variance risk premium is captured through the parameters η_v and η_u in equation (1.8). Based on the intuition of intertemporal asset pricing framework, the existing papers often assume a linear function of the volatility state as a specification of the variance risk premium in the affine stochastic volatility models (e.g., Heston 1993; Bollerslev, Gibson, and Zhou 2011). However, the form of a variance risk premium is not unambiguously defined because the market is incomplete under the existence of a stochastic volatility factor; there can be an infinite number of equivalent martingale measures in this case. My empirical tests show that the variance risk premium form adopted in

this paper fits the data better than a linear function of the volatility state. I also consider a jump risk premium by assuming $J_{1t}^{\mathbb{Q}} \neq J_{1t}$. Specifically, following the approach of Pan (2002), I posit that the whole jump risk premium is absorbed through the different mean jump sizes between the $\mathbb{P}$- and $\mathbb{Q}$-measures. That is, $J_{1t}^{\mathbb{Q}}$ is assumed to follow a normal distribution with mean $\mu_{1J}^{*} \neq \mu_{1J}$ and variance σ_{1J}. N_t and J_{2t} are assumed to be identical between the $\mathbb{P}$- and $\mathbb{Q}$-measures. The correlation between $J_{1t}^{\mathbb{Q}}$ and J_{2t} is the same as that between J_{1t} and J_{2t}.

Unfortunately, I cannot estimate the variance and jump risk premiums simultaneously. It is hard to disentangle the two types of risk premiums when the models are estimated by using the information on implied volatilities only. A difficulty of this kind is noted by Pan (2002). In fact, the jump risk premium is more likely to be implied by the prices of deep OTM put options or implied skewness (rather than the term structure of implied volatility). Through an array of empirical tests, I seek to decide which of the risk premiums to include in my applications. I consider the jump risk premium in the one-factor volatility models, as in Pan (2002), while I account for the variance risk premium in the two-factor volatility models.[11]

Finally, in my empirical work, I will apply an Euler scheme approximation to the above continuous-time model. For the physical model (and analogously for the risk-neutral model), the approximation is given by

$$\begin{aligned} y_{t+1} &= y_t + \mu - \frac{1}{2}h_t - \lambda\bar{\mu}_{1Jt} + \sqrt{h_t}\left[\varepsilon_{1,t+1} + \sum_{j>N_t}^{N_{t+1}} \xi_{1j}\right] \\ v_{t+1} &= v_t - \kappa_v v_t + \sigma_v\left[\varepsilon_{2,t+1} + \sum_{j>N_t}^{N_{t+1}} \xi_{2j} - \lambda\mu_{2J}\right] \\ u_{t+1} &= u_t - \kappa_u u_t + \sigma_u \varepsilon_{3,t+1}, \end{aligned} \tag{1.9}$$

where $\varepsilon_{1,t+1}$, $\varepsilon_{2,t+1}$, and $\varepsilon_{3,t+1}$ are standard normals with $\text{corr}(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}) = \rho_v$, $\text{corr}(\varepsilon_{1,t+1}, \varepsilon_{3,t+1}) = \rho_u$, and $\text{corr}(\varepsilon_{2,t+1}, \varepsilon_{3,t+1}) = 0$. ξ_{1j} and ξ_{2j} have the same distribution as the jumps in the continuous-time model as well. For computational purposes, I limit the maximum number of jumps occurring

[11] I find that accounting for the variance risk premium is more important in the two-factor volatility models than accounting for the jump risk premium.

on a single trading day to three. Given my estimate for jump intensity, this constraint does not affect the empirical results.

The model described above nests a range of model specifications with different parsimonies, including the Black-Scholes model. Model specifications used in this paper are described in Table 1.3.

1.4 Estimation Strategy

Unlike the GARCH option pricing models (e.g., Duan 1995; Heston and Nandi 2000), option pricing models with stochastic volatility pose a great deal of econometric challenge because volatility is unobservable. To address this issue, the existing papers in the option pricing literature often apply two-step procedures (e.g., Huang and Wu 2004; Broadie, Chernov, and Johannes 2007; Christoffersen, Heston, and Jacobs 2009). The idea of the two-step procedures is to estimate volatility factor(s) on each day using the day's cross-sectional option prices in the first step and to minimize the sum of option pricing errors over structural parameters by using the volatility factor(s) obtained in the first step. However, the methods are computationally unfeasible for the log volatility framework because there is no closed-form option pricing expression. While adopting a two-step procedure, I need to make a compromise. In the first step, I back out the short-run and long-run volatility factors on each day such that the models fit a pair of option-implied volatilities instead of a full cross-section of option prices; and in the second step, optimize the log likelihood using the returns and volatilities obtained in the first step. A detailed description of each step of the estimation follows.

1.4.1 Extracting the volatility factors

I estimate the physical and risk-neutral dynamics simultaneously by using the joint time series of returns and option-implied volatilities, building on the work of Chernov and Ghysels (2000), Pan (2002), and Ait-Sahalia and Kimmel (2007). The joint estimation is supported by empirical evidence that option-implied volatility is more informative than realized (or historical)

volatility in forecasting a future volatility of an underlying asset (e.g., Blair, Poon, and Taylor 2001; Christensen and Prabhala 1998; Fleming 1998; Jiang and Tian 2005). Besides, the physical and risk-neutral models should, in theory, be linked by sharing common parameters such as volatility of volatility and the leverage effect. In what follows, I provide a detailed description of this step for the two-factor volatility models.

The estimation starts by pinning down the two volatility factors on each day by matching two option-implied volatilities to their model-implied counterparts. Specifically, I use the time series of one- and three-month implied volatilities. For now, the problem at hand is interpreted as solving a system of two equations:

$$\text{Find } (v_t, u_t) \text{ such that } \mathrm{IV}_{t,\tau_1} = \mathrm{IV}^{\mathcal{M}}_{t,\tau_1} \text{ and } \mathrm{IV}_{t,\tau_2} = \mathrm{IV}^{\mathcal{M}}_{t,\tau_2}, \;\; t = 1,\ldots,N, \tag{1.10}$$

where $\mathrm{IV}_{t,\tau}$ and $\mathrm{IV}^{\mathcal{M}}_{t,\tau}$ refer to the option-implied and model-implied volatilities, respectively. While $\mathrm{IV}_{t,\tau}$ can be computed via the approach of Carr and Wu (2009), $\mathrm{IV}^{\mathcal{M}}_{t,\tau}$ can be obtained by integrating the quadratic variation of the log stock price, given a risk-neutral model and two volatility factors.[12] For the models with jumps in returns, for example, I get

$$\mathrm{IV}^{\mathcal{M}}_{t,\tau} = \sqrt{\frac{1}{\tau} E^{\mathbb{Q}}_t \left\{ \int_t^{t+\tau} h_\omega \Big[1 + \lambda(\mu_{1J}^2 + \sigma_{1J}^2) \Big] \, d\omega \right\}}.$$

The expectation in the above expression can be computed by means of Monte Carlo simulations. For the models with jumps in returns, for example, I get

$$\mathrm{IV}^{\mathcal{M}}_{t,\tau} = \sqrt{\frac{1}{\tau S} \sum_{s=1}^{S} \left\{ \int_t^{t+\tau} h_\omega^{(s)} \Big[1 + \lambda(\mu_{1J}^2 + \sigma_{1J}^2) \Big] \, d\omega \right\}}, \tag{1.11}$$

where for each s, $h_\omega^{(s)}$ is obtained by simulating a path from the risk-neutral analog of equation (1.9) for $t < \tau \leq t + \tau$, and S denotes the number of simulation paths.

[12] In the two-factor affine or CEV models, $\mathrm{IV}^{\mathcal{M}}_{t,\tau}$ can be expressed as a bilinear function of the two volatility states, so that equation (1.10) can be easily solved. However, this is not the case for the two-factor log volatility model.

Given the observed $\mathrm{IV}_{t,\tau}$ and the simulated $\mathrm{IV}^{\mathcal{M}}_{t,\tau}$, it is possible, in principle, to solve equation (1.10) for v_t and u_t. However, it is still computationally too intensive to repeat the work on each date in the entire sample. I address this obstacle by introducing a computationally manageable algorithm. Equation (1.11) implies that there exists the mapping $\Gamma = (\gamma_1, \gamma_2) : \mathbb{R}^2 \to \mathbb{R}^2$ from the volatility factors to the log implied volatilities:

$$\Gamma : (v_t, u_t) \to (\log \mathrm{IV}^{\mathcal{M}}_{t,\tau_1}, \log \mathrm{IV}^{\mathcal{M}}_{t,\tau_2}).$$

Γ relies on a set of parameters θ and a model $\mathcal{M}$, although θ and $\mathcal{M}$ are suppressed from the notation of Γ. As long as the mapping is one-to-one, there exists the implicit inverse mapping $\Gamma^{-1} = (\gamma_1^{-1}, \gamma_2^{-1}) : \mathbb{R}^2 \to \mathbb{R}^2$ from the log implied volatilities to the volatility factors:

$$\Gamma^{-1} : (\log \mathrm{IV}^{\mathcal{M}}_{t,\tau_1}, \log \mathrm{IV}^{\mathcal{M}}_{t,\tau_2}) \to (v_t, u_t).$$

The idea of a computationally efficient technique is to approximate the inverse mapping Γ^{-1} via a curve fitting technique. In particular, I posit that each of γ_1^{-1} and γ_2^{-1} is a bicubic polynomial function of $(\log \mathrm{IV}^{\mathcal{M}}_{t,\tau_1}, \log \mathrm{IV}^{\mathcal{M}}_{t,\tau_2})$. The detailed procedure is as follows.

Grid generation: It starts by generating a two-dimensional grid over v_t and u_t. Let $\{\widehat{v}_g, \widehat{u}_g\}_{g=1}^{G}$ be the grid, where G is the number of grid points and I use hats to indicate that these are grid points rather than data.

Simulation: For each $g = 1, \ldots, G$, then evaluate $(\log \widehat{\mathrm{IV}}^{\mathcal{M}}_{g,\tau_1}, \log \widehat{\mathrm{IV}}^{\mathcal{M}}_{g,\tau_2})$ using Monte Carlo methods as described above. Now I have $\{\log \widehat{\mathrm{IV}}^{\mathcal{M}}_{g,\tau_1}, \log \widehat{\mathrm{IV}}^{\mathcal{M}}_{g,\tau_2}\}_{g=1}^{G}$, which correspond to $\{\widehat{v}_g, \widehat{u}_g\}_{g=1}^{G}$. I use these collections of pairs as the basis for finding the approximate inverse mapping $\widehat{\Gamma}^{-1} = (\widehat{\gamma}_1^{-1}, \widehat{\gamma}_2^{-1})$.

Regression: Find $\widehat{\gamma}_1^{-1}$ by running a bicubic polynomial regression of $\{\widehat{v}_g\}_{g=1}^{G}$ on $\{\log \widehat{\mathrm{IV}}^{\mathcal{M}}_{g,\tau_1}, \log \widehat{\mathrm{IV}}^{\mathcal{M}}_{g,\tau_2}\}_{g=1}^{G}$. Repeat a similar procedure to find $\widehat{\gamma}_2^{-1}$.

In principle, there are many curve fitting schemes one could use. However, since I am using a gradient in a numerical optimization, it is best if the scheme results in smooth derivatives. I have empirically found that simply using bicubic polynomials works well; approximation errors are reasonably small. Once I have obtained $\widehat{\Gamma}^{-1}$, the hard work is done. Now, solving Equation (1.10) boils down to evaluating the approximating mapping $\widehat{\Gamma}^{-1}$ at a pair of option-implied volatilities $(\log \mathrm{IV}_{t,\tau_1}, \log \mathrm{IV}_{t,\tau_2})$. Repeat this for each $t = 1,\ldots,N$. The important thing to notice here is that computing $\widehat{\Gamma}^{-1}$, which is the costly step, needs to be done only once (for each set of candidate parameter values).

1.4.2 Maximum likelihood estimation

Having backed out the volatility factors $\{v_t, u_t\}_{t=1}^N$, we can apply an implied-state maximum likelihood estimation by treating those factors as observable. It is straightforward to compute the log likelihood for a set of parameters θ:

$$\log L(\{y_t, \mathrm{IV}_{t,\tau_1}, \mathrm{IV}_{t,\tau_2}\}_{t=1}^N; \theta) \approx \sum_{t=1}^{N-1} \left[\log p(y_{t+1}, v_{t+1}, u_{t+1} | y_t, v_t, u_t) + \log |J_{t+1}| \right], \tag{1.12}$$

where $|J_{t+1}|$ is the determinant of the Jacobian matrix $J_{t+1} = \frac{\partial(v_{t+1}, u_{t+1})}{\partial(\mathrm{IV}_{t+1,\tau_1}, \mathrm{IV}_{t+1,\tau_2})}$ and N is the number of business days. I make use of the approximating mapping $\widehat{\Gamma}^{-1}$ when calculating J_{t+1}. Since I allow for the possibility of multiple jumps a day, the local likelihood $p(y_{t+1}, v_{t+1}, u_{t+1} | y_t, v_t, u_t)$ can be computed by applying the law of total probability over the number of potential jumps:

$$p(y_{t+1}, v_{t+1}, u_{t+1} | y_t, v_t, u_t) = \sum_{i=0}^{\infty} p(y_{t+1}, v_{t+1}, u_{t+1} | y_t, v_t, u_t, \Delta N_{t+1} = i) p(\Delta N_{t+1} = i), \tag{1.13}$$

where $\Delta N_{t+1} = N_{t+1} - N_t$. In practice, I limit the maximum number of jumps on a single trading day to three. The term $p(\Delta N_{t+1} = i)$ is given by the Poisson density function with intensity λ. The term $p(y_{t+1}, v_{t+1}, u_{t+1} | y_t, v_t, u_t, \Delta N_{t+1} = i)$ is trivariate normal with mean and variance given by summing the means and variances of the diffusive and jump parts of the process, according to equation (1.9).

Parameter estimates, $\widehat{\theta}$, are obtained by using a BHHH optimizer (Berndt, Hall, Hall, and Hausman 1974),

$$\widehat{\theta} = \arg\max \log L(\{y_t, \mathrm{IV}_{t,\tau_1}, \mathrm{IV}_{t,\tau_2}\}_{t=1}^{N}; \theta).$$

1.5 Data

The application uses daily S&P 500 index options data from January 1, 1993 through December 31, 2008 ($N = 4029$). These data were obtained directly from the CBOE. The CBOE option market closes 15 minutes later than the S&P 500 index market. To address the issue of nonsynchronous trading hours, I make use of the spot and future prices computed using the put-call parity based on closing prices for at-the-money options (see, e.g., Ait-Sahalia and Lo 1998; Ait-Sahalia and Lo 2000). Options are European, so there is no issue regarding early exercise premium. Option prices are taken from the bid-ask midpoint on each day's close of the option market. Options with zero bid/ask prices or where the bid-ask midpoint is less than 0.125 are discarded. I also eliminate options violating the usual lower bound constraints. That is, I require

$$C(t,\tau,K) \geq \max(0, S_{t,\tau}\exp(-q_t\tau) - K\exp(-r_t\tau))$$

$$P(t,\tau,K) \geq \max(0, K\exp(-r_t\tau) - S_{t,\tau}\exp(-q_t\tau)),$$

where $C(t,\tau,K)$ and $P(t,\tau,K)$ are the time t prices of call and put options with time to maturity τ and strike price K, $S_{t,\tau}$ is the time t implied spot price of the S&P 500 index with time to maturity τ, q_t is the dividend payout rate, and r_t is the risk-free rate. Treasury bill rates (obtained from the Federal Reserve website) are interpolated to match options' times-to-maturity as a proxy for the risk-free rate. Dividend rates are obtained from Standard and Poor's information bulletin.

When extracting model-free volatility and skewness, I discard options with terms-to-maturity of less than eight calendar days, as in the calculation of the VIX index. Jiang and Tian (2007) report the possibility of large truncation and discretization errors in the VIX index. To reduce such errors, I use the scheme of interpolating/extrapolating the Black Scholes implied volatilities on a fine grid

across moneyness, similar to that of Carr and Wu (2009). Additionally, I require that valid prices exist for at least one (two) observation(s) for each of the OTM call and put options in calculating the implied volatility (skewness). Figure 1.4 plots the time series of returns, implied volatility, the term premium of implied volatility, and implied skewness.

Option prices can be decomposed into intrinsic and time values. Only the time value is affected by the (mis-)specification of option pricing models, so that I use OTM call and put options in testing option pricing performance. Besides, OTM options are more liquid than in-the-money options. However, extremely deep OTM options may not contain accurate pricing information because they are not actively traded and their prices are too small relative to bid-ask spreads. For these reasons, I consider OTM options with moneyness ranging between $-3.0 < d < 1.5$, where d is defined in Equation (1.5). This restraint eliminates 13% of OTM put options and 7% of OTM call options. Although it seems tight, the constraint is less strict than those used by Dumas, Fleming, and Whaley (1998) and Huang and Wu (2004). I require that term to maturity be between 10 calendar days and one year. Table 1.4 presents the summary statistics of options across different maturity bins. I split the sample into short-term (less than two months), mid-term (two to six months), and long-term (six to twelve months) options along the maturity dimension.

When performing a pricing error analysis, I make use of the high-frequency-based realized volatility and jump variation. The high-frequency data were obtained from TickData.com. Following Andersen, Bollerslev, Diebold, and Ebens (2001), Andersen, Bollerslev, Diebold, and Labys (2003), and Barndorff-Nielsen and Shephard (2002), I obtain daily realized volatility by summing the squared intraday returns over each day,

$$\mathrm{RV}_t^{(d)} \equiv \sum_{j=1}^{1/\Delta} \left(y_{t-1+j\Delta} - y_{t-1+(j-1)\Delta} \right)^2$$

where Δ is the sampling interval for the intraday data (I use five minute intervals). Monthly realized

volatility is obtained by summing the daily realized volatilities over the previous month,

$$\mathrm{RV}_t \equiv \sum_{i=0}^{21} \mathrm{RV}_{t-i}^{(d)}.$$

Following Carr and Wu (2009), I define the variance risk premium as the log difference between monthly realized variance and option-implied variance,

$$\mathrm{VRP}_t \equiv \log\left(\mathrm{RV}_t/\mathrm{VIX}_t^2\right),$$

where VIX_t is the VIX index, divided by $\sqrt{12}$ to get a monthly volatility measure comparable to RV_t. I use the log difference because I find that it provides a better measure than the difference in levels.

A measure of jump risk is obtained by using the approach of Barndorff-Nielsen and Shephard (2004). The bipower variation is given by

$$\mathrm{BV}_t^{(d)} \equiv \frac{\pi}{2} \sum_{j=2}^{1/\Delta} \left(y_{t-1+j\Delta} - y_{t-1+(j-1)\Delta}\right)\left(y_{t-1+(j-1)\Delta} - y_{t-1+(j-2)\Delta}\right).$$

The daily jump variation is defined by subtracting the daily bipower variation from the daily realized volatility,

$$\mathrm{JV}_t^{(d)} \equiv \max(\mathrm{RV}_t^{(d)} - \mathrm{BV}_t^{(d)}, 0).$$

And, finally, the monthly jump variation is obtained by summing the daily jump variations over the previous month,

$$\mathrm{JV}_t \equiv \sum_{i=0}^{21} \mathrm{JV}_{t-i}^{(d)}.$$

I typically refer to RV_t and JV_t as the realized volatility and jump variation hereafter (omitting the “monthly" specifier). Note that these variables are defined by using information available at time t.

1.6 Empirical results

Subsection 1.6.1 tests the models' ability to fit the term structure of model-free implied volatility. Subsection 1.6.2 investigates whether the models are capable of matching the term structure of implied skewness. These two preceding results imply option pricing performance, which is discussed in subsection 1.6.3. Lastly, in subsection 1.6.4, I analyze option pricing errors to identify the skewness factors implicit in option prices.

1.6.1 Term structure of volatility

While the two-factor volatility models are forced to fit two implied volatilities, more than two implied volatilities are observable on each day. The S&P 500 index option maturities have three near-term months followed by three additional months based on the March quarterly cycle. The daily average number of implied volatilities with less than one year to maturity is 5.83. Thus, the options market will show various movements in the term structure of volatility according to the economic nature of an arriving news event. This subsection examines whether the two-factor volatility models in question are good enough to capture the entire term structure of implied volatility.

I begin by investigating whether the two-factor (SV2) model provides flexibility in fitting the term structure of implied volatility. Figure 1.5 plots the term structure of model-implied volatility computed by means of Monte Carlo simulations, with some hypothetical values of v_t and u_t. The top panel is a representative case in which h_t is low (v_t and u_t are set in such a way that h_t is 10%). Notice that the SV2 model allows for the various term structures of volatility, depending on a combination of v_t and u_t. The figure shows that when h_t is low, the SV2 model is capable of generating both the increasing and decreasing term structures, while the one-factor (JYV1) model can generate only the increasing term structure. Similarly, the bottom panel is a representative case in which h_t is high (v_t and u_t are set in such a way that h_t is 30%). The SV2 model is able to produce both the increasing and decreasing term structures when h_t is high, as opposed to the

one-factor (JYV1) model.

Figure 1.6 exhibits actual fits in the term structure of volatility for two chosen days. The top panel is a case in which the term structure is sharply declining, even though short-term implied volatility is low. On the other hand, the bottom panel is a case in which the term structure is rapidly increasing, even though short-term implied volatility is high. The one-factor volatility models are, of course, unable to fit the data for these days. The figure exhibits that the SV2 (solid line), JY2 (dashed line), and JYV2 (dash-dot line) models are all good at capturing the observed term structures (marked cross).

Now I introduce the two metrics for evaluating the models' fit in the term structure of implied volatility:

$$\begin{aligned} \text{IV-RMSE} &= \sqrt{\frac{1}{\sum_{t=1}^{N} NV_t} \sum_{t=1}^{N} \sum_{\tau=1}^{NV_t} \left(\text{IV}_{t,\tau}^{\mathcal{M}} - \text{IV}_{t,\tau} \right)^2} \\ \text{RIV-RMSE} &= 100 \times \sqrt{\frac{1}{\sum_{t=1}^{N} NV_t} \sum_{t=1}^{N} \sum_{\tau=1}^{NV_t} \left(\frac{\text{IV}_{t,\tau}^{\mathcal{M}} - \text{IV}_{t,\tau}}{\text{IV}_{t,\tau}} \right)^2}, \end{aligned} \tag{1.14}$$

where N stands for the total number of business days and NV_t is the number of model-free implied volatilities at time t with less than one year to maturity. While IV-RMSE measures the root mean squared fitting errors in implied volatility, RIV-RMSE does this in a relative sense by accounting for levels of implied volatility.

The top panel of Table 1.6 reports the aggregate results of IV-RMSE and RIV-RMSE across the different models. First of all, the SV1 model badly mis-fits the term structure of implied volatility, with a IV-RMSE of 34.73% and a RIV-RMSE of 176.25%, because it generates the unrealistically increasing term structure of volatility. This undesirable result has also been found by Jones (2003), Pan (2002), and Chourdakis and Dotsis (2009), among others. This problem can be addressed by accounting for jumps in returns, as argued by Pan (2002). As in Pan's result, the JV1 model markedly improves the fitting errors in implied volatility over the SV1 model. In the JY1 model, IV-RMSE and RIV-RMSE are 2.73% and 15.37%, respectively. The JYV1 model is able to make a little bit more of improvement over the JV1 model. In the JYV1 model, IV-RMSE and RIV-RMSE are 2.48% and 13.58%, respectively.

All two-factor volatility models offer substantial improvements over any of the one-factor volatility models. In the two-factor volatility models, IV-RMSEs fall between 0.81% and 1.19%, depending on whether jumps in returns and volatility are included. The within-two-factor rankings are SV2 $>$ JY2 $>$JYV2. The best performing SV2 model gives a IV-RMSE of 0.81% and a RIV-RMSE of 3.46%. This size of errors should not be interpreted as disappointing, since I attempt to fit a long panel dataset of 16 years containing various volatility ups and downs with a single parameterization.

The bottom panel of Table 1.6 tabulates the disaggregate results of IV-RMSE and RIV-RMSE along varying levels of volatility. In the table, I define the low, mid, and high volatilities as days when the day's average implied volatility is less than 0.25 quantile, between 0.25 and 0.75 quantile, and larger than 0.75 quantile. IV-RMSE and RIV-RMSE are both large in periods of high volatility.

There has been debate on the appropriate number of volatility factors to describe option prices. The related literature includes Li and Zhang (2010), Lu and Zhu (2010), and Bondarenko (2007), among others. While the introduction of a third volatility factor can potentially reduce fitting errors in implied volatility, my empirical results imply that the improvement would not be large. Besides, the benefit of a third volatility factor comes at the cost of greater econometric complexity. Accordingly, I view two-factor volatility models as a reasonable compromise that attains both econometric convenience and tolerable accuracy.

1.6.2 Term structure of skewness

Section 1.2.2 shows that the term structure of skewness implicit in the option market is nearly flat on average. This stylized fact is associated with the similarity of volatility smirks. This subsection discusses whether the two-factor volatility models are capable of matching the option-implied skewness.

As far as a diffusive volatility process is concerned, three parameters are central for generating negative skewness: the leverage effect, the persistence of volatility, and the volatility of volatility. It is crucial to understand the role played by each of these parameters in generating

negative skewness.

Leverage effect: The leverage effect is widely recognized as one of the main mechanisms that are able to generate left-skewed return distributions. Let me clarify how the leverage effect is capable of generating such negative skewness in the case of discrete-time models. Under the existence of the leverage effect, the volatility of tomorrow's return depends on the realization of today's return. For example, if today's realized return is negative (positive), tomorrow's return will have higher (lower) volatility than today's return. This asymmetric response of tomorrow's volatility to today's return generates left-skewed return distributions over multiple horizons (even though the one-step return distribution is normal). With everything else equal, a stronger leverage effect will generate more left-skewed return distributions than a weaker one. For example, it is possible to derive an analytical form of skewness in the affine stochastic volatility model (see Das and Sundaram 1999). In this case, the skewness is proportional to the leverage effect.

Persistence of volatility: While the leverage effect generates negative skewness, it is not the only important variable to do this. Notice that the persistence of volatility determines how long the asymmetric response will hold up along an increasing horizon. If a volatility factor is weakly persistent, the weak persistence will dissipate the leverage effect (and thus the generated skewness) quickly along an increasing horizon. In contrast, a strongly persistent volatility factor delays the dissipation of the leverage effect along an increasing horizon, allowing for a substantial amount of negative skewness until a longer horizon. For example, in the affine stochastic volatility model, the absolute value of the skewness is proportional to the reciprocal of the persistence, as time to maturity approaches infinity.

