

Beyond Stochastic Volatility and Jumps in Returns and Volatility

Garland Durham¹ and Yang-Ho Park²

Leeds School of Business
University of Colorado

February 16, 2010

ABSTRACT

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. Although changes in the shape of the risk-neutral return distribution are related to the level of volatility, the relationship is weak. But, most models that attempt to fit both physical and risk-neutral models in a time-consistent manner include only a single state variable, volatility, and are thus capable of capturing little of the variation in implied skewness and kurtosis that is observed in the data. 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.

¹Leeds School of Business, University of Colorado, UCB 419, Boulder, CO. E-mail: garland.durham@colorado.edu

²Leeds School of Business, University of Colorado, UCB 419, Boulder, CO. E-mail: yangho@colorado.edu

Beyond Stochastic Volatility and Jumps in Returns and Volatility

ABSTRACT

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. Although changes in the shape of the risk-neutral return distribution are related to the level of volatility, the relationship is weak. But, most models that attempt to fit both physical and risk-neutral models in a time-consistent manner include only a single state variable, volatility, and are thus capable of capturing little of the variation in implied skewness and kurtosis that is observed in the data. 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.

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 (2009) 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 2007). 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 1 shows the time-series plots of option-

implied volatility, skewness, and kurtosis. Figure 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; Bates 2000; 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 2009). This is the basic approach used in this paper. But, unlike Christoffersen, Heston, and Jacobs (2009) and Santa-Clara and Yan (2009), 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 tack. 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 2009, 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 9.3%. VIX and VRP are significant, but not JV. 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 describes the models under consideration. In Section 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 4. Section 5 investigates the explanatory power of implied regime states from the models for risk-neutral implied-skewness and kurtosis. And, Section 6 concludes.

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

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

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 $J_{1t} \sim N(\mu_{1J}, \sigma_{1J}^2)$, $J_{2t} \sim N(\mu_{2J}, \sigma_{2J}^2)$, 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 $J_{1t}/\exp(v_t/2) \sim N(\mu_{1J}, \sigma_{1J}^2)$, $J_{2t}/\sigma(s_t) \sim N(\mu_{2J}, \sigma_{2J}^2)$, 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 $\bar{\mu}_{1J_t} = E(e^{J_{1t}} - 1)$ and $\bar{\mu}_{2J_t} = E(J_{2t})$ 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, $y(t) = \log(x_t)$,

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

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

$$\begin{aligned} dy_t &= \left[r_t - q_t - \bar{\mu}_{1Jt}\lambda_1 - \frac{1}{2}\exp(v_t) \right] dt + \exp(v_t/2)dW_{1t}^{\mathbb{Q}} + J_{1t}^{\mathbb{Q}}dN_{1t}^{\mathbb{Q}} \\ dv_t &= \left[k(\bar{v} - v_t) - \eta(s_t)v_t - \bar{\mu}_{2Jt}\lambda_1 \right] dt + \sigma(s_t)dW_{2t}^{\mathbb{Q}} + J_{2t}^{\mathbb{Q}}dN_{1t}^{\mathbb{Q}} \\ ds_t &= (1 - 2s_t)dN_{2t}^{\mathbb{Q}} \end{aligned} \tag{3}$$

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. The return risk premium is $\mu - (r_t - q_t)$.

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. We could equally well have set η equal to zero and introduced a jump risk

premium instead (this is the approach taken by Pan 2002). It would make essentially no difference in any of the results of the paper.

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 2009). 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 will use an Euler scheme approximation to the model. For the physical model (and analogously for the risk-neutral model),

the approximation is given by

$$\begin{aligned}
y_{t+1} &= y_t + \mu - \bar{\mu}_{1Jt}\lambda_1 - \frac{1}{2}\exp(v_t) + \exp(v_t/2)\varepsilon_{1,t+1} + \sum_{j>N_{1,t}}^{N_{1,t+1}} \xi_{1j} \\
v_{t+1} &= v_t + k(\bar{v} - v_t) - \bar{\mu}_{2Jt}\lambda_1 + \sigma(s_t)\varepsilon_{2,t+1} + \sum_{j>N_{1,t}}^{N_{1,t+1}} \xi_{2j}
\end{aligned} \tag{4}$$

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.

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.

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

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

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

simulations. For the SJ models, for example, we get

$$IV_t^T = \sqrt{\frac{1}{(T-t)S} \sum_{s=1}^S \left\{ \int_t^T e^{v_\tau^{(s)}} \left[1 + \lambda_1(\mu_{1J}^2 + \sigma_{1J}^2) \right] d\tau \right\}}, \quad (5)$$

where for each s , $v_\tau^{(s)}$ is obtained by simulating a path from the risk-neutral analog of Equation (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 (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,

$$\Gamma_i : SV \longrightarrow IV,$$

where SV denotes the spot volatility, i.e., $SV_t = \exp(v_t/2)$, and the subscript i denotes the conditioning on initial regime state. The simplest way to do this is to evaluate (5) on some grid of initial values for SV and then use some curve fitting technique to approximate Γ_i . That is, let $\widehat{SV}_1 < \widehat{SV}_2 < \dots < \widehat{SV}_G$ 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 $g = 1, \dots, G$, evaluate $\widehat{IV}_g = \Gamma_i(\widehat{SV}_g)$ 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 $\{(\widehat{SV}_g, \widehat{IV}_g)\}_{g=1}^G$. As long as this mapping is monotonic, it is equally straightforward to approximate the inverse, $\Gamma_i^{-1} : IV \longrightarrow SV$, 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 $\{(\widehat{SV}_g, \widehat{IV}_g)\}_{g=1}^G$ 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 $v_t = 2 \log(SV_t)$. Repeat this for each $t = 1, \dots, n$. 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.

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 $t = 0, \dots, n$, 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

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

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,

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

where NJ_t is the number of potential jumps on day t , NJ_{\max} is the maximum number of allowable jumps in a single day, and $p(NJ_t = k) = [\lambda_1]^k e^{-\lambda_1} / k!$ is given by the Poisson distribution with intensity λ_1 . The distribution of $(y_{t+1}, v_{t+1}^j | y_t, v_t^i, s_t = i, NJ_t = k)$ 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 (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 ,

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

3.3 Maximum likelihood estimation

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

$$\begin{aligned} \log L(\{y_t\}_{t=1}^n, \{IV_t\}_{t=1}^n; \theta) \approx & \sum_{t=1}^{n-1} \sum_{i=0}^1 \sum_{j=0}^1 \left[\log p(y_{t+1}, v_{t+1}^j | y_t, v_t^i, s_t = i) + \log p(s_{t+1} = j | s_t = i) \right. \\ & \left. + \log p(s_t = i) + \log J_{t+1}^j \right], \end{aligned} \quad (6)$$

where $J_{t+1}^j = |dv_{t+1}^j/dIV_{t+1}|$ is the Jacobian corresponding to regime state j . Recall that the mapping from volatility state, v_t^j , to IV_t is given by $IV_t = \Gamma_j[\exp(v_t^j/2)]$. 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,

$$\hat{\theta} = \arg \max \log L(\{y_t\}_{t=1}^n, \{IV_t\}_{t=1}^n; \theta).$$

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.

