ESSAYS ON OPTION VALUATION AND EMPIRICAL ASSET PRICING

A Dissertation

Presented to

The Faculty of the C.T. Bauer College of Business

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

by

Yuguo Liu

May, 2019

Acknowledgements

I want to express my sincere gratitude to my advisors, Kris Jacobs and Raul Susmel, for their invaluable time and guidance through the ups and downs of the journey. I am deeply indebted to my advisors, Kris and Raul. Without their support, I would not be able to complete the dissertation.

I am grateful to my dissertation committee members: Craig Pirrong, Hitesh Doshi, James Yae, Sang Byung Seo and Rex Yuxing Du for their encouragement and invaluable comments.

I am also thankful to Thomas George, Praveen Kumar, Stuart Turnbull, Nisan Langberg, Vijay Yerramilli, Kevin Roshak, Zack Liu, Paola Pederzoli and Rajkamal Vasu. They taught me fundamental knowledge of finance and empirical skills and provided supportive advice whenever I approach them.

Last but not least, I want to thank Mary Gould, Anesia Lamont, and Ieshia Deal for helping me deal with all the bureaucracy.

ESSAYS ON OPTION VALUATION AND EMPIRICAL ASSET PRICING

Abstract

Yuguo Liu

May, 2019

This dissertation consists of three essays on option valuation and empirical asset pricing. In the first essay, coauthored with Kris Jacobs, we propose an new approach based on Adaptive Metropolis-Hasting MCMC and the Implied Spot Variance Particle Filter to estimate dynamic models using large option panels. The estimation of option valuation models is challenging due to the complexity of the models and the richness of the option data. Many existing studies limit the time-series and especially the cross-sectional dimension of the option data, which may complicate the identification of model parameters. We address these computational constraints by filtering the state variables using particle weights based on model-implied spot volatilities rather than model prices. Some of our estimation results differ substantially from the existing literature. We show that samples restricted to at-the-money and especially short-maturity options may result in serious identification problems. The composition of the option sample also critically affects the relative importance of returns and options for parameter estimates when both are used in estimation.

In the second essay, coauthored with Guanglian Hu, we investigate the relation between variance risk premiums and option returns. Empirically, the out-of-the-money (OTM) S&P 500 call and put options have large negative average returns, and the literature interprets these results as inconsistent with asset pricing theory. We show that these negative OTM option returns are primarily due to the pricing of market volatility risk. With a negative volatility risk premium, expected option returns in a stochastic volatility model are consistent with average call and put option returns across all strikes. The volatility risk premium also predicts future option returns. The predictability of index option returns is both statistically and economically significant. Lastly, we find that some portion of OTM put option returns is attributable to the jump risk premium. Overall our results suggest that the pricing of volatility risk has a first-order effect on the cross-section of index option returns.

In the third essay, I propose a novel method to identify the physical measure parameters and the latent spot variances for dynamic asset pricing models from option data. It is well known that we can estimate physical parameters from underlying returns and risk-neutral

parameters from options, but the existing literature mainly focuses on the option price levels and therefore ignores that option price changes contain information about both the physical and risk-neutral distributions. I find that most physical parameters are better identified using options than using returns. The parameters estimated exclusively from options provide the best fit for the monthly option returns. Moreover, By re-examining the index option return puzzle based on the parameters estimated from options, I find that the option return puzzle disappears when jumps are included in the model.

Contents

List of Tables

List of Figures

Chapter 1

Estimation and Filtering with Big Option Data: Implications for Asset Pricing

1.1 Introduction

There is widespread agreement in the option valuation literature that relatively complex models with stochastic volatility and multiple volatility factors as well as jumps and tail factors are needed to explain the data.[1] It is also well-known that long time series are needed to reliably identify the volatility dynamics (Broadie, Chernov, and Johannes, 2007) and options with different moneyness and maturity may be needed to identify various model aspects. Moreover, to learn about risk premia and the structure of the stochastic discount factor, we need to estimate using both returns and options. Given the richness of the available option data, the resulting estimation of state-of-the-art option valuation models can be computationally very demanding.

[1] See Bates (2000) for evidence on multiple volatility factors. See Eraker (2004), Broadie, Chernov, and Johannes (2007), Pan (2002), Aït-Sahalia, Cacho-Diaz, and Laeven (2015a), Fulop, Li, and Yu (2015), Andersen, Fusari, and Todorov (2017), and Bates (2012) for evidence on jumps in returns and volatility. See Andersen, Fusari, and Todorov (2015b) for evidence on a tail factor in option prices.

This paper studies estimation and filtering dynamic option valuation models using big option data. It also documents the implications of restricting the size of the option panels. To illustrate the computational challenge resulting from big option panels, consider estimation using the OptionMetrics data used in many recent studies, and for simplicity consider using all available option data. 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 over time, but on average it is approximately 200. The full sample of option contracts based on daily OptionMetrics data for 1996-2015 therefore consists of roughly one million contracts.

Even much smaller, judiciously chosen, option samples may impose significant computational constraints. At least partly due to this computational cost, existing studies typically reduce the dimension of the data to make estimation computationally feasible.[2] In the time-series dimension, rather than using a short time series, often only one day per week is used (Bates, 2000, p.195). In the cross-sectional dimension, typically a small subset of the available option contracts is used. Some studies impose very stringent constraints on the cross-section of options and are effectively based on short-maturity at-the-money options.

The first objective of this paper is to adapt existing estimation methods to overcome these computational constraints. We apply Markov Chain Monte Carlo (MCMC) for the parameter search, but we filter the latent state variables using the particle filter (Johannes, Polson, and Stroud, 2009). In general, particle filtering is subject to significant computational constraints when assigning weights based on the model option price, but we address this by assigning weights based on model-implied spot volatilities instead. This enables us to estimate and test complex models using very large option panels. Because particle filtering using returns is relatively straightforward, we are also able to perform joint estimation of state-of-the-art

[2]Option samples may also be restricted for other reasons. For example, illiquid option contracts should be excluded from the sample. But regardless of these other reasons, it is well known and appreciated in the literature that computational constraints often necessitate restricting the sample size. For instance, Broadie, Chernov, and Johannes (2007, p.1461) state that "computational burdens severely constrain how much and what type of data can be used". See also Hurn, Lindsay, and McClelland (2015), Israelov and Kelly (2017), and Andersen, Fusari, and Todorov (2015a) for discussions of computational constraints and cost.

option pricing models based on returns and large option panels. This facilitates inference on model parameters characterizing tail events and risk premia. We show that the likelihood for our implementation of the particle filter is very similar to the likelihood associated with the traditional implementation of the particle filter, which is computationally much more intensive.

Our approach can in principle be applied to any parametric option pricing model. To illustrate its advantages, we implement our estimation method using daily data on a well-known class of models, the workhorse Heston (1993) square root and Duffie, Pan, and Singleton (2000) double-jump models. The limitations of these models are by now well understood, and several authors have suggested model improvements such as multiple volatility components, tail factors and alternative specifications of the jump process. Moreover, Bates (2018) shows that modeling of intraday returns may improve model fit. Related, Andersen et al. (2015b) use estimates of diffusive intraday volatility obtained using high-frequency data and Bandi and Reno (2016) use high-frequency data to estimate models with jumps in returns and volatility. We use daily data and the double-jump model because for this model implementation, the literature on dynamic option valuation models offers a wealth of evidence. By comparing our estimation results to existing studies and interpreting the differences, we are able to illustrate the advantages of our approach and the implications of restricting the sample. Moreover, the parameter estimates we obtain are useful in their own right.[3]

We use daily data for 1996-2015 in our empirical analysis. We do not explicitly address the optimal structure of the option sample. We start with the simple intuition that large cross-sections may be helpful to identify model parameters, especially those governing the tails of the distribution. However, it may not always be optimal or preferable to use the entire cross-section of options, for instance because of liquidity concerns. Moreover, one may prefer cross-sectional data that are balanced across maturity and moneyness, or balanced

[3]For example, Andersen, Fusari, and Todorov (2015a, p.1088) note that, prior to their study, the double-jump model had not been estimated in its full generality because of the computational constraints discussed above.

over time. We therefore do not use all available option contracts but instead we use a large benchmark option sample that is balanced cross-sectionally and over time.

The second contribution of this paper is to document the implications of the composition of the option sample. Because the reduction in the computational burden of our approach allows us to estimate the models using large option samples, we are able to document the consequences of restricting the option sample. We focus on two sets of implications. First, the composition of the option sample may affect inference, in the sense that the likelihoods for different option samples have different optima. A second implication is that the likelihood may be very flat for certain samples, which may lead to identification problems.