Volatility of volatility: Little is known about how volatility of volatility is related to skewness. Volatility of volatility cannot generate negative skewness in the absence of a leverage effect but is a key variable in determining the degree of negative skewness in the presence of a leverage effect. Given a level of the leverage effect, a high level of volatility of volatility

will generate more left-skewed return distributions than a low one. For example, in the affine stochastic volatility model, the skewness is (negatively) proportional to volatility of volatility.

Interestingly, short-run and long-run volatility factors play very different roles in generating the term structure of skewness. Table 1.5 shows the parameter estimates. Note that there are large differences in the persistence of volatility and volatility of volatility across the two volatility factors. For example, in the SV2 model, the short-run factor has $\kappa_v = 0.0938$ and $\sigma_v = 0.264$, while the long-run factor has $\kappa_u = 0.004$ and $\sigma_u = 0.075$. The values of κ_v and κ_u correspond to half-lives of 7.4 and 173 business days, respectively. The volatility of the short-run factor, σ_v, is 3.5 times larger than that of the long-run factor, σ_u. Because of these differences, the two volatility factors have opposing effects on skewness. The short-run factor can generate a large amount of negative skewness at short horizons thanks to the high volatility of volatility, but the leverage effect (and thus the skewness) dissipates quickly along an increasing horizon due to the weak persistence. In contrast, the long-run factor can generate only a small amount of negative skewness at short horizons due to the low volatility of volatility, but the leverage effect (and thus the skewness) holds up until a longer horizon due to the strong persistence.

To clarify the above, I look at the stand-alone effect of each volatility factor on skewness by simulating the term structures of skewness in three different ways; (i) simulate the stand-alone effect of the short-run factor on skewness with the long-run factor set at zero; (ii) simulate the stand-alone effect of the long-run factor on skewness with the short-run factor set at zero; (iii) simulate the combined effect of the both factors on skewness.[13] Figure 1.7 shows the term structures of the simulated skewness for the SV2 model. The short-run-only case (solid line) shows that the short-run factor is capable of generating a large amount of negative skewness at short horizons, but the skewness begins to disappear after the one-month horizon. In contrast, the long-run-only case (dashed line) shows that the long-run factor generates only a small amount of negative skewness at

[13] A simulation of this kind is feasible because I model the short-run and long-run factors as pure fluctuations around zero (under the $\mathbb{P}$-measure).

short horizons, but the skewness continues to decline with an increasing horizon (at least until the one-year horizon). The dash-dot line shows the combined effect of the two volatility factors.

While it is possible to compute the root mean squared fitting errors in skewness, I find it more informative to compare the *average* term structure of option-implied skewness to its model-implied counterpart. The average term structure of option-implied skewness is obtained by the kernel regression as shown in Figure 1.2. To compute the model-implied counterpart, I simulate the term structure of skewness by Monte Carlo methods on each day and take the average of those term structures over the entire sample period. The top panel of Figure 1.8 shows that all two-factor volatility models produce flat term structures much like those seen in the option market. The two volatility factors play different roles in generating flat patterns much like those observed in the data. The short-run factor is dominant in generating short-term skewness, while the long-run factor is dominant in generating long-term skewness.

However, all two-factor volatility models fall short of matching the observed term structure of skewness precisely. This mismatch will yield mispricings of OTM options, especially under-pricings of OTM puts and overpricings of OTM calls. In fact, this problem is not unique to this paper but has been widely reported in the existing option pricing literature. For example, Bakshi, Cao, and Chen (1997) and Christoffersen, Jacobs, and Mimouni (2010) exhibit similar problems, even though they estimate the models by minimizing option pricing errors. I now compare the proposed two-factor volatility models to the two-factor GARCH model of Christoffersen, Jacobs, Ornthanalai, and Wang (2008) (hereafter, GARCH2).[14] The physical GARCH2 model is estimated as

$$\begin{aligned} y_{t+1} &= y_t + r - d + 1.32h_{t+1} + \sqrt{h_{t+1}}z_{t+1} \\ h_{t+1} &= q_{t+1} + 0.728(h_t - q_t) + 8.26 \times 10^{-7} v_{1,t} \\ q_{t+1} &= 8.95 \times 10^{-7} + 0.990 q_t + 2.64 \times 10^{-6} v_{2,t}, \end{aligned} \tag{1.15}$$

[14] I re-estimate the GARCH2 model by using the time series of returns over the period from July 1962 to December 2008, as Christoffersen, Jacobs, Ornthanalai, and Wang (2008) use the period from July 1962 to December 2001 in their estimation.

where $v_{1,t} = (z_t^2 - 1) - 2 \times 748\sqrt{h_t} z_t$ and $v_{2,t} = (z_t^2 - 1) - 2 \times 69.7\sqrt{h_t} z_t$. In the GARCH2 model, q_t governs the (time-varying) long-run mean of volatility. The top panel of Figure 1.8 shows the average term structure of skewness generated by the risk-neutral GARCH2 model. Notice that all of the proposed two-factor volatility models do much better in fitting the term structure of skewness than the GARCH2 model, even though, like the GARCH2 model, they do not use any information on skewness in estimation.

It is rather puzzling that all models under consideration have trouble matching option-implied skewness. This problem is related to the empirical results of Bakshi, Cao, and Chen (1997) and Bates (2000) that the volatility of volatility and the absolute value of the leverage effect implied from option prices are unrealistically high relative to their time-series-based counterparts. Since the maximum likelihood estimation used in this paper optimizes the models' one-step forecasting ability, my estimates for the volatility of volatility and the absolute value of the leverage effect may not be large enough to rationalize the levels of skewness implicit in option prices.

One may think that this predicament can be solved by accounting for a jump risk premium (e.g., $\mu_{1J}^* \neq \mu_{1J}$). While a jump risk premium might be able to correct for the mis-fitting in short-term skewness, the mis-fitting in long-term skewness would remain as a puzzle because the jump effect on skewness dissipates quickly. To confirm this, I estimate the two-factor volatility models with a jump risk premium only (although not reported in this paper) and find that most of the jump effect on skewness disappears before the one-month horizon (at least in my application).

I now turn my analysis to the one-factor volatility models. The bottom panel of Figure 1.8 exhibits the average term structures of skewness for the one-factor volatility models. Unexpectedly, the term structures continue to decline with an increasing horizon. This feature stands out most for the SV1 model. This seemingly odd result arises because the joint estimation of the physical and risk-neutral one-factor volatility models produces a strong persistence rather than a weak persistence. In Table 1.5, the parameter κ_v is estimated as 0.0065 (SV1), 0.0086 (JY1), and 0.0081 (JYV1). These values correspond to half-lives of 107, 81, and 86 business days, respectively. This empirical result is consistent with the previous finding that a strongly persistent volatility factor

can generate more realistic levels of skewness at medium-to-long horizons.

In fact, other one-factor volatility models in the finance literature often produce a weak persistence of volatility rather than a strong persistence. As an example, I re-estimate the GARCH option pricing model of Heston and Nandi (2000) (hereafter, GARCH1) by using the time series of returns over the period from July 1962 to December 2008. The physical GARCH1 model is estimated as

$$\begin{aligned} y_{t+1} &= y_t + r - d + 1.34h_{t+1} + \sqrt{h_{t+1}}z_{t+1} \\ h_{t+1} &= w + 0.890h_t + 3.32 \times 10^{-6}(z_t - 145\sqrt{h_t})^2, \end{aligned} \tag{1.16}$$

where w is set at a small positive value to keep h_t from being negative, as in Christoffersen, Jacobs, Dorion, and Wang (2010). The half-life of h_t is $\frac{-\log(2)}{\log(0.890+3.32\times10^{-6}\times145^2)} = 17$ business days. The weak persistence of the GARCH1 model makes the skewness decay rather quickly along an increasing horizon, as is shown in the bottom panel of Figure 1.8.

The options data show that long-term skewness is as low as short-term skewness. While a jump process and a short-run factor have difficulty fitting this pattern of the data, a long-run volatility factor helps to generate more realistic levels of long-term skewness. Interestingly, the importance of a long-run volatility factor is consistent with the well-known empirical finding that stock returns have a long memory property (e.g., Ding, Granger, and Engle 1993; Bollerslev and Mikkelsen 1996; Baillie 1996; Gallant, Hsu, and Tauchen 1999). Studying option pricing models with the long memory property would be a promising research direction, especially for those who are interested in pricing long-term options including LEAPS (Long-term Equity AnticiPation Securities).

1.6.3 Option pricing performance

This subsection examines option pricing performance across the different models. The existing papers often use relatively short sample periods due to computational burden or econometric limitation. However, since a volatility surface varies over time, any empirical research on option

pricing can be dependent on the sample period used. For example, the empirical results obtained when a volatility smirk is steep may not be applicable to days when it is flat. In contrast, since my empirical results are based on a large-scale dataset of options, they are relatively robust to sample-specificity. Besides, the option pricing tests in this paper are effectively out-of-sample because I use only one or two constant-maturity implied volatilities on each day in estimation.

I use the following three metrics for measuring option pricing performance:

$$\begin{aligned}
\text{BS-RMSE} &= \sqrt{\frac{1}{\sum_{t=1}^{N} NO_t} \sum_{t=1}^{N} \sum_{i=1}^{NO_t} \left(\sigma_i^{\mathcal{M}} - \sigma_i^{BS}\right)^2} \\
\text{DL-RMSE} &= \sqrt{\frac{1}{\sum_{t=1}^{N} NO_t} \sum_{t=1}^{N} \sum_{i=1}^{NO_t} \left(O_i^{\mathcal{M}} - O_i\right)^2} \\
\text{FREQ} &= \frac{1}{\sum_{t=1}^{N} NO_t} \sum_{t=1}^{N} \sum_{i=1}^{NO_t} \mathbf{1}_{O_i^{\mathcal{M}} > O_i},
\end{aligned} \tag{1.17}$$

where $\sigma_i^{\mathcal{M}}$ and σ_i^{BS} refer to the modeled and observed Black-Scholes implied volatilities; $O_i^{\mathcal{M}}$ and O_i denote the modeled and observed option prices; and NO_t indicates the number of options at time t. BS-RMSE measures the root mean squared pricing errors in the Black-Scholes implied volatilities, which should be distinguished from IV-RMSE in subsection 1.6.1. DL-RMSE measures the root mean squared pricing errors in dollar option prices. FREQ computes how often a model overprices the observed option prices.

Table 1.7 reports option pricing performance across the different models. All of the two-factor volatility models produce smaller option pricing errors than any of the one-factor volatility models. For exposition, let me compare the best performing one-factor (JYV1) model to the best performing two-factor (JYV2) model. The JYV1 model yields a BS-RMSE of 2.32% and a DL-RMSE of \$4.08, while the JYV2 model gives a BS-RMSE of 1.96% and a DL-RMSE of \$2.95. These numbers imply that the JYV2 model improves the JYV1 model by 15.5% in BS-RMSE and 27.7% in DL-RMSE. The improvement is mainly due to the fact that the two-factor volatility models better fit the term structure of implied volatility (see subsection 1.6.1). However, the improvement is smaller than I expected, because the one-factor volatility models (especially, the JY1

and JYV1 models) can generate more realistic levels of skewness at medium-to-long horizons than the two-factor volatility models (see Figure 1.8).

In the preceding subsection, I find that all two-factor volatility models generate higher skewness than those observed in the option market. This undesirable result leads to underpricings of OTM puts and overpricings of OTM calls. To confirm this, I look at the frequency of overpricings and underpricings by using the metric FREQ. If FREQ is larger (smaller) than one-half, it means that the model tends to overprice (underprice) options. The two-factor volatility models have FREQs of 0.303-0.432 for OTM puts, implying that the models underprice the OTM puts more often than not. In contrast, FREQs fall between 0.914 and 0.936 for OTM calls in the two-factor volatility models, implying that the models overprice the OTM calls most of the time.

Now, let me compare the proposed two-factor volatility models to the competing GARCH2 model. The SV2 model gives a BS-RMSE of 2.18% and a DL-RMSE of $3.08, while the GARCH2 model gives a BS-RMSE of 6.44% and a DL-RMSE of $8.80. These numbers imply that the SV2 model outperforms the GARCH2 model by 66.2% in BS-RMSE and 65.0% in DL-RMSE. Of course, it may not be fair to compare the SV2 model to the GARCH2 model because while the former makes use of two implied volatilities in estimation, the latter does not. However, such a difference is not the only reason for the huge improvement. The preceding subsection shows that the two-factor volatility models do much better in fitting the term structure of skewness than the GARCH2 model (see the top panel of Figure 1.8).

1.6.4 Pricing error analysis

The remaining issue is associated with stochastic skewness. The options data have time-variation in skewness. It may be necessary to introduce a stochastic factor that is associated with negative skewness.[15] If a continuous diffusive process is important, one can think of randomizing

[15] An alternative approach to generating stochastic skewness is to adopt a two-factor affine or CEV model. Christoffersen, Jacobs, Ornthanalai, and Wang (2008) and Christoffersen, Heston, and Jacobs (2009) argue that the stochastic correlation implied by the two-factor affine model helps to improve option pricing performance. However, the two-factor volatility model would be imperfect in fitting the time-varying skewness because the dynamics of skewness are fundamentally different from those of volatility (see subsection 1.2.3).

the leverage effect or volatility of volatility. On the other hand, if a jump process is important, one may randomize the jump-related parameters such as a mean jump size or jump intensity. However, we do not know *a priori* which of these variables are associated with time-variation in skewness. Accordingly, the goal of this subsection is to test the possibility of each of the variables as a skewness factor.

The estimation strategy used in this paper makes it possible to identify skewness factors by analyzing option pricing errors. While I force the models to fit the term structure of volatility, the term structure of skewness remains unfit. As a result, a large fraction of option pricing errors occur along the moneyness dimension, implying that those errors should be associated with time-varying skewness factors. The filtered volatility of volatility, VOV_t, is used as a proxy for stochastic volatility of volatility, and the filtered leverage effect, LEV_t, is used as a proxy for stochastic leverage effect. These variables are obtained from Durham and Park (2010). I use the jump variation, JV_t, as a proxy for a jump-related parameter. Table 1.8 summarizes the theoretical predictions of the skewness factors on option pricing errors. Bearing this table in mind, I run regressions of the option pricing errors on the skewness factors:

$$\begin{aligned}\text{BS-RMSE}_t = \beta_0 + \beta_{\text{VOV}}\,\text{VOV}_t + \beta_{\text{LEV}}\text{LEV}_t + \beta_{\text{JV}}\log\text{JV}_t \\ + \beta_{\text{RV}}\log\text{RV}_t + \beta_{\text{CRSP}}\text{CRSP}_t + \beta_{\text{TMSP}}\text{TMSP}_t + \varepsilon_t,\end{aligned} \tag{1.18}$$

where BS-RMSE_t refers to the root mean squared pricing errors in Black-Scholes implied volatilities at time t. As control variables, I add the realized volatility RV_t, the credit spread CRSP_t (difference between 10-year AAA and BAA corporate yields), and the term spread of interest rates TMSP_t (difference between 10-year and 3-month Treasury yields). The correlations among the explanatory variables are provided in Table 1.9.

Table 1.10 reports the regression results, with Newey-West robust t-statistics over eight lags (Newey and West 1987) and adjusted R^2. I run the regressions on the control variables alone (top panel), the skewness factors alone (middle), and both (bottom). I first discuss the effects of the control variables on the pricing errors when all explanatory variables are included. The realized

volatility has statistically significant and positive coefficients with t-statistics of 6.77-6.91, with the implication that the pricing errors become larger in periods of higher volatility. The coefficients on the credit spread are marginally significant and positive with t-statistics of 1.92-2.48, implying that the pricing errors are large when credit risk is high.

The filtered volatility of volatility is significantly associated with the pricing errors, with the expected direction. This result is valid regardless of whether the control variables are included. When the control variables are included, the coefficients β_{VOV} fall between 1.59 and 1.70 with t-statistics of 12.9-13.5. Consistent with the theoretical implication of the volatility of volatility, the positivity of the coefficients indicates that the pricing errors are larger when volatility of volatility is higher. The reason is that the two-factor volatility models can reflect the fact that the option market tends to have more left-skewed return distributions in periods of high volatility of volatility. This variable is most significant among the variables considered in this paper. For robustness check, I also use the one-sided, rolling-based estimates for volatility of volatility, using the VIX index, and find that the results are qualitatively similar.

The filtered leverage effect is also significantly associated with the pricing errors, with the expected direction. This result is valid regardless of whether the control variables are included. When the control variables are included, the coefficients β_{LEV} range from 0.37 to 0.49 with t-statistics of 2.87-3.49. Consistent with the theoretical implication of the leverage effect, the positivity of the coefficients implies that the pricing errors are larger as the leverage effect gets stronger (more negative). The reason is that the two-factor volatility models can accommodate the fact that the option market tends to have more left-skewed return distributions as the leverage effect get stronger. For robustness check, I also use the one-sided, rolling-based estimate for the leverage effect, using the S&P 500 index and the VIX index, and find that the results are qualitatively similar.

However, evidence on jump risk is not decisive. The jump variation is significant without the control variables, but insignificant with the control variables. The realized volatility is highly correlated with the jump variation, with the correlation of 0.949 (see Table 1.9). The jump variation does not seem to deliver additional information on the pricing errors beyond the realized volatility.

The explanatory variables can account for a large fraction of variation in the pricing errors. The control variables alone are responsible for 32-36% of variation in the pricing errors. The regressions with the skewness factors alone yield adjusted R^2 of 47-49%. The explanatory variables all together can explain a significant portion of variation in the pricing errors, with adjusted R^2 of 54-56%. This high explanatory power suggests that the stochastic volatility models should be augmented with some skewness factors associated with time-variation in skewness. I therefore conclude that stochastic volatility of volatility is the most promising candidate as a skewness factor, while a stochastic leverage effect remains as a potential candidate.

1.7 Conclusion and future work

News events about a firm's future prospects have different term structure effects on volatility, depending on the market's perception of the persistence of each news event. This fact has important implications particularly for the options market because the term structure of volatility will react in various ways according to the economic nature of an arriving news event. The objective of this paper is, therefore, to examine the option pricing implications of short-run and long-run volatility factors, which are assumed to be driven by the short-run and long-run news events. To this end, I develop a class of log stochastic volatility models with short-run and long-run volatility factors, and then perform a term-structure-based analysis using an extensive dataset of S&P 500 index options over 1993-2008.

I find that the proposed two-factor volatility models have two desirable properties that help capture the term structures of option-implied volatility and skewness. First, the option data show evidence of time-variation in the long-run expectation of volatility, which may be caused by long-run news events. This evidence suggests that single-factor volatility models, which assume constancy in the long-run level of volatility, are clearly deficient for the purpose of option pricing. My empirical results demonstrate that the proposed two-factor volatility models excel at fitting the entire term structure of implied volatility, as opposed to single-factor volatility models. In particular, the inclusion of a long-run volatility factor is necessary to capture the dynamics of long-term

implied volatility. Because of this nice property, the two-factor volatility models improve option pricing performance over the one-factor volatility models.

Second, the option data reveal that the term structure of implied skewness is nearly flat on average. This feature is hard to reconcile with single-factor volatility models and jumps in returns. Single-factor volatility models are able to produce either a decreasing or an increasing term structure of skewness up to the one-year horizon, depending on the persistence of volatility, but not a flat term structure of skewness. Jumps in returns can be effective in producing short-term skewness but not long-term. But I find that the two-factor volatility models can generate flat term structures, as the two volatility factors involve very different levels of the skewness-related parameters. Specifically, while the short-run factor is dominant in generating short-term skewness, the long-run factor plays a pivotal role in generating long-term skewness. Besides, the proposed two-factor volatility models can generate far more realistic levels of skewness than the two-factor GARCH model of Christoffersen, Jacobs, Ornthanalai, and Wang (2008).

However, I find that the two-factor volatility models are still mis-specified in two dimensions. First, in accord with many of the prior studies, the two-factor volatility models fall short of matching the option-implied skewness precisely, causing underpricings of OTM puts and overpricings of OTM calls. Is this problem unique to the S&P index option market? A relevant research direction would be to apply the models and techniques developed in this paper to individual stock options and currency options that have less of volatility smirk. Second, I argue that any multi-factor stochastic volatility models including mine would be imperfect in capturing time-variation in skewness because the dynamics of skewness are loosely related to those of volatility. My investigation indicates that one potential extension of the two-factor volatility models would be to randomize the volatility of volatility in order to capture time-variation in skewness.

Table 1.1: Principal component analysis of implied volatilities.

The sample period covers January 1993 to December 2008. IVi refers to the i-month model-free implied volatility. PCi means the i-th principal component.

Summary statistics

	Mean	Min	Max	Std	Skewness	Kurtosis	AR(1)	AR(5)	AR(22)
$IV1_t$ (%)	19.22	7.76	81.09	8.56	2.20	11.75	0.983	0.941	0.786
$IV2_t$ (%)	19.61	9.54	74.21	7.91	1.91	9.77	0.990	0.958	0.819
$IV3_t$ (%)	19.79	10.37	69.86	7.46	1.77	8.92	0.991	0.963	0.833
$IV6_t$ (%)	20.09	11.28	62.15	6.68	1.53	7.52	0.994	0.971	0.856
$IV9_t$ (%)	20.15	12.15	57.09	6.14	1.40	6.79	0.994	0.974	0.867
$IV12_t$ (%)	20.25	12.39	53.74	5.88	1.27	6.06	0.995	0.976	0.878

Principal component analysis

	Proportion of variation		Factor loading					
PC	Marginal	Cumulative	IV1	IV2	IV3	IV6	IV9	IV12
PC1	97.73	97.73	0.484	0.454	0.429	0.382	0.348	0.329
PC2	2.10	99.83	0.626	0.255	0.034	-0.296	-0.425	-0.523
PC3	0.12	99.95	-0.561	0.354	0.569	0.158	-0.142	-0.436
PC4	0.02	99.98	0.158	-0.333	-0.164	0.551	0.393	-0.616
PC5	0.01	99.99	-0.166	0.610	-0.474	-0.290	0.509	-0.181
PC6	0.01	100.00	0.077	-0.347	0.489	-0.595	0.514	-0.126

Table 1.2: Principal component analysis of implied skewness.

The sample period covers January 1993 to December 2008. SKEW*i* refers to the *i*-month model-free implied skewness. PC*i* means the *i*-th principal component.

Summary statistics

	Mean	Min	Max	Std	Skewness	Kurtosis	AR(1)	AR(5)	AR(22)
$SKEW1_t$	-1.69	-3.75	-0.41	0.49	-0.54	3.24	0.902	0.804	0.627
$SKEW2_t$	-1.68	-3.95	-0.49	0.42	-0.24	3.25	0.921	0.850	0.682
$SKEW3_t$	-1.69	-3.89	-0.41	0.42	-0.31	3.53	0.915	0.827	0.664
$SKEW6_t$	-1.70	-2.86	-0.75	0.37	-0.16	2.52	0.969	0.936	0.848
$SKEW9_t$	-1.61	-2.80	-0.57	0.37	-0.25	2.66	0.982	0.957	0.875
$SKEW12_t$	-1.59	-2.91	-0.48	0.37	-0.25	2.78	0.979	0.939	0.824

Principal component analysis

	Proportion of variation		Factor loading					
PC	Marginal	Cumulative	SKEW1	SKEW2	SKEW3	SKEW6	SKEW9	SKEW12
PC1	77.47	77.47	-0.468	-0.434	-0.438	-0.387	-0.362	-0.346
PC2	13.07	90.54	0.635	0.271	0.044	-0.252	-0.435	-0.518
PC3	4.58	95.12	-0.564	0.346	0.671	-0.046	-0.181	-0.279
PC4	1.86	96.98	0.203	-0.761	0.490	0.314	-0.160	-0.124
PC5	1.68	98.66	-0.118	0.136	-0.318	0.739	0.112	-0.555
PC6	1.34	100.00	-0.064	0.146	-0.124	0.373	-0.780	0.460

Table 1.3: Model specifications.

Model	# of factors	Jumps	Variance risk premium	Jump risk premium
SV1	One	No jumps	Yes	N/A
JY1	One	Jumps in y_t	No	Yes
JYV1	One	Jumps in y_t and v_t	No	Yes
SV2	Two	No jumps	Yes	N/A
JY2	Two	Jumps in y_t	Yes	No
JYV2	Two	Jumps in y_t and v_t	Yes	No

Table 1.4: Summary statistics for options data.

The option pricing test in this paper considers out-of-the-money (OTM) put and call options. Moneyness d is defined as $d = \frac{\log(K/F_{t,\tau})}{\text{IV}_{t,\tau}\sqrt{\tau}}$, where $\text{IV}_{t,\tau}$ and $F_{t,\tau}$ refer to the model-free implied volatility and the implied future price at time t with time to maturity τ. BSIV stands for the Black-Scholes implied volatility

		Months-to-maturity			
		< 2M	2M-6M	6M-12M	Total
All OTM options	# of Options	197,631	199,170	177,005	573,806
$(-3.0 < d < 1.5)$	Average Price ($)	9.13	16.28	27.43	17.26
	Average BSIV (%)	21.67	21.44	21.37	21.50
OTM puts	# of Options	119,265	120,296	107,148	346,709
$(-3.0 < d < 0)$	Average Price ($)	8.92	15.81	26.41	16.71
	Average BSIV (%)	24.14	23.86	23.78	23.93
OTM calls	# of Options	78,366	78,874	69,857	227,097
$(0 < d < 1.5)$	Average Price ($)	9.47	17.00	28.98	18.09
	Average BSIV (%)	17.91	17.73	17.68	17.78

Table 1.5: Parameter Estimates.

The sample period covers January 1993 through December 2008 ($N = 4029$). Standard errors are in parentheses. Time is measured in business days.