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 $u_t = G_t(z_t|\mathcal{F}_{t-1})$ ($t = 1, \dots, n$). 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

$$\tilde{u}_t = \Phi^{-1}(u_t), \quad t = 1, \dots, n \quad (7)$$

where Φ is the standard normal distribution function. In this case, the transformed residuals $\{\tilde{u}_t\}$ 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 (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,

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

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 corresponding to price and volatility innovations separately,

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

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.

4 Empirical results

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 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 5). 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), daily realized volatility is obtained by summing the squared intraday returns over each day,

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

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,

$$RV_t \equiv \sum_{i=0}^{21} RV_{t-i}^{(d)}.$$

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

$$VRP_t \equiv \log \left(RV_t / VIX_t^2 \right),$$

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

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

The daily jump variation is defined by subtracting the daily bipower variation from

the daily realized volatility,

$$JV_t^{(d)} \equiv \max(RV_t^{(d)} - BV_t^{(d)}, 0).$$

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

$$JV_t \equiv \sum_{i=0}^{21} JV_{t-i}^{(d)}.$$

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 1. Figure 2 shows scatter plots of option-implied skewness and kurtosis versus option-implied volatility. Summary statistics are reported in Table 1.

4.2 Parameter estimates and model comparisons

Parameter estimates and log likelihoods for all of the models under consideration are shown in Table 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: $\text{SJ-RVL} > \text{SJ-RV} > \text{SJ} > \text{SV}$. 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 typically 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 $\text{SJ-RVL} > \text{SJ-RV} > \text{SJ-RL} > \text{SJ}$. 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 corresponds to a more negative variance risk premium. This is consistent with theory (Bollerslev, Gibson, and Zhou 2009). 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 3. For the models with regime switching in volatility of volatility, for example, σ takes on two possible values (σ^0 or σ^1). The filter described in Section 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 are thus given by $\hat{\sigma}_t = E_t[\sigma(s_t)] = p_t^0 \sigma^0 + p_t^1 \sigma^1$ ($t = 1, \dots, n$). An analogous procedure is used to compute $\hat{\rho}_t$ ($t = 1, \dots, n$) 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.

4.3 Diagnostics

QQ-plots for return and volatility residuals are shown for several models in Figures 4 and 5. Correlograms are shown in Figure 6. Jarque-Bera and Ljung-Box statistics are shown in Table 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. In the first place, 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.

5 Explanatory power for option-implied skewness and kurtosis

Sections 4.2 and 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 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

$$\begin{aligned} \text{SKEW}_t &= \beta_0 + \beta_{\text{RV}} \hat{s}_t^{\text{RV}} + \beta_{\text{RL}} \hat{s}_t^{\text{RL}} + \beta_{\text{VIX}} \log \text{VIX}_t + \beta_{\text{VRP}} \text{VRP}_t + \beta_{\text{JV}} \log \text{JV}_t + \varepsilon_t \\ \text{KURT}_t &= \beta_0 + \beta_{\text{RV}} \hat{s}_t^{\text{RV}} + \beta_{\text{RL}} \hat{s}_t^{\text{RL}} + \beta_{\text{VIX}} \log \text{VIX}_t + \beta_{\text{VRP}} \text{VRP}_t + \beta_{\text{JV}} \log \text{JV}_t + \varepsilon_t \end{aligned} \quad (8)$$

where \hat{s}_t^{RV} denotes the filtered state under the models with regime switching in volatility of volatility and \hat{s}_t^{RL} denotes the filtered state under the models with regime switching in leverage effect (as described in Section 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 (2009) 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 4.1. We first run the regressions on the control variables alone, and then add the regime state 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 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. The variance risk premium (VRP) is also highly significant for all models. The coefficient is always negative, suggesting that risk-aversion is positively related to skewness. That is, the risk-neutral distribution tends to be more left-skewed when risk aversion is low. This is consistent with the finding of Bakshi and Kapadia (2003) that there is a negative relation between delta-hedged gains (which are positively related to the variance risk premium) and skewness. We do not find jump variation (JV) 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 9.3% 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 4.2 and 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 distribution is substantially more left skewed in the high volatility of volatility state). The t -statistic is -9.94, corresponding to a p -value of around 10^{-22} . This model has good explanatory power, with an adjusted R^2 of 19.2% (versus 9.3% 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 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 distribution becomes substantially more left skewed in the state with the stronger leverage effect). The t -statistic is -7.93, 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 -11.43, 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 5 shows analogous 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. The variance risk premium (VRP) is positively related to kurtosis. That is, the risk-neutral distribution tends to be more fat-tailed when risk aversion is low. Jump variation (JV) is not significant when the other control variables are included.

Although VRP on its own is only marginally significant and has little explanatory power, when combined with the VIX it becomes highly significant. The regression on all three explanatory variables together has a surprisingly high R^2 of 17.0%.

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 18.4% for SJ-RL versus 22.8% 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 25.6%. In the full model, slope coefficients are 3.08 and 5.09 (with t -statistics of 5.38 and 8.86) 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 4.2 and 4.3. They also did not prove informative in terms of explanatory power for option-implied skewness or kurtosis, as discussed in this section. Our regression results do include the jump variation variable (JV) as a control. This variable, which serves as a proxy for jump risk, is not significant at even the 10% level in the regressions for either option-implied skewness or kurtosis when the other control variables are included.

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.

6 Conclusion

This paper proposes a new class of models that layers 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 are a prominent feature of observed data, we allow for regime switching in two parameters of the basic model: volatility of volatility and leverage effect. The both parameters are able to affect the skewness and kurtosis of returns. The application looks at SPX index data. To estimate the models, we rely upon data on only the time-series of price and option-implied volatility of the index. This allows us to use observations on option skewness and kurtosis for diagnostic purposes.

The models with regime switching fit the data much better than those without regime switching. Accounting for time-variation in volatility of volatility or leverage effect not only increases log likelihood significantly, but also provides improvements in some of our diagnostics of model fit. For example, the data show strong evidence of stochastic volatility of volatility. The regime-switching model in volatility of volatility does a good job of explaining this feature of the data. The best model in this dimension is the one with regime-switching in both the volatility of volatility and leverage effect.

