Extracting Dynamic Latent Factors from Large Option Panels*

Kris Jacobs Yuguo Liu
University of Houston

January 18, 2018

Abstract

Estimating option valuation models is challenging due to the complexity of the models and the richness of the available option data. Many existing studies therefore limit the time-series dimension and especially the cross-sectional dimension of the option data. This complicates the identification of model parameters from option data, especially the parameters characterizing the tails of the distribution. We propose new techniques to overcome these constraints, based on particle filtering with weights based on model-implied spot volatilities rather than model prices. We also use his approach to estimate option valuation models based on returns and options jointly. We provide an in-depth investigation of the double-jump models with jumps in returns and volatility. Both return and option data support models with jumps in volatility as well as jumps in returns. Using larger cross-sections of options results in model inference and parameter estimates that differ from the existing literature.

JEL Classification: G12

Keywords: Option Valuation; Particle Filter; Risk Premia; Option Panels.

*Correspondence to Kris Jacobs, Bauer College of Business, University of Houston, kjacobs@bauer.uh.edu.
We would like to thank Viktor Todorov and Peter Christoffersen for useful discussions and comments.

1 Introduction

There is widespread agreement in the literature on option valuation that relatively complex models with stochastic volatility, multiple volatility factors as well as jumps and tail factors are needed to explain the data. Given the richness of the available option data, the resulting estimation problems are computationally very intensive. Many existing studies therefore limit the time-series dimension and especially the cross-sectional dimension of the option data. As a result, effectively much of what we know about model parameters implied by option data is based on short-maturity at-the-money options. The use of these data may complicate the identification of model parameters, especially for the parameters characterizing the tails of the distribution.

We propose new techniques to overcome these constraints. Our approach is based on the particle filter. In general, particle filtering is also subject to significant computational constraints when assigning weights based on the model option price, but we overcome this by assigning weights based on model-implied spot volatilities instead. This enables us to estimate and test complex models using very large option data sets. Because particle filtering using returns is relatively straightforward, we are also able to estimate state-of-the-art option pricing models based on returns and large options data sets jointly. This is important to identify model parameters characterizing tail events and risk premia.

We illustrate these methods using the Heston (1993) square root option valuation model and the double-jump model of Duffie, Pan, and Singleton (2000). Our empirical findings contribute to the existing literature on empirical option pricing in several important ways. First, our parameter estimates and inference about model performance significantly differ from the existing literature. For example, when estimating model parameters from options, our estimates of the risk-neutral correlation between returns and variance are much higher than the existing literature, with the notable exception of Andersen, Fusari, and Todorov (2015a). Second, our estimates indicate the existence of jumps in returns and volatility. We confirm the findings of Broadie, Chernov, and Johannes (2007), Eraker (2004) and Eraker, Johannes, and Polson (2003) that jumps in volatility are important; in fact, our findings indicate an even more prominent role for jumps in volatility compared to the existing literature because they strongly improve model performance regardless of whether the model is estimated on returns, options, or both. Jumps in volatility are also more important than jumps in returns.

Finally, our estimation based on returns and large panels of option data results in improve identification of jump risk parameters. In models with return jumps, approximately 2% of the equity risk premium is due to jump risk, with the remaining 6 % due to diffusive risk.

The paper proceeds as follows. Section 2 provides a literature overview and motivates our approach. Section 3 presents the option valuation model and the return and option data used in estimation. Section 4 discusses the estimation of the model dynamics from returns and options based on the particle filter. Section 5 presents the empirical results and Section 6 concludes.

2 Literature Review

Bates (1996b) argues that an important challenge in the option valuation literature is to formulate models that can simultaneously fit returns (under the physical measure) and options (under the risk-neutral measure). Such models are extremely important to learn about risk premia and the structure of the stochastic discount factor.

Substantial progress has been made in meeting this challenge, but much remains to be done. The objective of this paper is to further contribute to this literature. Specifically, one of the main problems in responding to Bates challenge is that the estimation of state-of-the-art option valuation models from returns and options is computationally very demanding, especially when long time periods and large cross-section of options are used.

Ideally, when estimating dynamic models from option data, one wants to use as long a time-series as possible to identify the volatility dynamics, but also to use as much of the cross-section of options on every day in the sample, because this also helps to identify model parameters, especially those parameters characterizing the tails. To illustrate the computational challenge, consider for example estimation using the OptionMetrics data, which we use below and which have been used by many recent studies. The OptionMetrics data start in 1996. Based on data from 1996 to the end of 2015, we have more than 5000 trading days for which option contracts are available. The number of available index option series per day has significantly increased between 1996 and 2015, but on average it is approximately 200. A full sample of option contracts based on daily OptionMetrics data between 1996 and 2015 therefore consists of roughly one million contracts.

Even much smaller option data sets impose significant computational constraints. Existing

studies therefore make trade-offs and effectively reduce the dimension of the data to make estimation computationally feasible. In the time-series dimension, rather than using a short time series, often only one day per week is used. This effectively reduces the size of the data set by 80 percent. In the cross-sectional dimension, the trade-off typically involves using a subset of the available option series. Clearly both trade-offs imply a loss in efficiency.

Consider Table 1, which presents an overview of some important estimation results in the literature.[1] This table focuses on the Heston (1993) stochastic volatility model, which is relatively simple to estimate and for which we therefore have a wealth of existing evidence. We also consider more complex models in our empirical work below. Despite its relative simplicity, the computational constraints that are at the heart of our contribution apply even for this relatively simple model. Panel A of Table 1 indicates a large number of studies that estimate the Heston model under the physical measure based on index return data. In fact, we have limited ourselves to some of the more seminal contributions in this case and many other papers present related results. In contrast, relatively few papers present estimates under the risk-neutral measure based on option data, listed in Panel B, or joint estimates using both return and option data, listed in Panel C. Even those studies typically make important trade-offs due to computational complexity. For example, Aït-Sahalia and Kimmel (2007) estimate the model using return and option data using approximate maximum likelihood. They use daily data for 1990-2004, but on every day they only use the index return and the VIX, which effectively amounts to using one short-maturity at-the-money option every day. Pan (2002) uses implied-state Generalized Method of Moments (GMM) to estimate the model using return and option data for 1989-1996. Pan (2002) parameter estimates are obtained using a very limited cross-section of options, although model fit is also evaluated using a wider cross-section. Eraker (2004) also uses return and option data and uses a fully efficient Markov Chain Monte Carlo (MCMC) technique. The sample is limited to 1987-1990 and contains 3270 call options over 1006 trading days, averaging just over three options per day. Chernov and Ghysels (2000) use Efficient Method of Moments (EMM) for the period 1985-1994, but focus on short-maturity at-the-money calls.

We conclude that the studies listed in Panel C mainly address the computational constraints by reducing the size of the option sample. The studies listed in Panel B also reduce

[1] All parameters in Table 1 are annualized. Some of the papers cited instead report daily parameters or report parameters in percentages. We discuss this in more detail below.

the size of the option sample but often less drastically so. However, they make other trade-offs. Broadie, Chernov, and Johannes (2007) use a long time series to estimate the risk neutral parameters, but fix some parameters at their values implied from returns. Finally, Christoffersen, Jacobs, and Mimouni (2010) use a relatively long sample as well as a relatively sizeable cross-section of option series, but they use only one day per week and they further reduce computational constraints by simplifying the setup of the filtering problem.

In summary, Table 1 illustrates that existing estimates of the Heston model that use option data are heavily impacted by computational constraints, and that they address these computational constraints in various ways. Many studies reduce the cross-sectional dimension of the option data set. The objective of this paper is to introduce a method that can overcome these constraints and keep the time-series and cross-sectional dimension of the option data as large as possible. We achieve this by reducing the computational burden involved in filtering the state states from option data. Keeping the dimension of the option data as large as possible should be helpful in identifying risk premia as well as other model parameters.

Because we reduce the computational burden and are able to estimate option pricing models using large data sets, we can investigate the consequences of limiting the option sample for the estimates in the existing literature. For example, when estimating the Heston model on index returns, the parameter that captures skewness is identified from the robustly negative correlation between returns and variance. When estimating the model using options under the risk neutral measure, the parameter is again identified from the distributions skewness, which means that it is identified from out-of-the-money options. The same remark applies to the parameter that captures kurtosis. It is therefore likely that these parameters are not reliably estimated from at-the-money-options, which are predominant in many of the option samples used in the studies in Table 1. There is substantial variation in the estimates of these parameters in Table 1. These differences may of course be due to differences in the sample periods, but it is also possible that they reflect differences in the cross-section of the option series included in estimation.

The two studies most closely related to ours are Bates (2000) and Andersen, Fusari, and Todorov (2015a). Bates (2000) estimates models from option data by minimizing the sum of squares based on option prices using a large dynamic panel for 1988-1993, but does not use the model dynamic to filter the state variables. Instead, on each day a different spot volatility is estimated as a parameter. This approach can be seen as a generalization of the approach

of Bakshi, Cao, and Chen (1997), who estimate different model parameters on every day in the sample, also estimating the spot volatility as a parameter. Andersen, Fusari, and Todorov (2015a) proceed differently and impose consistency between the risk-neutral variance and a high-frequency volatility estimate from returns. They impose the model dynamic in filtering but the identification of the latent state dynamics is not exclusively based on the option data. Both studies are related to our approach, because they estimate option valuation models using large cross-sections of option prices. Our approach differs because we filter the latent states directly and exclusively from the option data.

3 Model and Data

We first discuss the return dynamic and the option valuation formula. Subsequently we discuss the return and option data.

3.1 The Model

Our empirical work is based on the SVCJ model with contemporaneous jump arrivals in return and variance (Duffie, Pan, and Singleton (2000)):

$$\frac{dS_t}{S_t} = (r_t - \delta_t + \gamma_t - \lambda\bar{\mu}_s)dt + \sqrt{V_t}dZ_t + (e^{J_t^s} - 1)dN_t \tag{1}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t + J_t^v dN_t \tag{2}$$

where S_t is the index level, r_t is the risk-free rate, δ_t is the dividend yield, γ_t is the total risk premium, κ denotes the speed of mean reversion, θ the unconditional mean variance, and σ determines the variance of variance. dZ_t and dW_t are Brownian motions with $corr(dZ_t, dW_t) = \rho$. N_t is a Poisson process with constant jump intensity λ, and J_t^s and J_t^v are the jump size parameters related to returns and variance, with correlation ρ_J. We assume $J_t^v \sim Exp(\mu_v)$ and $J_t^s|J_t^v \sim N(\mu_s + \rho_J J_t^v, \sigma_s^2)$. The term $\lambda\bar{\mu}_s$ is the compensation of jump component, with $\bar{\mu}_s = \frac{e^{(\mu_s+\sigma_s^2/2)}}{1-\rho_J\mu_v} - 1$.

We assume that the risk neutral dynamic is given by:

$$\frac{dS_t}{S_t} = (r_t - \delta_t - \lambda\bar{\mu}_s^Q)dt + \sqrt{V_t}dZ_t^Q + (e^{J_t^{sQ}} - 1)dN_t^Q \tag{3}$$

$$dV_t = \kappa^Q(\theta^Q - V_t)dt + \sigma\sqrt{V_t}dW_t^Q + J_t^{vQ}dN_t^Q \tag{4}$$

where we assume that the variance risk premium is equal to $\eta_v V_t$, and thus $\kappa^Q = \kappa - \eta_v$ and $\theta^Q = (\kappa\theta)/\kappa^Q$. We assume that the jump risk premiums are entirely attributable to the mean jump sizes of return and variance: $\eta_{J^s} = \mu_s - \mu_s^Q$ and $\eta_{J^v} = \mu_v - \mu_v^Q$.[2] The jump intensity λ and the standard deviation of the return jump size σ_s do not change across measures. Therefore, the total equity risk premium can be written as $\gamma_t = \eta_s V_t + \lambda(\bar{\mu}_s - \bar{\mu}_s^Q)$.

This specification nests several models in the existing literature. If we set $\lambda = 0$, it reduces to the SV model of Heston (1993). If we shut down the jump in variance, it becomes the SVJR model of Bates (1996a). It also nests a model with variance jumps only (SVJV) if we shut down the jumps in returns. Note that we do not estimate the more general SVSCJ model studied in Eraker, Johannes, and Polson (2003) and Pan (2002) for example. This model makes the jump intensity a function of volatility. Given the computational burden of estimating the models under consideration using option data, we keep the study of this model for future work.

The model price of a European call option $C^M(V_t, \Theta)$ with maturity τ and strike price K is given by:

$$C^M(V_t, \Theta) = e^{-r_t\tau}E^Q[max(S_T - K, 0)] \tag{5}$$

We use the Fast Fourier Transform (FFT) in Carr and Madan (1999) to compute option prices. All these models are affine and therefore quasi closed-form solutions for option prices are available. See Appendix A for a more detailed discussion. We use the superscript M to denote the model price, as opposed to the value of the option in the data. The model price is computed given the current state V_t and model parameters $\Theta(\kappa, \theta, \sigma, \rho, \lambda_J, ...)$. Note that V_t can be a scalar or a vector. In our application, we need to repeatedly calculate prices of options with different spot variances and the ability to vectorize the formula is computationally important.

[2] We use a simple structure of the jump risk premium because of concerns regarding identification, following Eraker (2004) and Pan (2002), for example. Broadie, Chernov, and Johannes (2007) also investigate more general entertain assumptions regarding the return jumps risk premium, but they use a very different empirical design.