Some of our findings simply confirm existing intuition. For example, we find that in the Heston model, it is difficult to reliably estimate the parameters that capture skewness and kurtosis when exclusively using at-the-money-options. We also confirm that it is critical to include options with different maturities to estimate the persistence of the variance. A more surprising finding is that exclusively relying on options with similar maturities also makes it more difficult to estimate the parameters characterizing the tails. We generally find that it is critical to include options with different maturities in the sample to reliably estimate the spot variance and variance mean reversion, which in turn impacts estimation of all other parameters, especially risk premia.

Regarding the moneyness dimension of the data and the tail parameters, we find that including in- and out-of-the-money options is even more critical to reliably estimate the skewness parameter than the kurtosis parameter. Including different maturities also facilitates inference on the skewness parameter and the parameters characterizing the jump processes. However, we find that inference on the price of return jump risk is challenging even when the sample includes large cross-sections of options.

When estimating the Heston (1993) and Duffie, Pan, and Singleton (2000) models using daily data for 1996-2015 and big cross-sections of options, the main differences and similarities between our parameter estimates and those in the available literature are as follows. We confirm the findings of Broadie, Chernov, and Johannes (2007), Eraker (2004) and Eraker, Johannes, and Polson (2003) that both jumps in variance and jumps in returns matter, but that variance jumps are relatively more important. Estimates of the risk-neutral correlation between returns and variance are more negative than the existing literature, with the notable exception of Andersen, Fusari, and Todorov (2015a). Related to this, our estimate of the variance of variance parameter based on options is smaller than the estimate based on returns, which means that the option-implied variance path is smoother than the return-implied variance path. Our findings suggest that the composition of the option sample is critically important for the relative weighting of returns and options in joint estimation. Some parameter estimates in the existing literature, such as the correlation between returns and variance, may largely reflect information from the underlying return series rather than from options, due to the composition of the option sample.[4]

Because we estimate the models using a joint likelihood based on options and returns, we can also compare our estimates of risk premia with the existing literature. Our estimates of the price of diffusive variance risk are smaller than existing ones in the literature, but the most important difference is that they are highly statistically significant. This finding is mainly due to the inclusion of options with different maturities in our sample. Related to the smaller estimates of the price of variance risk, estimates of variance persistence exceed existing estimates. Once again this finding is due to the composition of the option sample. Finally, our estimation based on returns and big panels of option data also allows us to reliably estimate 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. This finding is remarkably consistent with several existing studies.

To the best of our knowledge we are the first to explicitly address the implications of

[4]In joint estimation based on returns and options, the optimal relative weighting of returns and options in the likelihood is an important concern, because the abundance of option data may cause the weight of the returns in the likelihood to become negligible. We eliminate the impact of increasing numbers of option contracts over time by scaling the option likelihood by the inverse of the number of option contracts in the sample. See for example Bates (2003) for a discussion of this issue.

restricting the option sample. This is important because the majority of existing studies impose some restrictions, and it is not clear from the existing literature how this affects empirical results. This complicates comparisons across studies and conclusions about model performance. Moreover, our results also provide valuable insights on how to restrict the sample, if and when such restrictions are required for computational reasons. For instance, we find that for the models studied in this paper, certain restrictions in the moneyness dimension do not have very adverse consequences, while limiting the sample to a narrow range of maturities is likely to cause identification problems. This insight is not obvious a priori.

Our results are relevant for an extensive literature that estimates parametric dynamic option valuation models using return and/or option data. See Singleton (2006) for a discussion of this literature. Several seminal studies combine return and option data using different econometric techniques, which are all computationally demanding. Aït-Sahalia and Kimmel (2007) estimate the model using returns and option information using approximate maximum likelihood, but the information on options is exclusively based on the VIX. Pan (2002) uses implied-state Generalized Method of Moments (GMM) but uses one short-maturity at-the-money option per day in estimation. Chernov and Ghysels (2000) use Efficient Method of Moments (EMM) and use implied volatilities for the closest-to-maturity at-the-money calls. Eraker (2004) uses Markov Chain Monte Carlo (MCMC) technique and incorporates cross-sectional information by using on average approximately three options per day.

In the existing literature, Bates (2000) and Andersen, Fusari, and Todorov (2015a) are related to our approach because they use big cross-sections of index options in estimation. Bates (2000) uses a large dynamic panel for 1988-1993, but treats spot volatilities as parameters.[5] 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)

[5]Bates (2000) also uses a second estimation approach that imposes the dynamic constraint on the variance.

proceed differently and impose consistency between the spot volatility implied from options and a high-frequency volatility estimate from returns. Our approach differs because of our treatment of the filtering problem and the information used to filter the latent states.

Finally, several recent papers also address the need to include information from the entire cross-section of options in estimation and propose computational improvements that make the resulting estimation problems feasible. Feunou and Okou (2017) propose the use of option-implied moments to capture the cross-sectional information. Hurn, Lindsay, and McClelland (2015) document the advantages of parallel computing when using the particle filter to estimate complex option pricing models using large cross-sections of option prices.

The paper proceeds as follows. Section 1.2 presents the option valuation model used in our empirical work and the return and option data used in estimation. Section 3.3 discusses the estimation of the model dynamics from returns and options. Section 1.4 presents the empirical results and Section 1.5 discusses the implications of our estimates for parameter inference and identification, as well as the relative importance of returns and options for parameter estimates. Section 3.6 concludes.

1.2 Model and Data

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

1.2.1 The Model

Our approach can in principle be applied to any parametric option pricing model. To illustrate the advantages of our approach and the implications of restrictions on the option sample, we report on models that are well-known and that have been extensively studied. This allows us to demonstrate that our methods produce reliable estimation results, while we can also explain differences in estimation results by referring to the (larger) option data sets used in our analysis. We therefore base our empirical work on the SVCJ model with contemporaneous jump arrivals in return and variance (Duffie, Pan, and Singleton, 2000):

$$\frac{dS_t}{S_t} = (r_t - \delta_t + \gamma_t - \lambda\bar{\mu}_s)dt + \sqrt{V_t}dZ_t + (e^{J_t^s} - 1)dN_t \tag{1.1}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t + J_t^v dN_t \tag{1.2}$$

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 the 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:

$$\frac{dS_t}{S_t} = (r_t - \delta_t - \lambda\bar{\mu}_s^Q)dt + \sqrt{V_t}dZ_t^Q + (e^{J_t^{sQ}} - 1)dN_t^Q \tag{1.3}$$

$$dV_t = \kappa^Q(\theta^Q - V_t)dt + \sigma\sqrt{V_t}dW_t^Q + J_t^{vQ}dN_t^Q \tag{1.4}$$

where we assume that the diffusive 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 $\eta_{J^v} = \mu_v - \mu_v^Q$.[6] 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 $\gamma_t = \eta_s V_t + \lambda(\bar{\mu}_s - \bar{\mu}_s^Q)$, where $\eta_s V_t$ is the diffusive equity risk premium, which is assumed to be linear in V_t.

The SVCJ specification nests several models in the existing literature. See Singleton

[6]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) investigate more general entertain assumptions regarding the return jumps risk premium, but they use a very different empirical design.

(2006) for a detailed discussion. 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 (1996). 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), which makes the jump intensity a function of volatility. Given the computational burden of estimating the models under consideration using option data, we leave 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:

$$C^M(V_t, \Theta) = e^{-r_t \tau} E^Q[max(S_T - K, 0)] \tag{1.5}$$

We use the Fast Fourier Transform (FFT) in Carr and Madan (1998) to compute option prices. Because of the affine structure of the models, quasi closed-form solutions for option prices are available. See Appendix A for a more detailed discussion. We denote the model price by C^M as opposed to the option's market price C. The model price is computed given the current state V_t and model parameters $\Theta(\kappa, \theta, \sigma, \rho, \lambda, ...)$. In our application, we need to repeatedly calculate prices of options with different spot variances and the ability to vectorize the formula is important for computational reasons.

1.2.2 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 1.1 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 1.1 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 3.1 provides descriptive statistics on the index returns. They 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 sample contains more puts than calls, as expected. Moreover, the option data set obtained after imposing these filters 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.

We therefore create a more balanced panel. Put prices are converted into call prices based on put-call parity. We use six moneyness bins and five maturity bins. Moneyness is defined as strike price divided by index price (K/S). For each moneyness-maturity bin, we include only the most liquid option, defined as the option with the highest trading volume. 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 B of Table 3.1 provides sample sizes for these moneyness-maturity bins. Our analysis is based on this more balanced data set.[7] Panel B of Table 3.1 also reports average option prices and implied volatilities for these moneyness-maturity bins.

1.3 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 we address these problems. Finally we outline how we combine return and option data in estimation.