While the fact that our models do a better job of explaining the time-series data of price and volatility than existing models is useful, the models also demonstrate explanatory power for option skewness and kurtosis. To test this, we run regressions of option skewness and kurtosis on filtered regime states and several control variables

for each model. Since we do not use the information of skewness and kurtosis in model estimation, such regressions are meaningful.

Changes in volatility of volatility and leverage effect are each significantly associated with time-variations in option skewness and kurtosis. The results are robust to other possible determinants such as volatility state, variance risk premium (proxy for risk aversion), and jump variation (proxy for jump state). The best performing model is the one with regime switching in both volatility of volatility and leverage effect.

The coefficient on the regime state is in expected direction for all models. This result indicates that the option-implied return distributions tend to be more left-skewed and leptokurtic at a time of higher volatility of volatility or stronger leverage effect. Any model which makes an attempt to fit option skewness or minimize option pricing errors may be at risk of spuriously fitting option skewness, unless it consider the stochastic factors in question. A correct option pricing model should be able to account for the stochastic volatility of volatility and leverage effect as implied in the time-series data.

We also find that the specification of jump structures is important in fitting the time-series data of returns and volatility as well as matching the time-variation in skewness and kurtosis. The volatility-scaled jump model (SJ) outperforms the volatility-unscaled jump model (UJ) in these respects. The former not only improves log likelihood with the same number of parameters, but also provide far better diagnostic features of model fit than the latter. Besides, the SJ-based regime switching models do a better job of explaining option skewness and kurtosis than the UJ-based regime-switching models.

References

- 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.
- 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.
- 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.
- 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., S. F. GRAY, AND R. E. WHALEY (2000): “Regime Switching in Foreign Exchange Rates: Evidence From Currency Option Prices,” *Journal of Econometrics*, 94, 239–276.
- 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 (2009): “Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities,” Forthcoming in journal of econometrics.
- 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 (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.

- 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, AND K. MIMOUNI (2007): “Models for S&P500 Dynamics: Evidence from Realized Volatility, Daily Returns, and Option Prices,” Working paper, McGill University.
- 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.
- DUAN, J.-C. (2003): “A Specification Test for Time-Series Models by a Normality Transformation,” Working paper, University of Toronto.
- DUAN, J.-C., I. POPOVA, AND P. RITCHKEN (2002): “Option Pricing Under Regime Switching,” *Quantitative Finance*, 2, 116–132.
- 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.
- DUPIRE, B. (1994): “Pricing with a Smile,” *Risk*, 7, 18–20.

- 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.
- GRAY, S. F. (1996): “Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process,” *Journal of Financial Economics*, 42, 27–62.
- HAMILTON, J. D. (1989): “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle,” *Econometrica*, 57, 357–384.
- 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.
- 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 Economic Statistics*, 20, 390–411.
- JONES, C. (2003): “The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets,” *Journal of Econometrics*, 116, 181–224.
- KALIMIPALLI, M., AND R. SUSMEL (2004): “Regime-Switching Stochastic Volatility and Short-Term Interest Rate,” *Journal of Empirical Finance*, 11, 309–329.
- NEUBERGER, A. (1994): “The Log Contract,” *Journal of Portfolio Management*, 20, 74–80.
- NEWKEY, 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 (2009): “Crashes, Volatility, and the Equity Premium: Lessons from S&P500 Options,” Forthcoming in review of economics and statistics.
- SMITH, D. R. (2002): “Markov-Switching and Stochastic Volatility Diffusion Models of Short-Term Interest Rates,” *Journal of Business and Economic Statistics*, 20, 183–197.
- XING, Y., X. ZHANG, AND R. ZHAO (2007): “What does Individual Option Volatility Smirk Tell Us about Future Equity Returns,” Working paper, Rice University.

Table 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 30 days. $\text{VRP}_t \equiv \log(\text{RV}_t/\text{VIX}_t^2)$ denotes the variance risk premium. $\text{AR}(i)$ 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: 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 (1.27) | 3.70 (1.28) | 3.08 (1.25) | 3.03 (1.25) | 2.07 (1.25) | 3.42 (1.26) | 3.08 (1.23) | 3.00 (1.25) | 2.08 (1.25) |
| $\kappa \times 10^3$ | 4.29 (1.29) | 6.00 (1.35) | 6.69 (1.33) | 6.34 (1.32) | 11.07 (1.28) | 7.48 (1.35) | 7.36 (1.34) | 6.27 (1.32) | 11.04 (1.28) |
| \bar{v} | -9.71 (0.55) | -10.03 (0.38) | -9.70 (0.36) | -10.06 (0.32) | -9.63 (0.20) | -10.00 (0.32) | -9.80 (0.33) | -9.98 (0.33) | -9.63 (0.20) |
| $\eta^0 \times 10^3$ | 3.01 (0.25) | 2.84 (0.36) | 3.84 (0.35) | 2.86 (0.42) | 4.18 (0.36) | 2.49 (0.47) | 2.72 (0.44) | 3.01 (0.37) | 4.20 (0.36) |
| $\eta^1 \times 10^3$ | | | | 2.97 (0.59) | 2.46 (0.41) | 3.39 (0.40) | 4.27 (0.38) | 2.66 (0.54) | 2.46 (0.42) |
| $\sigma^0 \times 10^1$ | 1.32 (0.02) | 0.91 (0.03) | 1.23 (0.03) | 0.84 (0.03) | 1.18 (0.03) | 0.99 (0.03) | 1.25 (0.03) | 0.85 (0.03) | 1.18 (0.03) |
| $\sigma^1 \times 10^1$ | | | | 1.33 (0.06) | 2.12 (0.06) | | | 1.33 (0.05) | 2.12 (0.06) |
| ρ^0 | -0.74 (0.01) | -0.78 (0.01) | -0.82 (0.01) | -0.81 (0.01) | -0.91 (0.01) | -0.53 (0.04) | -0.61 (0.03) | -0.77 (0.01) | -0.91 (0.01) |
| ρ^1 | | | | | | -0.82 (0.01) | -0.85 (0.01) | -0.87 (0.02) | -0.91 (0.01) |
| ρ_J | | -0.71 (0.03) | -0.53 (0.04) | -0.70 (0.03) | -0.05 (0.10) | -0.77 (0.03) | -0.72 (0.04) | -0.69 (0.03) | -0.04 (0.10) |
| λ | | 0.47 (0.06) | 0.33 (0.04) | 0.55 (0.09) | 1.08 (0.14) | 0.30 (0.04) | 0.15 (0.02) | 0.46 (0.07) | 1.08 (0.15) |
| $\mu_{1J} \times 10^2$ | | -1.13 (6.59) | 0.05 (0.04) | -4.52 (6.38) | 0.08 (0.02) | -11.90 (8.78) | -0.05 (0.07) | -3.78 (6.98) | 0.08 (0.02) |
| $\sigma_{1J} \times 10^1$ | | 12.94 (0.75) | 0.06 (0.00) | 12.44 (0.77) | 0.03 (0.00) | 13.42 (0.85) | 0.08 (0.01) | 12.30 (0.77) | 0.03 (0.00) |
| $\mu_{2J} \times 10^1$ | | 2.79 (0.69) | 0.24 (0.07) | 3.39 (0.66) | 0.09 (0.04) | 3.59 (0.95) | 0.44 (0.13) | 3.04 (0.70) | 0.08 (0.04) |
| σ_{2J} | | 1.54 (0.07) | 0.15 (0.01) | 1.15 (0.07) | 0.05 (0.01) | 1.72 (0.09) | 0.21 (0.01) | 1.25 (0.07) | 0.05 (0.01) |
| π_0 | | | | 0.98 (0.01) | 0.98 (0.01) | 0.97 (0.01) | 0.96 (0.01) | 0.99 (0.00) | 0.98 (0.01) |
| π_1 | | | | 0.94 (0.01) | 0.94 (0.01) | 0.99 (0.00) | 0.98 (0.01) | 0.96 (0.01) | 0.94 (0.01) |
| $\log(L)$ | 38,491 | 38,851 | 38,810 | 38,961 | 38,960 | 38,883 | 38,858 | 38,969 | 38,960 |