3.2 Return and Option Data

We use S&P500 returns and option prices for the period January 1, 1996 to December 31, 2015, a total of 5031 trading days. The need to use a long sample to identify return dynamics for option valuation is emphasized by Eraker, Johannes, and Polson (2003), Eraker (2004) and Broadie, Chernov, and Johannes (2007), for example. We obtain index returns and risk-free rates from CRSP, and option prices, zero coupon yields, and dividend yields from OptionMetrics. Panel A of Figure 1 plots the time series of the daily returns. The financial crisis is readily apparent, and it is characterized by large negative as well as positive returns. Panel B of Figure 1 plots the squared returns. This figure clearly demonstrates the challenges in modeling the twenty-year sample period. In the financial crisis the variance spikes up, but mean reverts rather quickly. The same observation applies to other periods with large variance spikes. Panel A of Table 2 provides descriptive statistics on the index returns. The returns exhibit negative skewness and excess kurtosis.

We use both put and call index options and impose the following standard filters on the option data:

1. Discard options with fewer than 5 days and more than 365 days to maturity.
2. Discard options with implied volatility less than 5% and greater than 150%.
3. Discard options with volume or open interest less than 5 contracts.
4. Discard options with quotes that suggest data errors. We discard options for which the best bid exceeds the best offer, options with a zero bid price, and options with negative put-call parity implied price.
5. Discard options with price less than 50 cents.

After imposing these filters, the resulting data set contains 945,110 option contracts. The left column of Panel B of Table 2 indicates that the sample contains more puts than calls, as expected. Columns 2-4 provide sample sizes for successively smaller samples by eliminating options that are further out-of-the money. Moneyness is defined as strike price divided by index price (K/S).

The resulting option data set is not balanced over time. This imbalance is substantial: we have almost eight times more options in 2015 than in 1996. In principle, this is not a problem, but the results from a more balanced data set are easier to interpret. The resulting sample is also smaller, which provides computational advantages. We therefore create a more balanced

panel. We use six moneyness bins and five maturity bins. For each moneyness-maturity bin, we include only the most liquid option. The data set thus has thirty options per day unless options are not available for certain bins. This procedure yields a data set with 129,182 option contracts. Panel C of Table 2 provides sample sizes for these moneyness/maturity bins. Note that the data are still unbalanced because in the early years of the sample we do not have many observations for short maturities and moneyness larger than 1.10. Our analysis is based on this balanced data set. Panel D reports the average option price for these moneyness-maturity bins.

4 Estimation

We first discuss how the particle filter can be used to estimate models with latent states, and stochastic volatility models with jumps in particular. We then briefly discuss how this framework can be applied to estimate these models using return data. We discuss the computational problems that arise when applying these models to option data and how to address these problems. Finally we outline how to combine return and option data in estimation.

4.1 Estimation Framework and Notation

We first discuss the estimation method in general, because the algorithm is conceptually similar when different data sources are used. The implementation of teh algorithm however differs dependent on what is observable: returns, options, or both. We first need to time-discretize the continuous-time model. Several discretization methods are available and every scheme has certain advantages and drawbacks. We use the Euler scheme, which is easy to implement and has been found to work well for this type of applications (Eraker, 2004). Applying Ito's lemma and discretizing (1)-(2) gives

$$R_{t+1} = ln(\frac{S_{t+1}}{S_t}) = r_t - \delta_t - V_t/2 + \gamma_t - \lambda\bar{\mu}_s + \sqrt{V_t}z_{t+1} + J^s_{t+1}B_{t+1} \tag{6}$$

$$V_{t+1} - V_t = \kappa(\theta - V_t) + \sigma\sqrt{V_t}w_{t+1} + J^v_{t+1}B_{t+1} \tag{7}$$

where z_{t+1} and w_{t+1} are distributed standard normal. The discrete jump frequency B_{t+1} follows the Bernoulli distribution. For each time period, there is either no jump or one

jump. The corresponding discretized risk-neutral dynamics are identical but use the risk-neutral parameters. We implement the discretized model using daily returns, but we report annualized parameter estimates below.

We assume that observed option prices are equal to the model price plus error:

$$C_{t,h} = C_t^M(L_t|\Theta) + \epsilon_{t,h} \tag{8}$$

where $h = 1, 2, .., H_t$ and H_t is the total number of options at date t. We assume $\epsilon_{t,h}$ is normally distributed and $\epsilon_{t,h} \sim N(0, (\sigma_t^c)^2)$.

It is helpful to formulate these dynamics in a state-space representation. Denote L_{t+1} as the latent states that are used to generate the observables O_{t+1}. Based on the discretization, $L_{t+1} = (V_t, B_{t+1}, J_{t+1}^s, J_{t+1}^v)$ and $O_{t+1} = (\{C_t\}, R_{t+1})$.[3] Define the measurement density by $f_1(O_{t+1}|L_{t+1})$ and the transition densities by $f_2(L_{t+1}|L_t)$. The latent states evolve through the transition density function, while the observables are realizations conditional on the latent states and the measurement density. The state-space representation applies regardless of whether we observe returns, options, or both. When returns are observable, f_1 refers to equation (6); when options are observable, f_1 is given by equation (8).[4] For the persistent latent variance, V_t, f_2 represents equation (7) and for the non-persistent jump variable, f_2 is simply a random draw from the corresponding distribution.

The discretized dynamics can therefore be described as follows:

$$\begin{array}{ccccccc} (L_t) & \xrightarrow{f_2} & (L_{t+1}) & \xrightarrow{f_2} & (L_{t+2}) & \xrightarrow{f_2} & .... \\ \downarrow f_1 & . & \downarrow f_1 & . & \downarrow f_1 & . & . \\ O_t(\{C_{t-1}\}, R_t) & . & O_{t+1}(\{C_t\}, R_{t+1}) & . & O_{t+2}(\{C_{t+1}\}, R_{t+2}) & . & .... \end{array} \tag{9}$$

where $\{C_t\} = (C_{t,1}, ..., C_{t,H_t})$. Although we do not directly observe the latent states $L_t, L_{t+1}, ...$, we do observe the option prices and/or returns in each period.

We now discuss the estimation method in general, which can apply to the case where we observe returns, options, or both. We have two sets of unknowns: 1) parameters $\Theta(\kappa, \theta, \sigma, \rho, \lambda_J, ...)$

[3]For example: V_t generates the cross-section of option prices $\{C_t\}$; and V_t, together with B_{t+1} and J_{t+1}^s, produces the next period return R_{t+1}.

[4]Since the option price depends only on the current spot variance and the risk-neutral expected jumps rather than the realized jumps, $f_1(\{C_t\}|L_t)$ reduces to $f_1(\{C_t\}|V_t)$, which is equation (8).

and 2) latent states $L_{t+1}(V_t, B_{t+1}, J^s_{t+1}, J^v_{t+1})$. We use particle filtering to filter the latent states and adaptive Metropolis-Hastings sampling to perform the parameter search.

4.2 Particle Filtering

For now, think of the parameters Θ as given. A standard sampling-importance resampling (SIR) particle filter can be implemented at each time t using the following steps:

Step 1. Using the time t resampled particles L^i_t, $i = 1 : N$, where N is the total number of particles, for each particle i simulate $\tilde{L}^i_{t+1}$ from L^i_t according to f_2.

Step 2. Compute the weight for each particle:

$$\omega^i_{t+1} = f_1(O_{t+1}|L^i_{t+1}) \tag{10}$$

$$\pi^i_{t+1} = \omega^i_{t+1} / \sum_{j=1}^{N} \omega^j_{t+1} \tag{11}$$

Step 3. Resample the particles $\tilde{L}^i_{t+1}$ according to the normalized weights $\{\pi^i_{t+1}\}$, which gives L^i_{t+1}, $i = 1 : N$. Now go back to Step 1.

We can think of the weights π^i_{t+1} as constituting a discrete probability distribution for L_{t+1}. After resampling, the weight for each particle changes back to $1/N$. The SIR is extremely intuitive and simple to implement. However, since new particles are simulated blindly in step 1, it may lead to the well known sample impoverishment problem (Johannes, Polson, and Stroud (2009)), especially when N is not very large. Consider a scenario where we have an extremely large negative return at $t+1$. Particles with large V_t or large negative return jump occurrence will receive large weights, while other particles will be assigned weights close to zero. For randomly propagated particles, very few particles can be generated with large V_t or jump occurrence. As a result, the resampling might consist of repeated values of these few particles.

Pitt and Shephard (1999) introduce the Auxiliary Particle Filter (APF) to solve this problem by resampling before the propagation step. The steps of the APF can be described as follows:

Step 1. Compute the first stage weights based on the predictive likelihood for each particle L_t^i. We break up the computation in equation (10) in the SIR into two steps. We need to first calculate the time t expected values for each particle.

$$\hat{L}_{t+1}^i = E(f_2(L_{t+1}|L_t^i)) \tag{12}$$

Then we compute the weights evaluated at this conditional expectation

$$_1\omega_{t+1}^i = f_1(O_{t+1}|\hat{L}_{t+1}^i) \tag{13}$$

Finally we normalize of the weights, similar to what we do in the SIR.

$$_1\pi_{t+1}^i = {}_1\omega_{t+1}^i / \sum_{j=1}^{N} {}_1\omega_{t+1}^j \tag{14}$$

Step 2. Resample L_t^i according to the weights $\{_1\pi_{t+1}^i\}$, which gives $\breve{L}_t^i$.
Step 3. Simulate $L_{t+1}^i = f_2(L_{t+1}|\breve{L}_t^i, O_{t+1})$
Step 4. Compute the second stage weight for each particle:

$$_2\omega_{t+1}^i = f_1(O_{t+1}|L_{t+1}^i)/f_1(O_{t+1}|\hat{L}_{t+1}^i) \tag{15}$$

$$_2\pi_{t+1}^i = {}_2\omega_{t+1}^i / \sum_{j=1}^{N} {}_2\omega_{t+1}^j \tag{16}$$

Step 5. (Optional) Another resampling according to the weights $\{_2\pi_{t+1}^i\}$.

The APF differs from the standard SIR in two ways: 1) the APF resamples before simulating new latent states according to the predictive likelihood; 2) the APF takes new observations into account to generate new latent states. After resampling in Step 2, the variance particles are tilted toward those that more likely to have generated O_{t+1}. In Step 3, we first fix the variance $\hat{V}_t^i$ as well as the jump size $J_{t+1}^s = \mu_s$, and then simulate the jump occurrence B_{t+1}^i according to the likelihood of observing new observations: $p(B_{t+1}^i = 1) = f_1(O_{t+1}|\hat{V}_t^i, B_{t+1}^i = 1)$ and $p(B_{t+1}^i = 0) = f_1(O_{t+1}|\hat{V}_t^i, B_{t+1}^i = 0)$. Then we simulate the jump size conditional on B_{t+1}^i, $\hat{V}_t^i$ and O_{t+1}, where $\hat{V}_t^i$ is the conditional expected variance. Finally, we propagate V_t^i

conditional on everything else.

The literature contains alternative implementations of the APF, however, they all share the same basic approach: simulating particles by taking into account new observations.

4.3 Adaptive Metropolis-Hasting Sampling

The Metropolis-Hastings algorithm (MH) originally based on Metropolis et al. (1953) and Hastings (1970) is widely used in the existing literature due to its flexibility, especially when confronted with high-dimensional distributions. In general, MH randomly samples parameters from a proposed distribution $q_j(\Theta_j^p|\Theta_{j-1})$. Subsequently the algorithm moves to the new parameter set Θ_j^p with probability:

$$\alpha(\Theta_j^p, \Theta_{j-1}) = min(1, \frac{f_1(O_{1:T}|\Theta_j^p)p(\Theta_j^p)q_j(\Theta_j^p|\Theta_{j-1})}{f_1(O_{1:T}|\Theta_{j-1})p(\Theta_{j-1})q_j(\Theta_{j-1}|\Theta_j^p)}) \tag{17}$$

where f_1 denotes the (simulated) likelihood and $p(\Theta)$ is the prior for Θ, which can be uninformative.

The general idea is to search for optimal parameters by moving to a new parameter set with probability one if it generates a higher likelihood than the previous parameter set. To avoid getting stuck at local optima, the algorithm moves to new parameter set with a non-zero probability even if the new likelihood is lower.

Tuning of associated proposal variances of q_j is crucial to achieve estimation efficiency. Roberts and Rosenthal (2009) propose the adaptive random walk scheme, where q_j is estimated from previous iterations $(\Theta_1, ..., \Theta_{j-1})$ after every certain amount of iterations. For example, we start from some pre-specified variance for each dimension (parameter) of $q_j, j = 1 : 1000$ and run 1000 iterations. Then, we may adjust the new variance of $q_j, j = 1001 : 2000$ to be the variance calculated from the first 1000 iterations, and so on. Note that if q_j is set to be a fixed distribution q, it reduces to the basic MH algorithm.

According to Pitt (2002), the total likelihood of particle filtering conditional on a set of parameters can be expressed as a function of the unnormalized weights:

$$f_1(O_{1:T}|\Theta) = \sum_{t=1}^{T} \left\{ 1/N \sum_{j=1}^{N} {}_1\omega_{t+1}^j \right\} \left\{ 1/N \sum_{j=1}^{N} {}_2\omega_{t+1}^j \right\} \tag{18}$$

As a special case, in SIR we only have first stage weights while the second stage weights are $1/N$ for each particle. The total likelihood therefore reduces to:

$$f_1(O_{1:T}|\Theta) = \sum_{t=1}^{T} \left\{ 1/N \sum_{j=1}^{N} \omega_{t+1}^{j} \right\} \tag{19}$$

The total likelihood is simply a by-product of the particle filtering. Generally, in particle filtering we need to keep all the historical information (the ancestors) of every particular particle at time t. This becomes increasingly time consuming ($O(T^2)$) as T increases, especially when we have a large N. However, since the likelihood function (19) does not require us to record the latent variable path, we may simply search parameters first without recording latent processes, and then back out latent states with finalized parameters. In our application, we noticed that although the APF performed better for filtering latent states given a set of parameters, it performed worse than the SIR in parameter search due to the more volatile likelihood. Therefore, we apply SIR for the parameter search and then filter the latent states using APF. In our implementation, we use 10,000 particles.