1.3.1 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 the algorithm however differs dependent on the observables: 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 (3.1)-(3.2) gives

$$R_{t+1} = ln(\frac{S_{t+1}}{S_t}) = r_t - \delta_t - V_t/2 + \gamma_t - \lambda\bar{\mu}_s + \sqrt{V_t}z_{t+1} + J^s_{t+1}B_{t+1} \tag{1.6}$$

$$V_{t+1} - V_t = \kappa(\theta - V_t) + \sigma\sqrt{V_t}w_{t+1} + J^v_{t+1}B_{t+1} \tag{1.7}$$

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

[7] Note that the data are still unbalanced because in the early years of the sample we do not have many observations for short-maturity out-of-the-money calls and/or in-the-money puts.

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:

$$C_{t,h} = C^M_{t,h}(V_t|\Theta) + \epsilon_{t,h} \tag{1.8}$$

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 $\epsilon_{t,h} \sim N(0, \sigma_c^2)$.[8]

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, $L_{t+1} = (V_t, B_{t+1}, J^s_{t+1}, J^v_{t+1})$ and $O_{t+1} = (\{C_t\}, R_{t+1})$ where $\{C_t\} = (C_{t,1}, ..., C_{t,H_t})$.[9] Define the measurement density by $f_1(O_{t+1}|L_{t+1})$ and the transition density 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 the observables, f_1 refers to equation (3.21); when options are the observables, f_1 is given by equation (3.23).[10] For the persistent latent variance, V_t, f_2 represents equation (3.22) 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:

[8]We also investigated a time-varying σ_c, but this did not lead to significant improvements in option fit.

[9]Note the timing convention in the notation: V_t generates the cross-section of option prices $\{C_t\}$; and V_t, together with B_{t+1} and J^s_{t+1}, produces the next period return R_{t+1}.

[10]Since the option price depends only on the current spot variance and the risk-neutral expected jumps rather than the realized jumps, $f_1(\{C_t\}|L_{t+1})$ reduces to $f_1(\{C_t\}|V_t)$, which explains the timing convention in equation (3.23).

$$
\begin{array}{ccccccc}
(L_{t+1}) & \xrightarrow{f_2} & (L_{t+2}) & \xrightarrow{f_2} & (L_{t+3}) & \xrightarrow{f_2} & \ldots \\
\downarrow f_1 & & \downarrow f_1 & & \downarrow f_1 & & \\
O_{t+1}(\{C_t\}, R_{t+1}) & & O_{t+2}(\{C_{t+1}\}, R_{t+2}) & & O_{t+3}(\{C_{t+2}\}, R_{t+3}) & & \ldots
\end{array}
\tag{1.9}
$$

Although we do not directly observe the latent states $L_{t+1}, L_{t+2}, \ldots$, 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 $\Theta(\kappa, \theta, \sigma, \rho, \lambda, \ldots)$ 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.

1.3.2 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. Simulate the particles forward. 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 $\tilde{L}^i_{t+1} = f_2(L_{t+1}|L^i_t)$.

Step 2. Compute the weight for each particle and normalize:

$$
\omega^i_{t+1} = f_1(O_{t+1}|\tilde{L}^i_{t+1}) \tag{1.10}
$$

$$
\pi^i_{t+1} = \omega^i_{t+1} / \sum_{j=1}^{N} \omega^j_{t+1} \tag{1.11}
$$

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. 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 propagation. 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^i_t. We need to first calculate the time t expected values for each particle.

$$\hat{L}^i_{t+1} = E(f_2(L_{t+1}|L^i_t)) \tag{1.12}$$

Then we compute the weights evaluated at this conditional expectation

$${}_1\omega^i_{t+1} = f_1(O_{t+1}|\hat{L}^i_{t+1}) \tag{1.13}$$

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

$${}_1\pi^i_{t+1} = {}_1\omega^i_{t+1} / \sum_{j=1}^{N} {}_1\omega^j_{t+1} \tag{1.14}$$

Step 2. Resample L_t^i according to the weights $\{_1\pi_{t+1}^i\}$, which gives $\breve{L}_t^i$.

Step 3. Simulate particles forward by taking into account new observations, $\tilde{L}_{t+1}^i = f_2(L_{t+1}|\breve{L}_t^i, O_{t+1})$.

To simulate the latent states in $\tilde{L}_{t+1}^i = (\tilde{V}_t^i, \tilde{B}_{t+1}^i, \tilde{J}_{t+1}^{s,i}, \tilde{J}_{t+1}^{v,i})$, we simulate the jump arrivals $\tilde{B}_{t+1}^i$, the return jump size $\tilde{J}_{t+1}^{s,i}$, the spot variance $\tilde{V}_t^i$ and the variance jump size $\tilde{J}_{t+1}^{v,i}$ sequentially, while marginalizing out everything else. Note that $\tilde{J}_{t+1}^{v,i}$ does not affect the observable O_{t+1}, but only affects V_{t+1}^i in the next period. We therefore omit $\tilde{J}_{t+1}^{v,i}$ in the notation of the conditioning information set. We proceed as follows. First, $\tilde{B}_{t+1}$ takes the value of 1 with probability:

$$p(B_{t+1}^i = 1|\hat{J}_{t+1}^{s,i}, \hat{V}_t^i, O_{t+1}) \propto f_1(O_{t+1}|\hat{J}_{t+1}^{s,i}, \hat{V}_t^i, B_{t+1}^i = 1) \tag{1.15}$$

where $\hat{J}_{t+1}^{s,i} = \mu_s$ is the expected return jump size and $\hat{V}_t^i = E(f_2(V_t|\breve{V}_{t-1}^i, \breve{B}_t^i\breve{J}_t^{v,i}))$ is the expected spot variance at time t. Next, we simulate $\tilde{J}_{t+1}^{s,i}$ conditional on $\tilde{B}_{t+1}$, $\hat{V}_t^i$ and O_{t+1}, we simulate $\tilde{V}_t^i$ conditional on $\tilde{B}_{t+1}$, $\breve{B}_t^i\breve{J}_t^{v,i}$ and O_{t+1}, and we simulate $\tilde{J}_{t+1}^{v,i}$ conditional on $\tilde{J}_{t+1}^{s,i}$.

Step 4. Compute the second stage weight for each particle $\tilde{L}_{t+1}^i$:

$$_2\omega_{t+1}^i = f_1(O_{t+1}|\tilde{L}_{t+1}^i)/f_1(O_{t+1}|E(f_2(L_{t+1}|\breve{L}_t^i))) \tag{1.16}$$

$$_2\pi_{t+1}^i = {_2\omega_{t+1}^i}/\sum_{j=1}^{N} {_2\omega_{t+1}^j} \tag{1.17}$$

Step 5. Another resampling on $\tilde{L}_{t+1}^i$ according to the weights $\{_2\pi_{t+1}^i\}$, which gives L_{t+1}^i, $i = 1 : N$.

The APF differs from the standard SIR in two ways: 1) the APF resamples before simulating new latent states according to the predictive likelihood and thus only particles with high likelihood to have generated O_{t+1} move forward; 2) the APF takes new observations into account to propagate new latent states. These two key characteristics of the APF are crucial for capturing rare events and help correct potential sample impoverishment problems associated with the standard SIR particle filter. See Pitt and Shephard (1999) and Johannes, Polson, and Stroud (2009) for more detailed discussions.

1.3.3 Adaptive Metropolis-Hastings MCMC

We use the Adaptive Metropolis-Hastings MCMC (AMH-MCMC) for the parameter search. The AMH-MCMC technique is originally based on Metropolis et al. (1953) and Hastings (1970) and is widely used in the existing literature due to its flexibility, especially when dealing with high-dimensional distributions.[11] In general, AMH-MCMC randomly samples a parameter set from a proposal distribution $q_j(\Theta_j^p|\Theta_{j-1})$ and subsequently accepts a new parameter vector Θ_j^p with probability:

$$\alpha(\Theta_j^p, \Theta_{j-1}) = min(1, \frac{f_1(O_{1:T}|\Theta_j^p)p(\Theta_j^p)/q_j(\Theta_j^p|\Theta_{j-1})}{f_1(O_{1:T}|\Theta_{j-1})p(\Theta_{j-1})/q_j(\Theta_{j-1}|\Theta_j^p)}) \tag{1.18}$$

where f_1 denotes the (simulated) likelihood and $p(\Theta)$ is the prior for Θ, which can be uninformative. We assume a flat prior and thus $p(\Theta_j^p) = p(\Theta_{j-1})$.[12]

We assume the proposal distribution $q_j(\Theta_j^p|\Theta_{j-1})$ to be multivariate normal with mean zero. Tuning the parameter variances in this proposal distribution 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 a given amount of iterations. For example, we start from some pre-specified variance for each parameter and run 1000 iterations. Then, we adjust the new variance for each parameter 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, this implementation reduces to the standard MH algorithm.

[11] See Johannes and Polson (2009) for an in-depth discussion.