Test statistics are shown with p -values in parentheses.

| | 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) |

| | 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 (0.002) | 41 (0.004) | 41 (0.004) | 41 (0.004) | 40 (0.005) | 41 (0.004) | 40 (0.005) | 41 (0.004) | 40 (0.005) |
| Volatility | 116 (0.000) | 118 (0.000) | 116 (0.000) | 115 (0.000) | 100 (0.000) | 115 (0.000) | 114 (0.000) | 116 (0.000) | 100 (0.000) |
| Squared Vol. | 419 (0.000) | 552 (0.000) | 424 (0.000) | 128 (0.000) | 74 (0.000) | 559 (0.000) | 446 (0.000) | 118 (0.000) | 74 (0.000) |

Table 4: Regressions for option-implied skewness.

The table reports the results of the following regression,

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

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^{RV} and \hat{s}_t^{RL} for volatility of volatility and leverage effect respectively. The control variables are the VIX index, variance risk premium (VRP), and jump variation (JV). Results are shown for filtered states from both SJ (scaled jumps) and UJ (unscaled jumps) models.

| Constant | | $\widehat{s}_t^{\text{RV}}$ | | $\widehat{s}_t^{\text{RL}}$ | | $\log \text{VIX}_t$ | | VRP_t | | $\log \text{JV}_t$ | | 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% |
| -3.56 | (-9.34) | | | | | 0.67 | (4.65) | -0.11 | (-3.37) | 0.001 | (0.74) | 9.3% |
| 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% |
| -2.76 | (-16.78) | -0.57 | (-9.06) | | | 0.41 | (7.29) | | | -0.001 | (-1.10) | 16.3% |
| -3.80 | (-12.12) | -0.61 | (-9.92) | | | 0.81 | (7.00) | -0.14 | (-5.07) | 0.001 | (1.00) | 19.1% |
| -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% |
| -2.73 | (-11.68) | | | -0.63 | (-8.12) | 0.50 | (5.11) | | | -0.001 | (-0.80) | 17.7% |
| -3.22 | (-8.26) | | | -0.60 | (-7.89) | 0.68 | (4.43) | -0.07 | (-2.03) | 0.000 | (-0.11) | 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% |
| -2.73 | (-14.90) | -0.72 | (-11.55) | -0.78 | (-11.90) | 0.59 | (8.15) | | | -0.002 | (-1.89) | 31.1% |
| -3.44 | (-11.21) | -0.74 | (-11.96) | -0.74 | (-11.69) | 0.86 | (7.28) | -0.10 | (-3.59) | -0.001 | (-0.39) | 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% |
| -2.55 | (-14.94) | -0.44 | (-7.55) | | | 0.33 | (5.79) | | | -0.001 | (-1.15) | 14.3% |
| -3.58 | (-11.23) | -0.48 | (-8.50) | | | 0.73 | (6.24) | -0.14 | (-5.05) | 0.001 | (0.98) | 17.2% |
| -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% |
| -2.86 | (-12.77) | | | -0.50 | (-6.97) | 0.50 | (5.58) | | | -0.001 | (-0.73) | 14.2% |
| -3.50 | (-8.96) | | | -0.48 | (-6.82) | 0.74 | (4.89) | -0.09 | (-2.67) | 0.000 | (0.16) | 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% |
| -2.63 | (-14.76) | -0.58 | (-10.26) | -0.67 | (-10.38) | 0.51 | (7.79) | | | -0.002 | (-2.09) | 25.1% |
| -3.51 | (-11.36) | -0.61 | (-10.96) | -0.64 | (-10.45) | 0.85 | (7.30) | -0.12 | (-4.50) | 0.000 | (-0.06) | 27.2% |

Table 5: Regression for option-implied kurtosis.

The table reports the results of the following regression,

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

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^{RV} and \hat{s}_t^{RL} for volatility of volatility and leverage effect respectively. The control variables are the VIX index, variance risk premium (VRP), and jump variation (JV). Results are shown for filtered states from both SJ (scaled jumps) and UJ (unscaled jumps) models.