4.4 Estimation Using Returns Data

As mentioned above, the algorithm can be applied to different sources of information, which corresponds to different likelihoods $f_1(O_{t+1}|L_{t+1}^i)$. First consider estimation based on returns, where we simply use returns as the observables. Using equations (6) and (7) this gives:

$$\begin{aligned} f_1(O_{t+1}|L_{t+1}^i) &= f_1(R_{t+1}|L_{t+1}^i) \\ &= \frac{1}{\sqrt{2\pi V_t^i}} \exp\left\{ -\frac{1}{2} \frac{\left[R_{t+1} - (r - \delta_t - \frac{1}{2}V_t^i + \eta_s V_t^i - \lambda\bar{\mu}_s + J_{t+1}^s)\right]^2}{V_t^i} \right\} \end{aligned} \tag{20}$$

Our implementation of the particle filter uses 10,000 particles and 10,000 iterations. We set the first one fourth of the iterations as burn-in, and report the posterior mean and standard deviation for each parameter from the subsequent iterations. We provide additional details on the resulting particle filtering in the Appendix. See also Christoffersen, Jacobs, and Mimouni (2010) for a related implementation on the SV model.

4.5 Estimation Using Option Data

We now consider model estimation using option data only, that is, estimation that does not consider the underlying returns. The estimation of option pricing models using both returns and options is very important; indeed, it is one of the main objectives in the literature because it allows for the estimation of risk premia, as emphasized for example by Bates (1996a) and discussed in Section 2. However, our motivation is that prior to estimating the model using both returns and options, it is important to get a clear idea of the wedge in the data between the information under the physical measure (returns) and the information under the risk-neutral measure (options). This is an essential task prior to a meaningful analysis and discussion of the results of joint estimation.

The existing literature that estimates option pricing models using option data deals with latent states such as the spot variance broadly in two ways. The first approach is to extract the state variables from return data, either by filtering from daily returns or by calibrating from intra-day data. See for example Broadie, Chernov, and Johannes (2007) and Christoffersen, Jacobs, and Mimouni (2010). The other approach is to treat the spot variance as a parameter to be estimated along with other parameters (Bates, 2000). Both these estimation approaches are very useful, but it is important to realize that the resulting parameter estimates will either reflect P-measure information or ignore the dynamic of the spot variance. We now discuss an alternative approach that uses the particle filter to estimate model parameters exclusively based on option data.

4.5.1 Computational Constraints

The instantaneous variance follows the transition equation (7), but now the observables consist of a cross-section of H_t option prices for each day, denoted by $\{C_t\}$. The filtering problem therefore consists of evaluating the likelihood of observing the market option prices conditional on the latent states.

Conceptually this filtering procedure is just as straightforward as the one using returns, however, it encounters significant computational constraints. The measurement density now corresponds to equation (8). Rather than one return for each time period t, we now have hundreds of option prices available at time t. The likelihood for a single particle at time t can be calculated as:

$$
\begin{aligned}
f_1(O_{t+1}|L^i_{t+1}) &= \prod_{h=1}^{H_t} f_1(C_{t,h}|C^M_{t,h}(V^i_t|\Theta)) \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma^c_t}\right)^{H_t} exp\left(-\frac{\sum_{h=1}^{H_t}(C_{t,h}-C^M_{t,h}(V^i_t|\Theta))^2}{2(\sigma^c_t)^2}\right)
\end{aligned}
\tag{21}
$$

since the option price only depends on the spot variance, not jump realization, we can write V^i_t for L^i_{t+1}. The total likelihood for the entire sample summing over all particles is:

$$
f_1(O_{1:T}|\Theta) = \sum_{t=1}^{T}\left\{1/N\sum_{i=1}^{N}\left[\left(\frac{1}{\sqrt{2\pi}\sigma^c_t}\right)^{H_t} exp\left(-\frac{\sum_{h=1}^{H_t}(C_{t,h}-C^M_{t,h}(V^i_t|\Theta))^2}{2(\sigma^c_t)^2}\right)\right]\right\} \tag{22}
$$

When computing this likelihood, although a quasi closed-form solution for the option price is available in the affine models we consider, and each option price takes less than 0.01 seconds to evaluate, for each function evaluation we have to evaluate option prices along three dimensions: for each option, for each particle, and for each day. Our sample period consists of 5031 trading days and 200 options on average per day, and 10,000 particles leads to approximately 10,000,000,000 computations of the option price in each function evaluation, which is computationally infeasible.

4.5.2 The Implied Spot Variance Method

Given this computational burden, we propose a more efficient filtering algorithm based on the implied spot variance. The motivating idea can loosely be thought of as reducing the three-dimensional option evaluation computation into a (pseudo) two-dimensional computation.

According to equation (21), a particle's likelihood is inversely proportional to the sum of squared pricing errors (SSE) for that particle. Define the implied spot variance (ISV) on day t as the spot variance that results in the smallest SSE:

$$
ISV_t = \arg\min_{ISV_t} \sum_{h=1}^{H_t}(C_{t,h}-C^M_{t,h}(ISV_t|\Theta))^2 \tag{23}
$$

To illustrate how this sum of squared errors changes with particle values, we select a typical day in our sample, December 1, 2015. Figure 2 scatter plots the SSE against the particle

values using the following parameter values for the Heston model: $\kappa = 3$, $\theta = 0.25$, $\rho = -0.7$, and $\sigma = 0.4$. We have repeated this analysis for other days, models and parameter values and the conclusions are very similar.

By definition, the ISV is the lowest point on the parabola. The further a particle is away from ISV, the larger the SSE. Thus, we can fit this parabola and compute the SSE for any particle according to its distance to ISV. This simplifies an $O(T*N*H)$ problem to roughly an $O(T*H)$ problem, where $H = \frac{1}{T}\sum_{t=1}^{T} H_t$. In reality the reduction in computational complexity is less spectacular because we need an extra ISV search step. Fortunately, this step is extremely fast. For any given parameters, the ISV search step is performed by a two stage grid search: in the first stage, we divide the domain for the variance into 20 grids, and evaluate the SSE for each grid point; in the second stage, we divide the space on both sides of the lowest SSE grid point into another 20 grids, and we evaluate the SSE once more. The ISV is given by the grid point with the lowest SSE and we record the corresponding σ_t^c. The accuracy of the ISV search may of course be affected by the discreteness of the grid, but we can increase the number of grid points to improve accuracy. We have extensively experimented with this. We have also compared our approach to the gradient search method and found that 20 grids provide a satisfactory approximation. We then fit a parabola using the 20 grids in the second stage: $SSE_i = a_1 + a_2 * (V^i - ISV)^2$.

The likelihood for particle i can then be calculated as follows:

$$f_1(O_{t+1}|L_{t+1}^i) = \left(\frac{1}{\sqrt{2\pi}\sigma_t^c}\right)^{H_t} exp(-\frac{SSE_{ISV}(\hat{b}_t * (V_t^i - ISV_t)^2 + 1)}{2(\sigma_t^c)^2}) \tag{24}$$

$$f_1(O_{1:T}|\Theta) = \sum_{t=1}^{T}\left\{1/N\sum_{i=1}^{N}\left[\left(\frac{1}{\sqrt{2\pi}\sigma_t^c}\right)^{H_t} exp(-\frac{SSE_{ISV}(\hat{b}_t * (V_t^i - ISV_t)^2 + 1)}{2(\sigma_t^c)^2})\right]\right\} \tag{25}$$

where we use $SSE_{ISV} = a_1$ and $\frac{SSE_i}{SSE_{ISV}} = 1 + b(V^i - ISV)^2$ with $b = \frac{a_2}{a_1}$. Rather than computing option prices for $N = 10,000$ particles, we now only need to compute 40 grids in the ISV search step and then for each particle we calculate its likelihood based on its distance to ISV. In other words, we get rid of one dimension (the number of particles N) when computing option prices, thus saving more than 99% of computation time. Moreover, the ISV search step is sequentially independent, and therefore we can parallel the computation

on $t = 1 : T$. Note also that while the likelihood in equations (24) and (25) can be written in terms of the latent variable V_t, the filtered variance jumps are required in order to obtain the filtered variance path.

4.6 Joint Estimation Using Return and Option Data

In order to estimate the models using both sources of information, we combine the likelihood from returns and options. Because the likelihood on day t is comprised on a single return but a large cross-section of options, the options are weighted much more heavily if we simply sum up the likelihoods. These weights are important as they affect both the likelihood of a particle in the filtering step, as well as the total likelihood of a parameter set in the parameter search. In our implementation, we impose equal relative weights for returns and options. The resulting total likelihood for each particle on date t can therefore be expressed as:

$$f_1(O_{t+1}|L^i_{t+1}) = f_1(R_{t+1}|L^i_{t+1}) * (\prod_{h=1}^{H_t} f_1(C_{t,h}|C^M_{t,h}(V^i_t|\Theta)))^{1/H_t} \tag{26}$$

where the two components can be computed according to equations (20) and (21) respectively. The relative weights are equal in the sense that we constrain the information from returns to be equally important to the information from options, no matter how many options we have available on a given day. Given this likelihood function for each particle, the total likelihood can be calculated accordingly.

5 Empirical Results

We first present estimates obtained using option data only, using the approach from Section 4.5. We contrast these estimates with estimates from the existing literature, but also with return-based estimates as well as estimates obtained using options and returns jointly. We discuss the results on a model-by-model basis. We start with the simplest model, the Heston (1993) stochastic volatility model (SV). We then discuss the model with return jumps (SVJR), the model with variance jumps (SVJV), and the model with joint return and variance jumps (SVCJ).

5.1 Estimates of the SV Model

The first column of Table 3 presents estimates of the SV model from options, obtained using the estimation procedure in Section 4.5. These estimates can be compared with estimates in the existing literature in Panel B of Table 1, with the caveat that the estimates in Table 1 are obtained using different sample periods. Note that to facilitate comparisons, the estimates in Table 1 are all reported in annual units. This is similar to Pan (2002) but different from other studies such as Eraker (2004) and Broadie, Chernov, and Johannes (2007), for example.[5]

Estimates of θ of course depend on the sample period, but the estimate in Table 3 is well within the range of existing estimates in the literature in Panel B of Table 1. The estimate of κ in Table 3 is 1.916. It is smaller than the estimates in Broadie, Chernov, and Johannes (2007) and Christoffersen, Jacobs, and Mimouni (2010). It is similar to the estimate in Bates (2000). Recall that the estimates Bates (2000) are obtained using option data only. The estimates in Christoffersen, Jacobs, and Mimouni (2010) use an option-based objective function but the variance is filtered from returns. In Broadie, Chernov, and Johannes (2007), estimates are based on options but several parameters are constrained by returns-based estimates. We therefore conclude that when we exclusively use options in estimation, the variance process is estimated to be much more persistent.

The most striking difference between the estimates in Table 3 and the results from the existing literature in Panel B of Table 1 is that the estimate of the correlation between the return and variance innovations is -0.933, much more negative than in existing studies. This finding indicates that risk-neutral skewness may be even larger than commonly thought. However, Andersen, Fusari, and Todorov (2015a) report an estimate of ρ of -0.934 when estimating the SVCJ model. Therefore, once again our results are more consistent with the estimates from a study that largely relies on option-based information.

The estimate of σ is equal to 0.4091. This is small compared to the estimates in Panel B of Table 1. Note that in Broadie, Chernov, and Johannes (2007), the estimate of σ is simply constrained to be the return-based estimate. Finally, the estimate of η_v, the parameter that characterizes the diffusive variance risk premium, is equal to 0.801. This estimate has the

[5]Compared to the estimates reported in Eraker (2004) and Broadie, Chernov, and Johannes (2007), the estimates of κ and σ in Table 1 are multiplied by 2.52 (multiplied by 252 and divided by 100). The estimate of θ in Table 1 is multiplied by 0.0252 (multiplied by 252 and divided by 10,000). The estimate of ρ is the same.

expected sign and it is highly statistically significant.

Why are these estimates that are exclusively based on information from options so interesting? As highlighted before, the most important challenge in the option pricing literature is to build and estimate models that can provide a good fit to returns and options simultaneously. We agree with this and we take up this challenge below, but we believe that it is extremely important to first fully characterize how model estimates obtained from options differ from those obtained from returns, in order to characterize the challenge at hand and provide guidance on the pricing kernel need to bridge the gap between the two measures. Our results in Table 3 suggest that when exclusively using options in estimation, estimates from the simple SV models are different from most existing estimates in the literature, which are to some extent affected by return-based information.

We now turn to a characterization of the wedge between the physical and risk-neutral measure in our sample, by comparing the risk-neutral estimates for the SV model in the first column of Table 3 with return-based estimates for the same sample period. These results for the SV model are in the first column of Table 5. The mean reversion parameter κ is much larger when estimated from returns. It is 6.969 in Table 5, compared with 1.915 in Table 3. Note that many studies directly compare the physical mean reversion from returns with teh risk-neutral mean reversion from options. The risk-neutral mean reversion parameter $\kappa^Q = \kappa - \eta_v$ from Table 3 is equal to 1.915-0.801=1.114, much smaller than the physical mean reversion of 6.961 estimated from returns in Table 4. This findings is consistent with the estimates in the existing literature in Table 1: the risk-neutral variance is more persistent than the physical variance.