	One-factor			Two-factor		
	SV1	JY1	JYV1	SV2	JY2	JYV2
$\mu \times 10^4$	4.06	2.03	2.84	0.38	-1.61	-2.57
	(1.13)	(1.27)	(1.27)	(0.93)	(0.92)	(1.02)
$\kappa_v \times 10^2$	0.65	0.86	0.81	9.38	9.37	8.84
	(0.13)	(0.13)	(0.14)	(0.18)	(0.19)	(0.21)
$\kappa_u \times 10^2$				0.40	0.34	0.41
				(0.03)	(0.04)	(0.04)
$\eta_v \times 10^2$	-3.02			-1.77	-1.06	1.02
	(0.23)			(0.45)	(0.47)	(0.83)
$\eta_u \times 10^3$				4.72	4.92	3.74
				(1.00)	(0.99)	(1.48)
$\sigma_v \times 10^1$	1.33	1.35	0.97	2.64	2.63	1.89
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.05)
$\sigma_u \times 10^1$				0.75	0.73	0.76
				(0.01)	(0.01)	(0.01)
ρ_v	-0.74	-0.84	-0.77	-0.51	-0.58	-0.44
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
ρ_u				-0.58	-0.63	-0.71
				(0.01)	(0.01)	(0.01)
α	-9.49	-9.35	-9.84	-8.56	-8.33	-8.32
	(0.29)	(0.23)	(0.27)	(0.18)	(0.21)	(0.20)
λ		0.35	0.39		0.28	0.21
		(0.08)	(0.05)		(0.06)	(0.03)
μ_{1J}		0.38	-0.04		0.31	-0.11
		(0.06)	(0.08)		(0.06)	(0.07)
μ^*_{1J}		-1.12	-1.30			
		(0.13)	(0.08)			
σ_{1J}		0.78	1.35		0.82	1.42
		(0.05)	(0.08)		(0.06)	(0.09)
μ_{2J}			0.45			0.74
			(0.08)			(0.11)
σ_{2J}			1.47			1.88
			(0.07)			(0.11)
ρ_J			-0.73			-0.73
			(0.03)			(0.03)
log(L)	38,426	38,587	38,782	67,117	67,263	67,528

Table 1.6: Fitting errors in implied volatility across different models

I use the two metrics for evaluating the models' fit in the term structure of implied volatility:

$$\begin{aligned} \text{IV-RMSE} &= \sqrt{\frac{1}{\sum_{t=1}^{N} NV_t} \sum_{t=1}^{N} \sum_{\tau=1}^{NV_t} \left(\text{IV}_{t,\tau}^{\mathcal{M}} - \text{IV}_{t,\tau} \right)^2} \\ \text{RIV-RMSE} &= 100 \times \sqrt{\frac{1}{\sum_{t=1}^{N} NV_t} \sum_{t=1}^{N} \sum_{\tau=1}^{NV_t} \left(\frac{\text{IV}_{t,\tau}^{\mathcal{M}} - \text{IV}_{t,\tau}}{\text{IV}_{t,\tau}} \right)^2}, \end{aligned} \tag{1.19}$$

where N stands for the total number of business days; NV_t is the number of model-free implied volatilities at time t with times-to-maturity of less than one year; $\text{IV}_{t,\tau}$ and $\text{IV}_{t,\tau}^{\mathcal{M}}$ refer to the option-implied and model-implied volatilities, respectively. While IV-RMSE measures the root mean squared fitting errors in implied volatility, RIV-RMSE does this in a relative sense by accounting for levels of implied volatility.

Aggregate results

		IV-RMSE				RIV-RMSE			
		Months-to-maturity				Months-to-maturity			
		< 2M	2M-6M	6M-12M	Total	< 2M	2M-6M	6M-12M	Total
	SV1	2.18	14.53	56.69	34.73	9.82	70.05	288.63	176.25
	JY1	0.79	2.20	3.98	2.73	3.38	10.72	23.38	15.37
	JYV1	0.74	2.03	3.60	2.48	3.04	9.33	20.73	13.58
	SV2	0.63	0.69	1.03	0.81	2.74	2.62	4.55	3.46
	JY2	0.55	0.68	1.22	0.88	2.52	2.61	5.46	3.84
	JYV2	0.55	0.83	1.75	1.19	2.29	3.31	8.82	5.73

Disaggregate results by levels of implied volatility

		IV-RMSE				RIV-RMSE			
		Months-to-maturity				Months-to-maturity			
Volatility		< 2M	2M-6M	6M-12M	Total	< 2M	2M-6M	6M-12M	Total
Low	SV1	1.10	8.99	44.47	21.39	9.62	70.21	320.43	159.07
	JY1	0.39	1.60	4.49	2.40	3.40	12.66	32.64	18.10
	JYV1	0.33	1.33	3.92	2.08	2.93	10.46	28.49	15.64
	SV2	0.31	0.25	0.56	0.35	2.83	1.98	3.99	2.76
	JY2	0.32	0.24	0.66	0.38	2.92	1.93	4.75	2.99
	JYV2	0.26	0.32	1.33	0.65	2.35	2.52	9.57	4.91
Mid	SV1	1.72	13.47	55.61	35.51	9.65	71.82	292.12	194.30
	JY1	0.56	1.78	3.92	2.77	3.15	9.97	21.93	16.13
	JYV1	0.50	1.53	3.48	2.44	2.79	8.50	19.44	14.18
	SV2	0.49	0.47	0.79	0.62	2.66	2.37	4.16	3.31
	JY2	0.44	0.47	1.01	0.72	2.42	2.41	5.35	3.89
	JYV2	0.39	0.62	1.68	1.13	2.15	3.21	9.01	6.26
High	SV1	3.42	20.02	68.50	42.82	10.35	66.21	244.71	152.79
	JY1	1.31	3.23	3.54	2.95	3.77	10.03	12.63	9.91
	JYV1	1.25	3.17	3.49	2.90	3.56	9.72	12.29	9.61
	SV2	1.01	1.19	1.63	1.32	2.82	3.52	5.68	4.26
	JY2	0.83	1.16	1.85	1.37	2.31	3.46	6.27	4.45
	JYV2	0.91	1.38	2.19	1.62	2.48	4.12	7.58	5.34

Table 1.7: Option pricing performance across different models.

I use the following three metrics for measuring option pricing performance:

$$\begin{aligned}
\text{BS-RMSE} &= \sqrt{\frac{1}{\sum_{t=1}^{N} NO_t} \sum_{t=1}^{N} \sum_{i=1}^{NO_t} \left(\sigma_i^{\mathcal{M}} - \sigma_i^{BS}\right)^2} \\
\text{DL-RMSE} &= \sqrt{\frac{1}{\sum_{t=1}^{N} NO_t} \sum_{t=1}^{N} \sum_{i=1}^{NO_t} \left(O_i^{\mathcal{M}} - O_i\right)^2} \\
\text{FREQ} &= \frac{1}{\sum_{t=1}^{N} NO_t} \sum_{t=1}^{N} \sum_{i=1}^{NO_t} \mathbf{1}_{O_i^{\mathcal{M}} > O_i},
\end{aligned} \tag{1.20}$$

where $\sigma_i^{\mathcal{M}}$ and σ_i^{BS} refer to the modeled and observed Black-Scholes implied volatilities; $O_i^{\mathcal{M}}$ and O_i denote the modeled and observed option prices; and NO_t indicates the number of options at time t. BS-RMSE measures the root mean squared pricing errors in the Black-Scholes implied volatilities. DL-RMSE measures the root mean squared pricing errors in dollar option prices. FREQ computes how often a model overprices the observed option prices.

	BS-RMSE				DL-RMSE				FREQ			
	Months-to-maturity				Months-to-maturity				Months-to-maturity			
	< 2M	2M-6M	6M-12M	Total	< 2M	2M-6M	6M-12M	Total	< 2M	2M-6M	6M-12M	Total
All OTM options ($-3.0 < d < 1.5$)												
SV1	3.71	11.50	32.45	19.38	3.77	23.73	107.10	61.15	0.679	0.995	1.000	0.888
JY1	2.88	2.66	2.84	2.79	2.43	4.04	7.38	4.95	0.655	0.730	0.806	0.728
JYV1	2.36	2.20	2.40	2.32	1.99	3.28	6.13	4.08	0.643	0.639	0.704	0.660
SV2	2.53	1.97	1.98	2.18	2.06	2.57	4.32	3.08	0.634	0.567	0.421	0.545
JY2	2.56	2.05	1.87	2.19	1.94	2.61	3.98	2.92	0.611	0.590	0.514	0.574
JYV2	2.29	1.76	1.79	1.96	1.76	2.35	4.29	2.95	0.630	0.615	0.653	0.632
GARCH2	7.15	6.29	5.75	6.44	4.83	7.66	12.60	8.80	0.212	0.180	0.190	0.194
OTM puts ($-3.0 < d < 0$)												
SV1	2.84	10.69	32.05	18.97	2.48	21.12	98.15	55.98	0.494	0.992	1.000	0.823
JY1	2.23	2.09	2.78	2.37	1.50	2.99	6.56	4.14	0.446	0.588	0.745	0.587
JYV1	1.84	1.95	2.55	2.12	1.29	2.75	5.98	3.77	0.460	0.535	0.689	0.557
SV2	1.79	1.56	2.14	1.83	1.11	1.63	4.49	2.75	0.426	0.316	0.153	0.303
JY2	1.97	1.64	1.77	1.80	1.20	1.66	3.60	2.33	0.393	0.349	0.280	0.343
JYV2	1.66	1.33	1.51	1.51	1.08	1.49	3.25	2.11	0.436	0.409	0.455	0.432
GARCH2	8.51	7.61	6.95	7.74	5.25	8.51	13.98	9.75	0.087	0.060	0.083	0.077
OTM calls ($0 < d < 1.5$)												
SV1	4.74	12.64	33.05	19.98	5.13	27.23	119.54	68.28	0.961	1.000	1.000	0.987
JY1	3.66	3.34	2.92	3.34	3.39	5.26	8.48	5.98	0.974	0.948	0.898	0.942
JYV1	2.97	2.53	2.13	2.58	2.73	3.94	6.34	4.51	0.920	0.797	0.728	0.818
SV2	3.36	2.48	1.72	2.63	2.96	3.55	4.06	3.53	0.952	0.949	0.831	0.914
JY2	3.25	2.55	2.01	2.67	2.71	3.60	4.50	3.64	0.943	0.958	0.873	0.927
JYV2	3.00	2.26	2.14	2.51	2.46	3.25	5.52	3.89	0.925	0.930	0.955	0.936
GARCH2	4.33	3.38	3.10	3.66	4.11	6.15	10.11	7.10	0.401	0.364	0.353	0.373

Table 1.8: Theoretical predictions of pricing errors.

Variable	Proxy for	Effect on implied skewness	BS-RMSE	Sign
Filtered volatility of volatility (VOV_t)	stochastic volatility of volatility	The higher volatility of volatility, the more negative option-implied skewness.	larger	+
Filtered leverage effect (LEV_t)	stochastic leverage effect	The stronger leverage effect, the more negative option-implied skewness.	larger	+
Jump variation (JV_t)	stochastic jump	The higher jump variation, the more negative option-implied skewness.	larger	+

Table 1.9: Summary statistics for explanatory variables.

The sample period covers January 1993 to December 2008. VOV_t and LEV_t refer to the filtered volatility of volatility and filtered leverage effect obtained from Durham and Park (2010). JV_t and RV_t are the jump variation and realized volatility calculated using the five-minute high-frequency data over the past 22 business days. CRSP_t is the corporate credit spread (difference between 10-year AAA and BAA corporate yields). TMSP_t is the term spread of interest rates (difference between 10-year and 3-month Treasury yields).

Summary statistics

	VOV_t	LEV_t	$\text{JV}_t\,(\%^2)$	$\text{RV}_t\,(\%^2)$	$\text{CRSP}_t\,(\%)$	$\text{TMSP}_t\,(\%)$
Mean	0.26	0.71	3.33	21.79	0.88	1.50
Min	0.01	0.03	0.12	1.58	0.50	-0.95
Max	0.97	0.99	83.09	457.27	3.50	3.85
STD	0.26	0.27	6.89	40.45	0.36	1.19
Skewness	1.16	-0.88	7.19	6.72	3.78	0.10
Kurtosis	3.06	2.50	67.06	57.28	23.55	1.82
AR(1)	0.92	0.96	0.99	1.00	0.99	1.00
AR(5)	0.67	0.82	0.95	0.97	0.96	0.99
AR(22)	0.29	0.46	0.65	0.68	0.81	0.95

Correlation matrix

	VOV_t	LEV_t	$\text{JV}_t\,(\%^2)$	$\text{RV}_t\,(\%^2)$	$\text{CRSP}_t\,(\%)$	$\text{TMSP}_t\,(\%)$
VOV_t	1.000	-0.122	0.187	0.216	0.037	-0.098
LEV_t		1.000	0.034	0.062	0.183	-0.076
$\text{JV}_t\,(\%^2)$			1.000	0.949	0.762	0.168
$\text{RV}_t\,(\%^2)$				1.000	0.769	0.133
$\text{CRSP}_t\,(\%^2)$					1.000	0.290
$\text{TMSP}_t\,(\%)$						1.000

Table 1.10: Regression of pricing errors on skewness candidates.

The table reports the results of the following regression,

$$\text{BS-RMSE}_t = \beta_0 + \beta_{\text{VOV}}\ \text{VOV}_t + \beta_{\text{LEV}}\text{LEV}_t + \beta_{\text{JV}}\log \text{JV}_t + \beta_{\text{RV}}\log \text{RV}_t + \beta_{\text{CRSP}}\text{CRSP}_t + \beta_{\text{TMSP}}\text{TMSP}_t + \varepsilon_t.$$

Newey-West robust t-statistics over eight lags are shown in parentheses. BS-RMSE_t is the root mean squared pricing errors in the Black-Scholes implied volatilities. VOV_t and LEV_t refer to the filtered volatility of volatility and filtered leverage effect obtained from Durham and Park (2010). JV_t and RV_t are the jump variation and realized volatility calculated using the five-minute high-frequency data over the past 22 business days. CRSP_t is the corporate credit spread (difference between 10-year AAA and BAA corporate yields). TMSP_t is the term spread of interest rates (difference between 10-year and 3-month Treasury yields).

	Constant		VOV_t		LEV_t		$\log \text{JV}_t$		$\log \text{RV}_t$		CRSP_t		TMSP_t		$Adj.R^2$
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	
Control variables only															
SV2	0.19	(1.01)							0.520	(9.12)	0.337	(1.42)	-0.111	(-4.65)	35.5%
JY2	0.13	(.67)							0.508	(8.80)	0.416	(1.71)	-0.098	(-4.01)	36.0%
JYV2	0.19	(1.07)							0.447	(8.30)	0.287	(1.26)	-0.073	(-3.14)	32.0%
Skewness factors only															
SV2	0.54	(5.58)	1.83	(10.85)	0.59	(3.47)	0.387	(6.62)							47.4%
JY2	0.51	(5.28)	1.87	(10.99)	0.60	(3.48)	0.389	(6.38)							48.6%
JYV2	0.51	(5.80)	1.74	(11.14)	0.47	(3.01)	0.326	(5.89)							46.9%
All explanatory variables															
SV2	-0.80	(-3.65)	1.65	(12.87)	0.46	(3.29)	-0.185	(-2.60)	0.610	(6.91)	0.393	(2.05)	-0.067	(-3.60)	54.6%
JY2	-0.93	(-4.20)	1.70	(13.37)	0.49	(3.49)	-0.203	(-2.85)	0.611	(6.91)	0.478	(2.48)	-0.053	(-2.82)	56.2%
JYV2	-0.77	(-3.73)	1.59	(13.50)	0.37	(2.87)	-0.198	(-2.99)	0.560	(6.77)	0.348	(1.92)	-0.032	(-1.81)	54.0%

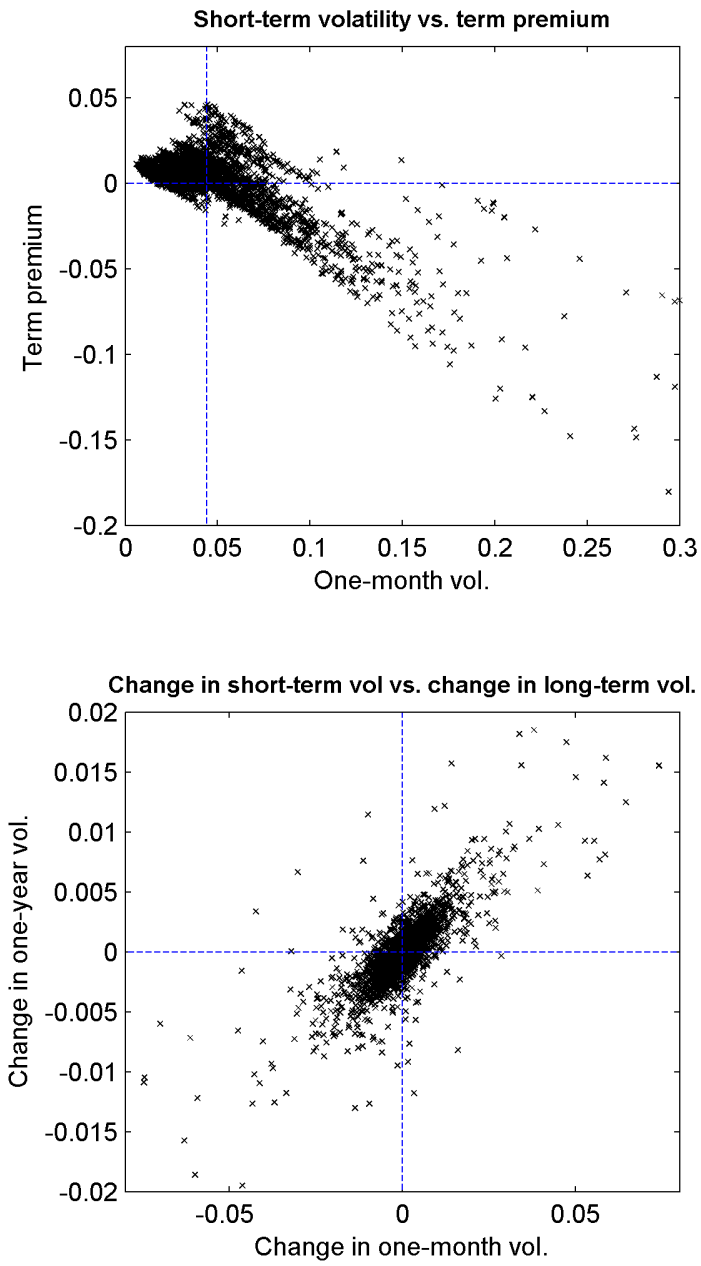

Figure 1.1: Scatter plots between option-implied volatilities.

The top panel is a scatter plot of the term premium (difference between one-year and one-month volatilities) against the short-term (one-month) implied volatility. The bottom panel is a scatter plot of changes in one-year implied volatility against changes in one-month implied volatility. Some outliers are deleted to take a closer look at the center of the scatter plots.

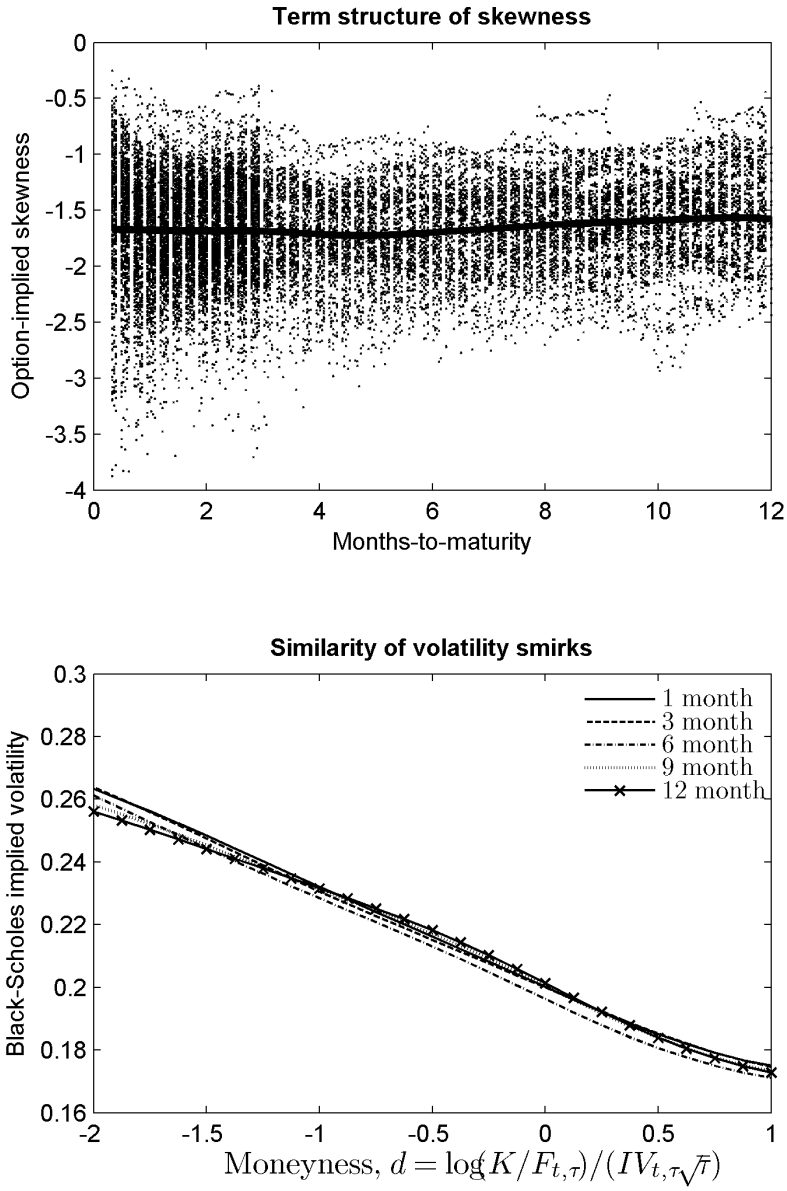

Figure 1.2: Term structure of skewness and the similarity of volatility smirks.

In the top panel, the thick solid line is the kernel regression of skewness on term to maturity with a Gaussian kernel and a manually chosen bandwidth. In the bottom panel, I compute the average volatility surface by running a kernel regression of Black-Scholes implied volatilities on term to maturity τ and moneyness d as defined in equation (1.5), and then obtain an array of volatility smirks by slicing the volatility surface at several times to maturity.

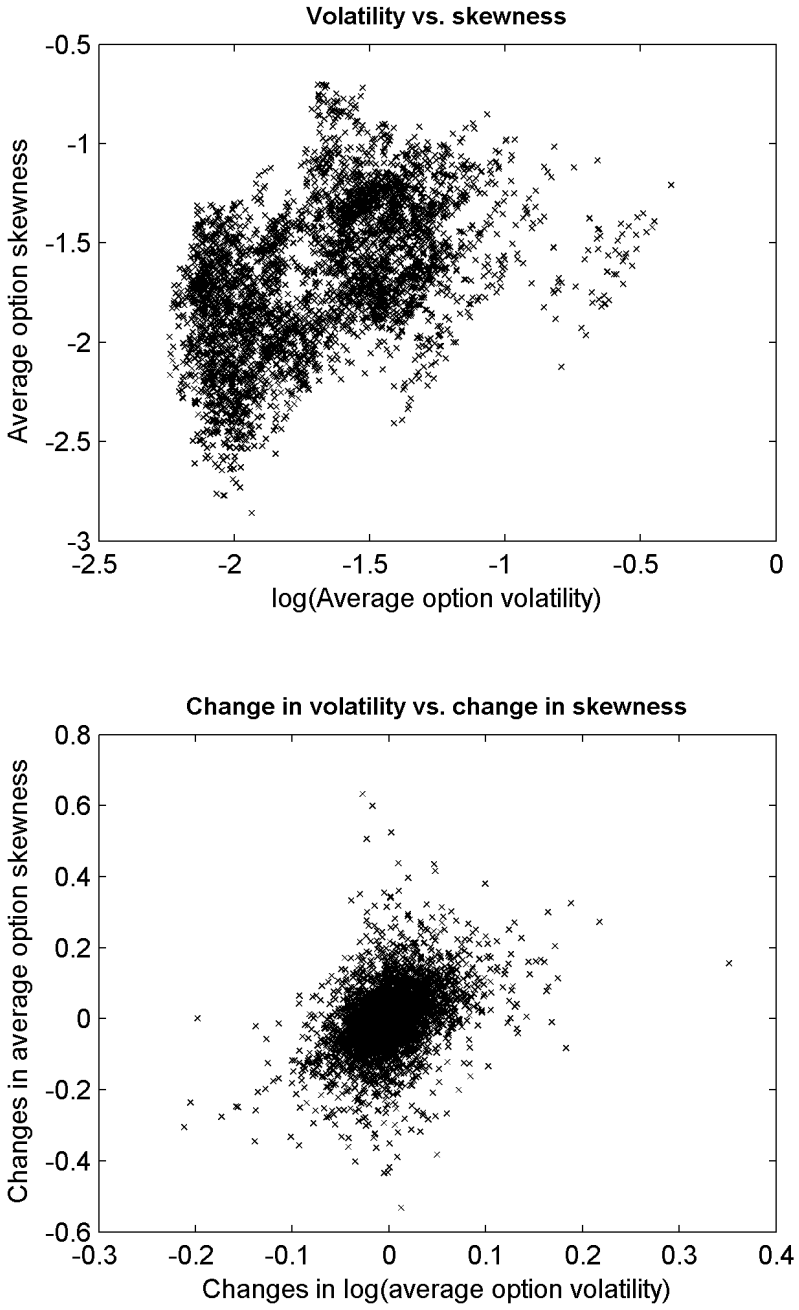

Figure 1.3: Scatter plots of option-implied skewness against option-implied volatility.

The top panel is a scatter plot of average implied skewness against average implied volatility, while the bottom panel is a scatter plot of changes in average implied skewness versus changes in average implied volatility. The averages are taken for options with less than one year to maturity.