| Constant | | $\widehat{s}_t^{\text{RV}}$ | | $\widehat{s}_t^{\text{RL}}$ | | $\log \text{VIX}_t$ | | VRP_t | | $\log \text{JV}_t$ | | 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% |
| 35.64 | (10.67) | | | | | -9.48 | (-7.59) | 2.24 | (8.09) | -0.02 | (-1.31) | 17.0% |
| 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% |
| 19.25 | (11.69) | 1.50 | (2.55) | | | -3.34 | (-6.18) | | | 0.02 | (2.90) | 9.4% |
| 36.51 | (11.75) | 2.18 | (3.87) | | | -10.00 | (-8.74) | 2.35 | (8.98) | -0.02 | (-1.47) | 18.4% |
| 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% |
| 19.04 | (10.22) | | | 5.29 | (7.87) | -4.38 | (-5.99) | | | 0.02 | (2.26) | 16.9% |
| 33.12 | (9.75) | | | 4.52 | (7.41) | -9.58 | (-7.26) | 1.92 | (7.02) | -0.01 | (-0.50) | 22.8% |
| 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% |
| 19.04 | (10.99) | 2.63 | (4.53) | 5.83 | (9.04) | -4.72 | (-7.21) | | | 0.027 | (3.11) | 18.9% |
| 34.03 | (10.99) | 3.08 | (5.38) | 5.09 | (8.86) | -10.32 | (-8.77) | 2.04 | (8.00) | -0.007 | (-0.48) | 25.6% |
| 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% |
| 18.94 | (11.06) | 0.66 | (1.20) | | | -3.17 | (-5.66) | | | 0.02 | (2.57) | 8.8% |
| 35.71 | (11.19) | 1.33 | (2.59) | | | -9.66 | (-8.24) | 2.33 | (8.85) | -0.02 | (-1.43) | 17.7% |
| 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% |
| 20.14 | (11.09) | | | 4.35 | (6.71) | -4.43 | (-6.50) | | | 0.02 | (2.35) | 14.3% |
| 35.14 | (10.33) | | | 3.81 | (6.58) | -10.06 | (-7.68) | 2.06 | (7.55) | -0.01 | (-0.74) | 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% |
| 19.48 | (11.06) | 1.68 | (3.15) | 4.82 | (7.58) | -4.46 | (-7.11) | | | 0.025 | (3.14) | 15.3% |
| 35.18 | (11.14) | 2.23 | (4.35) | 4.40 | (7.82) | -10.43 | (-8.83) | 2.19 | (8.52) | -0.011 | (-0.74) | 23.1% |

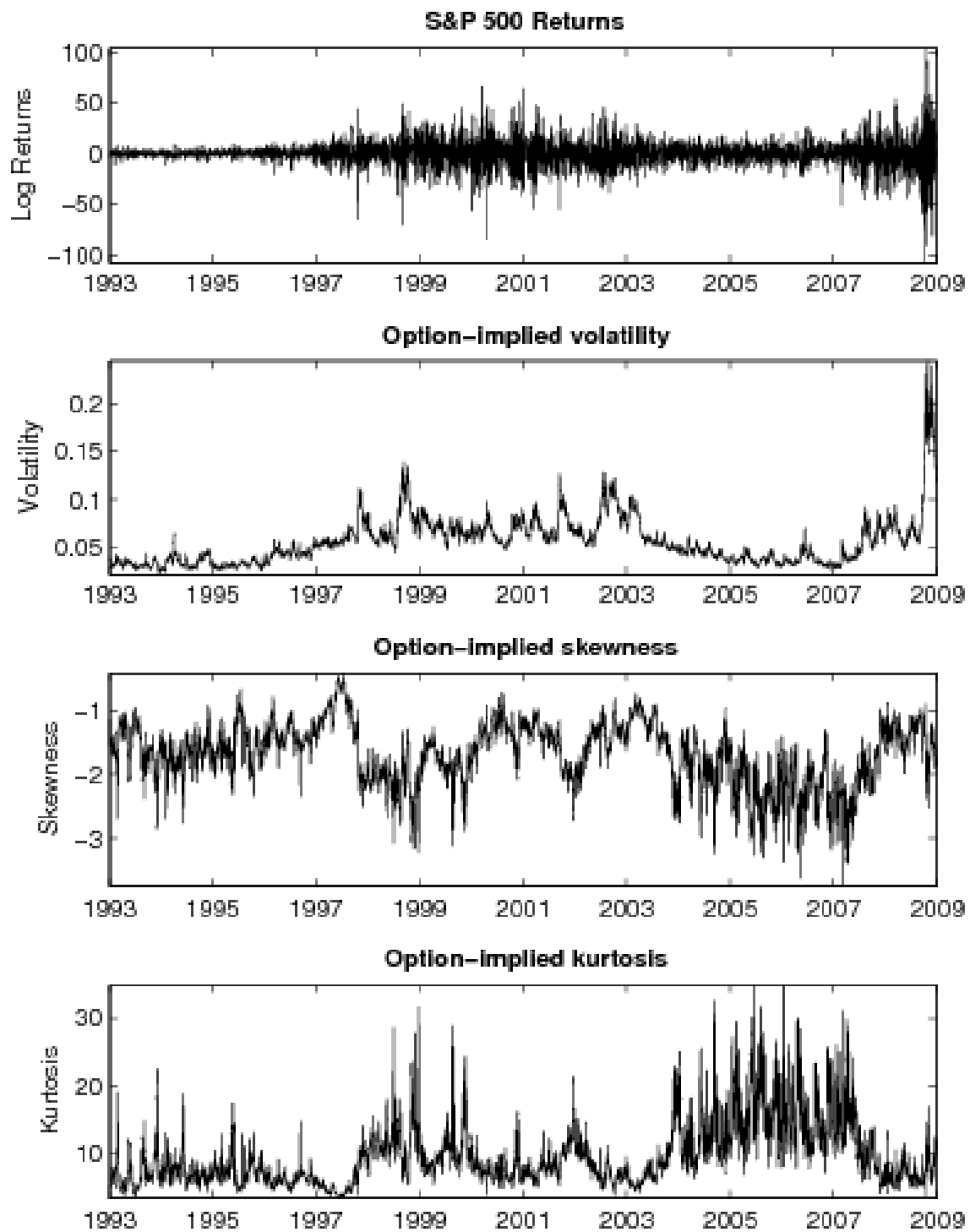


Figure 1: Time series of S&P 500 returns, option-implied volatility, option-implied skewness, and option-implied kurtosis.