The estimates of θ in Tables 3 and 5 are also consistent with the existing literature. The long-run physical variance of returns θ in Table 5 is estimated at 0.0359, close to the sample average of the return variance of 0.0384. The estimate of θ estimated from the option data in Table 3 is equal to 0.0396, which gives a risk-neutral long-run variance of $\theta^Q = (\kappa\theta)/\kappa^Q$= 0.0681. This confirms the finding in the literature that the risk-neutral variance exceeds the physical variance, consistent with the evidence from SV models in Table 1, as well as with nonparametric evidence, see for example Bollerslev, Tauchen, and Zhou (2009).

Figure 3 compares the filtered variance paths for the return-based estimates and the option-based estimates. Not only is the risk-neutral variance is on average higher than the physical variance, it exceeds the physical variance for most days in the sample. This is again consistent

with existing nonparametric evidence.

The SV estimate of ρ from returns in Table 5 is -0.796. This is a larger magnitude than existing return-based estimates of ρ in Panel A of Table 1. This finding may be partly due to our sample period, which unlike most of the studies cited in Table 1 includes the financial crisis. Despite this finding, note that this estimate of ρ is much smaller in absolute value than the estimate based on options in Table 3.

The return-based estimate of σ in Table 5 is 0.543, which is higher than the option-based estimate of 0.409 in Table 4. This difference is apparent from the filtered variance paths in Figure 3. Both estimates are well within the range of existing estimates in Table 3, so they do not seem surprising. However, the finding that the return-based σ exceeds the option-implied σ is unusual. For instance, from Table 1, Eraker's estimate of σ from returns is 0.277, much lower than the 0.554 estimate from returns and options in Panel C of Tables 3. Christoffersen, Jacobs, and Mimouni (2010) also find that the option-based σ exceeds the return-based σ. Our option-implied estimate of σ may therefore be surprising, but as with estimates of ρ, other studies do not offer a direct comparison of return-based and option-based estimates using the same sample period, so it is not straightforward to judge if the differences in the estimates of σ in Tables 3 and 5 are consistent with the existing literature.

We conclude that the most striking finding for the option-based and return-based estimates for the SV model in Tables 3 and 5 is the very high (in absolute value) estimate of ρ implied by the option data. This parameter determines the risk-neutral skewness, which seems to be more negative than implied by existing studies, with the notable exception of Andersen, Fusari, and Todorov (2015a). Despite an estimate of the physical skewness estimated from returns in Table 5 which is more negative than implied by existing studies, we confirm that risk-neutral skewness exceeds (in absolute value) physical skewness, consistent with economic intuition.

A potential explanation for why our results on skewness differ from the existing literature is that our option-based estimates of ρ are based on filtering from options data, as opposed to filtering in existing studies that is (partly) based on returns. This results in a relatively low option-implied estimate of σ as compared to the return-based estimates, which in turn affects the estimate of ρ.

Panel A of Table 4 reports on the dollar RMSE for the SV model. The overall RMSE is \$2.91. Table 2 shows that the average price of the options in our sample is \$67.66.

To assess the reliability of the adaptive Metropolis-Hastings procedure, Figure 3 plots the trace for the return-based estimation of the SV model. We use 10,000 iterations and the first 2,500 are treated as burn-in. The figure clearly indicates convergence towards the estimated parameters and toward the end of the iterations the fluctuations in the parameters are rather small.

5.2 Estimates of the SVJR Model

We now present our estimates of the stochastic volatility model with return jumps (SVJR) model. We highlight the differences between option-based and return-based estimates. We also comment on differences between our results and those in the existing literature. Column 2 of Table 3 presents our parameter estimates from options. Table 6 summarizes parameter estimates for this model from the existing literature. All parameters are annualized in both tables. Once again, parameters are often expressed differently in the papers referenced in these tables. For the SV parameters, see the discussion in the previous subsection. For the jump parameters, many studies express the mean and standard deviation of the jump in percentages, which means they are equal to the parameters in Table 6 multiplied by 100. For the jump intensity, some papers, such as Pan (2002) express it as in Table 6; others, such as Eraker, Johannes, and Polson (2003) express the intensity in daily units and need to be multiplied by 252 to obtain the estimates in Table 6.

First consider the SV parameters κ, θ, ρ, and σ. Note that jumps in returns capture higher moments in the return distribution, and it would therefore not be surprising to see changes in the estimates of the SV parameters ρ and σ. However, our findings from the SV model in Section 5.1 are largely confirmed for these parameters,.

The estimates of the jump parameters based on option data in Table 3 and return data in Table 5 are intuitively plausible. The risk neutral average jump size $\mu_s^Q = \mu_s - \eta_{J^s}$ in Table 3 is equal to -0.0376, whereas the physical jump size μ_s in Table 5 is equal to -0.0132. The jump risk premium is therefore equal to $0.0132 - 0.0376 = -0.0244$. These jumps occur on average less than once per year (λ=0.885 from options and λ=0.971 from returns). Our estimates therefore indicate the presence of relatively large, but infrequent jumps, which are negative on average, with risk-neutral average jump sizes that are larger than physical jump sizes. The standard deviation of the jumps σ_s is 4.92% from options and 2.00% from returns.

For comparison, Table 6 summarizes existing estimates of the SVJR model. The literature contains several estimation results for this model, but in most cases they use different information compared to our approach. First, as is the case for existing estimates of the SV model in Table 1, most of the existing estimates are obtained from returns. Panel A of Table 6 indicates that when estimating based on returns, some studies find jumps that occur more frequently, while others document larger but more infrequent jumps. Our results are closer to the latter group of studies. Consistent with our results, Eraker, Johannes, and Polson (2003) finds that the SVJR model leads to smaller return-based estimates of σ and κ compared to the SV model.

Relatively few studies offer evidence based on options (Panel B) or options and returns (Panel C). Pan (2000), using a method of moments technique based on options and returns data, finds evidence for relatively frequent jumps with a large risk-neutral mean and large standard deviation. Eraker finds a risk-neutral average jump size of -5% with a standard deviation of 16.7%. These jumps occur on average once every two years. Our results are close to those of Eraker (2004), but our estimated standard deviation of the jump size is smaller. Finally, Broadie, Chernov, and Johannes (2007) proceed somewhat differently. They estimate the jump parameters from (futures) options but keep the SV parameter constrained by theory at their values estimated from returns. Despite these differences in implementation, our results are also rather similar to those of Broadie, Chernov, and Johannes (2007).

We conclude that our estimates of the physical and risk-neutral dynamics of the SVJR model are intuitively plausible and in line with some of the existing estimates in the literature. We find evidence of relatively large but infrequent jumps. The risk premium, which is entirely attributed to the average size of the jump, is large in comparison to the physical estimate of the mean.

A comparison of the log likelihoods for the SV and SVJR models in Tables 5 and 6 indicates whether jumps in returns lead to statistically significant improvements in fit. For the return-based estimation, the improvement in the log likelihood is 16130-16123=7, which is statistically significant with a p-value of 0.0719 given the use of three extra parameters. For the option-based estimation in Table 5 the improvement in the log likelihood is 15459-15398 = 61, which is highly statistically significant with a p-value less than 0.0001. Improvements in likelihood are typically much larger for successful models when modeling options, which is due to the richness of the option data. Overall, we conclude that the existence of jumps in

returns is supported by these statistical tests.

Panel B of Table 4 reports the overall RMSE for the SVJR model as well as RMSEs by moneyness and maturity. While the option-based estimates of the SVJR model are strongly supported statistically over the SV model, it does not lead to large improvements in fit ($.2.86 for the SVJR model versus $2.91 for the SV model). This finding confirms the results of Eraker (2004), who notes that the loss function used in estimation does not necessarily lead to improvements in RMSE. This also applies to our empirical exercise. Moreover, the differences in fit between the model are a function of maturity and moneyness. This may be partly due to the emphasis on dollar loss in our objective function. We keep an investigation of alternative loss functions, such as loss functions based on implied volatilities, for future work.

5.3 Estimates of Jumps in Volatility

The third and fourth columns of Tables 3 and 5 present estimates of the SVJV and SVCJ models, respectively. The SVCJ model in column 4 contains jumps in returns and volatility that are correlated. Several existing studies report estimates for this model (see, among others, Eraker, Johannes, and Polson (2003), Eraker (2004), Broadie, Chernov, and Johannes (2007), Andersen, Fusari, and Todorov (2015a)). Table 7 reports some of the existing estimates of this model. We also report on the SVJV model in column 4 of Tables 3 and 5 because we encountered some identification problems when implementing the SVCJ model. The SVJV model is nested by the SVCJ model: it contains jumps in variance but not in returns, and as a result it has three fewer parameters. Moreover, several studies have argued that jumps in variance are more important than jumps in returns. We therefore want to highlight the relative importance of jumps in returns and jumps in variance for modeling options and returns.

First consider the results based on options in Table 3. For the SVJV model, the estimate of the risk neutral average jump size $\mu_v^Q = \mu_v - \eta_{J^v}$ is equal to 6.57%. The estimate of the risk premium η_{J^v} is small and negative, and the estimate of the frequency of the jumps λ is equal to 0.828, implying these jumps occur just less than once per year. When adding jumps in returns in the SVCJ model in column 4 of Table 3, the risk premium becomes somewhat larger and the frequency of the jumps decreases. The risk-neutral average returns jump size is larger (more negative) than in the SVJR model. The estimate of the correlation between the return and variance jump is negative, as expected, at -0.574.

The parameter estimates from returns in Table 5 indicate that the physical standard deviation of the variance jumps is significantly smaller. In the SVCJ model, the correlation between return and variance jumps is also smaller in absolute value compared to Table 3, at -0.4159.

Table 7 presents parameter estimates for the SVCJ model from existing studies. Our estimate of η_{J^v} is close to that of Andersen, Fusari, and Todorov (2015a) and well within the range of existing studies. The same remark applies to our correlation estimate. It is noteworthy that existing estimates of the variance risk parameters seem to be much more consistent across different studies, some of which use very different sample periods, compared to the existing estimates of jumps in returns in Table 5. One possible reason for this finding is that variance jumps are better identified in the data.

The log likelihoods in Tables 3 and 5 are very useful as indicators of the importance of return and variance jumps for modeling returns and options. First, using a log likelihood test, the more complex models are always statistically supported by the data. Second, the differences in the log likelihoods are again much larger for the option-based estimation in Table 3, which is not surprising. Our most interesting conclusions are with respect to the relative importance of return and variance jumps. The results in Tables 5 and 3 indicate that for the purpose of both return and option modeling, we need to account for jumps in returns as well as jumps in variance. In both tables, the SVCJ model achieves the highest log likelihood value, and the differences are statistically significant. Moreover, the SVJV model results in significant improvements in likelihood over the SV model, just like the SVJR model. It is tempting to compare the likelihoods for the SVJR and SVJV models, but please note that these models are not nested. The SVJV model results in larger improvements in the log likelihood in both Tables 3 and 5, but more importantly, when adding additional jumps in returns in the SVCJ model, the likelihood improves further.

However, Panels C and D in Table 4 indicate that despite the large improvements in log likelihood, the richer model with jumps do not outperform the simple SV model in terms of RMSE, again confirming the results of Eraker (2004). In fact, the overall RMSE of the SV model is lower than that of the SVJV and SVCJ models. The models with variance jumps perform particularly poorly for OTM long-maturity options. In future work we plan to explore how different loss functions affect these findings.

We conclude that the statistical evidence for both return and option data indicates the

presence of return jumps as well as variance jumps. This is largely consistent with the existing empirical literature. Our findings also confirm that while the statistical evidence supports the more elaborate models with returns and volatility jumps, these models do not necessarily lead to improved option fit, even in-sample. Moreover, the literature also finds that to fit options, return jumps are more important than volatility jumps. Finally, our estimates of variance jump parameters are overall consistent with existing estimates.

5.4 Joint Estimation Using Returns and Options

Table 8 presents estimation results based on the joint log likelihood function combining return and option data. While the focus of this study is on developing an estimation methodology that exclusively relies on option data, the joint study of returns and options is obviously of significant interest. Most importantly, it results in parameter estimates that are consistent with the mapping between the physical and risk-neutral dynamic, which is instructive about the implications of these theoretical constraints and risk premiums. Using both data sources in estimation may also help with identification.

Inspection of Table 8 shows that the parameters are closer to the parameters estimated from options in Table 3, despite the fact that we try to reduce the relative importance of the option data by using the weighted average likelihood in equation (26). The most important conclusion from Table 8 is that the log likelihoods tell a very important story. While the jumps in returns in the SVJR model lead to statistically significant improvements in fit compared with the SV model, it is the jumps in volatility in the SVJV and SVCJ models that generate very large improvements in the log likelihood. Most importantly, it can be seen from comparing Tables 8 and 3 that the volatility jumps are even more important when jointly modeling returns and options in Table 8. This suggests that the jumps are especially useful to capture and model risk premia. Existing studies also emphasize the usefulness of jump processes to capture risk premia (Pan (2002)), but the results in Table 8 clarify that these findings are to a large extent due to the risk premia on variance jumps.

Figures 4-7 provide additional evidence on the joint estimation results. The third panel in Figure 4 depicts the filtered variance path for the SV model estimated using returns and options jointly. The path is more similar to the option-implied path in the second panel, which is due to the estimate of σ from joint estimation, which is more similar to that obtained

from options. The first two panels of each of Figures 5-7 depict the diffusive variance implied by returns and joint returns and options for the SVJR, SVJV, and SVCJ models. All figures reflect the higher σ estimate from returns as compared to joint estimation. The variance paths look overall similar for the different models but there are some interesting differences. Consider for instance the return-based variance in the SVJV model, which contains larger spikes.