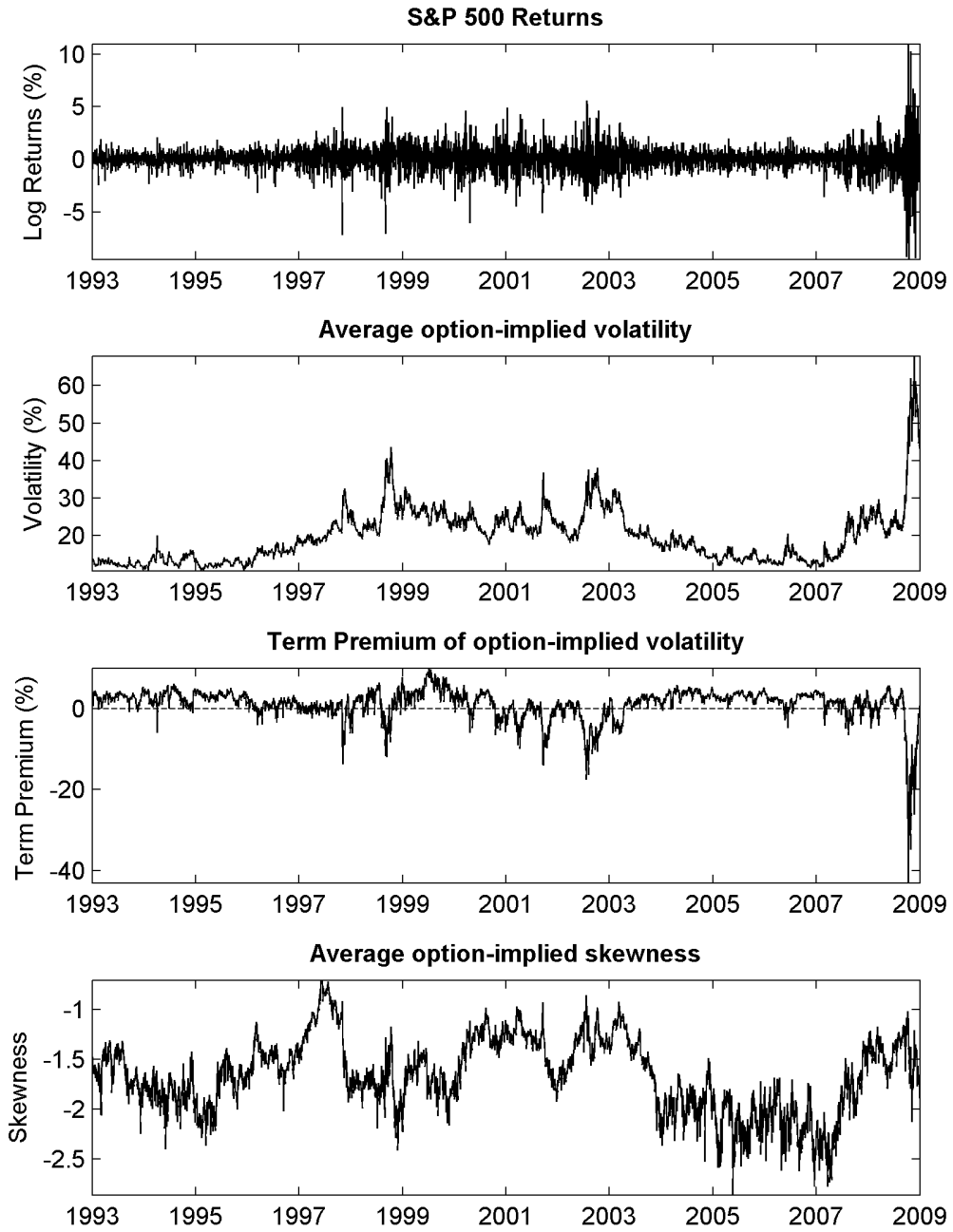

Figure 1.4: Time series of S&P 500 returns, average option-implied volatility, the term premium of option volatility, and average option-implied skewness.

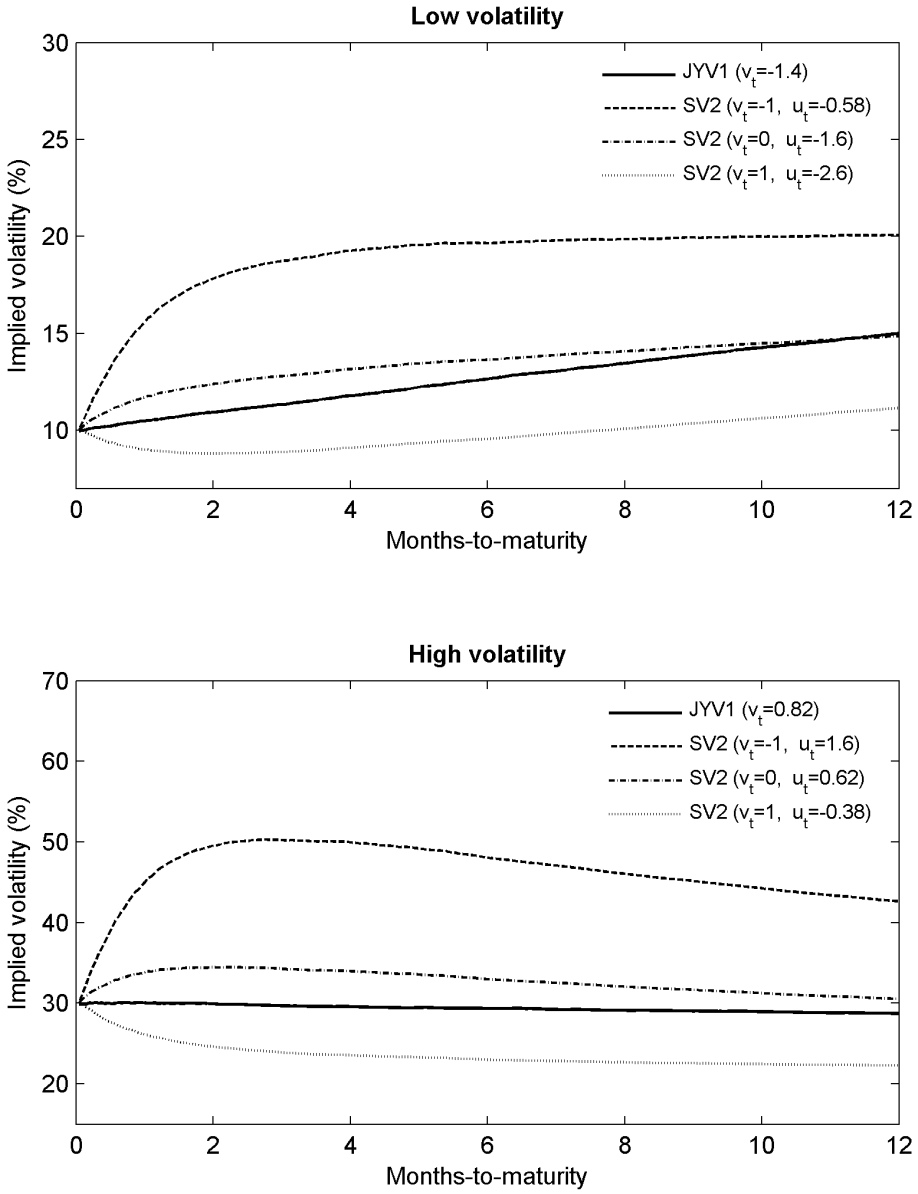

Figure 1.5: Term structures of implied volatility for the JYV1 and SV2 models.

The top panel represents a low volatility scenario with the instantaneous volatility set at 10%. The bottom panel represents a high volatility scenario with the instantaneous volatility set at 30%.

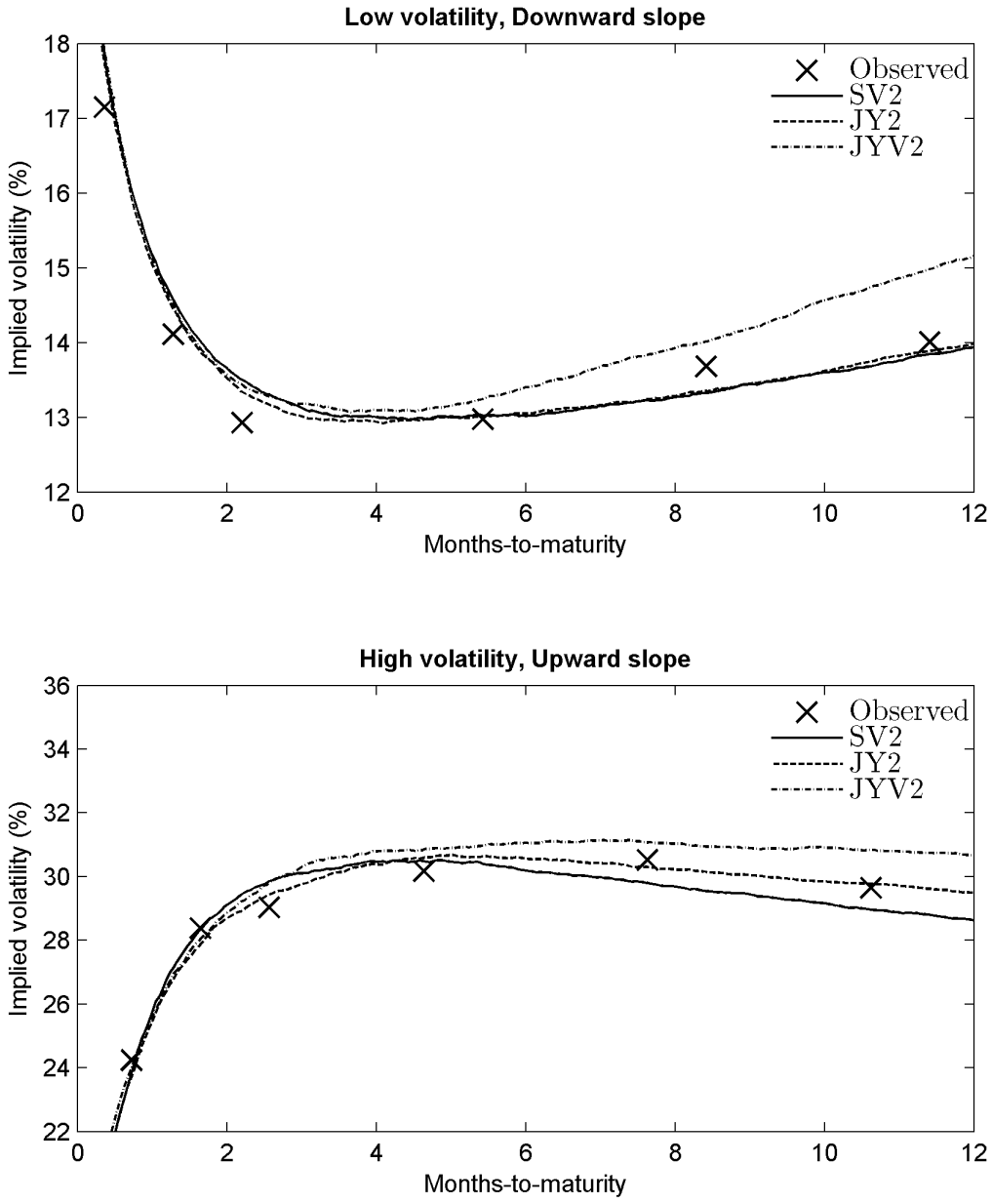

Figure 1.6: Actual fits in the term structure of implied volatility for the different two-factor volatility models.

The top panel is the term structure of implied volatility on January 29, 1999, when the term structure is declining even though short-term volatility is low. The bottom panel is the term structure of implied volatility on January 9, 1996, when the term structure is rising even though short-term volatility is high. Of course, the one-factor volatility models are unable to fit the data for these days.

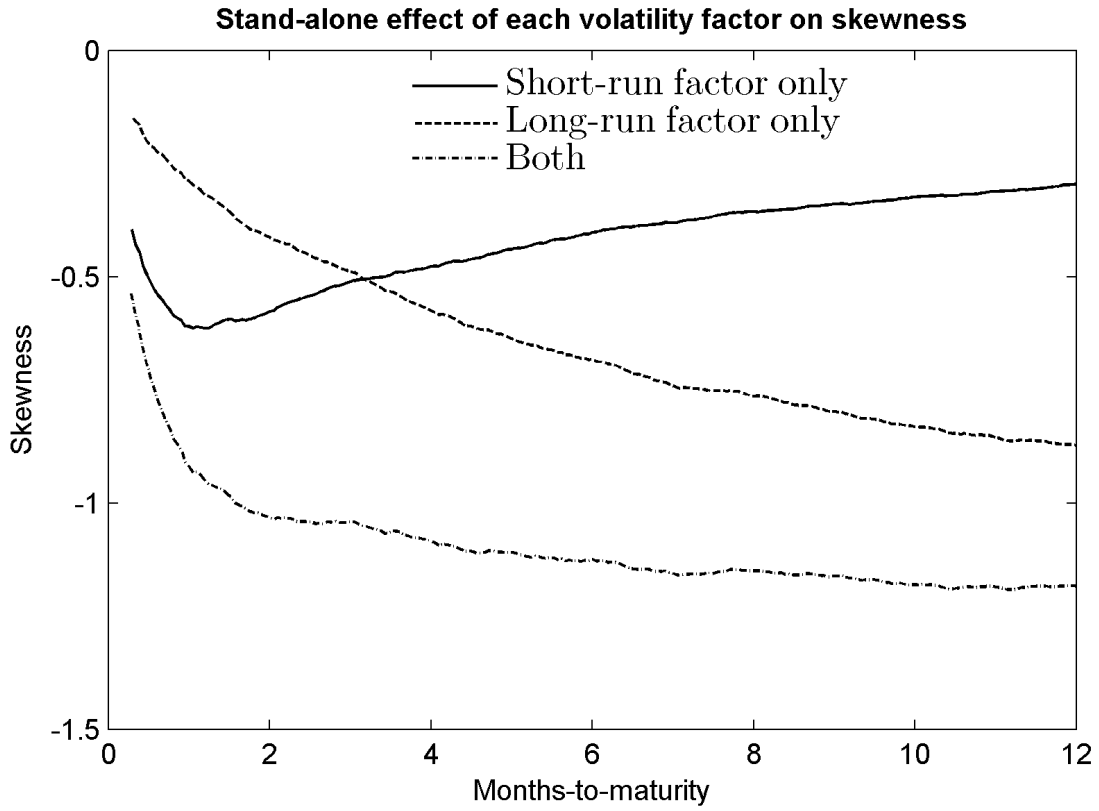

Figure 1.7: The stand-alone effect of each volatility factor on skewness.

I compute the term structure of skewness in three different ways for the SV2 model. The solid line is the term structure of skewness generated by the short-run factor only, with the long-run factor set at zero; the dashed line is the term structure of skewness generated by the long-run factor only, with the short-run factor set at zero; the dash-dot line is the term structure of skewness generated by the both factors.

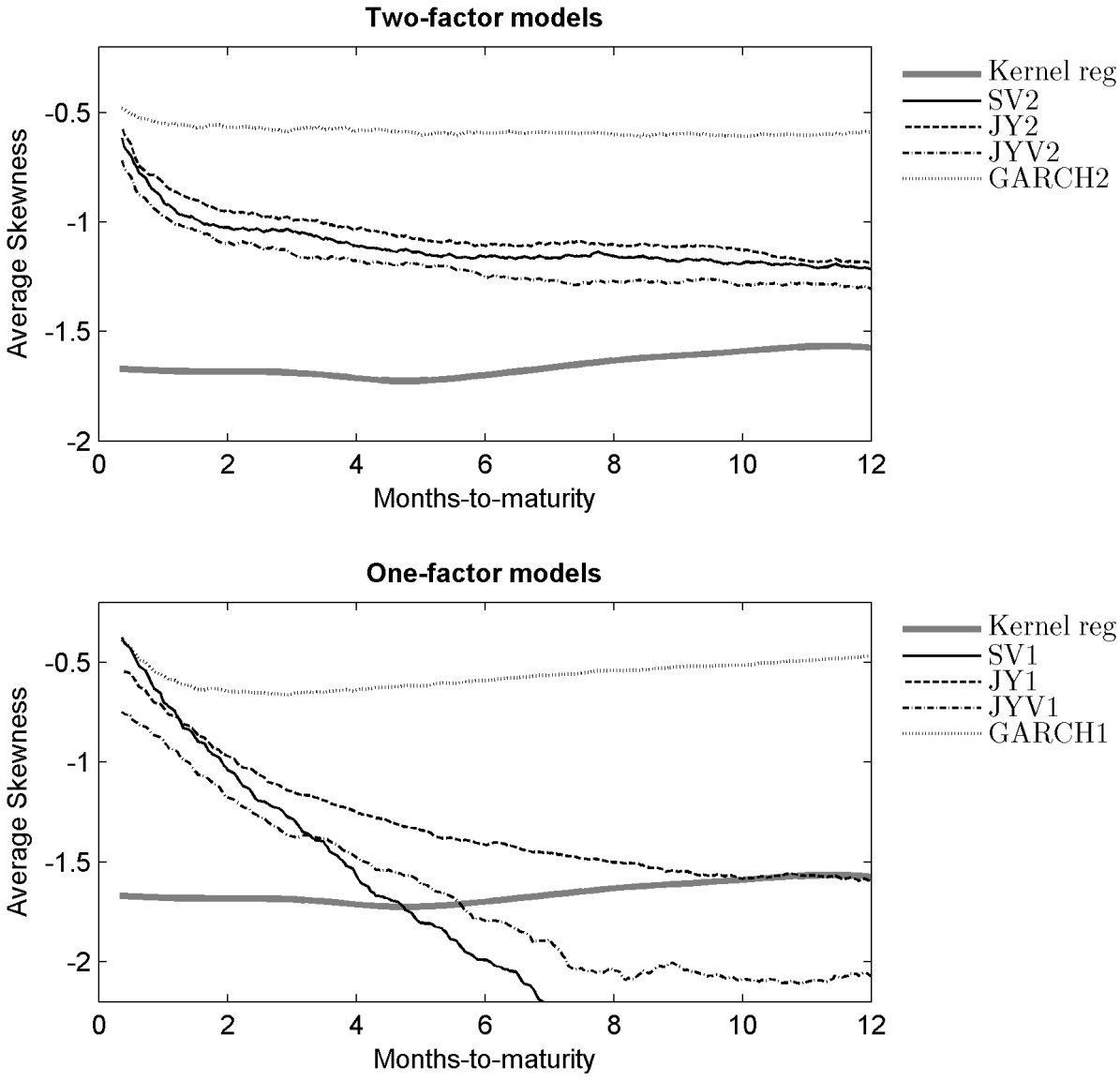

Figure 1.8: Comparison of the average term structures of skewness.

The top (bottom) panel shows the average term structures of skewness generated by the two-factor (one-factor) models, along with the observed average term structure of skewness (thick solid line).

Chapter 2

Beyond Stochastic Volatility and Jumps in Returns and Volatility

2.1 Introduction

Understanding volatility dynamics and improving option pricing have long been of interest to practitioners and academics. It is well-known that the volatility of many financial assets is time-varying, and an enormous amount of research has been devoted to studying this feature of financial data. But, there is strong empirical evidence suggesting that return distributions have time-varying skewness and kurtosis as well. For example, stochastic skewness in risk-neutral return distributions is implied by variation across time in the slope of the Black-Scholes implied volatility surface. Stochastic kurtosis is related to variation across time in the curvature of the Black-Scholes implied volatility surface. These are important features of observed option prices and are only weakly correlated with variation in option-implied volatility. Yet, modelling efforts that attempt to fit both physical and risk-neutral models simultaneously in a time-consistent manner typically include only a single state variable, volatility, and are thus capable of capturing little of the variation in implied skewness and kurtosis of returns that is observed in the data. Understanding variation in the shape of return distributions (and the shape of the implied volatility smile) is important in many applications, such as hedging and risk management. Any option pricing model that is unable to explain this feature of the data is clearly deficient.

The objective of this paper is to investigate model characteristics that are consistent with time-varying skewness and kurtosis in return distributions as is observed empirically in the option market. In particular, we look at models that allow for time-variation in volatility of volatility and

correlation between innovations in prices and volatility (leverage effect), both of which are able to generate variation in the shape of return distributions. We find strong evidence in favor of these features under the physical measure. An important aspect of our analysis is that the models are estimated using returns and implied volatility, but additional information about the shape of the return distribution embedded in option prices (i.e., under the risk-neutral measure) is not used in fitting them. By withholding this information from the model estimation, we are able to use it for diagnostic purposes. Toward this end, we look at some regressions to examine whether the implied state variables have explanatory power for option-implied skewness and kurtosis and find strong evidence that they do. That is, we find that information about the volatility of volatility and leverage effect extracted from the physical dynamics of returns and volatility has explanatory power for time-variation in the shape of returns distributions implied by option prices under the risk-neutral measure. This is important because it suggests that variation in the shape of risk-neutral return distributions (and of the Black-Scholes implied volatility smile) is not just due, for example, to changes in risk premia, but is associated with changes in related characteristics of the physical dynamics.

This paper adds to the existing literature on option-implied skewness. Dennis and Mayhew (2002) investigate the relation between option-implied skewness and firm characteristics, such as firm size, trading volume, and leverage. Han (2008) finds that market sentiment has some explanatory power for option-implied skewness. Harvey and Siddique (2000) look at asset pricing models that incorporate conditional skewness. They find that skewness has significant explanatory power for cross-sectional variation in expected returns across assets. Xing, Zhang, and Zhao (2007) find that option-implied skewness may be predictive of future returns. Bollen and Whaley (2004) look at the effect of net buying pressure in options markets on the shape of implied volatility smiles.

More closely related to this paper, there has also been work focused on understanding the relationship between option-implied skewness and the dynamics of the underlying asset price process. Das and Sundaram (1999) show that both volatility of volatility and correlation between the innovations in an asset's price and its volatility (leverage effect) affect the shape of the volatility

smirk and are thus possible candidates (see also Jones 2003). They show that the size and intensity of jumps in returns are also possibilities, though mostly affecting option-implied skewness at short terms to maturities. There has been some work toward implementing these ideas. For example, the two-factor stochastic volatility model of Christoffersen, Heston, and Jacobs (2009) is able to generate time-varying correlation, while Santa-Clara and Yan (2010) allow the jump intensity to be stochastic. In each of these papers, the authors find that option pricing performance is significantly improved.

Carr and Wu (2007) propose a stochastic skew model for foreign exchange rates with positive and negative jumps driven by independent Levy processes. Johnson (2002) looks at a stochastic volatility model with time-varying correlation between return and volatility innovations and shows that the model implies time-series variation in the shape of the implied-volatility smile. Jones (2003) proposes a constant elasticity of variance model that incorporates a time-varying leverage effect. Harvey and Siddique (1999) look at GARCH models that incorporate time-varying skewness.

In this paper, the underlying modelling framework is based on a standard single-factor stochastic volatility model. Although models of the affine (or affine-jump) class are often used in work of this kind due to their analytical tractability, these models have trouble fitting the data (e.g., Jones 2003, Ait-Sahalia and Kimmel 2007, Christoffersen, Jacobs, and Mimouni 2010). But since the techniques applied in this paper do not rely upon the analytical tractability of the affine models, we are able to choose among classes of models based on performance instead. We have found that log volatility models provide a useful starting point. We allow for contemporaneous jumps in both returns and volatility. We build on this framework by adding a regime-switching feature for the parameters corresponding to volatility of volatility and leverage effect. This idea is motivated by the fact that changes in either of these two variables, at least under the risk-neutral measure, are capable of generating variation in the shape of the Black-Scholes implied volatility smirk.

As pointed out by Das and Sundaram (1999), time-varying jump dynamics provide another

possible mechanism for generating variation in the shape of return distributions. Although we tested models including regime-switching in jump parameters, such models did not turn out to be empirically useful. We do not report the results for these experiments to keep the presentation manageable.

We make no effort to attach a particular economic interpretation to the regime states. We regard this modelling framework simply as a convenient way of generating a flexible family of models that is capable of behavior that captures an important feature of the data that we are interested in.

The regime-switching framework is useful because it is the simplest possible approach that incorporates time-variation in the parameters of interest. Our empirical results demonstrate that it is also sufficient to capture the essential characteristics of observed behavior that are of interest here. More sophisticated approaches involving continuous state spaces are of course possible. However, such models require additional assumptions not needed here regarding the form of the processes describing the evolution of the states. The analysis is also substantially less transparent.

Our empirical work uses S&P 500 index (SPX) option data. Option-implied volatility, skewness, and kurtosis are estimated using the model-free approach of Bakshi, Kapadia, and Madan (2003). Figure 2.1 shows the time-series plots of option-implied volatility, skewness, and kurtosis. Figure 2.2 shows the scatter plots of option-implied skewness and kurtosis versus volatility. It is evident from these plots that there are substantial and persistent variations in skewness and kurtosis, and that these variations are weakly correlated with the level of volatility. Models with only a single volatility state are hard to reconcile with the empirical features of these data. In particular, any option pricing model with only a single state variable implies that skewness and kurtosis are perfectly correlated with volatility, in stark contrast to the data.

Jones (2003) looks at a CEV model with different elasticity parameters for the parts of volatility innovations that are correlated/uncorrelated with price innovations. In this model, the strength of the leverage effect depends upon the level of the volatility state. Jones argues that this helps to capture time variation in the shape of return distributions. Pan (2002) includes a time-

varying jump risk premium that is proportional to the level of the volatility state. Both papers find that option-pricing performance is significantly improved. But, these papers include only a single state variable, volatility, which means that they are able to capture only a small part of the time variation in option-implied skewness and kurtosis.

One approach toward resolving this shortcoming is to recalibrate the model parameters on a daily basis using the observed panel of option prices with varying moneyness and time to maturity (e.g., Rubinstein 1994; Bakshi, Cao, and Chen 1997; Duffie, Pan, and Singleton 2000). However, this approach is not time-consistent and does not provide much help in understanding the dynamics of the underlying asset price nor the relationship between these dynamics and observed option prices.

An alternative approach is to include in the model an additional stochastic factor that is related to skewness and kurtosis (e.g., Christoffersen, Heston, and Jacobs 2009 and Santa-Clara and Yan 2010). This is the basic approach used in this paper. But, unlike Christoffersen, Heston, and Jacobs (2009) and Santa-Clara and Yan (2010), we do not use cross-sectional option prices or option-implied skewness and kurtosis in the estimation step. Given observations of option-implied volatility and skewness (or kurtosis), it is generally easy for any two-factor model to match both exactly, even if the model is badly misspecified. With this procedure, however, it is unclear if the implied dynamics of the states are actually present under the physical measure or are just artifacts of forcing the model to match the shape of the volatility smile. Differences between physical and risk-neutral dynamics are typically attributed to risk-premia. But, if the model is misspecified, these are also likely to be artifacts.

Our study takes a somewhat different task. While we match option-implied volatility, we do not force the model to match option-implied skewness or kurtosis exactly. We also do not look for risk premia which cause the models to fit the shape of the implied volatility smirk on average. Rather, we fit the models using only information from returns and implied volatility. By then regressing option-implied skewness and kurtosis on the implied states (and some control variables), we are able to draw useful conclusions about the extent to which the implied states are

informative about variation across time in the shape of the risk-neutral return distribution.