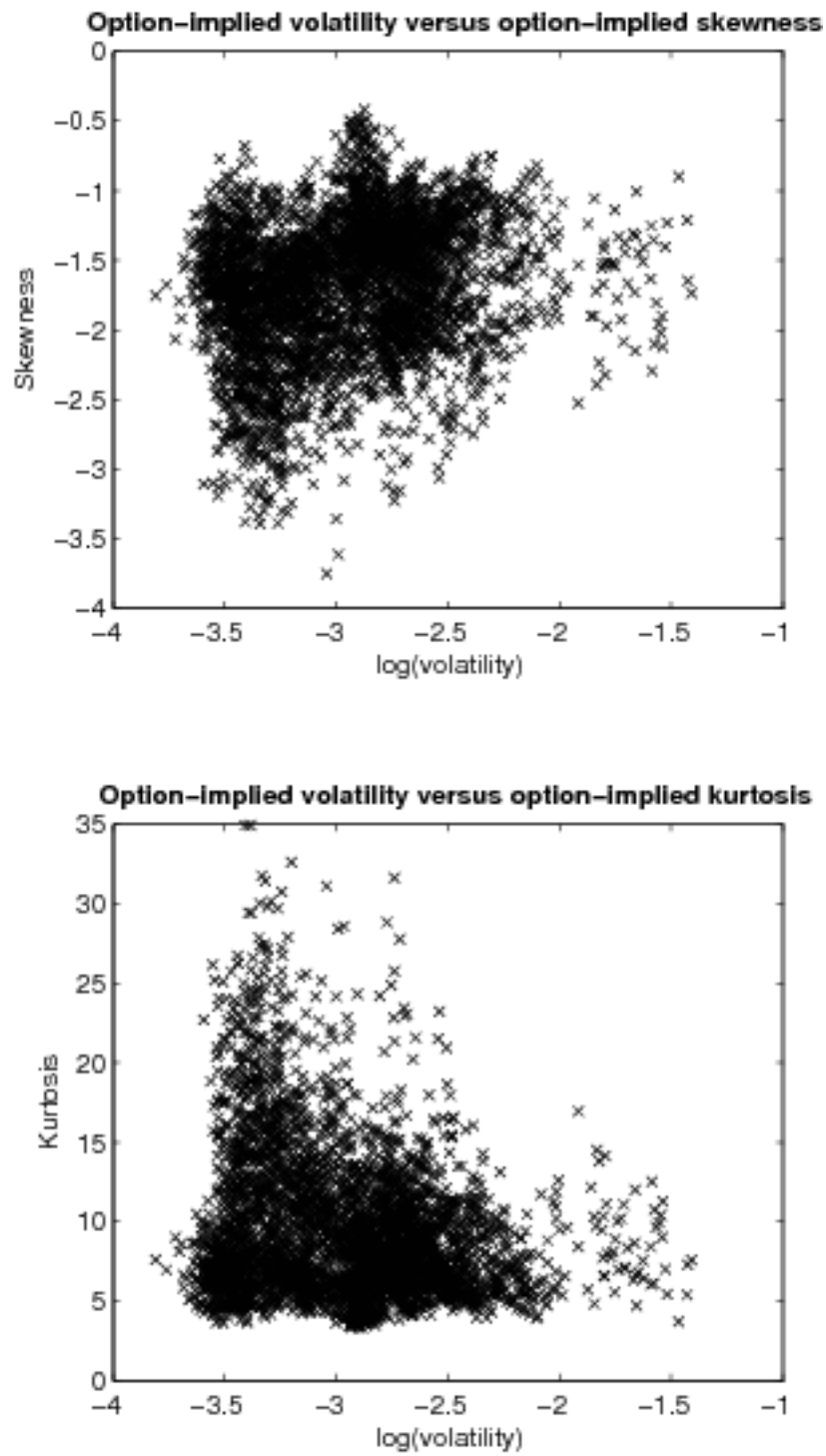


Figure 2: Scatter plots of option-implied skewness and kurtosis against option-implied volatility.

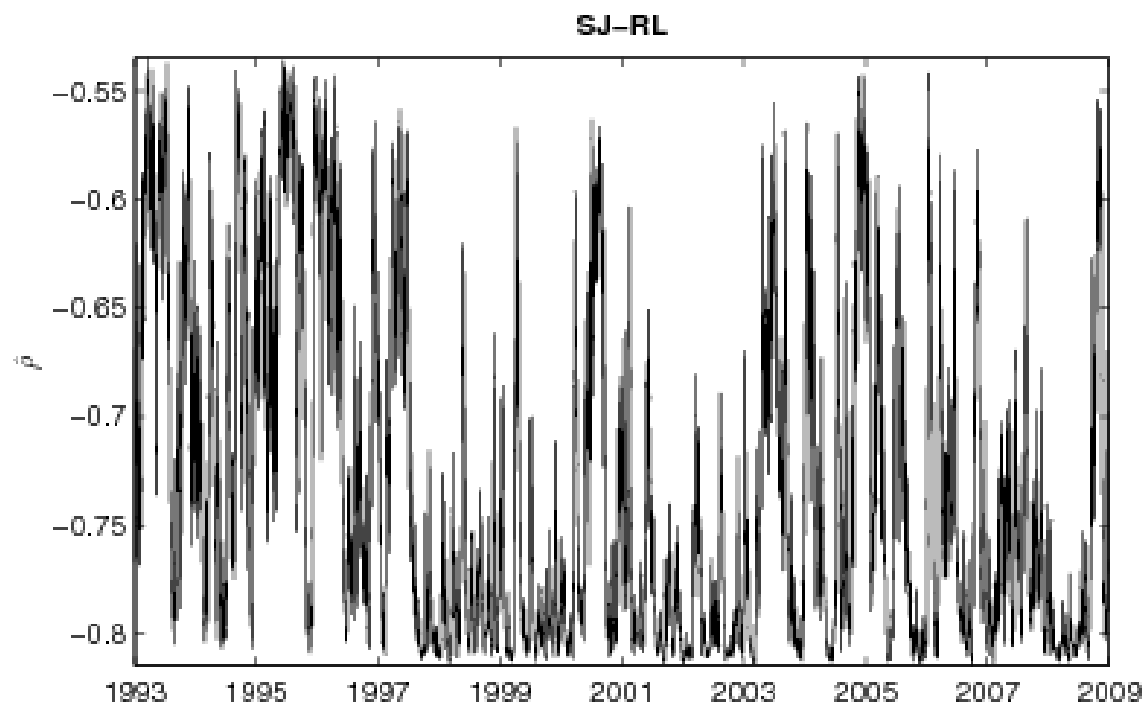
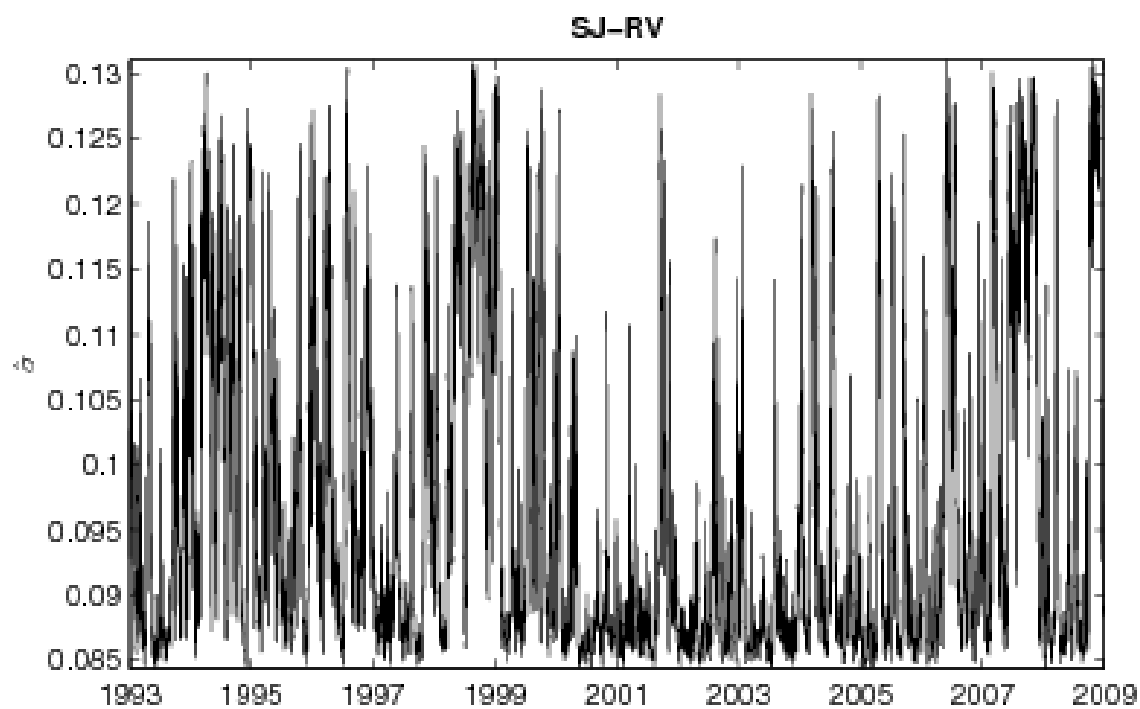