The most interesting aspect of Figures 5-7 are the jumps filtered from the returns data and the joint data. All figures indicate that the return and variance jumps are infrequent and rather large, regardless of whether they are obtained from returns or from a joint estimation. Figure 5 clearly indicate that the risk-neutral (negative) return jumps are larger than the physical jumps.

Our estimates also provide insight into the estimates of the diffusive variance risk premium η_v. The estimates are qualitatively similar for the option-based results in Table 3 and the joint estimation results in Table 8. In the SV model, the parameter is positive and statistically very significant. When adding return jumps, the estimate is still positive and statistically significant, but smaller. When adding volatility jumps, the sign of η_v may change and the statistical significance decreases. We therefore conclude that identification of the diffusive variance risk premium is not a problem, unless variance jumps are included, which make it difficult to separately identify diffusive and jump variance risk premiums.

Estimation based on the joint likelihood data allows us to compute the risk premiums. The bottom two rows in Table 8 present the average equity risk premium for the different models. Recall that the total risk premium is given by $\gamma_t = \eta_s V_t + \lambda(\bar{\mu}_s - \bar{\mu}_s^Q)$, where $\eta_s V_t$ is due to diffusive risk and $\lambda(\bar{\mu}_s - \bar{\mu}_s^Q)$ is due to jump risk. For the SV model, the equity risk premium is 8.78 %, close to the sample average of 7.97 %. The models that contain returns jumps imply a larger average risk premium. For these two models, the decomposition of the overall average equity risk premium in diffusive and jump components is similar. Approximately 2% of the equity risk premium is due to jump risk, with the rest due to diffusive risk. Broadie, Chernov, and Johannes (2007) report that in their sample, price jump risk premia contribute about 3% per year to an overall equity premium of 8%. These results are obtained using futures data, a different sample and an entirely different approach, but they are remarkably similar. Pan reports that price jump risk premia contribute about 3.5% per year to an overall equity premium of 9%. Note that a variance decomposition shows that in the SVJR model,

approximately 6% of the variation in returns is due to jumps, while it is approximately 8% in the SVCJ model. Consistent with the existing literature, these computations indicate that jumps are relatively more important for risk premiums than for explaining overall return variation.

6 Conclusion

Estimating state-of-the art option valuation models is challenging due to the complexity of the models and the richness of the available option data. We propose new techniques to overcome these constraints, using particle filtering with weights based on model-implied spot volatilities rather than model price. We also use his approach to estimate option valuation models based on returns and options jointly.

We provide an in-depth investigation of the double-jump models with jumps in returns and volatility (Duffie, Pan, and Singleton (2000)) using twenty years of daily data, and almost thirty option contracts per day with different maturities and widely different moneyness. The resulting model estimates are very different from many existing studies that mainly use short-maturity at-the-money options. The most important differences obtain for the parameters that identify that tails of the distribution, which cannot reliably be identified from at-the-money options. We find strong evidence for volatility jumps and volatility jump risk premia, which are important even when allowing for jumps in returns.

Our approach can be used to study a variety of other important questions in the option valuation literature. The most pressing question is to study extensions of the model analyzed in this paper, using for instance the SVSCJ model, multiple volatility factors as in Bates (2000) or a tail factor as in Andersen, Fusari, and Todorov (2015b).

References

Aït-Sahalia, Yacine, and Robert Kimmel, 2007, Maximum Likelihood Estimation of Stochastic Volatility Models, *Journal of Financial Economics* 83, 413–452.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015a, Parametric Inference and Dynamic State Recovery From Option Panels, *Econometrica* 83, 1081–1145.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015b, The Risk Premia Embedded in Index Options, *Journal of Financial Economics* 117, 558–584.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical Performance of Alternative Option Pricing Model, *Journal of Finance* 52, 2003–2049.

Bates, David S., 1996a, Jump and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options, *Review of Financial Studies* 9, 69–107.

Bates, David S., 1996b, Testing Option Pricing Models, *Handbook of Statistics* 14, 567—-611.

Bates, David S., 2000, Post-'87 Crash Fears in S&P˜500 Futures Options, *Journal of Econometrica* 94, 181–238.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22, 4463–4492.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model Specification and Risk Premia: Evidence From Futures Options, *Journal of Finance* 62, 1453–1490.

Carr, Peter, and Dilip B Madan, 1999, Option Valuation Using the Fast Fourier Transform, *Journal of Computational Finance* 2, 61–73.

Chernov, Mikhail, and Eric 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.

Christoffersen, Peter, Kris Jacobs, and Karim Mimouni, 2010, Volatility Dynamics for the S&P500: Evidence From Realized Volatility, Daily Returns and Option Prices, *Review of Financial Studies* 23, 3141–3189.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* 68, 1343–1376.

Eraker, Bjørn, 2004, Do Stock Prices and Volatility Jump? Evidence From Spot and Option Prices, *Journal of Finance* 59, 1367–1403.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The Impact of Jumps in Equity Index Volatility and Returns, *The Journal of Finance* 58, 1269–1300.

Hastings, W K, 1970, Monte Carlo Sampling Methods Using Markov Chains and Their Applications, *Biometrica* 57, 97.

Heston, Steven L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *The Review of Financial Studies* 6, 327–343.

Johannes, Michael S., Nicholas G. Polson, and Jonathan R. Stroud, 2009, Optimal filtering of jump diffusions: Extracting latent states from asset prices, *Review of Financial Studies* 22, 2759–2799.

Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller, 1953, Equation of State Calculations by Fast Computing Machines, *The Journal of Chemical Physics* 21, 1087–1092.

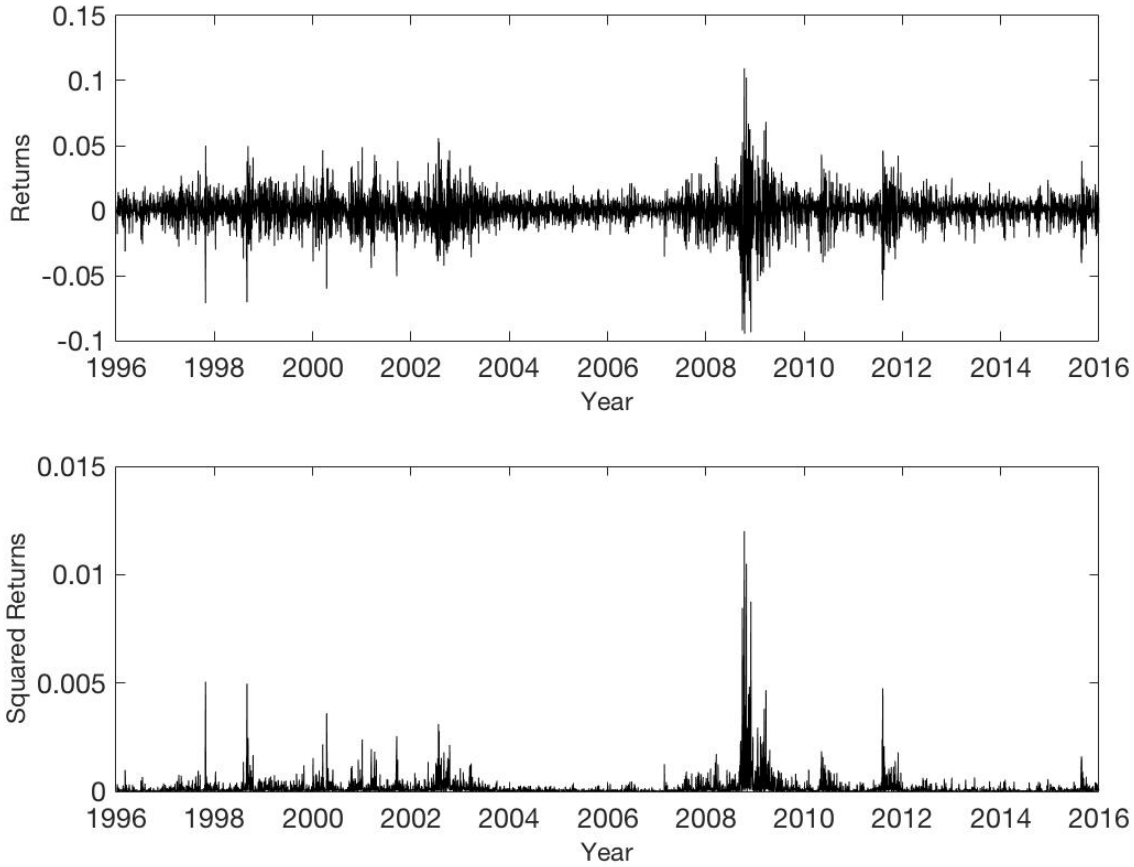
Pan, Jun, 2002, The Jump-risk Premia Implicit in Options: Eevidence from an iItegrated Time-series Study, *Journal of Financial Economics* 63, 3–50.

Pitt, Michael K., 2002, Smooth Particle Filters for Likelihood Evaluation and Maximization.

Pitt, Michael K., and Neil Shephard, 1999, Filtering via Simulation: Auxiliary Particle Filters, *Journal of the American Statistical Association* 94, 590–599.

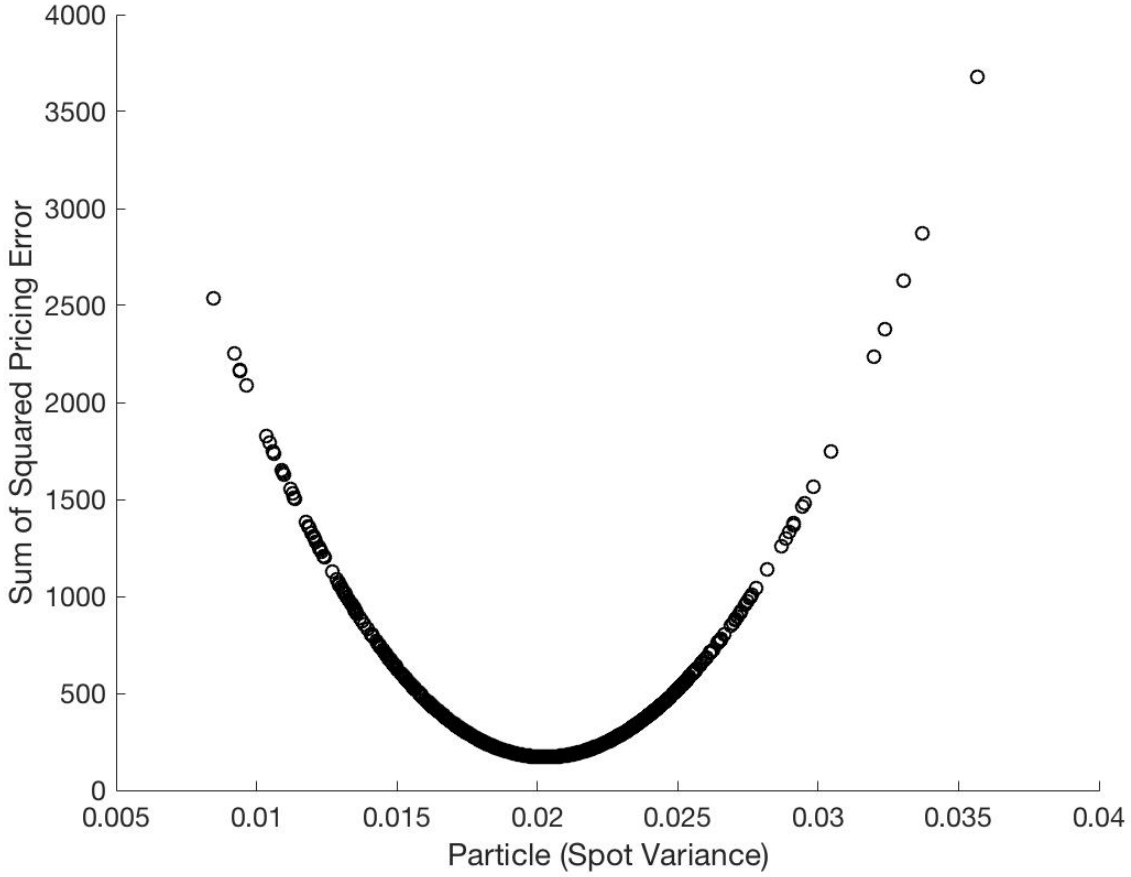
Roberts, Gareth O., and Jeffrey S. Rosenthal, 2009, Examples of Adaptive MCMC, *Journal of Computational and Graphical Statistics* 18, 349–367.