[12] We found it difficult to distinguish between frequent small return jumps and more infrequent large return jumps, especially when estimating using returns only. Following Eraker, Johannes, and Polson (2003), we therefore impose an informative prior on jump frequency to favor infrequent but large return jumps when index returns are the only observables.

In short, the idea behind AMH-MCMC 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 local optima, the algorithm moves to new parameter set with a non-zero probability even if the new likelihood is lower than the previous one.

The total likelihood in equation (1.18) is simply a by-product of particle filtering. According to Malik and Pitt (2011), the total likelihood associated with the particle filter conditional on a set of parameters can be expressed as a function of the unnormalized weights:

$$f_1(O_{1:T}|\Theta) = \prod_{t=1}^{T} \left\{ \left(1/N \sum_{j=1}^{N} {}_1\omega_{t+1}^{j} \right) \left(1/N \sum_{j=1}^{N} {}_2\omega_{t+1}^{j} \right) \right\} \tag{1.19}$$

As a special case, for the SIR we only have first stage weights. The total likelihood therefore reduces to:

$$f_1(O_{1:T}|\Theta) = \prod_{t=1}^{T} \left\{ 1/N \sum_{j=1}^{N} \omega_{t+1}^{j} \right\} \tag{1.20}$$

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 (1.20) does not require us to record the latent variable path, we proceed by simply conducting the parameter search first without recording the latent processes, and then back out latent states with the result of the parameter search. 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.

1.3.4 Estimation Using Returns Data

As mentioned above, the algorithm can be applied to different sources of information, which corresponds to different likelihoods $f_1(O_{t+1}|L^i_{t+1})$. First consider estimation based on returns, where we simply use returns as the observables. Using equations (3.21) and (3.22) this gives:

$$\begin{aligned} f_1(O_{t+1}|L^i_{t+1}) &= f_1(R_{t+1}|L^i_{t+1}) \\ &= \frac{1}{\sqrt{2\pi V^i_t}} \exp\left\{ -\frac{1}{2} \frac{\left[R_{t+1} - (r_t - \delta_t - \frac{1}{2}V^i_t + \eta_s V^i_t - \lambda\bar{\mu}_s + J^s_{t+1}B_{t+1})\right]^2}{V^i_t} \right\} \end{aligned} \tag{1.21}$$

The Appendix provides additional details on particle filtering based on returns data. See also Christoffersen, Jacobs, and Mimouni (2010) for a related implementation on the SV model.

1.3.5 Estimation Using Option Data

We now consider model estimation using option data only, that is, estimation that does not consider the underlying returns.

The existing literature that estimates option pricing models using large panels of options 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 Andersen, Fusari, and Todorov (2015a) 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 viable, but the resulting parameter estimates will either reflect return-based 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.

1.3.5.1 Computational Constraints

The instantaneous variance follows the transition equation (3.22), 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 as straightforward as the one using returns; however, it encounters significant computational constraints. The measurement density now corresponds to equation (3.23). Rather than one return for each time period t, we now have (in our sample) up to thirty option prices available at time t. The likelihood for the i^{th} particle at time t can be calculated as:

$$\begin{aligned} f_1(O_{t+1}|L^i_{t+1}) &= \left(\prod_{h=1}^{H_t} f_1(C_{t,h}|C^M_{t,h}(V^i_t|\Theta)) \right)^{1/H_t} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma_c} \right) exp \left(- \frac{\sum_{h=1}^{H_t} (C_{t,h} - C^M_{t,h}(V^i_t|\Theta))^2}{2\sigma_c^2 H_t} \right) \end{aligned} \tag{1.22}$$

We take the square root of H_t in order to normalize the likelihood with respect to the different numbers of options on each day. The total likelihood for the entire sample summing over all particles is:

$$f_1(O_{1:T}|\Theta) = \prod_{t=1}^{T} \left\{ 1/N \sum_{i=1}^{N} \left[\left(\frac{1}{\sqrt{2\pi}\sigma_c} \right) exp \left(- \frac{\sum_{h=1}^{H_t} (C_{t,h} - C^M_{t,h}(V^i_t|\Theta))^2}{2\sigma_c^2 H_t} \right) \right] \right\} \tag{1.23}$$

When computing this likelihood, 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. Nonetheless, this computation encounters significant computational constraints. 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 up to 30 options per day, and 10,000 particles leads to approximately 1,500,000,000 computations of the option price in each function evaluation, which is computationally infeasible.

1.3.5.2 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 (1.22), 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:

$$ISV_t = \underset{ISV_t}{\arg\min} \sum_{h=1}^{H_t} (C_{t,h} - C_{t,h}^M(ISV_t|\Theta))^2 \tag{1.24}$$

To illustrate how this sum of squared errors changes with particle values, we select a typical day in our sample, December 1, 2015. Figure 1.2 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 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. 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.

After we obtain ISV_t, we fit a parabola using the grid points, and then we compute the SSE for each particle: $SSE_{i,t} = a_{1,t} + a_{2,t} * (V_t^i - ISV_t)^2$. The likelihood for particle i can now be calculated as follows:

$$f_1(O_{t+1}|L_{t+1}^i) = \left(\frac{1}{\sqrt{2\pi}\sigma_c}\right) exp(-\frac{SSE_{ISV,t}(\hat{b}_t * (V_t^i - ISV_t)^2 + 1)}{2\sigma_c^2 H_t}) \tag{1.25}$$

which gives

$$f_1(O_{1:T}|\Theta) = \prod_{t=1}^{T} \left\{ 1/N \sum_{i=1}^{N} \left[\left(\frac{1}{\sqrt{2\pi}\sigma_c}\right) exp(-\frac{SSE_{ISV}(\hat{b}_t * (V_t^i - ISV_t)^2 + 1)}{2\sigma_c^2 H_t}) \right] \right\} \tag{1.26}$$

where we use $SSE_{ISV,t} = a_{1,t}$ and $\frac{SSE_{i,t}}{SSE_{ISV,t}} = 1 + b_t(V_t^i - ISV_t)^2$ with $b_t = \frac{a_{2,t}}{a_{1,t}}$. 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 parallelize the computation on $t = 1 : T$. Note also that while the likelihood in equations (1.25) and (1.26) 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.

1.3.5.3 How Reliable is the ISV Particle Filter?

While our implementation of the particle filter is computationally much less expensive, the question arises how accurately it approximates the log likelihood for a conventional implementation of the particle filter. The problem is of course that to investigate the quality of this approximation, the computational costs of the conventional particle filter again become an obstacle.

We therefore investigate the quality of the approximation using a subsample. We limit ourselves to the option sample for the year 2015. Figure 1.3 compares the log likelihood for our implementation using ISV and the traditional implementation of the particle filter. We plot the likelihood as a function of each of the five parameters in the SV model. We use 1,000 particles in this exercise, which makes the likelihood less smooth but makes the exercise computationally feasible. The results suggest that our approach is very reliable and provides a good approximation to a traditional, much more computationally expensive, implementation of the particle filter. We have also conducted smaller-scale comparisons using different sample periods and the results are very similar.

1.3.6 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. The relative weights of returns and options in the likelihood are critically important, especially because on every day, we have a single return but a large cross-section of options. The weights 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. Recall that in our option likelihood, we normalize the likelihood with respect to the number of options on each day. Therefore we effectively give equal weight to options and returns in the joint likelihood. This allows us to abstract from the size of the cross-section of options and focus on the composition of the option data. This implementation is somewhat ad-hoc, but the specification of any joint likelihood of options and returns is not entirely guided by theory and therefore to some extent ad-hoc.

Using this approach, the resulting total likelihood for each particle on date t can be expressed as:

$$f_1(O_{t+1}|L^i_{t+1}) = f_1(R_{t+1}|L^i_{t+1}) * (\prod_{h=1}^{H_t} f_1(C_{t,h}|C^M_{t,h}(V^i_t|\Theta)))^{1/H_t} \tag{1.27}$$

where the two components can be computed according to equations (1.21) and (1.22) 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 similar to the approach in Sections 1.3.4 and 1.3.5.

1.4 Empirical Results

We present and discuss estimates from returns and option data jointly, using the approach in Section 1.3.6. We also obtained estimates exclusively based on return data, as well as estimates exclusively based on option data, using the approaches in Sections 1.3.4 and 1.3.5 respectively. To save space we do not discuss these results in detail. Because we occasionally refer to them, we include them in the Appendix.

We discuss our 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 correlated return and variance jumps (SVCJ). We compare our results with estimates from the existing literature. See also Singleton (2006, chapter 15) for a discussion of existing results.

Subsequently we discuss and compare model fit. We present results for the moments implied by the estimated option pricing models, and we use these moments to illustrate important differences between the models. Finally, we discuss preference parameters implied by our estimates.