Figure 3: Time series of filtered values of state-dependent parameters.

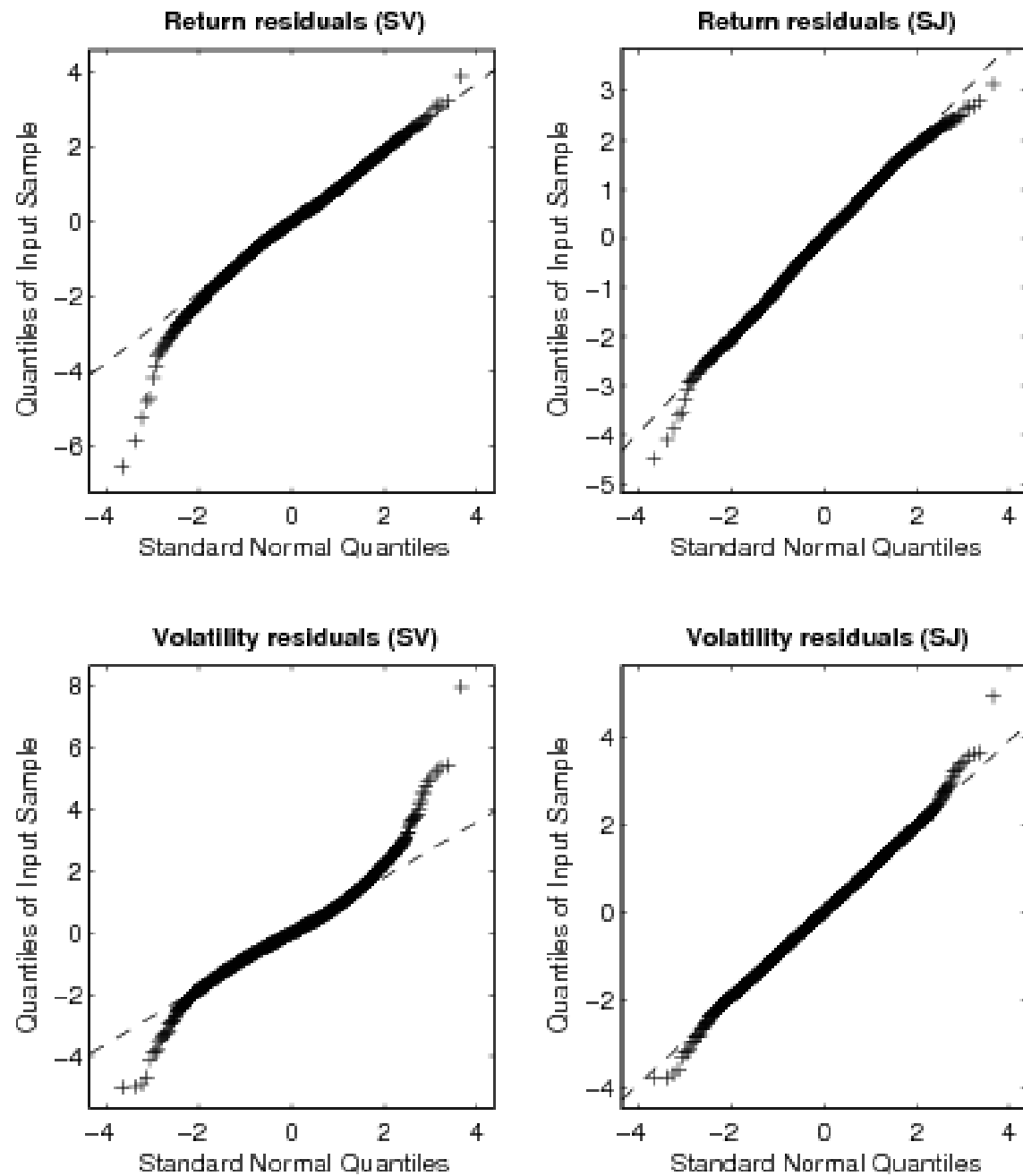


Figure 4: QQ plots for generalized residuals, SV and SJ models.