Figure 1: Daily Returns 1996-2015



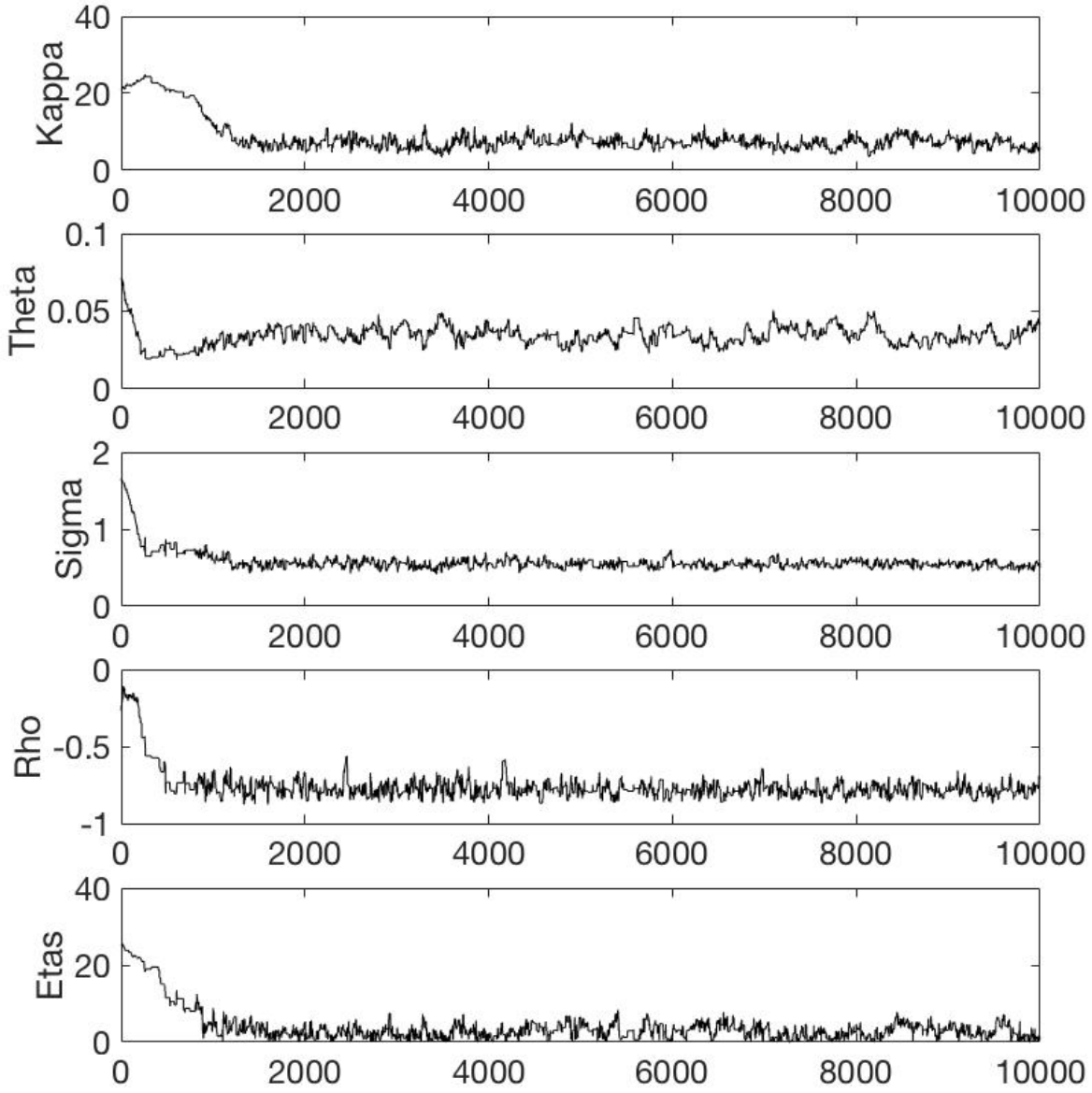
Notes: We plot daily log returns in the top panel. In the bottom panel, we plot squared daily log returns. The sample period is from January 1, 1996 until December 31, 2015.

Figure 2: Sum of Option Pricing Errors vs Particles



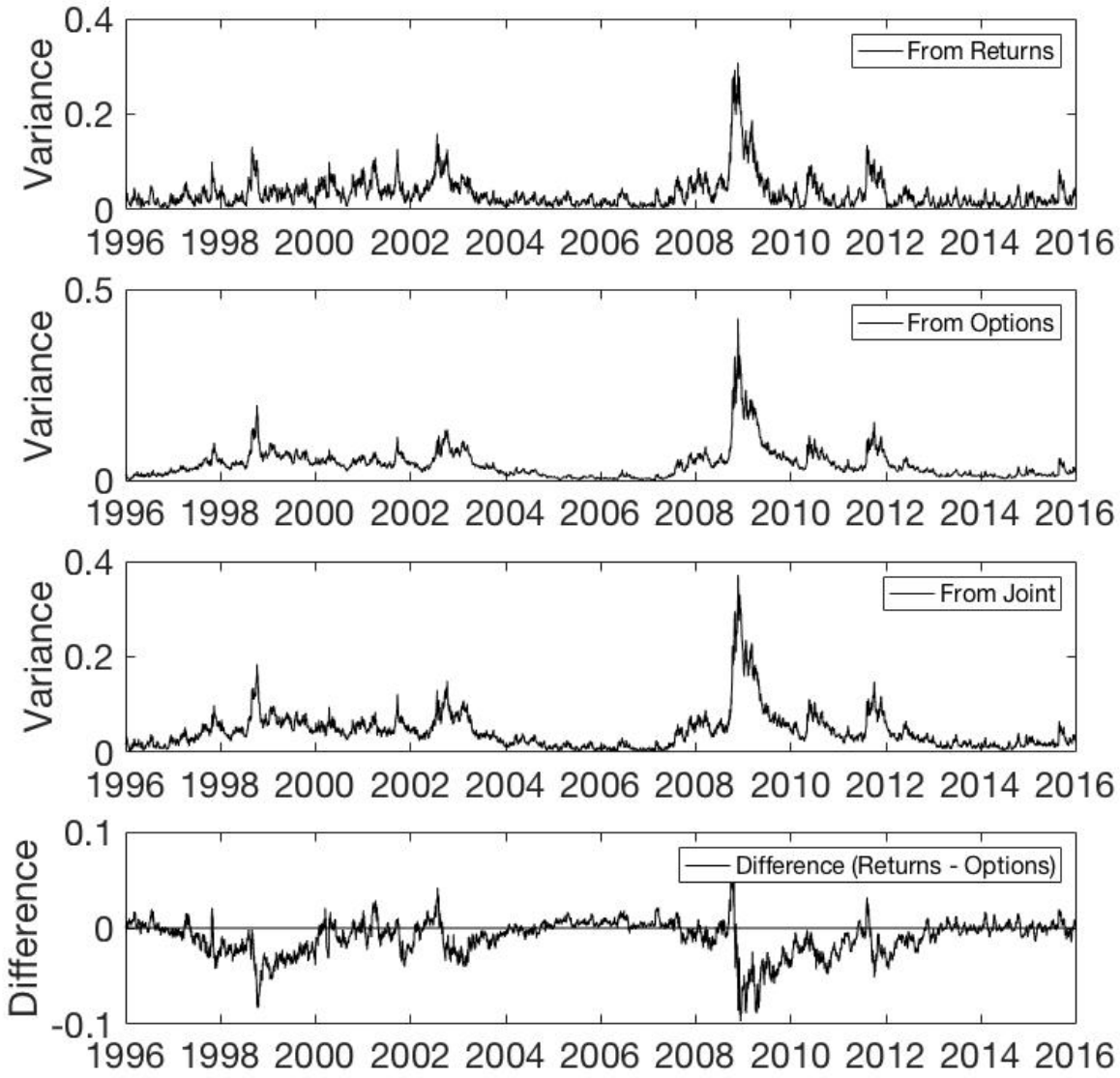
Notes: We show how the sum of squared pricing errors changes with the particle values in the case of the SV model. For each of the particle values (spot variances), we calculate the sum of squared pricing errors. We use option data for December 1, 2015 and the following parameter values: $\kappa = 3$, $\theta = 0.25$, $\sigma = 0.4$, and $\rho = -0.7$.

Figure 3: Parameter Trace for SV Model Parameters. Return-Based Estimation



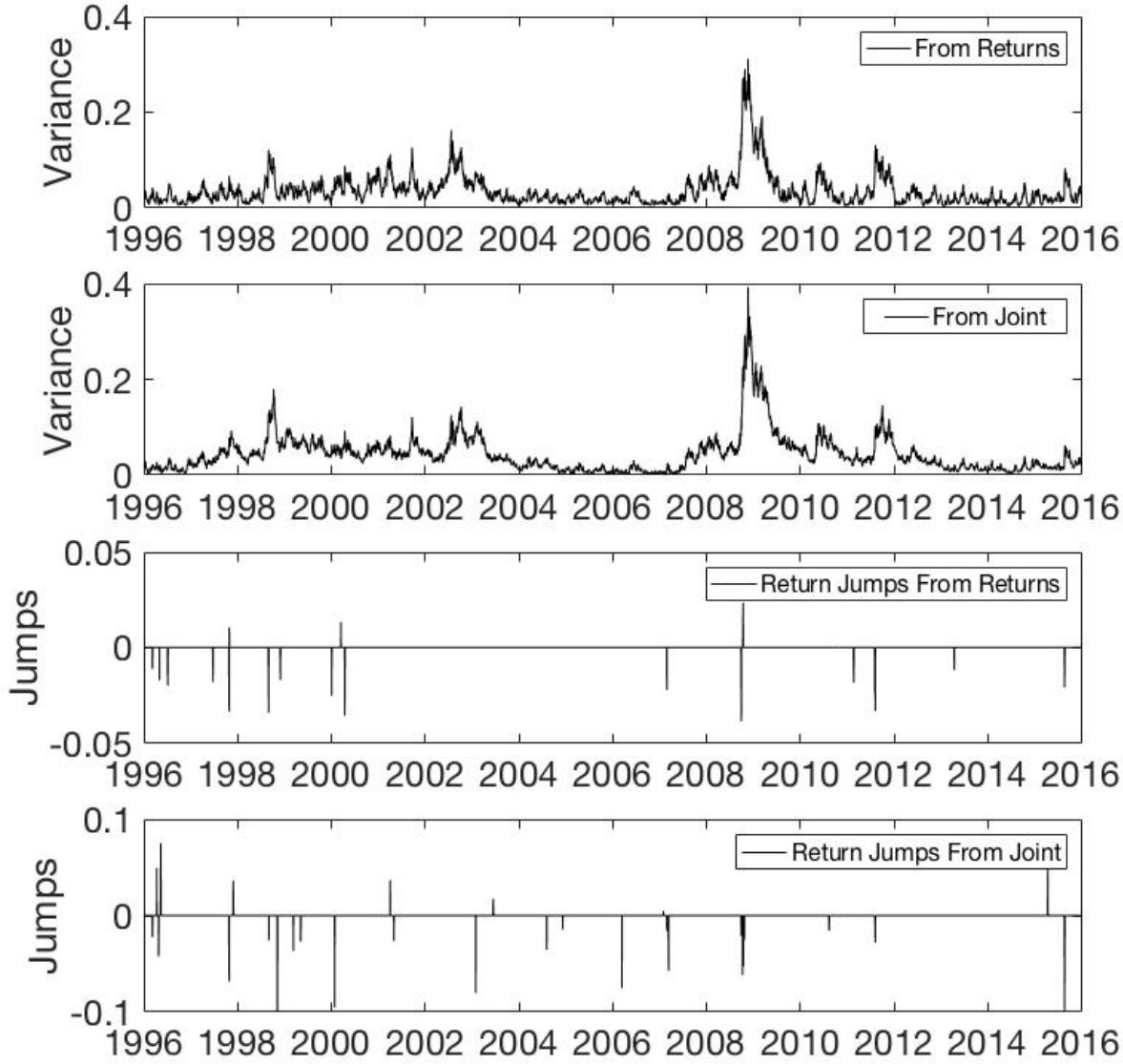
Note: We plot the full traces for each parameter in the SV model. We use 10,000 iterations. The first 1/4 of the iterations are treated as burn-in.

Figure 4: Filtered Variance Paths. SV Model



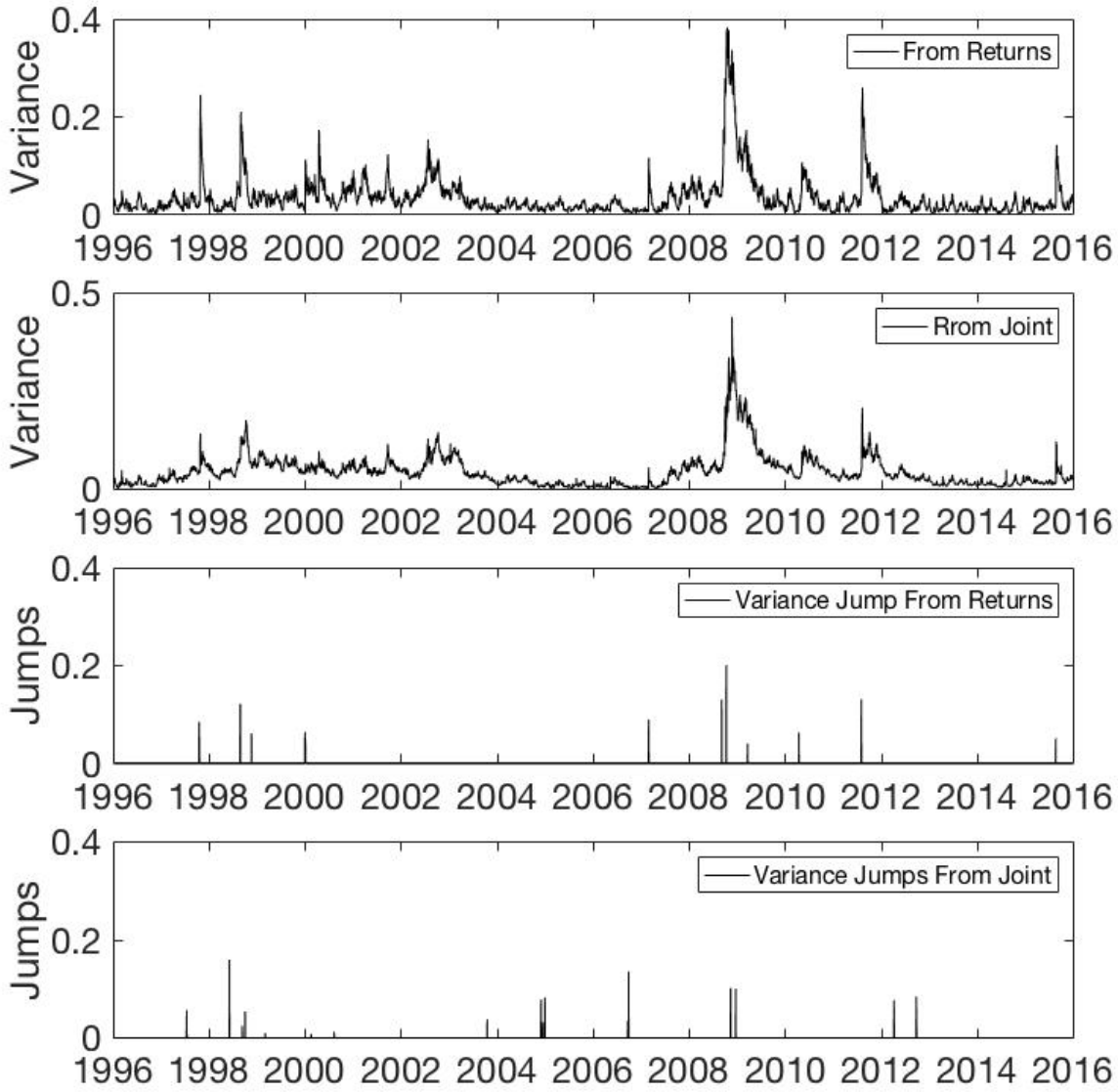
Notes: We plot the filtered variance path estimated from returns in the top panel, the variance estimated from options in the second panel, and the variance from joint estimation in the third panel. The bottom panel plots the difference between the variance estimated from returns and the variance estimated from options.