We also examine two different specifications for the jump structure, depending on whether jumps are scaled by the volatility state or not. The unscaled jump models (UJ) assume that jumps innovations are identically distributed across time. This form has been commonly used in the existing literature (e.g., Eraker, Johannes, and Polson 2003; Eraker 2004; Broadie, Chernov, and Johannes 2007). In the scaled jump models, on the other hand, jumps are larger on average when the volatility state is higher. Such models are potentially able to generate dynamics that more realistically reflect the data. Our results indicate a strong preference for the scaled jump form.

We do not demonstrate potential improvements in fitting observed option prices by using models such as those proposed in this paper. If one were actually interested in fitting observed option prices, it would be desirable to use the full panel of option prices observed each day to back out implied states. The model would be able to match perfectly on a day-by-day basis both the volatility and skewness of the risk-neutral distribution implied by observed option prices (corresponding roughly to the level and slope of the Black-Scholes implied-volatility smile), with an accompanying improvement in fit to the option prices themselves. But, as described above, this is not the purpose of the paper.

To summarize, the analysis is comprised of two main parts. First, we fit the models using SPX prices and option-implied volatility. We compare models based on log likelihoods and other diagnostics. The second step involves testing the explanatory power of the implied regime states for option-implied skewness and kurtosis. Critically, information about the shape of the risk-neutral distribution is not used in the estimation step, which makes this diagnostic useful.

Regarding the first step of the analysis (model estimation), including jumps in the model (whether scaled or unscaled) provides a huge improvement relative to the base model with no jumps. The log likelihood increases by over 300 points. Other diagnostics of model fit are also greatly improved. The models with scaled jumps heavily dominate those with unscaled jumps. For the models without regime-switching, the difference is over 40 points in log likelihood.

Including the regime-switching feature provides additional improvements that are highly

significant. The model with regime switching in volatility of volatility and scaled jumps provides an increase of 110 point in log likelihood relative to the model without regime switching, with improvements in the other diagnostics of model fit as well. The regime switching model nests the model without regime switching and includes four additional parameters. A standard likelihood ratio test indicates that a difference of 110 points in log likelihood corresponds to a p-value of around 10^{-22}. These results provide overwhelming evidence against models in which volatility of volatility and leverage effect are constant.

In the second step of the analysis, regressions testing whether the implied states have explanatory power for option-implied skewness and kurtosis are also decisive. The slope coefficients are strongly significant and in the expected directions. Option-implied skewness tends to be more negative when volatility of volatility is high or the leverage effect is more pronounced (more negative correlation between price and volatility innovations). Option-implied kurtosis tends to be more positive when volatility of volatility is high or the leverage effect is more pronounced. We also include several control variables that could be related to variation in the shape of return distributions. The VIX index is included to control for correlation between option-implied volatility and option-implied skewness and kurtosis. We also include a nonparametric measure of the variance risk premium (VRP) based on the difference between the VIX index and realized volatility (Bollerslev, Gibson, and Zhou 2011, Carr and Wu 2009). Finally, the jump variation (JV) measure of Barndorff-Nielsen and Shephard (2004) is used as a proxy for jump risk. Our regression results are robust to inclusion of these control variables.

As in the estimation step, models with scaled jumps always dominate models with unscaled jumps. For option-implied skewness, the regression on the three control variables alone gives an adjusted R^2 of only 9.2%. VIX and JV are significant, but not VRP. The best regression model includes both volatility of volatility and leverage states (in addition to the control variables). This model has an R^2 of over 32% and slope coefficients for the states have t-statistics with absolute value greater than 11 (corresponding to p-values of less than 10^{-27}). Results for the implied kurtosis regressions are qualitatively similar (though weaker).

The remainder of the paper is organized as follows. Section 2.2 describes the models under consideration. In Section 2.3, we develop the estimation strategy, focusing on how to deal with the unobservable volatility and regime states. Model estimates and diagnostics are presented in Section 2.4. Section 2.5 investigates the explanatory power of implied regime states from the models for risk-neutral implied-skewness and kurtosis. And, Section 2.6 concludes.

2.2 Models

The modeling framework used in this paper is based on a standard stochastic volatility model, variants of which have appeared frequently in the existing literature. We simultaneously fit both physical and risk-neutral models. Although affine models are often used in work of this kind due to their analytical tractability, these models have trouble fitting the data. However, since the techniques applied in this paper do not rely upon the nice analytical features of the affine models, we are able to consider other possibilities. We have found that log volatility models work well and use these as the basis for our modelling framework.

A large body of literature documents the significance of jumps in returns and volatilities. Bakshi, Cao, and Chen (1997), Bates (2000), Pan (2002), and Andersen, Benzoni, and Lund (2002) find jumps in returns to be important. Duffie, Pan, and Singleton (2000), Eraker, Johannes, and Polson (2003), Eraker (2004), and Broadie, Chernov, and Johannes (2007) argue in favor of including jumps in volatility as well. In light of this evidence, we allow for jumps in both prices and volatility.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and information filtration $\{\mathcal{F}_t\}$, the ex-dividend stock price, x_t, is assumed to evolve as

$$\begin{aligned} dx_t/x_t &= \left[\mu - \bar{\mu}_{1J_t}\lambda_1\right]dt + \exp(v_t/2)dW_{1t} + (e^{J_{1t}} - 1)dN_{1t} \\ dv_t &= \left[\kappa(\bar{v} - v_t) - \bar{\mu}_{2J_t}\lambda_1\right]dt + \sigma(s_t)dW_{2t} + J_{2t}dN_{1t} \\ ds_t &= (1 - 2s_t)dN_{2t} \end{aligned} \tag{2.1}$$

where v_t and s_t are the volatility state and the regime state, respectively. The regime state is either

0 or 1. W_{1t} and W_{2t} are standard Brownian motions with regime-dependent correlation $\rho(s_t)$. N_{1t} and N_{2t} are Poisson processes with intensity λ_1 and $\lambda_2(s_t)$, respectively.

We consider two different forms for the jump structure, depending upon whether jumps are scaled by the same factor as the Brownian motion or unscaled. The unscaled model (UJ) assumes that jump innovations are identically distributed across time. In this case, the jumps are bivariate normal with $J_{1t} \sim N\left(\mu_{1J}, \sigma_{1J}^2\right)$, $J_{2t} \sim N\left(\mu_{2J}, \sigma_{2J}^2\right)$, and $\text{corr}(J_{1t}, J_{2t}) = \rho_J$. This form has been commonly used in the existing literature (e.g., Eraker, Johannes, and Polson 2003; Eraker 2004; Broadie, Chernov, and Johannes 2007). In contrast, the scaled model (SJ) assumes that jumps scale in proportion to the volatility of the diffusion component of the process. That is, J_{1t} and J_{2t} are bivariate normal with $J_{1t}/\exp(v_t/2) \sim N\left(\mu_{1J}, \sigma_{1J}^2\right)$, $J_{2t}/\sigma(s_t) \sim N\left(\mu_{2J}, \sigma_{2J}^2\right)$, and $\text{corr}(J_{1t}, J_{2t}) = \rho_J$. By generating larger jumps when volatility is higher, the SJ model is potentially capable of providing more realistic dynamics (this hypothesis is indeed confirmed in the empirical section). Let $\bar{\mu}_{1Jt} = E(e^{J_{1t}} - 1)$ and $\bar{\mu}_{2Jt} = E(J_{2t})$ denote the mean jump sizes (note that these are time-varying for the SJ models).

The regime-dependent parameters, σ and ρ, allow for variation across time in the volatility of volatility and leverage effect. These are the mechanisms by which it is possible to generate time-varying skewness and kurtosis. The regime-dependence of λ_2 lets the regimes differ in persistence.

For simplicity, we only look at models with two possible regimes, although extending our techniques to more regimes is straightforward (at the cost of more free parameters to estimate). The regime state process is the continuous-time analog of a discrete-time Markov switching model in which the probability of switching from state s to state $1-s$ is $\lambda_2(s)\Delta t$ $(s = 0, 1)$. If the discretization interval is short relative to the intensity of the Poisson process, the discrete-time approximation will be good. This is the case for all of the models estimated in this paper.

It is often useful to transform the model into log prices, $y(t) = \log(x_t)$,

$$\begin{aligned} dy_t &= \left[\mu - \bar{\mu}_{1Jt}\lambda_1 - \frac{1}{2}\exp(v_t)\right]dt + \exp(v_t/2)dW_{1t} + J_{1t}dN_{1t} \\ dv_t &= \left[\kappa(\bar{v} - v_t) - \bar{\mu}_{2Jt}\lambda_1\right]dt + \sigma(s_t)dW_{2t} + J_{2t}dN_{1t} \\ ds_t &= (1 - 2s_t)dN_{2t}. \end{aligned} \tag{2.2}$$

While the dynamics of the underlying asset are described by the above model, options are priced according to a risk-neutral measure, $\mathbb{Q}$. We assume that under this measure the model takes the form

$$\begin{aligned} dy_t &= \left[r_t - q_t - \bar{\mu}_{1Jt}\lambda_1 - \frac{1}{2}\exp(v_t)\right]dt + \exp(v_t/2)dW_{1t}^{\mathbb{Q}} + J_{1t}^{\mathbb{Q}}dN_{1t}^{\mathbb{Q}} \\ dv_t &= \left[\kappa(\bar{v} - v_t) - \eta(s_t)v_t - \bar{\mu}_{2Jt}\lambda_1\right]dt + \sigma(s_t)dW_{2t}^{\mathbb{Q}} + J_{2t}^{\mathbb{Q}}dN_{1t}^{\mathbb{Q}} \\ ds_t &= (1 - 2s_t)dN_{2t}^{\mathbb{Q}} \end{aligned} \tag{2.3}$$

where r_t and q_t denote the risk-free rate and the dividend rate, respectively. $W_{1t}^{\mathbb{Q}}$ and $W_{2t}^{\mathbb{Q}}$ are standard Brownian motions with regime-dependent correlation $\rho(s_t)$ under the risk-neutral measure.

Jump parameters are the same under physical and risk neutral measures. In other words, we do not attempt to identify any jump risk premia. This is because the models and data used in this paper have limited power to separately identify jump risk premium and diffusive volatility premia. As is well-known, option-implied volatility provides a biased estimate of the volatility of observed returns. Jump risk and diffusive volatility premia are two possible mechanisms for accounting for this difference. The effects of these can be disentangled either by looking at cross-sections of option prices across moneyness or by looking at options with varying times to maturity (see Pan 2002 for a detailed discussion of this issue). Studying the characteristics of jump risk premia is certainly an interesting issue, and has been looked at by Pan (2002) and Broadie, Chernov, and Johannes (2007) among others. However, it is not the topic of this paper, and our models and data choices were made to address other issues. We identify a volatility risk premium, η, which accounts for the difference between physical and option-implied volatility, absorbing any potential jump risk premium as well.

The market variance risk premium is generally found to be negative (e.g., Bakshi and Kapadia 2003; Coval and Shumway 2001; Carr and Wu 2009). This finding can be understood in the framework of classic capital asset pricing theory (e.g., Heston 1993; Bakshi and Kapadia 2003; Bollerslev, Gibson, and Zhou 2011). But, according to theory, the variance risk premium should be dependent on the volatility of volatility and the leverage effect (correlation between returns and changes in volatility state). For example, the variance risk premium should be more negatively priced as the volatility of volatility increases or the correlation decreases. We allow for this possibility by letting the volatility risk premium parameter $\eta(s_t)$ be regime-dependent.

In the application, we look at several variants of this model:

SV Stochastic volatility, no jumps, no regime switching.

SJ Stochastic volatility, volatility-scaled jumps, no regime switching.

UJ Stochastic volatility, volatility-unscaled jumps, no regime switching.

SJ-RV Stochastic volatility, volatility-scaled jumps, regime switching for σ.

UJ-RV Stochastic volatility, volatility-unscaled jumps, regime switching for σ.

SJ-RL Stochastic volatility, volatility-scaled jumps, regime switching for ρ.

UJ-RL Stochastic volatility, volatility-unscaled jumps, regime switching for ρ.

SJ-RVL Stochastic volatility, volatility-scaled jumps, regime switching for both σ and ρ.

UJ-RVL Stochastic volatility, volatility-unscaled jumps, regime switching for both σ and ρ.

And finally, in our empirical work, we use an Euler scheme approximation to the model. For the physical model (and analogously for the risk-neutral model), the approximation is given by

$$\begin{aligned} y_{t+1} &= y_t + \mu - \bar{\mu}_{1J_t}\lambda_1 - \frac{1}{2}\exp(v_t) + \exp(v_t/2)\varepsilon_{1,t+1} + \sum_{j>N_{1,t}}^{N_{1,t+1}} \xi_{1j} \\ v_{t+1} &= v_t + \kappa(\bar{v} - v_t) - \bar{\mu}_{2J_t}\lambda_1 + \sigma(s_t)\varepsilon_{2,t+1} + \sum_{j>N_{1,t}}^{N_{1,t+1}} \xi_{2j} \end{aligned} \tag{2.4}$$

where $\varepsilon_{1,t+1}$ and $\varepsilon_{2,t+1}$ are standard normals with correlation $\rho(s_t)$, and ξ_{1j} and ξ_{2j} have the same distribution as the jumps in the continuous-time model. The regime state, s_t, follows the discrete-time Markov process with $p(s_{t+1} = i|s_t = j) = P_{ij}$, corresponding to the transition matrix

$$P = \begin{pmatrix} \pi_0 & 1-\pi_1 \\ 1-\pi_0 & \pi_1 \end{pmatrix}.$$

For computational purposes, we need to set some upper bound for the maximum number of jumps possible in a single day. In our application we set this upper bound to five jumps. Given our estimates for jump intensity, the approximation error due to imposing this constraint is negligible.

2.3 Methodology

The introduction of the regime-switching feature means that the models used in this paper require the development of new estimation techniques. While the estimation strategies used in similar work often rely heavily on computationally costly simulation methods, the approach we propose in this paper can be executed in several minutes on a typical desktop PC. Our approach consists of three steps: (1) back out volatility states from observed option prices, (2) filter regime states using a Bayesian recursive filter, and (3) optimize the likelihood function using the volatility and regime states obtained in the previous two steps. As a by-product of the algorithm, we obtain a series of generalized residuals which we make use of for model diagnostics. A detailed description of each step of the procedure is provided below.

2.3.1 Extracting the volatility states

Building on the work of Chernov and Ghysels (2000), Pan (2002), Ait-Sahalia and Kimmel (2007), and others, we make use of observed option prices in addition to the price of the underlying asset to estimate the model. The panel of option prices observed on any particular day implies a risk-neutral distribution of returns for the underlying asset (see, e.g., Derman and Kani 1994,

Dupire 1994, Rubinstein 1994, and Ait-Sahalia and Lo 1998). Neuberger (1994) and Carr and Madan (1998) provide a model-free approach to estimating the integrated variance of returns under the risk-neutral measure using observed option prices (see also Demeterfi, Derman, Kamal, and Zou 1999 and Carr and Wu 2009). Bakshi, Kapadia, and Madan (2003) build on this work and extend it to higher moments.

Given estimates for the risk-neutral variance it is possible to back out spot volatility and regime states. The idea of using model-free estimates of option-implied volatility in this manner has been used in similar applications by Jones (2003), Duan and Yeh (2007), and Ait-Sahalia and Kimmel (2007), among others.

Volatility can also be estimated using high-frequency data or based on returns data alone. However, Blair, Poon, and Taylor (2001), Christensen and Prabhala (1998), and Fleming (1998) argue that, at least for the S&P 100 index, option prices are more informative than realized volatility in forecasting future volatility. Jiang and Tian (2005) find that the model-free implied volatility offers the most efficient information for volatility forecasting, and that it subsumes the information contained in the Black-Scholes implied volatility and the realized volatility.

The problem of how to back out the corresponding volatility and regime states remains. Following Bakshi, Kapadia, and Madan (2003), let IV_t^T denote the square root of the variance of log returns on the interval $(t,T]$ under the risk-neutral measure. Given a risk-neutral model for stock price dynamics and values for the volatility and regime states, IV_t^T can be obtained by integrating the quadratic variation of the log stock price. For the SJ models, for example, we get

$$\mathrm{IV}_t^T = \sqrt{\frac{1}{T-t} E_t^{\mathbb{Q}} \left\{ \int_t^T e^{\nu_\tau} \left[1 + \lambda_1 (\mu_{1J}^2 + \sigma_{1J}^2) \right] d\tau \right\}}.$$

The expectation in the above expression can be computed by means of Monte Carlo simulations.

For the SJ models, for example, we get

$$\mathrm{IV}_t^T = \sqrt{\frac{1}{(T-t)S}\sum_{s=1}^{S}\left\{\int_t^T e^{v_\tau^{(s)}}\left[1+\lambda_1(\mu_{1J}^2+\sigma_{1J}^2)\right]d\tau\right\}}, \tag{2.5}$$

where for each s, $v_\tau^{(s)}$ is obtained by simulating a path from the risk-neutral analog of Equation (2.4) for $t < \tau \leq T$, and S denotes the number of simulation paths.

However, we actually want to go in the reverse direction. That is, observed values for IV_t^T are available and we need to obtain the corresponding volatility and regime states, v_t and s_t, by inverting Equation (2.5). We begin by showing how to do this conditional on the regime state (the issue of backing out the regime state is addressed in the next subsection).

Given an initial value for the regime state, $s_t = i$, the first step is to obtain an approximation to the mapping from spot to integrated volatility,

$$\Gamma_i : \mathrm{SV} \longrightarrow \mathrm{IV},$$

where SV denotes the spot volatility, i.e., $\mathrm{SV}_t = \exp(v_t/2)$, and the subscript i denotes the conditioning on initial regime state. The simplest way to do this is to evaluate (2.5) on some grid of initial values for SV and then use some curve fitting technique to approximate Γ_i. That is, let $\widehat{\mathrm{SV}}_1 < \widehat{\mathrm{SV}}_2 < \cdots < \widehat{\mathrm{SV}}_G$ be the grid, where G is the number of grid points and we use hats to indicate that these are grid points rather than data. For each $g = 1, \ldots, G$, evaluate $\widehat{\mathrm{IV}}_g = \Gamma_i(\widehat{\mathrm{SV}}_g)$ using Monte Carlo methods as described above (note that while the initial regime state is given, it evolves randomly thereafter). Then approximate Γ_i based on the collection of pairs $\{(\widehat{\mathrm{SV}}_g, \widehat{\mathrm{IV}}_g)\}_{g=1}^G$. As long as this mapping is monotonic, it is equally straightforward to approximate the inverse, $\Gamma_i^{-1} : \mathrm{IV} \longrightarrow SV$, which is what we are really interested in. Let $\widehat{\Gamma}_i^{-1}$ denote the approximation.

In principle, there are many curve fitting schemes one could use. However, since we are using the result in a numerical optimization, it is best if the scheme results in smooth derivatives. We have found that simply fitting a cubic polynomial to the collection $\{(\widehat{\mathrm{SV}}_g, \widehat{\mathrm{IV}}_g)\}_{g=1}^G$ using

nonlinear least squares works well. For all of the models used in this paper, Γ_i is close to linear, so a cubic polynomial gives plenty of flexibility. We also tried higher order polynomials, splines, and various other interpolation schemes, but none provided noticeable improvements. Approximation errors are negligible for any reasonable scheme.

Once we have obtained $\widehat{\Gamma}_i^{-1}$, the hard work is done. Given an observed value for option-implied integrated volatility, IV_t and assuming $s_t = i$, evaluate $\widehat{\Gamma}_i^{-1}$ to obtain SV_t. Then, the volatility state itself is given by $v_t = 2\log(\mathrm{SV}_t)$. Repeat this for each $t = 1,\ldots,n$. The important thing to notice here is that computing $\widehat{\Gamma}_i$, which is the costly step, need only be done once (for each set of candidate parameter values). Once this is accomplished, the evaluation step is fast.

2.3.2 Filtering the regime states

Given an observed value of IV_t, we now have two possible values for SV_t (and v_t), one for each regime state. Let v_t^j denote the volatility state corresponding to regime j. The second step of the estimation involves applying a filter to compute $p_t^j = p(s_t = j|\mathcal{F}_t) = p(v_t = v_t^j|\mathcal{F}_t)$.

The filter is constructed recursively using standard techniques. Let $p_t = (p_t^0, p_t^1)'$ for each $t = 0,\ldots,n$, and initialize the filter by setting p_0^j equal to the marginal probability of state j ($j = 0,1$). Now, suppose that p_t is known. The problem is to compute p_{t+1}. This is given by

$$p_{t+1}^j = \frac{\sum_{i=0}^1 p(y_{t+1}, v_{t+1}^j|y_t, v_t^i, s_t = i)\cdot p(s_{t+1} = j|s_t = i)\cdot p_t^i}{\sum_{k=0}^1\sum_{i=0}^1 p(y_{t+1}, v_{t+1}^k|y_t, v_t^i, s_t = i)\cdot p(s_{t+1} = k|s_t = i)\cdot p_t^i}, \qquad j = 0,1.$$

The third factor in the summand is known from the previous step of the recursion. The second factor is determined by the Markov transition matrix of the regime state process. But, some attention is required in computing the first factor. Since we allow for the possibility of more than a single jump per day, it is necessary to sum over the potential number of jumps,

$$p(y_{t+1}, v_{t+1}^j|y_t, v_t^i, s_t = i) = \sum_{k=0}^{\mathrm{NJ}_{\max}} p(y_{t+1}, v_{t+1}^j|y_t, v_t^i, s_t = i, \mathrm{NJ}_t = k)p(\mathrm{NJ}_t = k)$$

where NJ_t is the number of potential jumps on day t, $\mathrm{NJ}_{\max}$ is the maximum number of allowable jumps in a single day, and $p(\mathrm{NJ}_t = k) = [\lambda_1]^k e^{-\lambda_1}/k!$ is given by the Poisson distribution with intensity λ_1. The distribution of $\left(y_{t+1}, v_{t+1}^j | y_t, v_t^i, s_t = i, \mathrm{NJ}_t = k\right)$ is bivariate normal with mean and variance given by summing the means and variances of the diffusive part of the process and k jumps, according to the equation (2.4).

It is sometimes useful to speak of the filtered regime state. By this we mean the expected value of s_t conditional on information available at time t,

$$\hat{s_t} = E_t(s_t) = p_t^0 \cdot 0 + p_t^1 \cdot 1 = p_t^1.$$

2.3.3 Maximum likelihood estimation

Having backed out volatility states and computed filtered regime state probabilities, computing the log likelihood is straightforward:

$$\begin{aligned}\log L(\{y_t\}_{t=1}^n, \{\mathrm{IV}_t\}_{t=1}^n; \theta) \approx \sum_{t=1}^{n-1} \sum_{i=0}^{1} \sum_{j=0}^{1} \Big[& \log p(y_{t+1}, v_{t+1}^j | y_t, v_t^i, s_t = i) + \log p(s_{t+1} = j | s_t = i) \\ & + \log p(s_t = i) + \log J_{t+1}^j \Big],\end{aligned} \tag{2.6}$$

where $J_{t+1}^j = \left| dv_{t+1}^j / dIV_{t+1} \right|$ is the Jacobian corresponding to regime state j. Recall that the mapping from volatility state, v_t^j, to IV_t is given by $\mathrm{IV}_t = \Gamma_j[\exp(v_t^j/2)]$. The Jacobian is obtained from the derivative of the inverse of this. As in the preceding subsection, $p(y_{t+1}, v_{t+1}^j | y_t, v_t^i, s_t = i)$ must be computed by summing across the number of potential jumps.

Parameter estimates are obtained by numerical optimization,

$$\widehat{\theta} = \arg\max \log L(\{y_t\}_{t=1}^n, \{\mathrm{IV}_t\}_{t=1}^n; \theta).$$

We use a BHHH optimizer (Berndt, Hall, Hall, and Hausman 1974), but the criterion function is well-behaved and nearly any optimizer will work fine.

2.3.4 Specification Tests

Our diagnostic tests are based on the idea of generalized residuals (Bai 2003, Duan 2003, and others). Let $\{z_t\}_{t=1}^n$ be a sequence of random vectors where z_t has distribution $G_t(z|\mathcal{F}_{t-1})$. Let $u_t = G_t(z_t|\mathcal{F}_{t-1})$ $(t = 1,\ldots,n)$. If the model is correctly specified, $\{u_t\}$ should be *i.i.d.* uniform$(0,1)$. The hypothesis that $\{G_t(z_t|\mathcal{F}_{t-1};\theta)\}$ is the true data generating process for $\{z_t\}$ can be tested by performing diagnostics on $\{u_t\}$.

But, it is often more useful to instead perform diagnostics on

$$\widetilde{u}_t = \Phi^{-1}(u_t), \qquad t = 1,\ldots,n \tag{2.7}$$

where Φ is the standard normal distribution function. In this case, the transformed residuals $\{\widetilde{u}_t\}$ should be *i.i.d.* standard normal under the hypothesis of correct model specification. It is these that we shall refer to as generalized residuals.

For the models in this paper, the generalized residuals are computed in a manner similar to equation (2.6),

$$u_{t+1} = \sum_{i=0}^{1}\sum_{j=0}^{1} P(y_{t+1}, v_{t+1}^j | y_t, v_t^i, s_t = i)\cdot p(s_{t+1} = j | s_t = i)\cdot p_t^i,$$

where $P(\cdot)$ denotes a cdf. As before, this cdf must be computed by summing across the number of potential jumps,

$$P(y_{t+1}, v_{t+1}^j | y_t, v_t^i, s_t = i) = \sum_{k=0}^{\mathrm{NJ}_{\max}} P(y_{t+1}, v_{t+1}^j | y_t, v_t^i, s_t = i, \mathrm{NJ}_t = k) p(\mathrm{NJ}_t = k).$$

These residuals correspond to the joint distribution of price and volatility innovations. Although one could certainly study these, we have found it more useful to study marginal residuals corre-

sponding to price and volatility innovations separately,

$$u_{y,t+1} = \sum_{i=0}^{1} P(y_{t+1}|y_t, v_t^i, s_t = i) \cdot p(s_t = i)$$
$$u_{v,t+1} = \sum_{i=0}^{1}\sum_{j=0}^{1} P(v_{t+1}^j|y_t, v_t^i, s_t = i) \cdot p(s_{t+1} = j|s_t = i) \cdot p(s_t = i).$$

In the diagnostics reported in our application, we always use the generalized residuals obtained by applying the inverse normal cdf to these, $\tilde{u}_{y,t} = \Phi^{-1}(u_{y,t})$ and $\tilde{u}_{v,t} = \Phi^{-1}(u_{v,t})$.