1.4.1 The SV Model

We start by discussing our estimates for the Heston (1993) square root stochastic volatility (SV) model, which is an important building block in the option valuation literature. The first column of Table 1.2 presents estimates of the SV model. These estimates can be compared with estimates from the existing literature in Panel C of Table 1.3, with the caveat that the estimates in Table 1.3 are obtained using different sample periods.[13] Note that to facilitate comparisons, the estimates in Table 1.3 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.[14]

To further motivate our empirical results, consider these existing estimates of the Heston model in Table 1.3. Panel A of Table 1.3 indicates that a large number of studies report results for the Heston model under the physical measure based on index return data. 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. The studies in Panel C typically face 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. However, note that while the parameter estimates in Pan (2002) are obtained using a very limited cross-section of options, model fit is also evaluated using a wider

[13] In Table 1.3, Panel A contains existing estimates for the Heston model based on returns and Panel B contains estimates based on options.

[14] Compared to the estimates reported in Eraker (2004) and Broadie, Chernov, and Johannes (2007), the estimates of κ and σ in Table 1.3 are multiplied by 2.52 (multiplied by 252 and divided by 100). The estimate of θ in Table 1.3 is multiplied by 0.0252 (multiplied by 252 and divided by 10,000).

cross-section. Eraker (2004) uses Markov Chain Monte Carlo (MCMC) based on return and option data. 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, and focus on short-maturity at-the-money calls.

We conclude that existing estimates of the Heston model that use option data face computational constraints, and that they reduce the cross-sectional dimension of the option data set, presumably partly due to these constraints. See Hurn et al. (2015) and Broadie et al. (2007) for related discussions. Clearly some of these constraints can be addressed with additional computational resources, but in itself this is not sufficient. We propose methods that help overcome these constraints and allow us to keep the time-series and cross-sectional dimension of the option data as large as possible.

We show that these methods are helpful to estimate risk premia as well as other model parameters. Consider our estimate of the variance risk premium. The posterior mean of η_v, the parameter that characterizes the diffusive variance risk premium, is equal to 1.0836 in Table 1.2. This estimate has the expected sign and it is highly statistically significant. The existing literature does not contain many estimates of η_v that are based on options and returns jointly, but nevertheless our estimate differs from the existing literature in important ways. It is smaller and statistically significant. Several existing studies report estimates of η_v that are not statistically significant, see for instance Eraker (2004) and Broadie et al. (2007). Pan (2002) and Hurn et al. (2015) report statistically significant estimates. The sample used in Hurn et al. (2015), who also obtain statistically significant estimates of η_v, is similar to ours in the sense that it includes options with widely different moneyness and maturity. Broadie et al. (2007) note that insignificant estimates in the existing literature may be due to a flat volatility term structure of volatility and/or the absence of options with longer maturities in the data. Panel B of Table 3.1 indicates that the term structure of implied volatility in our sample varies by moneyness.

Our estimate of η_v is also smaller than existing estimates based on options and returns.[15] The small magnitude of the estimate of η_v has important implications because the variance risk premium η_v defines the relation between physical and risk neutral mean reversion and long-run variance. The estimate of the long-run physical variance of returns θ in Table 1.2 is 0.0351. This gives a risk-neutral long-run variance of $\theta^Q = \kappa\theta/\kappa^Q == \kappa\theta/(\kappa - \eta_v) = 0.0706$. Consistent with the literature, the risk-neutral variance considerably exceeds the physical variance due to a positive estimate of η_v. This finding is also consistent with nonparametric evidence, see for example Bollerslev, Tauchen, and Zhou (2009).

Now consider mean reversion in the variance. The estimate of κ in Table 1.2 is 2.156. The risk-neutral mean reversion parameter $\kappa^Q = \kappa - \eta_v$ from Table 1.2 is thus equal to 2.156-1.084=1.072. The finding that the physical and risk-neutral mean reversion estimated from options are smaller than the physical mean reversion from returns confirms existing findings. However, our estimate of physical mean reversion based on joint estimation is much smaller than the estimates in Panel C of Table 1.2.[16] Note that this difference is unlikely to be due to the sample period: the return-based estimate of κ is consistent with existing estimates in the literature. The option-based estimate suggests that our estimate of κ in Table 1.2 is due to our option sample. One possible explanation of our findings is that in a joint estimation exercise, the variance process is estimated to be much more persistent under the physical measure when option samples include options with longer maturities.

We conclude that our estimate of the price of diffusive variance risk η_v differs from the existing literature both with respect to its magnitude and statistical significance, and this has important implications for the estimates of the mean reversion in variance. Moreover, our estimates of κ^Q and θ^Q are more plausible than in some existing studies.

Our estimates also differ from existing results with respect to the parameters characterizing the tails of the distribution. In Table 1.2, the estimate of ρ, the correlation between the

[15] Our estimate is also smaller than the estimate in Bates (2000), which is obtained using options only, but additionally imposes a dynamic constraint on the spot variance.

[16] Note that κ^Q is negative in some existing papers, which implies a negative long-run variance θ^Q.

return and variance innovations, is -0.9161, much more negative than estimates in existing studies in Panels B and C of Table 1.3. Once again the estimates based on returns and options separately show that this finding is driven by the option sample. Andersen, Fusari, and Todorov (2015a) report an estimate of ρ of -0.934 when estimating the SVCJ model. Our findings are thus more consistent with estimates from a study that mainly relies on option-based information.

The estimate of σ in column 1 of Table 1.2 is equal to 0.4262. For this parameter, the difference with the existing literature is somewhat more subtle. Note that the estimate of σ based on options is 0.4091 and the estimate from returns is 0.5430. The estimate of σ based on options and returns jointly in Table 1.2 is therefore closer to the option-based estimate, which is smaller than the return-based estimate. Now consider the estimates of σ in the existing literature in Table 1.3. The option-based and joint estimates in Panels B and C are generally larger than the return-based estimates in Panel A. In studies that allow for a direct comparison, such as Eraker (2004), Bakshi, Cao, and Chen (1997) and Christoffersen, Jacobs, and Mimouni (2010), the option-implied variance process is more variable than the return-implied variance process. Our findings suggest that if the variance dynamics are imposed and large option panels are used in estimation, this finding does not necessarily obtain. This finding is consistent with the results in Andersen, Fusari, and Todorov (2015a), who find that the option-implied variance path is less variable than the physical variance path estimated nonparametrically using high-frequency returns data. Note that these findings on σ are related to our findings on ρ. Skewness is determined by ρ but kurtosis is increasing in $|\rho|$ and σ. Hence, our more negative estimates of ρ also make the smaller estimates of σ possible. This again confirms the findings in Andersen, Fusari, and Todorov (2015a).

Note from Table 1.2 that most parameters are rather precisely estimated, not only the parameter η_v which characterizes the variance risk premium. Tables 3.3 indicates that we also get relatively small posterior standard deviations when estimating the model using options only, but that the confidence intervals are wider when using returns only in estimation.

Figures 3.1 and 1.5 provide more insight into this finding. These figures plot the trace for the return-based and option-based estimation respectively of the SV model. Our implementation of the particle filter uses 10,000 particles. For the return-based estimation, we use 50,000 iterations. For the option-based estimation, we use 2,500 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. The figure clearly indicates rather rapid convergence for both estimation exercises and for all parameters. The figures also clearly indicate that the estimation problem is much more stable for the option-based estimation, which explains the smaller posterior standard deviations.

Figure 3.2 provides additional evidence on the SV model. It compares the filtered variance paths based on returns, options and joint estimation. The filtered variance path from options is smoother than that from returns, which is consistent with the lower values of option-based κ and σ. The filtered variance path based on returns and options jointly in the third panel is more similar to the option-implied path in the second panel.

We conclude that our estimates of the parameters characterizing the tails of the distribution in the Heston model, σ and ρ, also differ in important ways from much of the existing literature.

1.4.2 The SVJR Model

We now discuss our estimates of the stochastic volatility model with return jumps (SVJR) model in column 2 of Table 1.2. The parameters in both tables are annualized.

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 if the tail parameters ρ and σ were different from the SV model. However, the estimates of these parameters are similar to those for the SV model in column 1.

Our estimates indicate the presence of relatively large, but infrequent jumps, which are negative on average, with risk-neutral average jump sizes that are more negative than physical jump sizes. The physical jump size μ_s is equal to -0.0134, the jump risk premium η_{J^s} is equal to 0.0236, and the risk-neutral average jump size $\mu_s^Q = \mu_s - \eta_{J^s}$ is therefore equal to -0.0370. These jumps occur on average less than once per year (λ=0.895). The standard deviation of the jumps σ_s is 4.91%.

The literature contains several estimation results for this model, but in most cases they use different information compared to our approach. First, as with existing estimates of the SV model, many of the existing estimates are obtained from returns. 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.