Figure 5: Filtered Variances and Jumps. SVJR Model



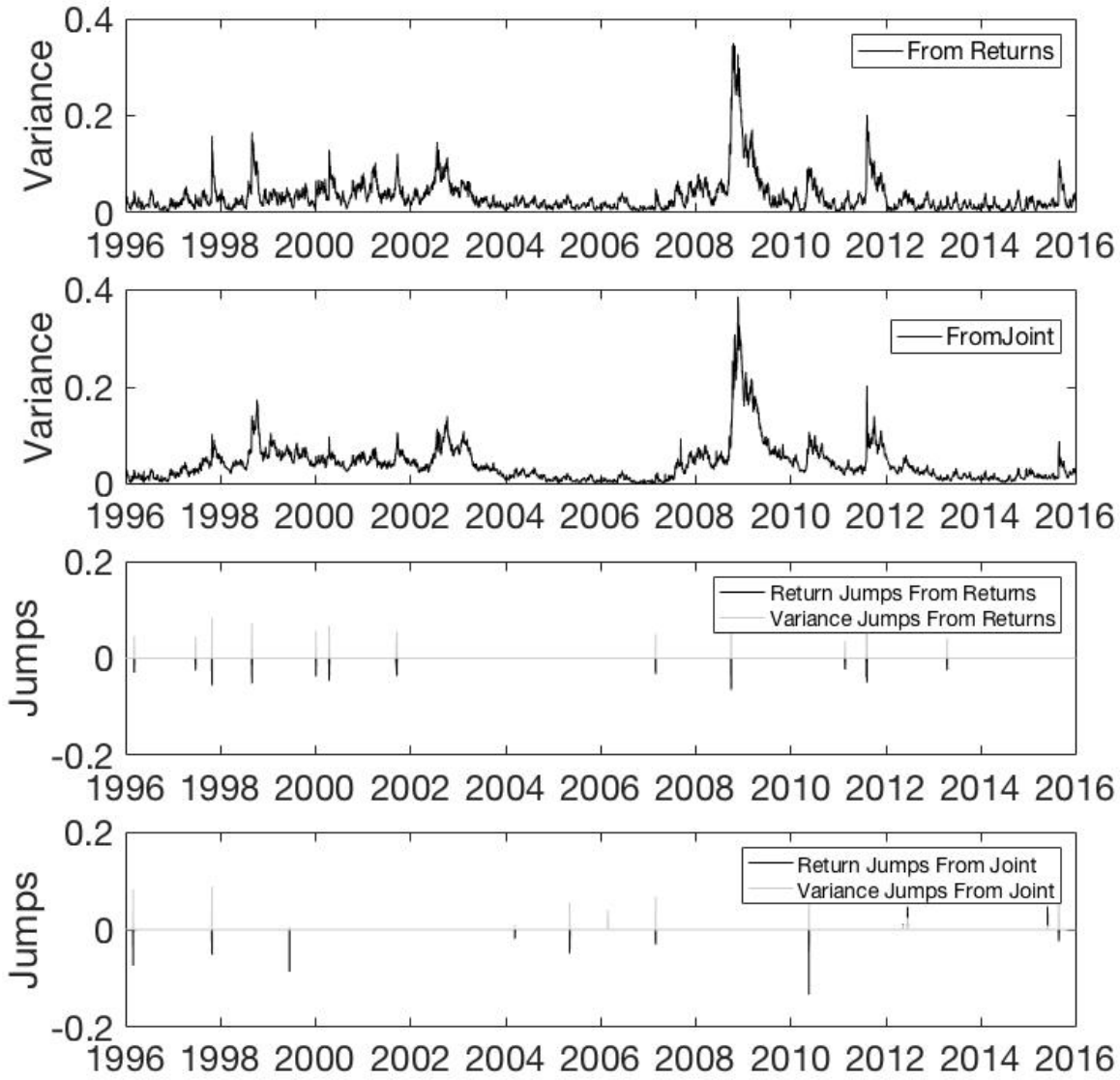
Notes: We plot the variance estimated from returns and the variance from joint estimation in the top two panels. The bottom two panels plot the return jumps filtered from return-based and joint estimation respectively.

Figure 6: Filtered Variances and Jumps. SVJV Model



Notes: We plot the variance estimated from returns and the variance from joint estimation in the top two panels. The bottom two panels plot the return jumps filtered from return-based and joint estimation respectively.

Figure 7: SVCJ Filtered Variances and Jumps. SVCJ Model



Notes: We plot the variance estimated from returns and the variance from joint estimation in the top two panels. The bottom two panels plot the return jumps filtered from return-based and joint estimation respectively.

Table 1: Parameter Estimates in Existing Studies: The Heston SV Model.

(a) Based on Returns

Author	Period	κ	θ	σ	ρ
ABL	1953-1996	4.032	0.017	0.202	-0.380
Benzoni	1996-1997	3.931	0.013	0.197	-0.597
CV	1980-2000	14.282	0.033	5.193	-0.629
CGGT	1953-1999	3.276	0.015	0.151	-0.279
EJP	1980-2000	5.821	0.023	0.361	-0.397
Jones	1986-2000	3.704	0.026	0.524	-0.603
Eraker	1987-1990	4.284	0.022	0.277	-0.373
Bates[1]	1953-1996	5.940	0.016	0.315	-0.579
CJM	1996-2004	6.520	0.035	0.460	-0.771

(b) Based on Options

Author	Period	κ	θ	σ	ρ
BCC	1988-1991	1.150	0.040	0.390	-0.640
Bates[2]	1988-1993	1.490	0.067	0.742	-0.571
BCJ	1987-2003	7.056	0.019	0.361	-0.397
CJM	1996-2004	2.879	0.063	0.537	-0.704

(c) Based on Returns and Options

Author	Period	κ	θ	σ	ρ	η_v
Pan	1989-1996	7.100	0.014	0.320	-0.530	7.600
Eraker	1987-1990	4.788	0.049	0.554	-0.569	2.520
ASK	1990-2004	5.070	0.046	0.480	-0.767	

Notes: We report parameters for the SV model in existing studies. Estimates with a star (*) indicate risk-neutral values. All parameters are annualized. BCC: Bakshi, Cao and Chen (1997),based on the S&P 500; ABL: Andersen, Benzoni, and Lund (2002), based on the S&P 500; Benzoni: Benzoni (2002), based on the S&P 500; CV: Chacko, and Viceira (2003), based on the S&P 500; CGGT: Chernov, Gallant, Ghysels, and Tauchen (2003), based on the DJIA; EJP: Eraker, Johannes, and Polson (2003), based on the S&P 500; Jones: Jones (2003), based on the S&P 100; Eraker: Eraker (2004), based on the S&P 500; Bates[1]: Bates (2006), based on the S&P 500; Bates[2]: Bates (2000), based on the S&P 500; CJM: Christoffersen, Jacobs, and Mimouni (2010), based on the S&P 500; BCJ: Broadie, Chernov, and Johannes (2005), based on the S&P 500; Pan: Pan (2000), based on the S&P 500; ASK: Ait-Sahalia and Kimmel (2007), based on the S&P 500.

Table 2: Return and Option Data

(a) Panel A: Return Data

	Mean	StdDev	Skewness	Kurtosis	Max	Min
Index Returns	0.0797	0.1958	-0.0535	10.7980	0.1158	-0.0904

(b) Panel B: All Option Data

		Moneyness		
	All	0.7-1.3	0.9-1.1	0.95-1.05
All	945,110	845,093	597,313	378,714
Call	368,693	365,778	305,641	197,224
Put	576,417	479,315	291,672	181,490

(c) Panel C: Balanced Option Data. Number of Option Contracts

	Maturity (days)					
Moneyness	5-30	30-60	60-90	90-180	180-365	All
0.85-0.90	4312	4750	4249	4621	4186	22118
0.90-0.95	4377	4824	4532	4839	4540	23112
0.95-1.00	4383	4838	4626	4922	4823	23592
1.00-1.05	4383	4842	4633	4949	4823	23630
1.05-1.10	4046	4741	4380	4774	4596	22537
1.10-1.15	1432	2639	2443	3711	3968	14193
All	22933	26634	24863	27816	26936	129182

(d) Panel D: Balanced Option Data. Average Call Price

	Maturity (days)					
Moneyness	5-30	30-60	60-90	90-180	180-365	All
0.85-0.90	149.35	150.81	158.71	165.44	184.66	161.79
0.90-0.95	100.03	105.41	113.95	125.17	146.97	118.31
0.95-1.00	40.72	50.48	60.28	76.45	101.39	65.87
1.00-1.05	11.16	20.82	30.91	45.14	71.75	35.95
1.05-1.10	2.62	4.92	9.42	19.37	40.57	15.38
1.10-1.15	2.47	2.92	5.09	9.29	23.44	8.64
All	51.06	55.89	63.06	73.48	94.80	67.66

Notes: Panel A reports descriptive statistics for the sample of index returns. The mean and standard deviation are annualized. We report the total available number of contracts in our sample period in Panel B. Panel C shows the number of contracts and average call price in the balanced option data set where we choose the most liquid option within each moneyness-maturity range. Moneyness is defined as K/S. Due to the fact that OTM options are generally more heavily traded than ITM options, this data set mainly consists of OTM call and OTM put options.

Table 3: Parameter Estimates. Option-Based Estimation

	SV	SVJR	SVJV	SVCJ
κ	1.9153	1.3812	1.2853	0.6183
	(0.0431)	(0.0758)	(0.0781)	(0.3120)
θ	0.0396	0.0394	0.0274	0.0264
	(0.0009)	(0.0007)	(0.0012)	(0.0005)
σ	0.4091	0.3922	0.3741	0.2940
	(0.0120)	(0.0221)	(0.0170)	(0.0121)
ρ	-0.9330	-0.9562	-0.9515	-0.9518
	(0.0253)	(0.0236)	(0.0777)	(0.0441)
η_v	0.8010	0.4961	0.0468	-0.3640
	(0.0448)	(0.0684)	(0.0375)	(0.0515)
λ		0.8554	0.8284	0.6719
		(0.0414)	(0.0543)	(0.0585)
$\mu_s - \eta_{J^s}$		-0.0376		-0.0474
		(0.0011)		(0.0014)
σ_s		0.0492		0.0396
		(0.0012)		(0.0015)
μ_v			0.0644	0.0623
			(0.0043)	(0.0015)
η_{J^v}			-0.0013	-0.0025
			(0.0023)	(0.0014)
ρ_J				-0.5740
				(0.0051)
Loglikelihood	-15459	-15398	-15308	-15239

Notes: We report parameters estimated using option data only for the SV, SVJR, SVJV and SVCJ models. Parameters are annualized and under the physical measure. Since μ_s is not identified from option prices, we fix this parameter at the value estimated from returns. In parentheses, we report the standard error for each parameter.

Table 4: Root Mean Squared Error. Option-Based Estimation

	Maturity (days)					
Moneyness	5-30	30-60	60-90	90-180	180-365	All
Panel A: SV						
0.85-0.90	2.94	3.12	3.08	2.70	4.13	3.19
0.90-0.95	2.83	2.81	2.44	2.00	3.95	2.81
0.95-1.00	3.00	2.85	2.36	1.56	3.97	2.75
1.00-1.05	2.88	3.34	3.04	2.18	3.81	3.05
1.05-1.10	2.32	2.52	2.91	2.97	3.39	2.82
1.10-1.15	2.78	2.42	2.61	2.81	3.64	2.85
All	2.79	2.84	2.74	2.37	3.82	2.91
Panel B: SVJR						
0.85-0.90	2.88	3.05	2.99	2.71	3.99	3.12
0.90-0.95	2.68	2.65	2.30	1.97	3.58	2.64
0.95-1.00	2.91	2.90	2.50	1.70	3.38	2.68
1.00-1.05	2.92	3.45	3.18	2.34	3.34	3.05
1.05-1.10	2.32	2.53	2.94	2.95	3.30	2.81
1.10-1.15	2.78	2.39	2.58	2.76	3.82	2.86
All	2.75	2.83	2.75	2.41	3.57	2.86
Panel C: SVJV						
0.85-0.90	2.95	3.17	3.17	2.84	4.38	3.30
0.90-0.95	2.86	2.87	2.52	2.13	4.10	2.89
0.95-1.00	3.01	2.83	2.33	1.53	3.85	2.71
1.00-1.05	2.84	3.29	3.06	2.26	3.61	3.02
1.05-1.10	2.30	2.44	2.72	2.75	3.49	2.74
1.10-1.15	2.77	2.45	2.66	2.70	4.13	2.94
All	2.79	2.84	2.75	2.37	3.93	2.93
Panel D: SVCJ						
0.85-0.90	2.73	2.87	2.81	2.67	4.43	3.10
0.90-0.95	2.55	2.53	2.23	2.10	4.11	2.70
0.95-1.00	2.88	2.82	2.42	1.74	3.73	2.72
1.00-1.05	2.89	3.41	3.18	2.40	3.39	3.06
1.05-1.10	2.33	2.57	3.05	3.43	3.73	3.02
1.10-1.15	2.78	2.47	2.73	2.86	4.13	2.99
All	2.69	2.78	2.74	2.54	3.92	2.93

Notes: We report the Root Mean Squared Error (RMSE) by moneyness and data-to-maturity with option-based parameters for each of the models. RMSE is measured by dollar amount.