Having constructed these generalized residuals, testing can proceed using standard time series techniques. There is a wide variety of techniques available to test that the residuals are independent and normally distributed. In this paper, we look at Jarque-Bera tests and normal-quantile plots to test normality, and Ljung-Box tests and correlograms to test for autocorrelation. We look at Ljung-Box tests and correlograms for both the residuals as well as squared residuals. Tests based on the squared residuals allow us to look for unexplained stochastic volatility in returns and stochastic volatility of volatility. We have found these diagnostics to be helpful.

2.4 Empirical results

2.4.1 Data

The application uses daily S&P 500 index (SPX) option data from Jan 1, 1993 through Dec 31, 2008 ($N = 4025$). These data were obtained directly from the CBOE. To address the issue of nonsynchronous closing times for the SPX index and option markets, SPX close prices are computed using put-call parity based on closing prices for at-the-money options (see, e.g., Ait-Sahalia and Lo 1998). Option prices are taken from the bid-ask midpoint at each day's close. Options with zero bid/ask prices or where the the bid-ask midpoint is less than 0.125 are discarded. We also eliminate options violating the usual lower bound constraints. That is, we require $C(t,\tau,K) \geq \max(0, x_t \exp(-q_t\tau) - K\exp(-r_t\tau))$ and $P(t,\tau,K) \geq \max(0, K\exp(-r_t\tau) - x_t\exp(-q_t\tau))$ where $C(t,\tau,K)$ and $P(t,\tau,K)$ are the time t prices of call and put options with time-to-maturity τ and

strike price K, x is the index price, q is the dividend payout rate, and r is the risk-free rate. Finally, we require that valid prices exist for at least two out-of-the-money call and put options for each day. Options are European, so there is no issue regarding early exercise premium.

Time-series of one-month risk-neutral volatility, skewness, and kurtosis are computed using SPX option prices following the model-free approach of Bakshi, Kapadia, and Madan (2003). Following Carr and Wu (2009), we use the two closest times to maturity greater than eight days and linearly interpolate to construct 30 day constant maturity series. Jiang and Tian (2007) report the possibility of large truncation and discretization errors in the VIX index. To reduce such errors, we follow the approach of Carr and Wu (2009) in interpolating/extrapolating option prices on a fine grid across moneyness.

Five-minute intraday S&P 500 index returns are used to compute measures of the variance risk premium and jump risk. Since these variables are possibly related to option-implied skewness and kurtosis, we include them as control variables in the regressions of option-implied skewness and kurtosis on the regime state (Section 2.5). The high-frequency data were obtained from Tick-Data.com.

Following Andersen, Bollerslev, Diebold, and Ebens (2001), Andersen, Bollerslev, Diebold, and Labys (2003), and Barndorff-Nielsen and Shephard (2002), daily realized volatility is obtained by summing the squared intraday returns over each day,

$$\mathrm{RV}_t^{(d)} \equiv \sum_{j=1}^{1/\Delta} \left(y_{t-1+j\Delta} - y_{t-1+(j-1)\Delta} \right)^2$$

where Δ is the sampling interval for the intraday data (we use five minute intervals). Monthly realized volatility is obtained by summing the daily realized volatilities over the previous month,

$$\mathrm{RV}_t \equiv \sum_{i=0}^{21} \mathrm{RV}_{t-i}^{(d)}.$$

Following Carr and Wu (2009), we define the variance risk premium as the log difference

between monthly realized variance and option-implied variance,

$$\mathrm{VRP}_t \equiv \log\left(\mathrm{RV}_t/\mathrm{VIX}_t^2\right),$$

where VIX_t is the VIX index, divided by $\sqrt{12}$ to get a monthly volatility measure comparable to RV_t. We use the log difference because we find that it provides a better measure than the difference in levels.

A measure of jump risk is obtained using the approach of Barndorff-Nielsen and Shephard (2004). The bipower variation is given by

$$\mathrm{BV}_t^{(d)} \equiv \frac{\pi}{2}\sum_{j=2}^{1/\Delta}\left(y_{t-1+j\Delta} - y_{t-1+(j-1)\Delta}\right)\left(y_{t-1+(j-1)\Delta} - y_{t-1+(j-2)\Delta}\right).$$

The daily jump variation is defined by subtracting the daily bipower variation from the daily realized volatility,

$$\mathrm{JV}_t^{(d)} \equiv \max(\mathrm{RV}_t^{(d)} - \mathrm{BV}_t^{(d)}, 0).$$

And, finally, the monthly jump variation is obtained by summing the the daily jump variations over the previous month,

$$\mathrm{JV}_t \equiv \sum_{i=0}^{21}\mathrm{JV}_{t-i}^{(d)}.$$

We shall typically refer to RV_t and JV_t as the realized volatility and jump variation hereafter (omitting the “monthly" specifier). Note that these variables are defined using information available at time t.

One- and three-month Treasury bill rates (obtained from the Federal Reserve website), interpolated to match option maturity, are used as a proxy for the risk-free rate. Dividend rates are obtained from the Standard and Poor’s information bulletin. Actual dividend payouts are used as a proxy for expected payouts.

Throughout, time is measured in trading days. SPX returns and option-implied moments

are plotted in Figure 2.1. Figure 2.2 shows scatter plots of option-implied skewness and kurtosis versus option-implied volatility. Summary statistics are reported in Table 2.1.

2.4.2 Parameter estimates and model comparisons

Parameter estimates and log likelihoods for all of the models under consideration are shown in Table 2.2. As found in previous work, including jumps in returns and volatility gives a large improvement in log likelihood relative to a model with no jumps (over 300 points in log likelihood). The SJ model (jumps scaled by volatility state) is strongly preferred over the UJ model (unscaled jumps). The improvement is over 40 points in log likelihood with the same number of parameters. In the scaled jump model, jumps are larger on average when overall volatility is high. In the unscaled jump model, in contrast, jump sizes are identically distributed across time and unaffected by the volatility state. Although models with unscaled jumps have been commonly used in previous work, (e.g., Pan 2002; Eraker, Johannes, and Polson 2003; Broadie, Chernov, and Johannes 2007), this form is not consistent with the data.

Including regime switching provides additional large improvements over the models with no regime switching. The SJ model with regime switching in volatility of volatility (SJ-RV) offers an improvement of 110 points in log likelihood relative to the corresponding model without regime switching (SJ) at the cost of four additional parameters. Including regime switching in leverage effect (SJ-RL) provides less of a gain, but still gives an improvement of 32 points in log likelihood relative to the model without regime switching (SJ) at the cost of four additional parameters. Including regime-switching in both volatility of volatility and leverage effect (SJ-RVL) does slightly better than either alone, with an improvement of 118 points in log likelihood relative to the SJ model and five additional parameters.

To the extent that the models are nested, comparisons can be done by standard likelihood ratio tests. For example, the following models nest: SJ-RVL $>$ SJ-RV $>$ SJ $>$ SV. In pairwise comparisons among these models, the larger model always rejects the smaller model, with p-values ranging between .0001 and 10^{-153}. The overall preferred model, SJ-RVL, rejects its non-regime-

switching counterpart, SJ, with a p-value of about 10^{-48}. This would typically be regarded as convincing evidence. Note that while SJ does not include regime switching, it does include jumps and is itself vastly better than the models with unscaled jumps (or no jumps at all) commonly used in the existing literature.

A potentially more useful way to compare models is by using some form of information criterion. Akaike and Schwarz information criteria are common choices. These are based on comparison of log likelihood minus some penalty based on the number of free parameters in the model. The Akaike information criterion uses a penalty equal to the number of free parameters, while the Schwarz criterion (also known as the Bayesian information criterion) uses a penalty equal to the number of free parameters times $\ln(n)/2$, where n is the number of observations. For either of these, the results are the same: SJ models are always preferred over their UJ counterparts. Among SJ models, the ranking is SJ-RVL $>$ SJ-RV $>$ SJ-RL $>$ SJ. The rankings are all decisive.

Our joint analysis of the physical and risk-neutral models allows us to estimate the parametric variance risk premium. Since the variance risk premium may be dependent on regime states, we allow the parameter that determines this premium, η, to be regime-dependent. For the SJ model with regime switching in volatility of volatility (SJ-RV), η is larger in the high volatility of volatility state than in the low one, i.e., $\eta^1 > \eta^0$. Since log volatility, v_t, is always negative, this implies that $\eta^1 v_t < \eta^0 v_t$. That is, higher volatility of volatility corresonds to a more negative variance risk premium. This is consistent with theory (Bollerslev, Gibson, and Zhou 2011). Similarly, in the models with regime switching in leverage effect (both SJ-RL and UJ-RL), the stronger leverage effect state (more negative correlation between returns and volatility innovations) corresponds to a more negative variance risk premium. This is also consistent with theory.

Time series plots of filtered values for the regime-dependent parameters in several models are shown in Figure 2.3. For the models with regime switching in volatility of volatility, for example, σ_t takes on two possible values (σ^0 or σ^1). The filter described in Section 2.3.2 provides probabilities for each state conditional on available information, $p_t^j = p(s_t = j|\mathcal{F}_t)$. Filtered estimates of σ_t are thus given by $\hat{\sigma}_t = E_t[\sigma(s_t)] = p_t^0\sigma^0 + p_t^1\sigma^1$ $(t = 1,\ldots,n)$. An analogous procedure

is used to compute $\hat{\rho}_t$ $(t = 1, \ldots, n)$ for the models with regime-switching leverage effect.

The jump models can be thought of as special cases of the regime-switching model where the regime state (jump or no jump) does not depend on past information (i.e., the state variable is *i.i.d.*). In our regime-switching models, in contrast, the state is quite persistent. Indeed, it is precisely this persistence that accounts for the large improvements in log likelihood. With SJ-RV, for example, the estimated persistence parameters are $\pi_0 = .98$ and $\pi_1 = .94$ (the probability of staying in regime 0 or regime 1, respectively, from one day to the next). The expected duration of stays is 50 days for regime 0 and 17 days for regime 1. By not accounting for persistence in the state, the jump models leave a great deal of information on the table.

Expectations of future volatility of volatility and leverage effect are dramatically different depending on the current regime. The estimated volatility of volatility parameters are 0.084 for state 0 versus 0.133 for state 1 in the SJ-RV model. The estimated leverage parameters are -0.53 for state 0 versus -0.82 for state 1 in the SJ-RL model. Because of the high degree of persistence in regime states, these differences remain even over relatively long time horizons. This is not of purely theoretical interest. Any investors interested in the dynamics of volatility will find this information useful. For example, volatility options and swaps are highly dependent on the volatility of volatility. As discussed below, these persistent differences also affect the shape of the volatility smirk.

2.4.3 Diagnostics

QQ-plots for return and volatility residuals are shown for several models in Figures 2.4 and 2.5. Correlograms are shown in Figure 2.6. Jarque-Bera and Ljung-Box statistics are shown in Table 2.3.

Beginning with the qq-plots and Jarque-Bera statistics, including jumps in the model provides an enormous improvement over the model with no jumps, consistent with the previous literature. This is true regardless of the form of the jumps (SJ or UJ). However, the scaled jump models (SJ) do much better than those with unscaled jumps (UJ), suggesting that the SJ models

do a much better job of capturing non-normality in return and volatility innovations. For the SJ models, including regime switching provides additional small improvements (for the UJ models, only regime switching in leverage effect helps). The outperformance of the SJ models over the UJ models is maintained even when the regime switching feature is added.

In any event, all of the models are rejected by the Jarque-Bera test at conventional significance levels. Standard jump models, even with the addition of the regime switching feature, are simply not flexible enough to capture accurately the true shape of the distribution of innovations. Perhaps models with multiple jump processes might do better. Possible alternatives might be to try using more flexible regime-switching models or modeling the bivariate distribution of return-volatility innovations using mixtures of normals.

Turning now to the correlograms, all of the models do relatively well at eliminating autocorrelation in return residuals. On the other hand, all of the models fail with respect to the volatility residuals, which show significant negative autocorrelation for the first several lags. This is suggestive of the existence of a second volatility factor omitted by the model (e.g., Gallant, Hsu, and Tauchen 1999; Chacko and Viceira 2003; Christoffersen, Heston, and Jacobs 2009).

All of the models except those with regime switching in volatility of volatility also show a large amount of unexplained autocorrelation in squared volatility residuals. This is suggestive of stochastic volatility of volatility. Although the regime-switching model implements a very simple form of stochastic volatility of volatility, it is a substantial improvement over the other models.

The fact that none of the models are able to pass the full set of diagnostics may seem disappointing, but should not really be viewed as such. The models proposed in this paper still perform dramatically better than other commonly-used models. The data set is very rich, with 4025 observations on both price and volatility. The idea that one might fully capture the behavior of this data in a model with just a dozen or so parameters is highly optimistic. Furthermore, the diagnostics provide useful information on failure modes which can help to direct future work. Strong diagnostics are a good thing, not a bad thing.

2.5 Explanatory power for option-implied skewness and kurtosis

Sections 2.4.2 and 2.4.3 show that the regime switching models provide dramatic improvements in fitting the joint dynamics of returns and the volatility state. Both time-varying leverage effect and stochastic volatility of volatility are important. Furthermore, the regimes corresponding to high and low states of both variables are highly persistent, implying that there is considerable predictability regarding future regime states. In addition, each of these variables has a direct effect on the shape of return distributions. As noted by Das and Sundaram (1999), standard jump models can add skewness and kurtosis to the distribution of returns, but primarily at short horizons. The effect dissipates relatively quickly. Increasing either the correlation between return and volatility innovations (leverage effect) or volatility of volatility also adds skewness and kurtosis to return distributions. The effects of these variables are weaker at short horizons but more persistent.

Option prices provide an alternative source of information about the distribution of returns at various horizons, but under the risk-neutral rather than physical measure. Features of these distributions that are of potential interest include variance, skewness, and kurtosis, as well as the term structure of variance. These characteristics all exhibit considerable variation across time. A great deal of attention has been devoted to models that describe time-variation in the variance in returns, but less attention has been paid to the other features.

The issue of interest here is whether time-varying characteristics of the physical dynamics of returns have explanatory power for time-variation in the shape of the risk-neutral distribution (or equivalently, in the shape of the volatility smirk). An alternative explanation is that time-variation in the shape of the risk-neutral distribution is due largely or entirely to changes in risk premia.

Based on observed data, option-implied skewness and kurtosis are related to the level of the volatility state, but the relationship is weak. This suggests that models with volatility as the only state variable will be able to explain only a small part of the time-variation exhibited by the shape of the risk-neutral distribution (or equivalently, the implied volatility smirk). Including jumps and jump risk premia in various configurations can help match the *average* shape of the implied

volatility smirk, but will never be able to explain *time variation* that is independent of the volatility state unless an additional state variable is introduced.

In this section, we examine whether the regime states implied by our regime-switching models have any explanatory power for the shape of the risk neutral distribution. Recall that option-implied skewness and kurtosis are not used in the estimation, which makes this diagnostic meaningful.

Given a model for the price of an asset and values for any relevant state variables, one can easily obtain the model-implied distribution of returns at any horizon using Monte Carlo methods. We will refer to these as model-implied skewness and kurtosis. Given time series for option-implied volatility and skewness (or kurtosis) and a model with two state variables (such as our regime-switching models), it is not difficult to find model parameters and values for the state variables such that model-implied and option-implied characteristics match exactly. To the extent that implied physical and risk-neutral model parameters differ, these differences would typically be attributed to risk premia. However, since this approach confounds the information from physical dynamics and option prices, it is difficult to tell if the implied variation in states (e.g., leverage effect or volatility of volatility) is truly evident in the physical model or just an artifact of trying to match the shape of the volatility smile. The problem here is a pervasive one in the option-pricing literature. It is possible to fit even a badly misspecified model to the data. The fitted model is likely to generate spurious risk premia, but the exact source and extent of the misspecification are difficult to diagnose.

The approach that we take in this paper is different. While we match option-implied volatility exactly, we do not use the information on option-implied skewness and kurtosis in the estimation. In particular, we make no effort to use the information in option prices to come up with model parameters and regime states such that option-implied and model-implied skewness or kurtosis match exactly. We also do not look for risk premia which cause the models to fit the shape of the implied volatility smirk on average. Rather, we fit the models using only information from returns and implied volatility. We are then able to make a clear and unambiguous examination of the extent

to which the implied states are informative about variation across time in observed option-implied skewness and kurtosis.

One possible hypothesis is that time variation in the shape of the risk-neutral distribution is due exclusively to time-varying risk premia. We are able to reject this hypothesis decisively.

The idea is to regress option-implied skewness and kurtosis on the regime state, controlling for other possible determinants. As described in Section 2.3, it is not possible based on observed data to determine the regime state conclusively. The available information consists of filtered probabilities, $p_t^j = p(s_t = j|\mathcal{F}_t)$ $(j = 0, 1)$. The operational regression equations are thus

$$\begin{aligned} \text{SKEW}_t &= \beta_0 + \beta_{\text{RV}}\hat{s}_t^{\text{RV}} + \beta_{\text{RL}}\hat{s}_t^{\text{RL}} + \beta_{\text{VIX}}\log\text{VIX}_t + \beta_{\text{JV}}\log\text{JV}_t + \beta_{\text{VRP}}\text{VRP}_t + \varepsilon_t \\ \text{KURT}_t &= \beta_0 + \beta_{\text{RV}}\hat{s}_t^{\text{RV}} + \beta_{\text{RL}}\hat{s}_t^{\text{RL}} + \beta_{\text{VIX}}\log\text{VIX}_t + \beta_{\text{JV}}\log\text{JV}_t + \beta_{\text{VRP}}\text{VRP}_t + \varepsilon_t \end{aligned} \tag{2.8}$$

where $\hat{s}_t^{\text{RV}}$ denotes the filtered state under the models with regime switching in volatility of volatility and $\hat{s}_t^{\text{RL}}$ denotes the filtered state under the models with regime switching in leverage effect (as described in Section 2.3.2). We also include several control variables. VIX_t is the VIX index, which serves as a proxy for the volatility state. VRP_t is the variance risk premium, which Bollerslev, Gibson, and Zhou (2011) argue reflects information on variation across time in investors' risk aversion. Option-implied skewness and kurtosis may also be affected by the market expectation of jump risk. We proxy for this by using the past one-month jump variation, JV_t. A more detailed description of these variables may be found in Section 2.4.1. We first run the regressions on the control variables alone, and then add the regime states to the regressions to see if there is a significant increase in explanatory power. We report Newey-West robust t-statistics over eight lags (Newey and West 1987) and adjusted R^2. For completeness, results are reported for the regressions based on filtered regime states from both SJ and UJ models.

Table 2.4 shows the regressions for option-implied skewness. We first discuss the contributions of the control variables toward explaining option-implied skewness. The volatility state (VIX) is highly significant for all models and regardless of whether the regime state is included in the regression. The coefficient is always positive, meaning that low volatility states are associated

with more left-skewed return distributions. This is consistent with Dennis and Mayhew (2002), who report a similar finding for individual stocks. Jump risk (JV) is also highly significant for all models. The coefficient is always negative, suggesting that greater jump risk is associated with more left-skewed return distributions under the risk-neutral measure. This is consistent with intuition. In contrast, we do not find the variance risk premium (VRP) to be significant when the other control variables are included.

Although the volatility state is highly significant, it has relatively low explanatory power for option-implied skewness, with an adjusted R^2 of 7.6%. All of the control variables together are able to explain only 9.2% of the time-variation in option-implied skewness. This low explanatory power suggests that models with only a single state variable (volatility) will not be able to match time-varying patterns in the Black-Scholes volatility smirk realistically.

Including the filtered regime state in the regressions provides a dramatic improvement in explanatory power. This is true regardless of the particular model used. SJ models always outperform the corresponding UJ models. Including the regime state from the models with regime-switching in volatility of volatility (SJ-RV and UJ-RV) is slightly better than including the regime state from the models with regime-switching in leverage effect (SJ-RL and UJ-RL). But the clear winner is the model which includes regime states from both regime switching models (RV and RL). The states corresponding to volatility of volatility and regime switching each have substantial explanatory power when included in the regression alone. But, these two variables serve as complementary sources of information. For the regressions using states generated from the SJ models (together with all control variables), the R^2 is 18-19% when either the RV or RL states are included individually, but jumps to over 32% when both are included. The rankings here are consistent with those implied by the model comparisons described in Sections 2.4.2 and 2.4.3. But, it is important to recognize that this result is not necessarily expected *a priori* since the dependent variable here, implied skewness, was not involved in fitting those models. The entire point of performing the study in this manner is to enable regressions such as those under examination in this section as meaningful diagnostics.

For the SJ-RV model (regime-switching in volatility of volatility), the coefficient on the regime state is highly significant and in the expected direction. For the full regression (including regime state and all control variables), the estimated slope coefficient for the regime state is -0.61, indicating that a change from state 0 to state 1 is associated with a 0.61 decrease in skewness (i.e., the distibution is substantially more left skewed in the high volatility of volatility state). The t-statistic is -9.95, corresponding to a p-value of around 10^{-22}. This model has good explanatory power, with an adjusted R^2 of 19.0% (versus 9.2% for the control variables alone). These results are both statistically and economically significant.

The SJ-RL model (regime switching in leverage effect) provides comparable performance. The coefficient on the regime state is highly significant and in the expected direction. For the full regression (including regime state and all control variables), the estimated slope coefficient is -0.60, indicating that a change from state 0 to state 1 is associated with a 0.60 decrease in skewness (i.e., the distibution becomes substantially more left skewed in the state with the stronger leverage effect). The t-statistic is -7.17, corresponding to a p-value of around 10^{-15}. This model has an adjusted R^2 of 18.3%.

We also ran the regression using regime states from the models with regime switching in both volatility of volatility and leverage effect (SJ-RVL and UJ-RVL), but do not report the results in the table out of space considerations. For the SJ-RVL regression (including all control variables), the slope coefficient for the regime state is -0.65 with a t-statistic of -11.50, indicating a slightly stronger effect than in the other models. The adjusted R^2 is 23.0%, which indicates slightly better explanatory power than the more constrained models.

But the regressions including regime states from both RV and RL models do substantially better than any of those which include only a single regime state. It is important to understand the difference between these regression and the ones involving regime states from the RVL models. In the RVL models, both volatility of volatility and leverage effect are state dependent, but there is only a single state variable controlling both. In the regression including regime states from both RV and RL models, there are two different regime state variables involved. *A priori*, it is unknown

whether the two states contain largely the same or significantly different information. It turns out that the states provide different information. The slope coefficients change little when both states are included in the model (relative to including them individually), they are even more significant (based on t-statistics), and the increase in explanatory power is substantial.

Table 2.5 shows analagous regressions for option-implied kurtosis. The results are qualitatively similar to those for option-implied skewness. The volatility state (VIX) is highly significant for all models and regardless of whether the regime state is included in the regression. The coefficient is negative, implying that low volatility states are associated with fatter-tailed return distributions. Jump risk (JV) is positively related to kurtosis. That is, the risk-neutral distribution tends to be more fat-tailed when jump risk is high, consistent with intuition. Although the variance risk premium (VRP) is highly significant when jump risk is omitted from the regression, it has little explanatory power when jump risk is included.

Including the regime state in the regression provides additional improvement in explanatory power. As with option-implied skewness, SJ models always outperform the corresponding UJ models. Slope coefficients for the regime state have the expected signs. That is, higher volatility of volatility and stronger leverage effect are both associated with more kurtotic return distributions.

Including the regime state corresponding to regime switching in leverage effect does better here than including the regime state corresponding to regime switching in volatility of volatility (adjusted R^2 of 22.9% for SJ-RL versus 18.1% for SJ-RV). As is the case with implied skewness, including both regime states in the regression is better yet, with an R^2 of 25.8%. In the full model, slope coefficients are 3.11 and 5.19 (with t-statistics of 5.52 and 8.66) for the SJ-RV and SJ-RL states respectively. These are both economically and statistically significant.

Pan (2002) and Broadie, Chernov, and Johannes (2007) argue that the existence of a jump risk premium could have a significant effect on the shape of the volatility smirk. However, their models include only a single state variable and are not able to fit *time variation* in the shape of the volatility smirk that is independent of the volatility state. The goal of this paper is very different from those. In particular, our objective is to identify model characteristics that could be responsible

for time variation in the shape of the volatility smirk. Also, whereas those papers rely substantially on risk premia in jump sizes to fit the smirk, our focus is on whether there is information present in the physical measure that is related to changes in the shape of the smirk. The issue of whether information exists in the physical measure that is related to features of observed option prices is important. Any differences between the dynamics implied by the physical measure and those implied by option prices will be ascribed to risk premia. But if models are misspecified, then the risk premia “discovered" in this way are likely to be just artifacts.

On the other hand, we make no effort to investigate factors underlying the average shape of the smirk, and thus have nothing to say about the issue of whether the smirk is, *on average*, due to jumps or some other factor. Among other things, the relative importance of various potential factors toward explaining the shape of the smirk will depend on the time-horizon under consideration. As pointed out by Das and Sundaram (1999), jumps can have a large effect on the skewness of return distributions over short horizons, but the effects dissipate relatively quickly. The leverage effect and volatility of volatility have less of an effect at short horizons, but the effects are more persistent. Sorting out the relative importance of these factors would require looking at the term structure of the smirk (which this paper makes no attempt to do).

We did, nonetheless, investigate models with regime switching in jump intensity, but did not find such models to be useful for the purposes of this paper. These models were heavily dominated by the other regime switching models in terms of log likelihood and the model diagnostics discussed in Sections 2.4.2 and 2.4.3. They also did not prove informative in terms of explanatory power for option-implied skewness or kurtosis, as discussed in this section. On the other hand, our regression results do include the jump variation variable (JV) as a control. We find this variable to have useful explanatory power.

In a nutshell, we find that the regime states corresponding to both volatility of volatility and leverage effect have strong explanatory power for option-implied return distributions. Models used in simultaneous estimation of physical and risk-neutral measures typically use only a single state variable, volatility. While there is some linkage between the volatility state and the shape of the

implied-volatility smirk, the relationship is relatively weak. If such models are taken seriously, one must conclude that most of the time-variation in the shape of the implied-volatility smirk is due to unmodelled changes in risk premia. The results of this study are important because they demonstrate a model in which observable phenomena in the physical measure have strong explanatory power for time-variation in the shape of the implied-volatility smirk.