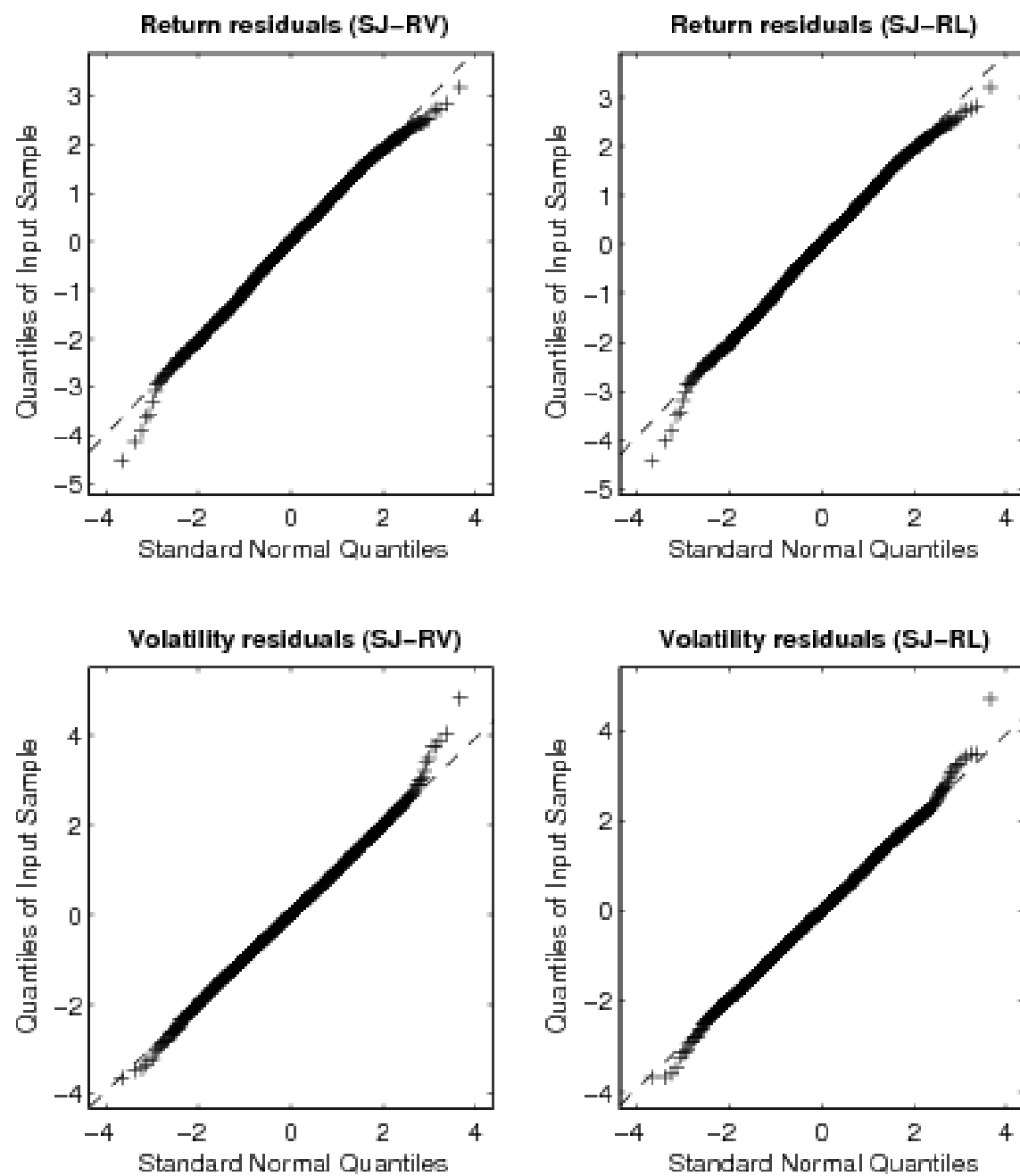


Figure 5: QQ plots for generalized residuals, SJ-RV and SJ-RL models.

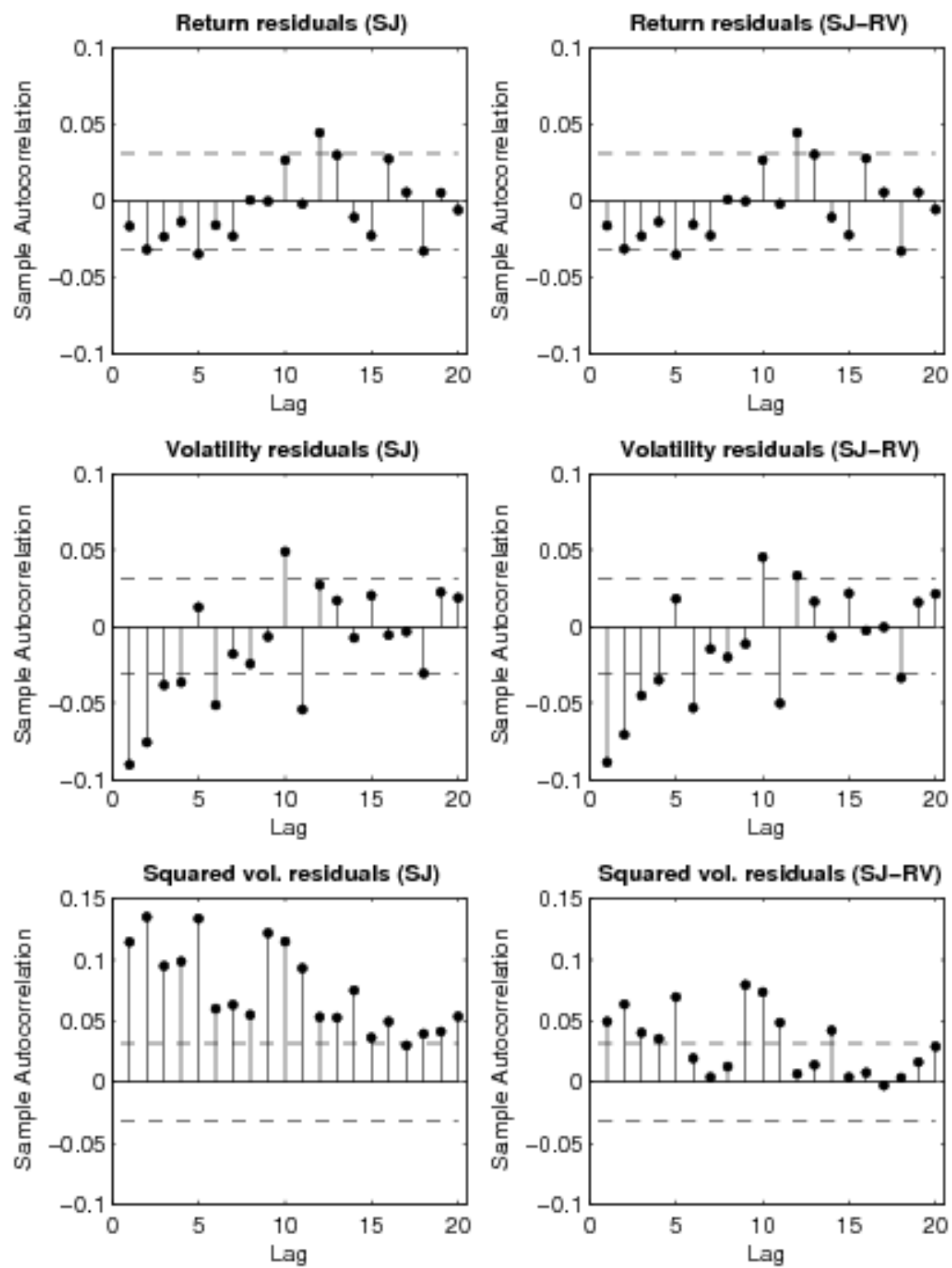


Figure 6: Correlograms for SJ and SJ-RV models. Dotted lines show the 95% confidence band.