Relatively few studies offer evidence based on options or options and returns. Pan (2002), 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 (2004) 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 the 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 findings on return jumps are rather similar to those of Broadie, Chernov, and Johannes (2007).

Table 1.2 also presents 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. The average risk premium in the SVJR model is larger than in the SV model. In the SVJR model, approximately 2% of the equity risk premium is due to jump risk, with the remainder 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 also reports roughly similar estimates, with price jump risk premia that contribute about 3.5% per year to an overall equity premium of 9%.

A variance decomposition shows that in the SVJR model, approximately 6% of the variation in returns is due to jumps. Consistent with the existing literature, our results therefore indicate that return jumps are relatively more important for risk premiums than for explaining overall return variation.

1.4.3 Jumps in Variance

The third and fourth columns of Table 1.2 present estimates of the SVJV and SVCJ models, respectively. The SVCJ model in column 4 contains jumps in returns and variance 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). We also report on the SVJV model in column 3 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. This model also allows us to further comment on the relative importance of jumps in returns and jumps in variance for modeling returns and options.

For the SVJV model, the estimate of the risk neutral average jump size $\mu_v^Q = \mu_v - \eta_{J^v}$ is equal to 6.46%. 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 0.919, implying these jumps occur just less than once per year. When adding jumps in returns in the SVCJ model in column 4, the jump size is very similar but the frequency of the jumps decreases. The estimate of the correlation between the return and variance jump is negative, as expected, at -0.5030.

Our estimate of μ_v^Q is very 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. Furthermore, existing estimates of variance jumps seem to be much more consistent in the literature, even though these studies use very different sample periods, compared to the existing estimates of jumps in returns. A plausible reason for the robustness of these estimates is that variance jumps are better identified in the data.

Our estimates also provide additional insight into the diffusive variance risk premium η_v. Recall that in the SV model, we find that while our positive estimate is consistent with the literature, it is also estimated precisely, which contradicts many existing findings. When adding return jumps, our estimate of η_v is still positive and statistically significant, but smaller. However, when adding variance jumps, the sign of η_v may change and the statistical significance decreases. This suggests that identification of the diffusive variance risk premium is not a problem in the absence of variance jumps. However, it is difficult to separately identify diffusive and jump variance risk premiums. This conclusion may be due to our use of daily data. Bandi and Reno (2016) show that identification in models with jumps in returns and volatility is facilitated when using high-frequency intraday data.

We conclude that our estimates of variance jump parameters are overall consistent with existing estimates. The decomposition of the average equity risk premium in diffusive and jump components in the SVCJ model is similar to the one in the SVJR model. Approximately 2% of the equity risk premium is due to jump risk, and 7% is due to diffusive risk.

1.4.4 Model Fit

The log likelihoods in Table 1.2 are very useful as indicators of the importance of return and variance jumps for modeling returns and options. Using a likelihood ratio test, the more complex models are always statistically supported by the data. We also need to account for jumps in returns as well as jumps in variance when returns and options are modeled separately. These findings confirm the importance of jumps in returns (Pan, 2002; Bates, 2000) and jumps in variance (Broadie et al., 2007).

Table 1.2 indicates that 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 variance in the SVJV and SVCJ models that result in very large improvements in the log likelihood. Moreover, Jumps are especially important when jointly modeling returns and options. This suggests that these jumps are especially useful to model risk premia, consistent with the findings in Pan (2002).

Finally, it is worth noting that despite the large improvements in log likelihood, the richer models with jumps do not substantially outperform the simple SV model in terms of RMSE. The RMSE for the SV model is approximately \$3, and the jump models only improve the fit by 3 cents. This confirms the results of Eraker (2004). Singleton (2006) notes that these findings may be due to the fact that we optimize the likelihood rather than minimizing root mean squared error. We leave a more detailed investigation for future research.

1.4.5 Option-Implied Moments

A useful way to highlight the differences between the estimated models is to study the conditional moments. We present results for the conditional variance of returns, the conditional covariance between returns and variance, and the conditional variance of variance. These moments are provided in Panel A of Table 1.4.

Panel B of Table 1.4 presents the averages over our sample for the daily moments for the SV, SVJR, SVJV and SVCJ models. Clearly the differences between the models are small for the conditional variance and conditional covariance between returns and variance. However, the models with jumps in variance are characterized by a much higher variance of variance.

It is also useful to consider how these moments change over time. Figure 1.7 presents the time series of the conditional variance of variance, which peaks in the financial crisis for all four models.[17] Figure 1.7 indicates that while there are large differences across models

[17] We do not report results for the time series of the conditional variance and the conditional covariance between returns and variance to save space. We simply note that these paths look very similar across models. The path of the conditional variance can be gleaned from Figure 3.2. For the conditional covariance, note

in the average of the conditional variance of variance, as evident from Panel B of Table 1.4, the paths of the conditional variance of variance are highly correlated across models.

1.4.6 Implied Preference Parameters

We now provide a more structural interpretation of the estimated equity and variance risk premiums. This can be done for both the diffusive and jump risk; here we limit our analysis to the diffusive risk. Recall that the diffusive equity premium in the models we estimate is given by $\gamma_t = \eta_s V_t$. This equity risk premium and the corresponding risk neutralization can be motivated by a pricing kernel derived from the preferences of a representative agent.

In the case of a representative agent with power utility, the diffusive parameter η_s that characterizes the equity risk premium is equal to the rate of relative risk aversion. Our estimate of η_s for the SV model in Table 1.2 is equal to 2.502. This estimate is similar in magnitude to estimates obtained in other studies that use both options and returns, such as Hurn et al. (2015) and Eraker (2004), for example. However, it is also well-known (see for example Bates, 2000) that in the power utility case, the parameter η_v that characterizes the reduced form diffusive risk premium $\eta_v V_t$ is given by $\eta_v = -\phi\rho\sigma$, where ϕ is the rate of relative risk aversion. It is therefore possible to use the model estimates to calibrate another, variance-based, estimate of the risk aversion parameter, which should in principle be equal to that implied by the equity premium. Using our estimates of η_v, ρ, and σ for the SV model in Table 1.2, we obtain an estimate of risk aversion equal to 2.775, very close to the 2.502 estimate implied by the equity premium. We repeated this exercise using the estimates in other studies, and in all cases we found that the estimates of risk aversion implied by the diffusive variance risk premium are a multiple of the estimate implied by the diffusive equity risk premium. This finding suggests that our estimates may benefit from improved identification due to the larger cross-sections of options used in estimation.

Christoffersen et al. (2013) consider a more general utility specification in which the

from the equations in Table 1.4 that the time variation in the path is determined by the conditional variance. The conditional covariance therefore reaches a minimum during the financial crisis for all models.

utility of the representative agent is directly affected by the variance. In this setup the reduced form η_s and η_v parameters are related to the rate of relative risk aversion ϕ and the positive parameter that captures the variance aversion ξ as follows:

$$\eta_s = \phi - \xi\sigma\rho \tag{1.28}$$

$$\eta_v = -\eta_s\sigma\rho + (1-\rho^2)\sigma^2\xi \tag{1.29}$$

Solving for ϕ and ξ using the parameter estimates for the SV model in Table 1.2, we find that the implied ϕ and ξ are equal to 1.073 and 3.660 respectively, which means that both parameters have the theoretically expected sign. The values are also plausible. We repeated this exercise using parameter estimates from existing papers. Most existing estimates imply that one of the parameters has a sign that is inconsistent with economic theory. This finding again suggests that our estimates may benefit from improved identification due to the presence of options with different moneyness and maturity in the sample.

1.5 The Information in Options and Returns

Using the ISV particle filter, we were able to obtain model estimates using big panels of options. We now discuss how restricting the option sample in the moneyness and/or maturity dimension affects parameter inference and identification, and how it impacts the relative weighting of options and returns.

1.5.1 Likelihoods for Different Option Samples

The motivation for our newly proposed implementation of the particle filter is that while cross-sectional information is helpful for identification of certain model parameters, conventional implementations of the particle filter, as well as other existing estimation methods, are computationally very expensive when using large option panels. As previously discussed, this is one of the reasons why existing studies have restricted the time-series and cross-sectional dimension of the data. A natural question is if restricting the data set in this way affects inference, model properties, and identification.

These implications can in principle be studied in various ways. One approach is to use Monte Carlo techniques: simulate repeated option data samples for a given model and then investigate if we can retrieve the known parameters. This approach is very instructive but also has some clear disadvantages. We therefore use a more direct approach. We simply inspect the log likelihood as a function of the parameters of interest. Because of the high cost of computing the log likelihood, this approach is not feasible using many existing estimation techniques, such as the conventional implementation of the particle filter. However, with our implementation of the particle filter, it is relatively straightforward to evaluate the likelihood.