2.6 Conclusion

This paper proposes a new class of models that layer regime switching on top of a standard stochastic volatility model with jumps in both returns and volatility. Motivated by the time-varying nature of option-implied skewness and kurtosis that is a prominent feature of observed data, we allow for regime switching in two parameters of the basic model: volatility of volatility and leverage effect. Both parameters play important roles in determining the skewness and kurtosis of returns. The regime-switching framework is useful because it is the simplest possible approach that incorporates time-variation in the parameters of interest. We make no effort to attach a particular meaning to the regime states, but regard this simply as a convenient framework for generating a flexible family of models capable of capturing features of the data that are of interest.

The application looks at SPX index and option data. In the first step of the analysis, we estimate the models relying only upon observations of the index price and option-implied volatility. This allows us to use observations on option-implied skewness and kurtosis for diagnostic purposes.

The models with regime switching fit the data dramatically better than those without regime switching. Accounting for time-variation in the volatility of volatility or leverage effect not only increases log likelihood significantly, but also provides improvements in other diagnostics of model fit. The best model in this dimension is the one with regime-switching in both the volatility of volatility and leverage effect.

While this first step of the analysis provides strong evidence of time-variation in model characteristics that are related to the shape of return distributions under the physical measure,

the second step investigates the relationship between these characteristics and the skewness and kurtosis of returns under the risk-neutral measure (i.e, implied by observed option prices). This part of the study is carried out by running regressions of option-implied skewness and kurtosis on the regime states extracted in the estimation step (and several control variables).

We find that the models have strong explanatory power. In other words, there is a strong relationship between model characteristics related to the shape of return distributions under the physical measure and the skewness and kurtosis of return distributions implied by option prices. Although commonly-used models in the existing literature include volatility as the only state variable, this variable has an adjusted R^2 of only 7.6% for option-implied skewness. Other control variables that we look at include nonparametric measures of jump risk and volatility risk premium. These are both obtained from high-frequency data (5 minute intraday returns). All of the control variables together explain only 9.2% of the variation in implied skewness. In contrast, regressions that include either the regime state corresponding to volatility of volatility or leverage effect (together with the control variables) have an adjusted R^2 of 18-19%. When both states are included in the regression, the adjusted R^2 is over 32%. Slope coefficients on the regime states are highly significant and in the expected direction. For example, t-statistics are greater than 11 (in absolute value) for the regression of implied skewness on both regime states and all of the control variables (SJ form of jumps). Option-implied return distributions tend to be more left-skewed and leptokurtic when volatility of volatility is high or leverage effect is strong (i.e., correlation between returns and volatility innovations is more negative). Results for option-implied kurtosis are qualitatively similar though weaker.

We also find that the specification of jump form is important both in fitting the time-series data of returns and volatility as well as in explaining variation in option-implied skewness and kurtosis. Volatility-scaled jumps (SJ) outperform unscaled jumps (UJ) in both respects.

Table 2.1: Summary statistics.

The sample period covers January 1993 to December 2008. $\Delta\log(\text{SPX}_t)$ refers to S&P 500 index log returns. VIX_t is the VIX index, divided by $\sqrt{12}$ to get a monthly volatility measure for comparison. IV_t, SKEW_t, and KURT_t denote the one-month option-implied volatility, skewness, and kurtosis, computed using the model-free approach of Bakshi, Kapadia, and Madan (2003). RV_t and JV_t are the realized volatility and the jump variation, calculated using five-minute high-frequency data over the past 22 trading days. $\text{VRP}_t \equiv \log(\text{RV}_t/\text{VIX}_t^2)$ denotes the variance risk premium. AR(i) means the i-lagged autocorrelation.

	Mean	Min	Max	STD	Skewness	Kurtosis	AR(1)	AR(5)	AR(22)
$\Delta\log(\text{SPX}_t)$	0.00	-0.09	0.11	0.01	-0.24	13.11	-0.06	-0.04	0.05
VIX_t $(\%)$	5.73	2.69	23.34	2.40	2.23	11.91	0.98	0.94	0.78
IV_t $(\%)$	5.62	2.22	24.46	2.55	2.26	12.31	0.98	0.94	0.79
SKEW_t	-1.69	-3.75	-0.41	0.49	-0.54	3.25	0.90	0.80	0.63
KURT_t	9.55	3.34	34.94	4.56	1.62	6.22	0.83	0.73	0.57
RV_t $(\%^2)$	21.79	1.58	457.27	40.45	6.72	57.28	1.00	0.97	0.68
JV_t $(\%^2)$	3.33	0.12	83.09	6.89	7.19	67.06	0.99	0.95	0.65
VRP_t	-0.88	-2.16	0.82	0.42	0.16	3.07	0.96	0.83	0.50

Table 2.2: Parameter Estimates.

The sample period covers January 1993 through December 2008 ($N = 4025$). Standard errors are in parentheses. Time is measured in trading days.

	No regime switching			RV model		RL model		RVL model	
	SV	SJ	UJ	SJ-RV	UJ-RV	SJ-RL	UJ-RL	SJ-RVL	UJ-RVL
$\mu \times 10^4$	4.69	3.70	3.08	3.03	2.07	3.42	3.08	3.00	2.08
	(1.27)	(1.28)	(1.25)	(1.25)	(1.25)	(1.26)	(1.23)	(1.25)	(1.25)
$\kappa \times 10^3$	4.29	6.00	6.69	6.34	11.07	7.48	7.36	6.27	11.04
	(1.29)	(1.35)	(1.33)	(1.32)	(1.28)	(1.35)	(1.34)	(1.32)	(1.28)
$\bar{v}$	-9.71	-10.03	-9.70	-10.06	-9.63	-10.00	-9.80	-9.98	-9.63
	(0.55)	(0.38)	(0.36)	(0.32)	(0.20)	(0.32)	(0.33)	(0.33)	(0.20)
$\eta^0 \times 10^3$	3.01	2.84	3.84	2.86	4.18	2.49	2.72	3.01	4.20
	(0.25)	(0.36)	(0.35)	(0.42)	(0.36)	(0.47)	(0.44)	(0.37)	(0.36)
$\eta^1 \times 10^3$				2.97	2.46	3.39	4.27	2.66	2.46
				(0.59)	(0.41)	(0.40)	(0.38)	(0.54)	(0.42)
$\sigma^0 \times 10^1$	1.32	0.91	1.23	0.84	1.18	0.99	1.25	0.85	1.18
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$\sigma^1 \times 10^1$				1.33	2.12			1.33	2.12
				(0.06)	(0.06)			(0.05)	(0.06)
ρ^0	-0.74	-0.78	-0.82	-0.81	-0.91	-0.53	-0.61	-0.77	-0.91
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.03)	(0.01)	(0.01)
ρ^1						-0.82	-0.85	-0.87	-0.91
						(0.01)	(0.01)	(0.02)	(0.01)
ρ_J		-0.71	-0.53	-0.70	-0.05	-0.77	-0.72	-0.69	-0.04
		(0.03)	(0.04)	(0.03)	(0.10)	(0.03)	(0.04)	(0.03)	(0.10)
λ		0.47	0.33	0.55	1.08	0.30	0.15	0.46	1.08
		(0.06)	(0.04)	(0.09)	(0.14)	(0.04)	(0.02)	(0.07)	(0.15)
$\mu_{1J} \times 10^2$		-1.13	0.05	-4.52	0.08	-11.90	-0.05	-3.78	0.08
		(6.59)	(0.04)	(6.38)	(0.02)	(8.78)	(0.07)	(6.98)	(0.02)
$\sigma_{1J} \times 10^1$		12.94	0.06	12.44	0.03	13.42	0.08	12.30	0.03
		(0.75)	(0.00)	(0.77)	(0.00)	(0.85)	(0.01)	(0.77)	(0.00)
$\mu_{2J} \times 10^1$		2.79	0.24	3.39	0.09	3.59	0.44	3.04	0.08
		(0.69)	(0.07)	(0.66)	(0.04)	(0.95)	(0.13)	(0.70)	(0.04)
σ_{2J}		1.54	0.15	1.15	0.05	1.72	0.21	1.25	0.05
		(0.07)	(0.01)	(0.07)	(0.01)	(0.09)	(0.01)	(0.07)	(0.01)
π_0				0.98	0.98	0.97	0.96	0.99	0.98
				(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
π_1				0.94	0.94	0.99	0.98	0.96	0.94
				(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
log(L)	38,491	38,851	38,810	38,961	38,960	38,883	38,858	38,969	38,960

Table 2.3: Diagnostic tests for generalized residuals.

Test statistics are shown with p-values in parentheses.

Jarque-Bera Test

	No regime switching			RV model		RL model		RVL model	
	SV	SJ	UJ	SJ-RV	UJ-RV	SJ-RL	UJ-RL	SJ-RVL	UJ-RVL
Return	445	20	113	17	245	10	50	19	246
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.006)	(0.000)	(0.000)	(0.000)
Volatility	2,367	22	121	20	318	8	30	15	320
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.018)	(0.000)	(0.001)	(0.000)

Ljung-Box Test (with 20 lags)

	No regime switching			RV model		RL model		RVL model	
	SV	SJ	UJ	SJ-RV	UJ-RV	SJ-RL	UJ-RL	SJ-RVL	UJ-RVL
Return	43	41	41	41	40	41	40	41	40
	(0.002)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Volatility	116	118	116	115	100	115	114	116	100
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Squared Vol.	419	552	424	128	74	559	446	118	74
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 2.4: Regressions for option-implied skewness.

The table reports the results of the following regression,

$$\text{SKEW}_t = \beta_0 + \beta_{\text{RV}}\hat{s}_t^{\text{RV}} + \beta_{\text{RL}}\hat{s}_t^{\text{RL}} + \beta_{\text{VIX}}\log\text{VIX}_t + + \beta_{\text{JV}}\log\text{JV}_t + \beta_{\text{VRP}}\text{VRP}_t + \varepsilon_t.$$

Newey-West robust t-statistics over eight lags are shown in parentheses. The sample period covers January 1993 to December 2008. SKEW_t denotes the one-month option-implied skewness. Filtered regime states are $\hat{s}_t^{\text{RV}}$ and $\hat{s}_t^{\text{RL}}$ for volatility of volatility and leverage effect respectively. The control variables are the VIX index, jump variation (JV), and variance risk premium (VRP). Results are shown for filtered states from both SJ (scaled jumps) and UJ (unscaled jumps) models.

Constant		$\hat{s}_t^{\text{RV}}$		$\hat{s}_t^{\text{RL}}$		$\log\text{VIX}_t$		$\log\text{JV}_t$		VRP_t		Adj. R^2
coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	
Control variables only												
-2.78	(-16.43)					0.37	(6.53)					7.6%
-3.38	(-12.02)					0.60	(5.94)	-0.10	(-3.21)			9.2%
-2.75	(-13.84)					0.36	(4.97)			0.000	(-0.32)	7.6%
-3.39	(-11.76)					0.60	(5.37)	-0.10	(-2.19)	0.008	(0.11)	9.2%
Control variables plus filtered states from SJ models												
-2.82	(-18.07)	-0.56	(-9.02)			0.44	(8.48)					16.2%
-3.63	(-14.57)	-0.61	(-9.98)			0.74	(8.42)	-0.13	(-5.01)			19.0%
-3.17	(-16.19)	-0.59	(-9.62)			0.51	(8.87)			-0.167	(-3.73)	17.9%
-3.62	(-14.45)	-0.61	(-9.95)			0.74	(7.72)	-0.12	(-3.18)	-0.007	(-0.11)	19.0%
-2.82	(-15.39)			-0.62	(-7.37)	0.54	(7.15)					17.5%
-3.25	(-12.37)			-0.60	(-7.17)	0.69	(7.10)	-0.07	(-2.23)			18.3%
-3.01	(-15.20)			-0.61	(-7.31)	0.57	(7.87)			-0.092	(-1.91)	18.0%
-3.25	(-11.86)			-0.60	(-7.17)	0.69	(6.18)	-0.07	(-1.45)	-0.007	(-0.10)	18.3%
-2.87	(-18.05)	-0.71	(-11.53)	-0.76	(-10.66)	0.65	(10.36)					30.6%
-3.52	(-16.28)	-0.74	(-12.04)	-0.74	(-10.77)	0.89	(11.23)	-0.10	(-4.02)			32.3%
-3.17	(-18.84)	-0.73	(-11.88)	-0.75	(-10.90)	0.71	(12.00)			-0.143	(-3.35)	31.8%
-3.50	(-15.73)	-0.74	(-12.03)	-0.74	(-10.77)	0.87	(9.82)	-0.09	(-2.40)	-0.028	(-0.45)	32.3%
Control variables plus filtered states from UJ models												
-2.62	(-16.06)	-0.44	(-7.48)			0.36	(6.81)					14.2%
-3.41	(-13.30)	-0.48	(-8.59)			0.66	(7.33)	-0.13	(-4.97)			17.0%
-2.96	(-14.77)	-0.47	(-8.14)			0.43	(7.36)			-0.168	(-3.68)	15.9%
-3.41	(-13.12)	-0.48	(-8.56)			0.66	(6.65)	-0.13	(-3.14)	-0.007	(-0.10)	17.0%
-2.93	(-15.91)			-0.50	(-6.57)	0.53	(7.43)					14.0%
-3.45	(-12.68)			-0.48	(-6.56)	0.73	(7.19)	-0.08	(-2.77)			15.2%
-3.16	(-15.27)			-0.49	(-6.61)	0.58	(8.05)			-0.110	(-2.32)	14.8%
-3.45	(-12.27)			-0.48	(-6.55)	0.72	(6.39)	-0.08	(-1.82)	-0.005	(-0.07)	15.2%
-2.76	(-16.64)	-0.57	(-10.09)	-0.65	(-9.82)	0.56	(9.35)					24.7%
-3.52	(-15.05)	-0.61	(-11.07)	-0.64	(-10.25)	0.85	(10.04)	-0.12	(-4.79)			27.2%
-3.11	(-17.14)	-0.60	(-10.71)	-0.65	(-10.21)	0.64	(10.76)			-0.169	(-3.88)	26.4%
-3.50	(-14.63)	-0.62	(-11.03)	-0.64	(-10.22)	0.83	(8.92)	-0.11	(-2.91)	-0.028	(-0.44)	27.2%

Table 2.5: Regressions for option-implied kurtosis.

The table reports the results of the following regression,

$$\mathrm{KURT}_t = \beta_0 + \beta_{\mathrm{RV}}\hat{s}_t^{\mathrm{RV}} + \beta_{\mathrm{RL}}\hat{s}_t^{\mathrm{RL}} + \beta_{\mathrm{VIX}}\log \mathrm{VIX}_t + \beta_{\mathrm{JV}}\log \mathrm{JV}_t + \beta_{\mathrm{VRP}}\mathrm{VRP}_t + \varepsilon_t.$$

Newey-West robust t-statistics over eight lags are shown in parentheses. The sample period covers January 1993 to December 2008. KURT_t denotes the one-month option-implied kurtosis. Filtered regime states are $\hat{s}_t^{\mathrm{RV}}$ and $\hat{s}_t^{\mathrm{RL}}$ for volatility of volatility and leverage effect respectively. The control variables are the VIX index, jump variation (JV), and variance risk premium (VRP). Results are shown for filtered states from both SJ (scaled jumps) and UJ (unscaled jumps) models.

Constant		$\hat{s}_t^{\mathrm{RV}}$		$\hat{s}_t^{\mathrm{RL}}$		$\log \mathrm{VIX}_t$		$\log \mathrm{JV}_t$		VRP_t		Adj. R^2
coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat	
Control variables only												
20.22	(12.46)					-3.66	(-6.96)					8.4%
32.99	(13.34)					-8.40	(-9.68)	2.02	(8.02)			16.6%
19.23	(11.34)					-3.21	(-5.64)			0.02	(2.22)	8.7%
32.49	(13.23)					-7.98	(-8.74)	1.72	(4.94)	0.65	(1.08)	16.7%
Control variables plus filtered states from SJ models												
20.31	(12.67)	1.44	(2.44)			-3.81	(-7.45)					9.0%
33.86	(14.27)	2.18	(3.90)			-8.91	(-10.83)	2.14	(8.93)			18.0%
26.82	(14.40)	1.99	(3.49)			-5.18	(-9.70)			3.03	(7.22)	15.5%
33.33	(14.19)	2.19	(3.93)			-8.46	(-9.80)	1.82	(5.38)	0.70	(1.18)	18.1%
20.51	(12.11)			5.14	(7.03)	-5.00	(-7.57)					16.2%
31.98	(13.75)			4.60	(7.13)	-9.14	(-10.79)	1.82	(7.12)			22.8%
26.15	(14.73)			4.84	(7.30)	-6.06	(-10.03)			2.65	(6.33)	21.2%
31.39	(13.39)			4.62	(7.16)	-8.65	(-9.55)	1.47	(4.20)	0.76	(1.33)	22.9%
20.69	(12.63)	2.50	(4.31)	5.63	(8.11)	-5.40	(-8.68)					18.0%
33.09	(15.13)	3.09	(5.47)	5.17	(8.60)	-9.95	(-12.69)	1.96	(8.16)			25.6%
26.81	(15.80)	2.97	(5.22)	5.40	(8.74)	-6.62	(-11.87)			2.86	(7.03)	23.8%
32.44	(14.75)	3.11	(5.52)	5.19	(8.66)	-9.41	(-11.24)	1.57	(4.71)	0.85	(1.51)	25.8%
Control variables plus filtered states from UJ models												
20.00	(12.07)	0.59	(1.07)			-3.64	(-6.94)					8.5%
33.07	(13.62)	1.34	(2.64)			-8.58	(-10.20)	2.12	(8.74)			17.3%
26.18	(13.80)	1.14	(2.19)			-4.91	(-9.05)			2.99	(7.06)	14.8%
32.55	(13.52)	1.36	(2.68)			-8.13	(-9.20)	1.80	(5.28)	0.69	(1.15)	17.4%
21.47	(12.52)			4.21	(6.24)	-5.00	(-7.85)					13.7%
33.56	(13.99)			3.91	(6.57)	-9.43	(-10.79)	1.93	(7.63)			21.1%
27.40	(14.89)			4.08	(6.64)	-6.16	(-10.24)			2.80	(6.73)	19.3%
32.99	(13.74)			3.93	(6.59)	-8.95	(-9.73)	1.58	(4.56)	0.75	(1.32)	21.3%
21.01	(12.26)	1.54	(2.88)	4.63	(7.10)	-5.09	(-8.34)					14.6%
33.79	(14.66)	2.25	(4.48)	4.49	(7.87)	-9.89	(-12.00)	2.07	(8.58)			23.0%
27.21	(15.19)	2.10	(4.10)	4.64	(7.87)	-6.36	(-11.31)			3.00	(7.33)	20.9%
33.16	(14.40)	2.27	(4.54)	4.52	(7.90)	-9.35	(-10.81)	1.69	(5.06)	0.84	(1.48)	23.2%

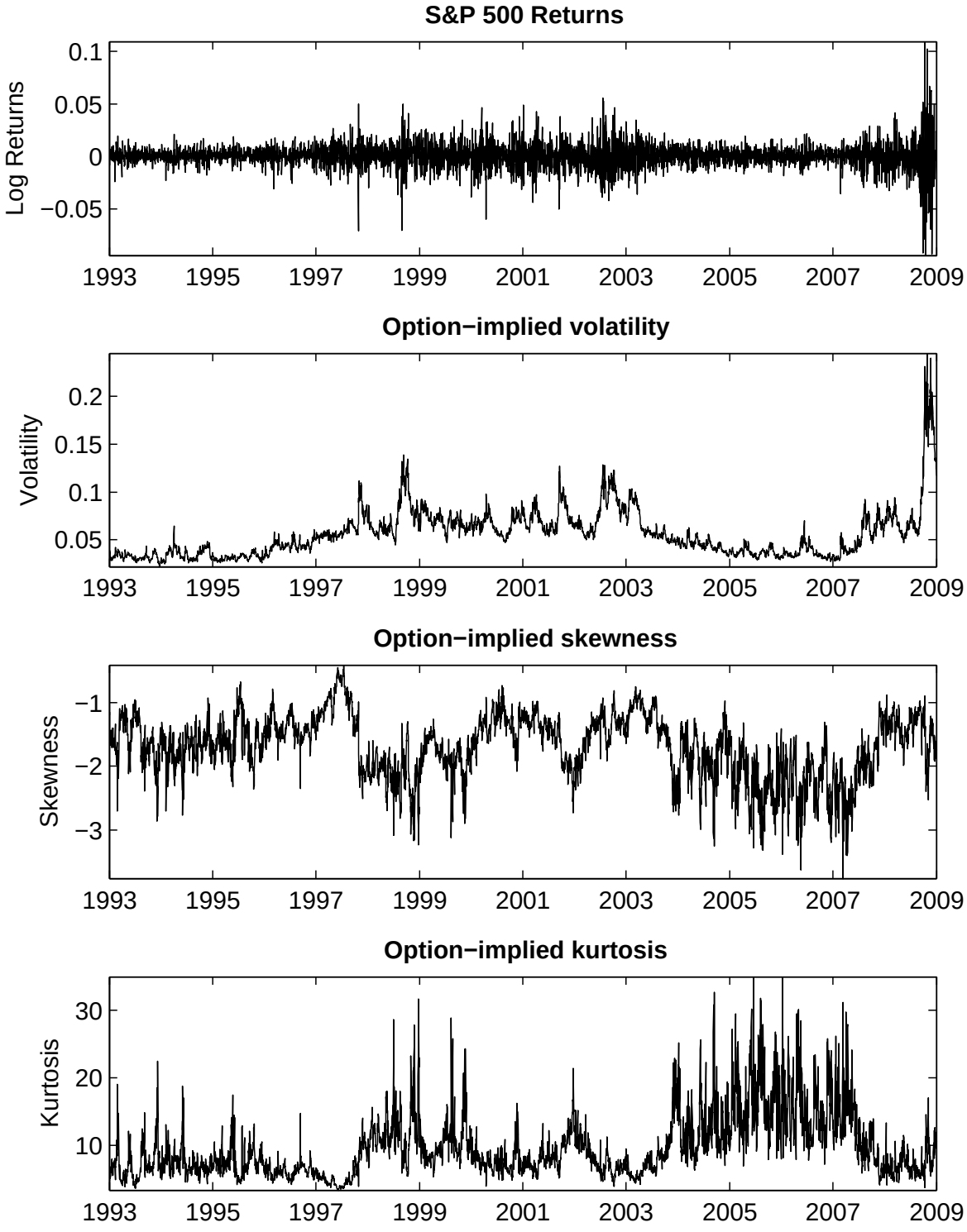

Figure 2.1: Time series of S&P 500 returns, option-implied volatility, option-implied skewness, and option-implied kurtosis.

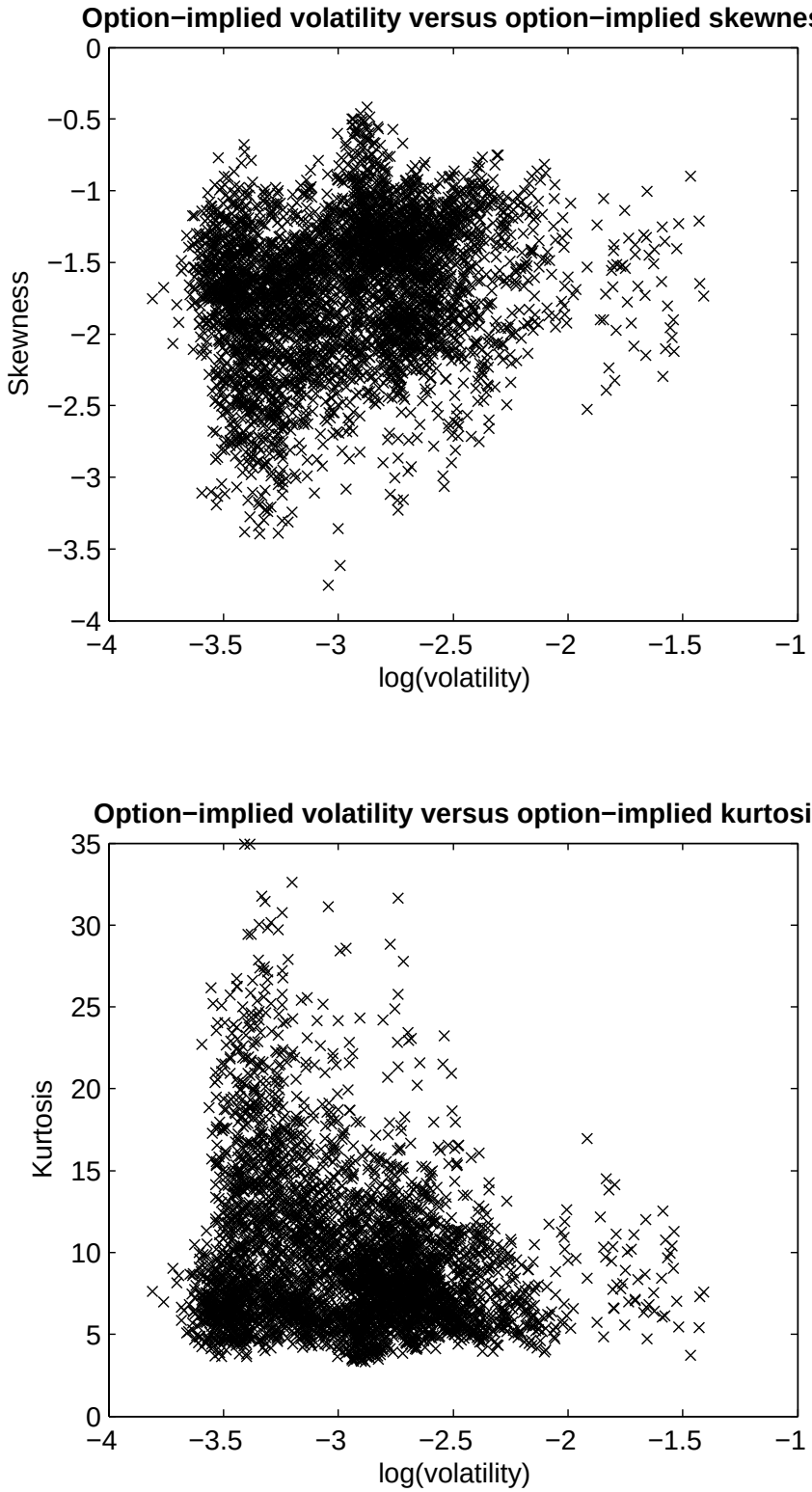

Figure 2.2: Scatter plots of option-implied skewness and kurtosis against option-implied volatility.