Table 5: Parameter Estimates. Return-Based Estimation

	SV	SVJR	SVJV	SVCJ
κ	6.9691	6.3308	8.0033	7.5660
	(0.9034)	(1.0469)	(1.1433)	(1.1116)
θ	0.0359	0.0350	0.0273	0.0271
	(0.0034)	(0.0042)	(0.0031)	(0.0029)
σ	0.5430	0.5290	0.4754	0.4680
	(0.0286)	(0.0332)	(0.0344)	(0.0337)
ρ	-0.7906	-0.7967	-0.8217	-0.8214
	(0.0240)	(0.0296)	(0.0279)	(0.0294)
η_s	2.5374	2.3737	2.5311	3.1167
	(1.1318)	(1.2476)	(1.2247)	(1.2206)
λ		0.9713	0.9845	0.9953
		(0.1257)	(0.1103)	(0.1189)
μ_s		-0.0132		-0.0101
		(0.0074)		(0.0076)
σ_s		0.0203		0.0212
		(0.0087)		(0.0102)
μ_v			0.0662	0.0547
			(0.0110)	(0.0103)
ρ_J				-0.4159
				(0.1067)
Loglikelihood	16123	16130	16135	16141

Notes: We report parameters estimated using returns only for the SV, SVJR, SVJV and SVCJ models. Parameters are annualized. In parentheses, we report the standard error for each parameter.

Table 6: Parameter Estimates in Existing Studies: The SVJR Model

(a) Based on Returns

Author	Period	κ	θ	σ	ρ	λ	μ_s	σ_s
ABL	1953-1996	3.704	0.013	0.184	-0.620	5.040	0.000	0.012
CV	1980-2000	12.187	0.019	4.274	-0.271	0.372	0.051	$exp(+)$
						5.379	0.026	$exp(-)$
CGGT	1953-1999	2.790	0.016	0.101	0.430	1.764	-0.030	0.018
EJP	1980-2000	3.226	0.021	0.240	-0.467	1.512	-0.025	0.041
Eraker	1987-1990	3.024	0.021	0.202	-0.468	0.756	-0.037	0.066
Bates[1]	1953-1996	4.380	0.014	0.244	-0.612	0.744	-0.010	$exp(+)$
						5.379	0.026	$exp(-)$
CJM	1996-2004	6.589	0.032	0.450	-0.777	2.790	-0.013	0.013

(b) Based on Options

Author	Period	κ	θ	σ	ρ	λ	μ_s	σ_s
BCC	1988-1991	2.030	0.040	0.380	-0.570	0.590	-0.050	0.070
BCJ	1987-2003	5.796	0.012	0.240	-0.467	1.512	-0.100	0.041
CJM	1996-2004	2.638	0.063	0.448	-0.782	2.832	-0.015	0.006

(c) Based on Returns and Options

Author	Period	κ	θ	σ	ρ	λ	μ_s	σ_s
Pan	1989-1996	6.400	0.015	0.300	-0.530	12.300	-0.008	0.039
		3.300*	0.030*				-0.192*	
Eraker	1987-1990	4.788	0.042	0.512	-0.586	0.504	-0.010	0.167
		2.772*	0.072*				-0.050*	

Notes: We report parameter estimates for the SVJR model from existing studies. Estimates with a star (*) indicate risk-neutral values. All parameters are annualized. BCC: Bakshi, Cao and Chen (1997),based on the S&P 500; ABL: Andersen, Benzoni, and Lund (2002), based on the S&P 500; CV: Chacko, and Viceira (2003), based on the S&P 500; CGGT: Chernov, Gallant, Ghysels, and Tauchen (2003), based on the DJIA; EJP: Eraker, Johannes, and Polson (2003), based on S&P 500; Eraker: Eraker (2004), based on S&P 500; Bates[1]: Bates (2006), based on S&P 500; CJM: Christoffersen, Jacobs, and Mimouni (2010), based on the S&P 500; BCJ: Broadie, Chernov, and Johannes (2005), based on the S&P 500; Pan: Pan (2000), based on the S&P 500.

Table 7: Parameter Estimates in Existing Studies: The SVCJ Model

(a) Based on Returns

Author	Period	κ	θ	σ	ρ	λ	μ_s	σ_s	μ_v	ρ_J
EJP	1980-2000	6.552	0.014	0.199	-0.484	1.663	-0.018	0.029	0.037	-0.601
Eraker	1987-1990	4.032	0.014	0.146	-0.461	1.008	-0.032	0.049	0.032	0.312

(b) Based on Options

Author	Period	κ	θ	σ	ρ	λ	μ_s	σ_s	μ_v	ρ_J
BCJ	1987-2003	14.112	0.006	0.199	-0.484	1.663	-0.066	0.029	0.108	-0.601
AFT	1996-2010	2.049	0.033	0.354	-0.934	4.435	0.005	0.004	0.052	-0.502

(c) Based on Returns and Options

Author	Period	κ	θ	σ	ρ	λ	μ_s	σ_s	μ_v	ρ_J
Eraker	1987-1990	5.796	0.034	0.411	-0.582	0.504	-0.061	0.036	0.041	-0.693
		2.772*	0.071*				-0.075*			

Notes: We report parameter estimates for the SVCJ model from existing studies. Estimates with a star (*) indicate risk-neutral values. All parameters are annualized. EJP: Eraker, Johannes, and Polson (2003), based on the S&P 500; Eraker: Eraker (2004), based on the S&P 500; BCJ: Broadie, Chernov, and Johannes (2005), based on the S&P 500; AFT: Andersen, Fusari, and Todorov (2015), based on the S&P 500.

Table 8: Parameter Estimates Based on Joint Estimation Using Returns and Options

	SV	SVJR	SVJV	SVCJ
κ	2.1564	1.5531	1.0158	1.1248
	(0.1790)	(0.1384)	(0.1712)	(0.1316)
θ	0.0351	0.0359	0.0282	0.0241
	(0.0017)	(0.0015)	(0.0016)	(0.0015)
σ	0.4262	0.4152	0.3892	0.3450
	(0.0176)	(0.0157)	(0.0110)	(0.0127)
ρ	-0.9161	-0.9378	-0.9417	-0.9237
	(0.0181)	(0.0194)	(0.0130)	(0.0127)
η_s	2.5016	2.3513	2.7833	3.0401
	(0.3341)	(0.3081)	(0.3077)	(0.3486)
η_v	1.0836	0.5753	-0.1376	0.0498
	(0.1842)	(0.1441)	(0.0895)	(0.0355)
λ		0.8949	0.9188	0.6005
		(0.0239)	(0.0363)	(0.0359)
μ_s		-0.0134		-0.0104
		(0.0008)		(0.0005)
σ_s		0.0491		0.0426
		(0.0005)		(0.0006)
η_{J^s}		0.0236		0.0361
		(0.0007)		(0.0003)
μ_v			0.0594	0.0608
			(0.0011)	(0.0013)
η_{J^v}			-0.0052	0.0018
			(0.0012)	(0.0015)
ρ_J				-0.5030
				(0.0076)
ERP_Diffusion	0.0878	0.0844	0.0785	0.0733
ERP_Jump		0.0206		0.0210
Loglikelihood	1265	1281	1352	1397

Notes: We report parameter estimates jointly estimated using both returns and options for the SV, SVJR, SVJV and SVCJ models. Parameters are annualized and under the physical measure. In parentheses, we report the standard error for each parameter. ERP_Diffusion and ERP_Jump represent the equity risk premium from diffusion and jump.

A Appendix

A.1 Option Pricing Model

We use the Fast Fourier Transform (FFT) method based on Carr and Madan (1999). The time-t price of a call option with strike K and maturity τ can be written as:

$$C = \int_k^\infty e^{-r\tau}(e^{s_T} - e^k) f_s(s_T) ds_T \tag{A.1}$$

where $s_T = log(S_T)$ and $k = log(K)$. In the above formula, as k approaches $-\infty$, C_k converges to S_t rather than zero and thus is not square-integrable. Carr and Madan introduce a dampening factor α to solve the problem and let:

$$c = e^{\alpha k} C \tag{A.2}$$

Now the Fourier transform can be applied to c:

$$\begin{aligned}
\psi(u) &= \int_{-\infty}^{\infty} e^{iuk} c dk \\
&= \int_{-\infty}^{\infty} e^{iuk} e^{\alpha k} \int_k^\infty e^{-r\tau}(e^{s_T} - e^k) f_s(s_T) ds_T dk \\
&= \int_{-\infty}^{\infty} e^{-r\tau} f_s(s_T) \int_{-\infty}^{s_T} [e^{s_T + (\alpha+iu)k} - e^{(\alpha+1+iu)k}] dk ds_T \\
&= \int_{-\infty}^{\infty} e^{-r\tau} f_s(s_T) [\frac{e^{s_T(\alpha+1+iu)}}{(\alpha+iu)(\alpha+1+iu)}] ds_T \\
&= \frac{e^{-r\tau}}{(\alpha+iu)(\alpha+1+iu)} \int_{-\infty}^{\infty} e^{(\alpha+1+iu)s_T} f_s(s_T) ds_T \\
&= \frac{e^{-r\tau}\phi_{s_T}(\alpha+1+iu)}{(\alpha+iu)(\alpha+1+iu)}
\end{aligned} \tag{A.3}$$

here ϕ_{s_T} denotes the risk neutral characteristic function of log-price. The call option value is given by:

$$
\begin{aligned}
C &= e^{\alpha k}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iuk}\psi(u)du \\
&= \frac{e^{-\alpha k}}{\pi}\int_{-\infty}^{\infty} Re[e^{-iuk}\psi(u)]du
\end{aligned}
\tag{A.4}
$$

The second equality holds because the imaginary part of $\psi(u)$ is odd while the real part is even.

$$
C^M = \frac{e^{-\alpha k}}{\pi}\int_{0}^{\infty} Re[e^{-iuk}\psi(u)]du \tag{A.5}
$$

A.2 Particle Filtering Using Returns

The particle filtering algorithm relies on the approximation of the true density of the state L_{t+1} by a set of N discrete points or particles that are updated iteratively through equation (7). Here we outline how SIR particle filtering is implemented using the return data.

Step 1: Simulating the State Forward

For $i = 1 : N$, we first simulate all shocks from their corresponding distribution:

$$
(z_{t+1}, w_{t+1}, B_{t+1}, J^s_{t+1}, J^v_{t+1})^i \tag{A.6}
$$

where the correlation between the state variables needs to be taken into account.[6] Then, new particles are simulated according to equation (7):

$$
V_t = V_{t-1} + \kappa(\theta - V_{t-1}) + \sigma\sqrt{V_{t-1}}w_t + J^v_t B_t \tag{A.7}
$$

Note that $t+1$ shocks affect R_{t+1} and V_{t+1}, and thus to simulate V_t, we in fact need w_t, J^v_t and B_t from the previous period. We then record w_{t+1}, J^v_{t+1} and B_{t+1} for the next period for each particle.

[6] The Brownian shocks z_{t+1} and w_{t+1} are correlated with coefficient ρ, and J^s_{t+1} and J^v_{t+1} are correlated with coefficient ρ_J.

Step 2: Computing and Normalizing the Weights

Now, we compute the weights according to the likelihood for each particle $i = 1 : N$:

$$\begin{aligned}\omega_{t+1}^i &= f_1(O_{t+1}|L_{t+1}^i) \\ &= \frac{1}{\sqrt{2\pi V_t^i}} \exp\left\{-\frac{1}{2}\frac{\left[R_{t+1} - (r - \delta_t - \frac{1}{2}V_t^i + \eta_s V_t^i - \lambda\bar{\mu}_s + J_{t+1}^s)\right]^2}{V_t^i}\right\}\end{aligned} \tag{A.8}$$

The normalized weights π_{t+1}^i are calculated as:

$$\pi_{t+1}^i = \omega_{t+1}^i / \sum_{j=1}^{N} \omega_{t+1}^j \tag{A.9}$$

Step 3: Resampling

The set $\left\{\pi_{t+1}^i\right\}_{i=1}^N$ can be viewed as a discrete probability distribution of $L_{t+1} = (V_t, J_{t+1}^v, B_{t+1})$ from which we can resample. The resampled $\left\{L_{t+1}^i\right\}_{i=1}^N$ as well as its ancestors are stored for the next period.

The filtering for period $t+1$ is now done. The filtering for period $t+2$ starts over from step 1 by simulating based on resampled particles and shocks for period $t+1$. By repeating these steps for all $t = 1 : T$, particles that are more likely to generate the observed return series tend to survive till the end, yielding a discrete distribution of filtered spot variances for each day.

The implementation for the APF is quite similar, except some extra simulation and weighting steps are needed, as discussed in Section 4.