Figure 1.8 illustrates these implications using the parameters of the SV model. We present the log likelihood as a function of the parameters in the SV model: κ, θ, σ, ρ, and the price of diffusive variance risk η_v. In each of the panels in Figure 1.8, the solid line represents the log likelihood for the option sample we use in our empirical exercise in Section 1.4, which we refer to as ALL. This dataset is described in Panel B of Table 3.1. The two other lines present the log likelihoods as a function of the parameters for more restricted option samples. Following the existing literature, we restrict either the moneyness dimension or the maturity dimension. ATM represents the sample with ATM options but all maturities and SM represents the sample with short maturity options but all moneyness. For the ATM sample, we exclude options with moneyness outside the 0.98-1.02 range. We select up to six options for each of the five maturity bins in Table 3.1. For the SM sample, we rely on a sample with options with maturity less than 30 days in Panel B of Table 3.1.[18] Note that the ATM and SM datasets are somewhat smaller than the one used in our benchmark analysis, but because the likelihood is scaled back by the number of options in the dataset, this does

[18]Because the sample does not contain options with maturity less than 5 days, if options with maturity less than 30 days are not available we include the options for the shortest maturity longer than 30 days.

not affect the comparison in Figure 1.8. The critical difference between the datasets is the characteristics of the options in the sample. Panel C of Table 3.1 presents the sample sizes as well as descriptive statistics.

The top two panels of Figure 1.8 show that the log likelihood for the ATM sample is very similar to the log likelihood for the ALL sample as a function of the κ and θ parameters. However, when only using short maturities (the SM sample), the log likelihood is very flat, which may complicate inference.[19] The two panels in the middle row of Figure 1.8 show the log likelihood as a function of the σ and ρ parameters. Overall, the results confirm our intuition. The likelihood is flatter for the ATM sample than for the ALL sample. However, Figure 1.8 clearly indicates that this problem is much more serious for the skewness parameter ρ than the kurtosis parameter σ. Moreover, including different maturities also clearly facilitates inference on the skewness parameter.

Finally, the bottom panel presents the results for the price of diffusive variance risk η_v. The results are consistent with those for the other parameters.

Figure 1.9 reports on a similar exercise for the jump parameters. The two top panels report on jumps in returns in the SVJR model. The two bottom parameters report on jumps in variance in the SVJV model. In each case, the left panel reports on jump frequency and the right panel reports on jump size. The most important observation is that Figure 1.9 confirms the conclusions from Figure 1.8. Limiting the option sample makes it more difficult to conduct inference on the jump parameters. This is especially the case when using options with similar (short) maturities. Figure 1.9 also suggests that the parameters associated with variance jumps are easier to identify than the parameters associated with return jumps. The risk premium associated with the return jump is especially hard to identify.

These findings have important implications for the existing literature. As mentioned

[19]This finding is not necessarily due to the use of short maturities, but rather to the absence of multiple maturities in the sample. We verified that similar problems obtain when only long-maturity options are included in the sample. When using a single maturity, it is difficult to estimate mean reversion, which also affects the estimation of other parameters. We present results for samples with short maturities only because of the existing literature.

above, many existing studies use samples that consist exclusively of short-maturity and/or at-the money options. One of the reasons for these restrictions is that estimation is prohibitively costly with large option panels. Our findings show that sample composition impacts inference, and using larger samples requires new estimation approaches such as the one we propose. Moreover, our findings suggest that with very restricted samples, identification of some of the parameters may be a problem. This may explain the substantial differences between parameter estimates in the existing literature, although differences in sample period as well as other variations in the empirical setup may also play a role.

1.5.2 Returns and the Cross-Section of Options

It is well-known that when using both return and option data in a joint estimation exercise, in some cases the data heavily weigh the information in the options. In the existing literature, this is mainly due to the fact that we have only one return per day but a large number of options. However, recall that our approach effectively gives equal weights to returns and options to avoid this imbalance. In our implementation, it is therefore the cross-sectional composition of the option sample that determines the relative weight of returns and options. Figure 1.10 illustrates the relative information in the option and return data in a joint estimation exercise, and the impact that the choice of option sample can have on this trade-off. We present the log likelihood as a function of four parameters in the SV model: κ, θ, σ, and ρ. In each of the panels, the broken line represents the log likelihood for the option sample we use in our empirical exercise in Section 1.4. The solid line represents the likelihood for returns.

Figure 1.10 indicates that the option data are much more informative about the model parameters than the return data. The return likelihood is flatter than the option likelihood for all parameters, with the exception of ρ. The return likelihood clearly indicates that values of ρ close to -1 are not supported by the return data, whereas the option data have difficulty differentiating between values of ρ between -0.9 and -1.

Now consider the likelihoods in Figure 1.8, based on samples that are restricted in the moneyness or maturity dimension. These samples result in much flatter option likelihoods as a function of some of the parameters, and this will affect the relative importance of the return and option data in joint estimation. This may explain why existing studies based on returns and options often report estimates of ρ closer to -0.6, whereas our estimate in Table 1.2 is approximately -0.9.

1.6 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 implement an estimation approach that partly overcomes these constraints, by filtering the state variables using weights based on model-implied spot volatilities rather than model prices.

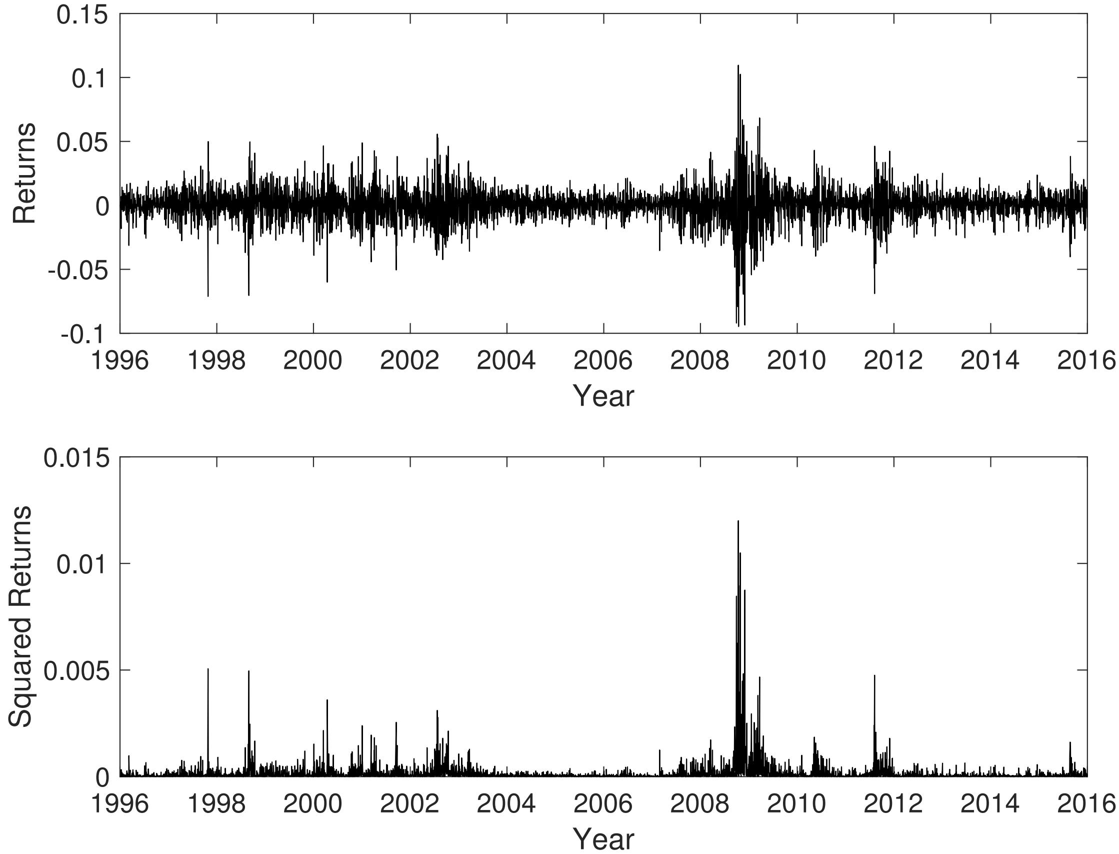
We first demonstrate that this approach is reliable. Then we illustrate our approach by estimating the class of models with jumps in returns and variance (Duffie, Pan, and Singleton, 2000) using twenty years of daily data, and almost thirty option contracts per day with different maturities and widely different moneyness. We confirm that both return and variance jumps are important, but variance jumps more so. Some of our parameter estimates are consistent with the existing literature, but we obtain more precise estimates of the diffusive price of variance risk, and the diffusive parameters characterizing the tails and the persistence of the variance process also differ from most existing studies.