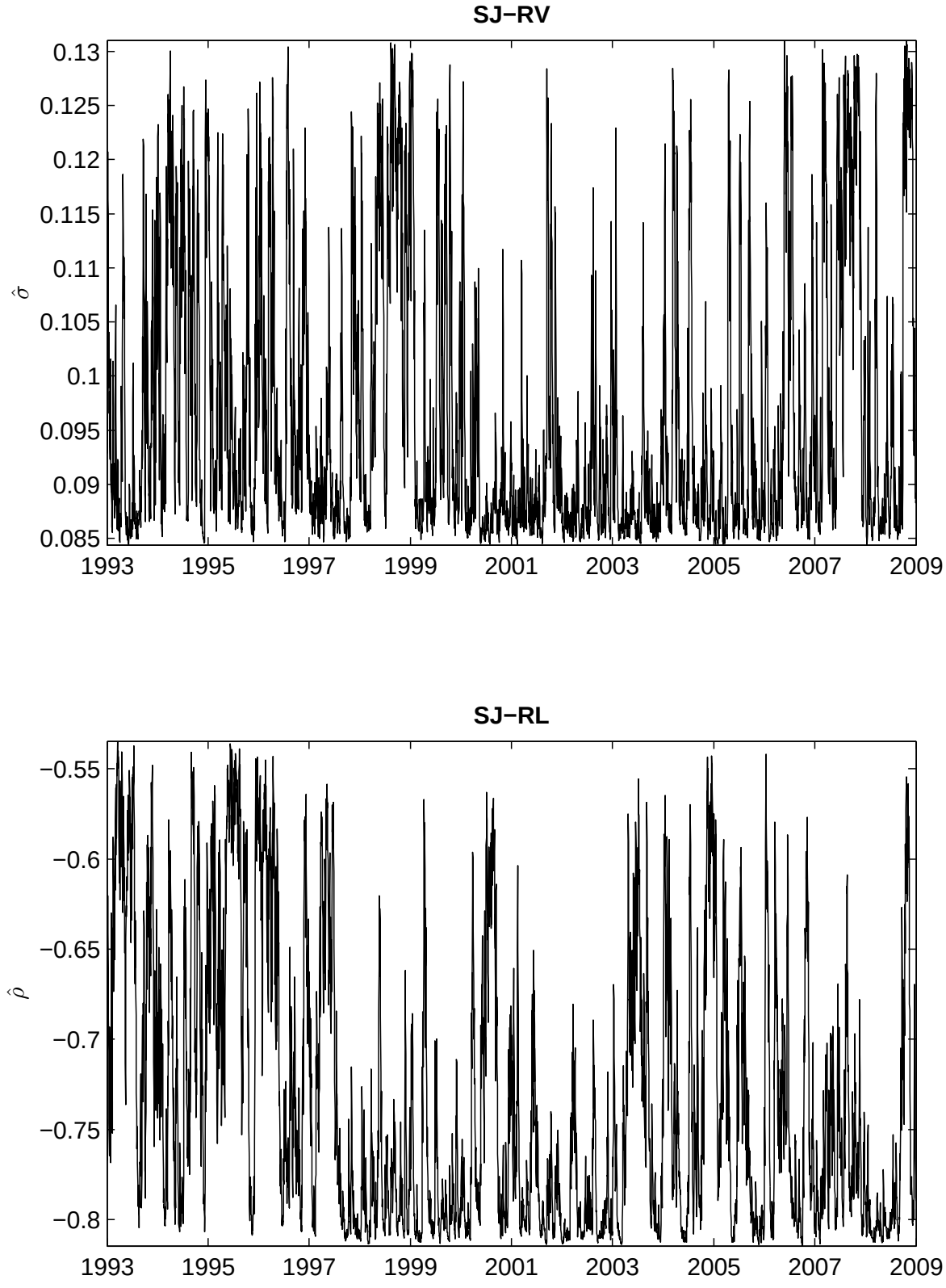

Figure 2.3: Time series of filtered values of state-dependent parameters.

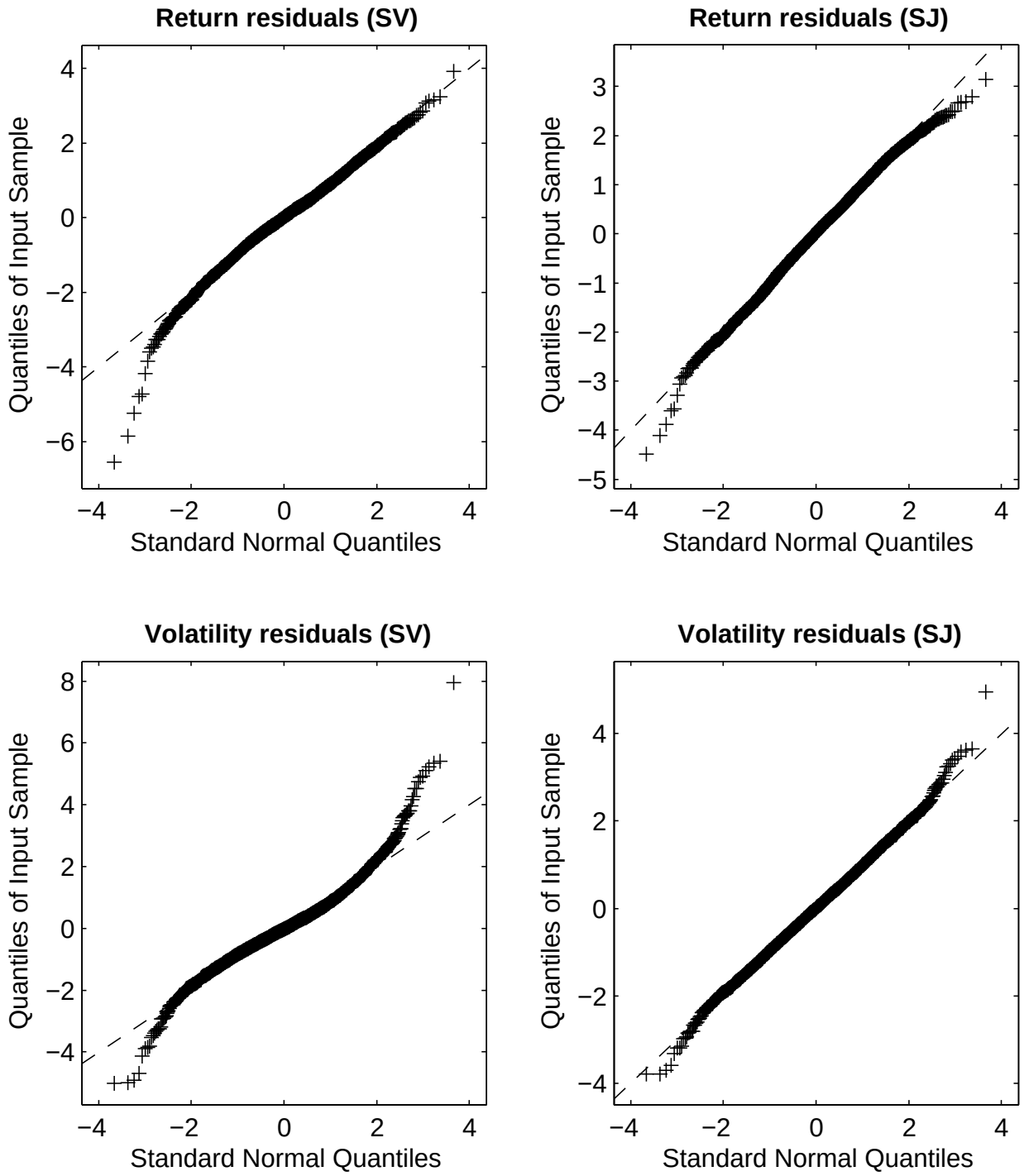

Figure 2.4: QQ plots for generalized residuals, SV and SJ models.

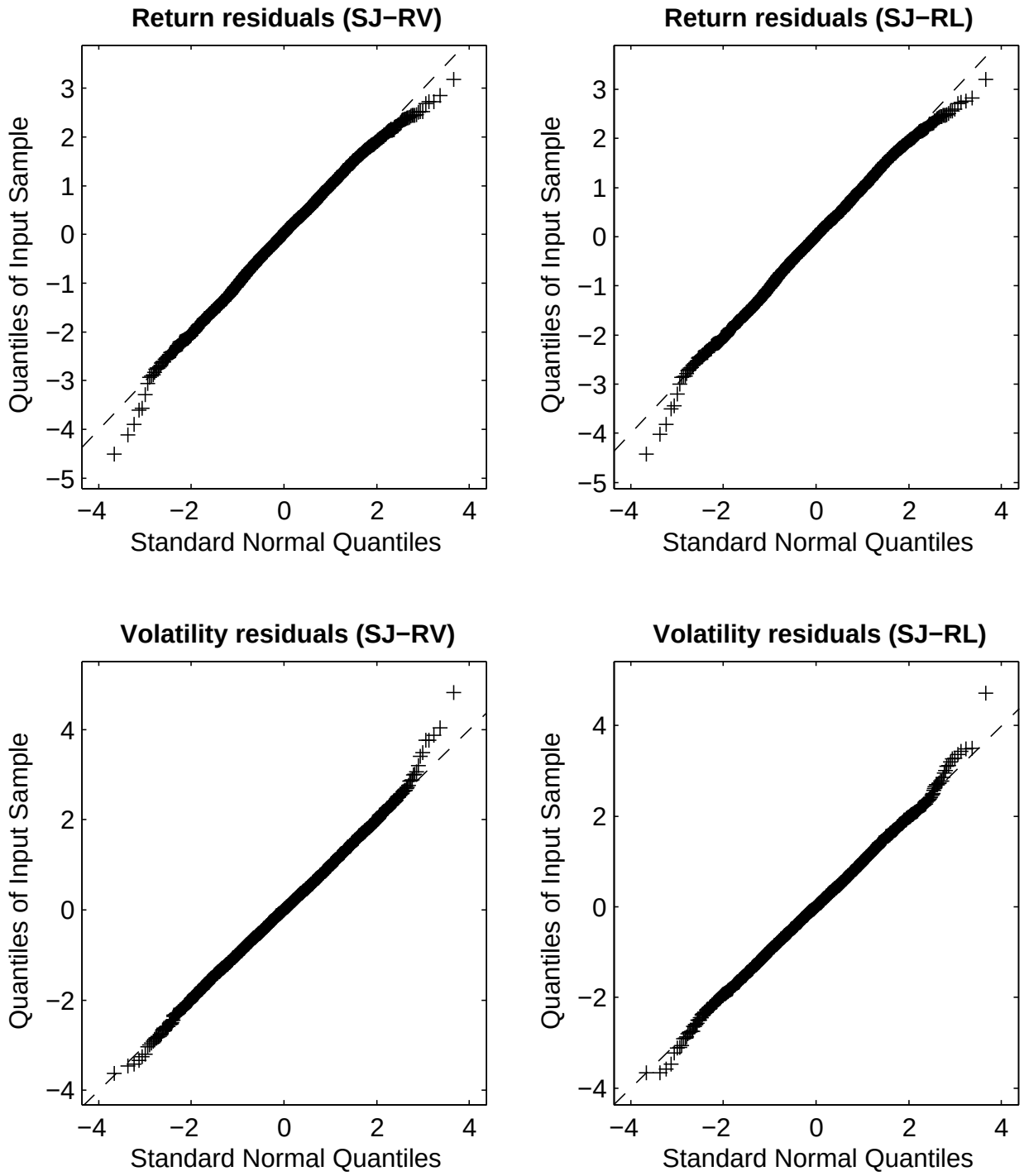

Figure 2.5: QQ plots for generalized residuals, SJ-RV and SJ-RL models.

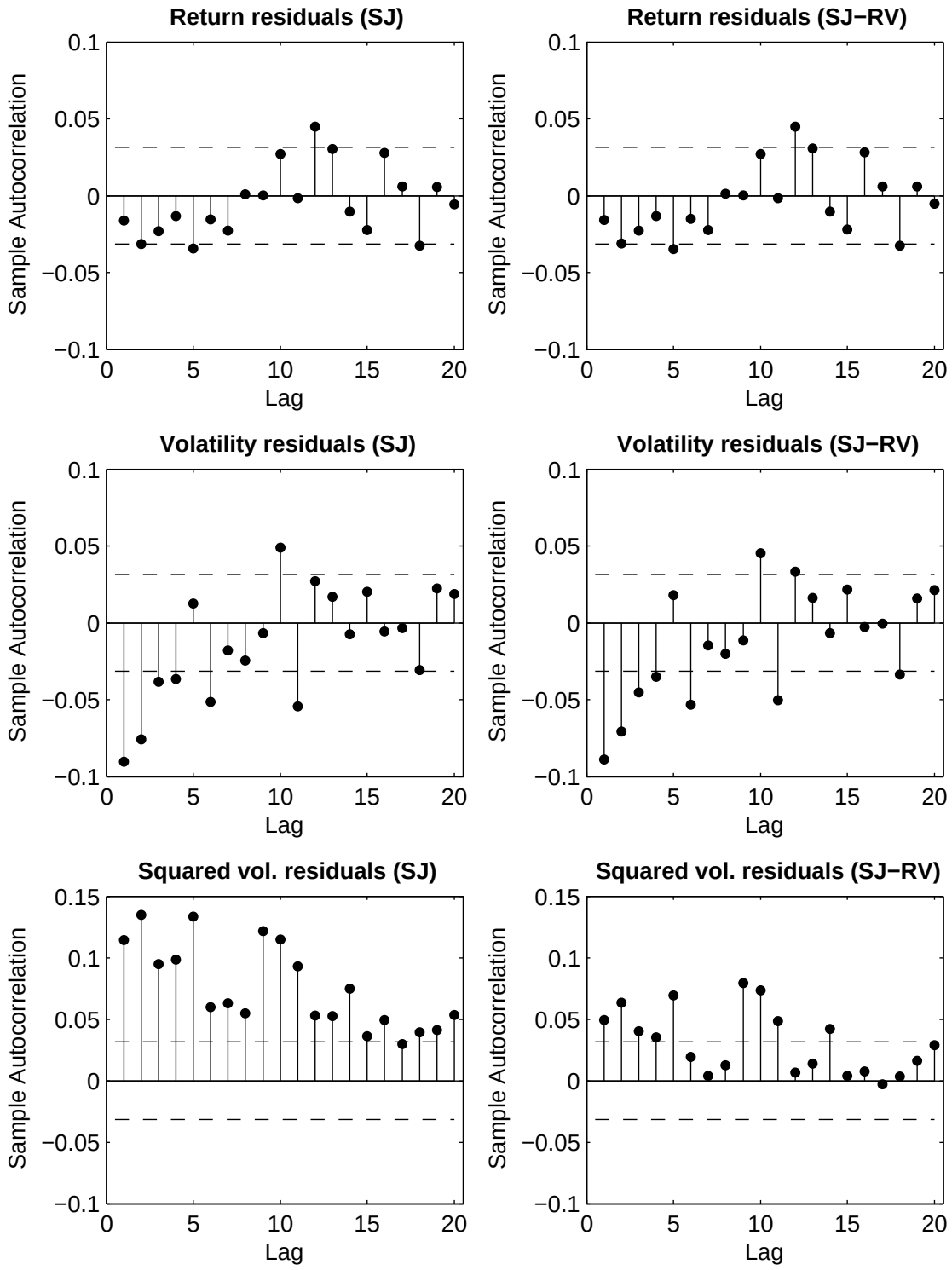

Figure 2.6: Correlograms for SJ and SJ-RV models.

Dotted lines show the 95% confidence band.

Bibliography

ADRIAN, T., AND J. ROSENBERG (2008): "Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk," *Journal of Finance*, 63, 2997–3030.

AIT-SAHALIA, Y., AND R. KIMMEL (2007): "Maximum Likelihood Estimation of Stochastic Volatility Models," *Journal of Financial Economics*, 83, 413–452.

AIT-SAHALIA, Y., AND A. W. LO (1998): "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," *Journal of Finance*, 53, 499–547.

——— (2000): "Nonparametric Risk Management and Implied Risk Aversion," *Journal of Econometrics*, 94, 9–51.

ANDERSEN, T., L. BENZONI, AND J. LUND (2002): "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57, 1239–1284.

ANDERSEN, T., T. BOLLERSLEV, F. DIEBOLD, AND H. EBENS (2001): "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43–76.

ANDERSEN, T., T. BOLLERSLEV, F. DIEBOLD, AND P. LABYS (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

BAI, J. (2003): "Testing Parametric Conditional Distributions of Dynamic Models," *Review of Economics and Statistics*, 85, 531–549.

BAILLIE, R. (1996): "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5–59.

BAKSHI, G., C. CAO, AND Z. CHEN (1997): "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 527–566.

BAKSHI, G., AND N. KAPADIA (2003): "Delta-Hedged Gains and the Negative Market Volatility Risk Premium," *Review of Financial Studies*, 16, 527–566.

BAKSHI, G., N. KAPADIA, AND D. MADAN (2003): "Stock Return Characteristics, Skew Laws, and Differential Pricing of Individual Equity Options," *Review of Financial Studies*, 16, 101–143.

BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (2002): "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of Royal Statistical Society B*, 64, 253–280.

——— (2004): "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

BARONE-ADESI, G., R. F. ENGLE, AND L. MANCINI (2008): "A GARCH Option Pricing Model with Filtered Historical Simulation," *Review of Financial Studies*, 21, 1223–1258.

BATES, D. (2000): "Post-'87 Crash Fears in S&P500 Futures Options," *Journal of Econometrics*, 94, 181–238.

BERNDT, E., B. HALL, R. HALL, AND J. HAUSMAN (1974): "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, 3, 653–665.

BLAIR, B. J., S. POON, AND S. J. TAYLOR (2001): "Forecasting S&P 100 Volatility: the Incremental Information Content of Implied Volatilities and High-Frequency Index Returns," *Journal of Econometrics*, 105, 5–26.

BOLLEN, N., AND R. WHALEY (2004): "Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?," *Journal of Finance*, 59, 711–753.

BOLLERSLEV, T., M. GIBSON, AND H. ZHOU (2011): "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities," *Journal of Econometrics*, 160, 235–245.

BOLLERSLEV, T., AND H. O. MIKKELSEN (1996): "Modeling and Pricing Long Memory in Stock Market Volatility," *Journal of Econometrics*, 73, 151–184.

BONDARENKO, O. (2007): "Nonparametric Test of Affine Option Models," Working paper, University of Illinois at Chicago.

BRITTEN-JONES, M., AND A. NEUBERGER (2000): "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance*, 55, 839–866.

BROADIE, M., M. CHERNOV, AND M. JOHANNES (2007): "Model Specification and Risk Premia: Evidence from Futures Options," *Journal of Finance*, 62, 1453–1490.

CARR, P., AND D. MADAN (1998): "Towards a Theory of Volatility Trading," in *Volatility: new estimation techniques for pricing derivatives*, ed. by R. Jarrow. RISK Publications, London.

CARR, P., AND L. WU (2003): "The Finite Moment Log Stable Process and Option Pricing," *Journal of Finance*, 58, 753–777.

——— (2007): "Stochastic Skew in Currency Options," *Journal of Financial Economics*, 86, 213–247.

——— (2009): "Variance Risk Premiums," *Review of Financial Studies*, 22, 1311–1341.

CHACKO, G., AND L. VICEIRA (2003): "Spectral GMM Estimation of Continuous-Time Processes," *Journal of Econometrics*, 116, 259–292.

CHERNOV, M., AND E. GHYSELS (2000): "A Study towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 56, 407–458.

CHOURDAKIS, K., AND G. DOTSIS (2009): "Maximum Likelihood Estimation and Dynamic Asset Allocation with Non-Affine Volatility Processes," Working paper, University of Essex.

CHRISTENSEN, B. J., AND N. R. PRABHALA (1998): "The Relation between Implied and Realized Volatility," *Journal of Financial Economics*, 50, 125–150.

CHRISTOFFERSEN, P., S. HESTON, AND K. JACOBS (2009): "The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well," *Management Science*, 55, 1914–1932.

CHRISTOFFERSEN, P., K. JACOBS, C. DORION, AND Y. WANG (2010): "Volatility Components, Affine Restrictions, and Non-Normal Innovations," *Journal of Business and Eoconomic Statistics*, Forthcoming.

CHRISTOFFERSEN, P., K. JACOBS, AND K. MIMOUNI (2010): "Models for S&P500 Dynamics: Evidence from Realized Volatility, Daily Returns, and Option Prices," *Review of Financial Studies*, 23, 3141–3189.

CHRISTOFFERSEN, P., K. JACOBS, C. ORNTHANALAI, AND Y. WANG (2008): "Option Valuation with Long-Run and Short-Run Volatility Components," *Journal of Financial Economics*, 90, 272–297.

COVAL, J., AND T. SHUMWAY (2001): "Expected Option Returns," *Journal of Finance*, 56, 983–1009.

DAS, S. R., AND R. K. SUNDARAM (1999): "Of Smiles and Smirks: A Term Structure Perspective," *Journal of Financial and Quantitative Analysis*, 34, 211–239.

DEMETERFI, K., E. DERMAN, M. KAMAL, AND J. ZOU (1999): "More Than You Ever Wanted to Know about Volatility Swaps," Goldman sachs quantitative strategies research notes, Goldman Sachs.

DENNIS, P., AND S. MAYHEW (2002): "Risk-Neutral Skewness: Evidence from Stock Options," *Journal of Financial and Quantitative Analysis*, 37, 471–493.

DERMAN, E., AND I. KANI (1994): "Riding on the Smile," *Risk*, 7, 32–39.

DING, Z., C. GRANGER, AND R. ENGLE (1993): "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1, 83–106.

DIZ, F., AND T. J. FINUCANE (1993): "Do the Options Markets Really Overreact?," *Journal of Futures Markets*, 13, 299–312.

DUAN, J.-C. (1995): "The GARCH Option Pricing Model," *Mathematical Finance*, 5, 13–32.

——— (2003): "A Specification Test for Time-Series Models by a Normality Transformation," Working paper, University of Toronto.

DUAN, J.-C., AND C. YEH (2007): "Jump and Volatility Risk Premiums Implied by VIX," Working paper, University of Toronto.

DUFFIE, D., J. PAN, AND K. SINGLETON (2000): "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343–1376.

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

DUPIRE, B. (1994): "Pricing with a Smile," *Risk*, 7, 18–20.

DURHAM, G. B. (2006): "Monte Carlo Methods for Estimating, Smoothing, and Filtering One- and Two-Factor Stochastic Volatility Models," *Journal of Econometrics*, 133, 273–305.

——— (2008): "Risk-Neutral Modelling with Affine and Non-Affine Models," Working paper, University of Colorado at Boulder.

DURHAM, G. B., AND Y.-H. PARK (2010): "Beyond Stochastic Volatility and Jumps in Returns and Volatility," Working paper, University of Colorado at Boulder.

EGLOFF, D., M. LEIPPOLD, AND L. WU (2010): "Variance Risk Dynamics, Variance Risk Premia, and Optimal Variance Swap Investments," *Journal of Financial and Quantitative Analysis*, Forthcoming.

ERAKER, B. (2004): "Do Equity Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," *Journal of Finance*, 59, 1367–1403.

ERAKER, B., M. JOHANNES, AND N. POLSON (2003): "The Impact of Jumps in Volatility and Returns," *Journal of Finance*, 58, 1269–1300.

FLEMING, J. (1998): "The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices," *Journal of Empirical Finance*, 5, 317–345.

GALLANT, A. R., C. HSU, AND G. TAUCHEN (1999): "Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance," *Review of Economics and Statistics*, 81, 617–631.

HAN, B. (2008): "Investor Sentiment and Option Prices," *Review of Financial Studies*, 21, 387–414.

HARVEY, C. R., AND A. SIDDIQUE (1999): "Autoregressive Conditional skewness," *Journal of Financial and Quantitative Analysis*, 34, 465–487.

——— (2000): "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, 55, 1263–1295.

HESTON, S. L. (1993): "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327–343.

HESTON, S. L., AND S. NANDI (2000): "A Closed-Form GARCH Option Valuation Model," *Review of Financial Studies*, 13, 585–625.

HUANG, J.-Z., AND L. WU (2004): "Specification Analysis of Option Pricing Models Based on Time-Changed Levy Processes," *Journal of Finance*, 59, 1405–1439.

JIANG, G., AND Y. TIAN (2005): "The Model-Free Implied Volatility and Its Information Content," *Review of Financial Studies*, 18, 1305–1342.

——— (2007): "Extracting Model-Free Volatility from Option Prices: An Examination of the VIX Index," *Journal of Derivatives*, 14, 35–60.

JOHNSON, T. C. (2002): "Volatility, Momentum, and Time-Varying Skewness in Foreign Exchange Returns," *Journal of Business and Eoconomic Statistics*, 20, 390–411.

JONES, C. (2003): "The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets," *Journal of Econometrics*, 116, 181–224.

LI, G., AND C. ZHANG (2010): "On the Number of State Variables in Options Pricing," *Management Science*, 56, 2058–2075.

LU, Z., AND Y. ZHU (2010): "Volatility Components: the Term Structure of Dynamics of VIX Futures," *Journal of Futures Markets*, 30, 230–256.

NEUBERGER, A. (1994): "The Log Contract," *Journal of Portfolio Management*, 20, 74–80.

NEWEY, W. K., AND K. D. WEST (1987): "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

PAN, J. (2002): "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3–50.

RUBINSTEIN, M. (1994): "Implied Binomial Trees," *Journal of Finance*, 49, 771–818.

SANTA-CLARA, P., AND S. YAN (2010): "Crashes, Volatility, and the Equity Premium: Lessons from S&P500 Options," *Review of Economics and Statistics*, 87, 435–451.

SCHWERT, G. W. (1989): "Why Does Stock Market Volatility Change Over Time?," *Journal of Finance*, 44, 1115–1153.

STEIN, J. (1989): "Overreactions in the Options Market," *Journal of Finance*, 44, 1011–1023.

XING, Y., X. ZHANG, AND R. ZHAO (2007): "What does Individual Option Volatility Smirk Tell Us about Future Equity Returns," Working paper, Rice University.

XU, X., AND S. J. TAYLOR (1994): "The Term Structure of Volatility Implied by Foreign Exchange Options," *Journal of Financial and Quantitative Analysis*, 29, 57–74.