We use our approach to show that sample composition affects estimation results. Restrictive option samples may give rise to flat likelihoods, which complicates inference. When both returns and options are used in estimation, the composition of the option sample may affect the relative importance of returns and options for parameter estimates. While some of these findings are not necessarily surprising, it is important to characterize the implications of restricting the option sample. Such restrictions may affect not only comparisons between existing studies, but also our overall assessment of these models and the option valuation literature.

Moreover, our results provide guidance on the optimal way to restrict the sample, if and when restrictions are required. For instance, we find that for the models studied in this paper, certain restrictions in the moneyness dimension do not have serious consequences, while limiting the sample to a narrow range of maturities is likely to cause identification problems. However, we leave a more detailed study of the optimal composition of the option sample for future research.

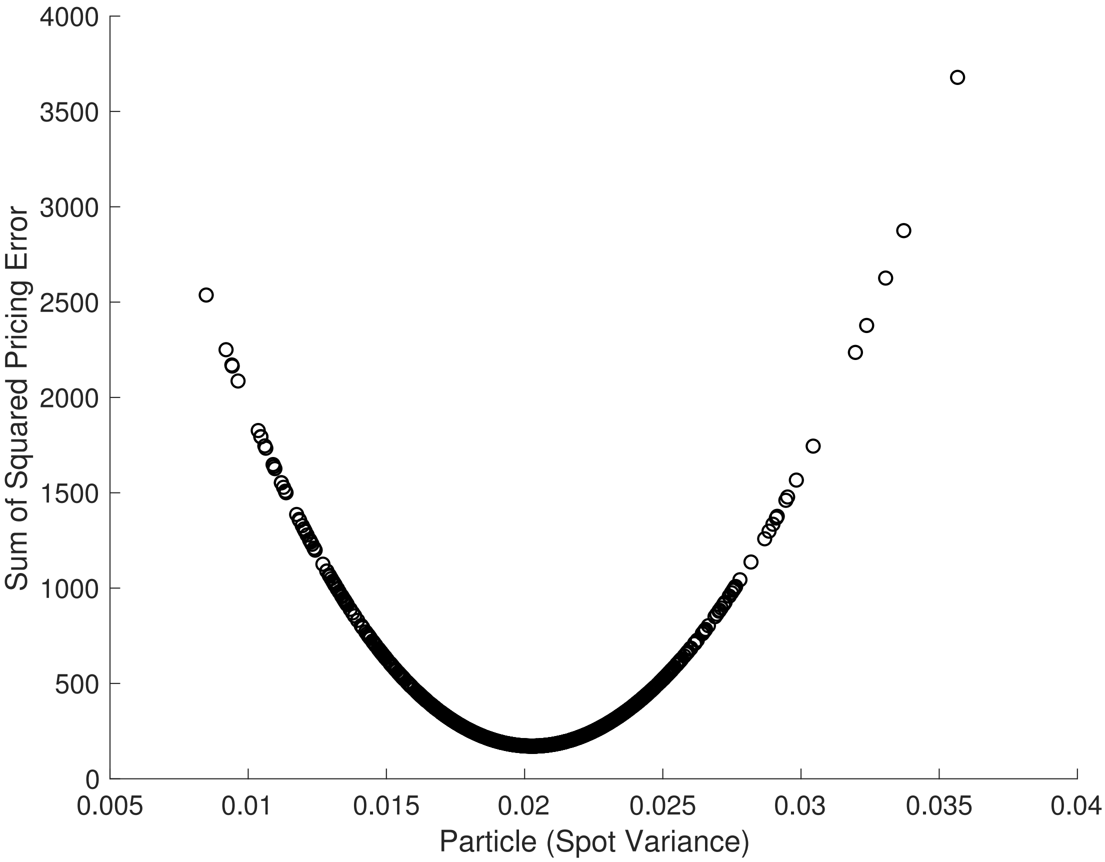
Our results suggest several other extensions. In future work we intend to use the computational advantage of our approach to study richer models with multiple volatility factors (Bates, 2000), time varying jump intensities (Pan, 2002), tail factors (Andersen, Fusari, and Todorov, 2015a), and different parametric specifications of the jump processes (Bates, 2012; Andersen, Fusari, and Todorov, 2017). Our approach can also be combined with the use of high-frequency returns (Bates, 2018) or volatility estimates based on high-frequency data (Andersen, Fusari, and Todorov, 2015b). A more detailed comparison of the computational efficiency and properties of our approach and that of existing methods is also needed. Finally, we plan to investigate the implications of the choice of error specification and loss function (Hurn et al., 2015) for parameter estimates and model performance.

Figure 1.1: Daily Returns and Squared Returns 1996-2015



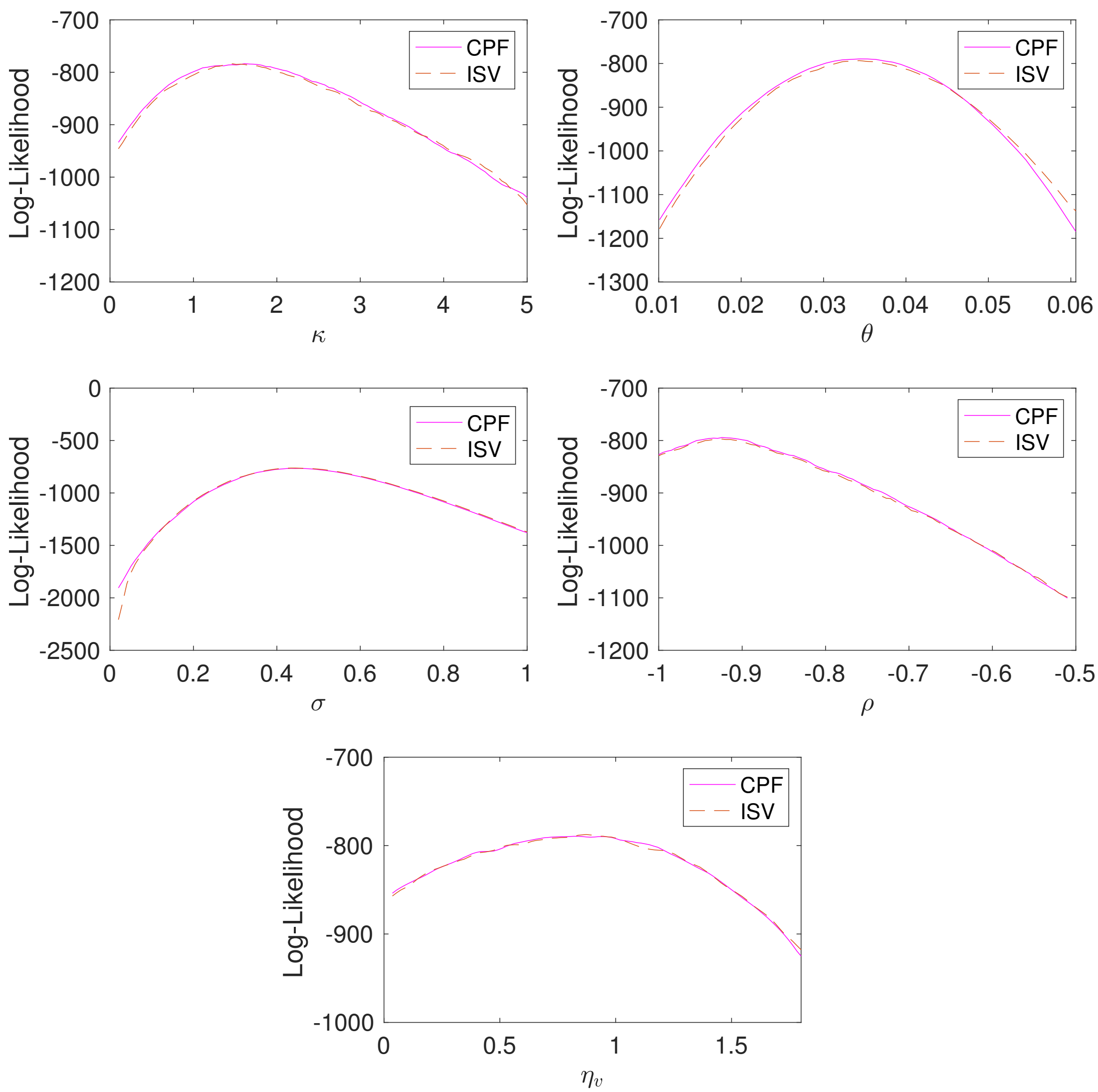
Notes: We plot daily log returns in the top panel. In the bottom panel, we plot squared daily log returns. The sample period is from January 1, 1996 until December 31, 2015.

Figure 1.2: Sum of Option Pricing Errors vs Particles



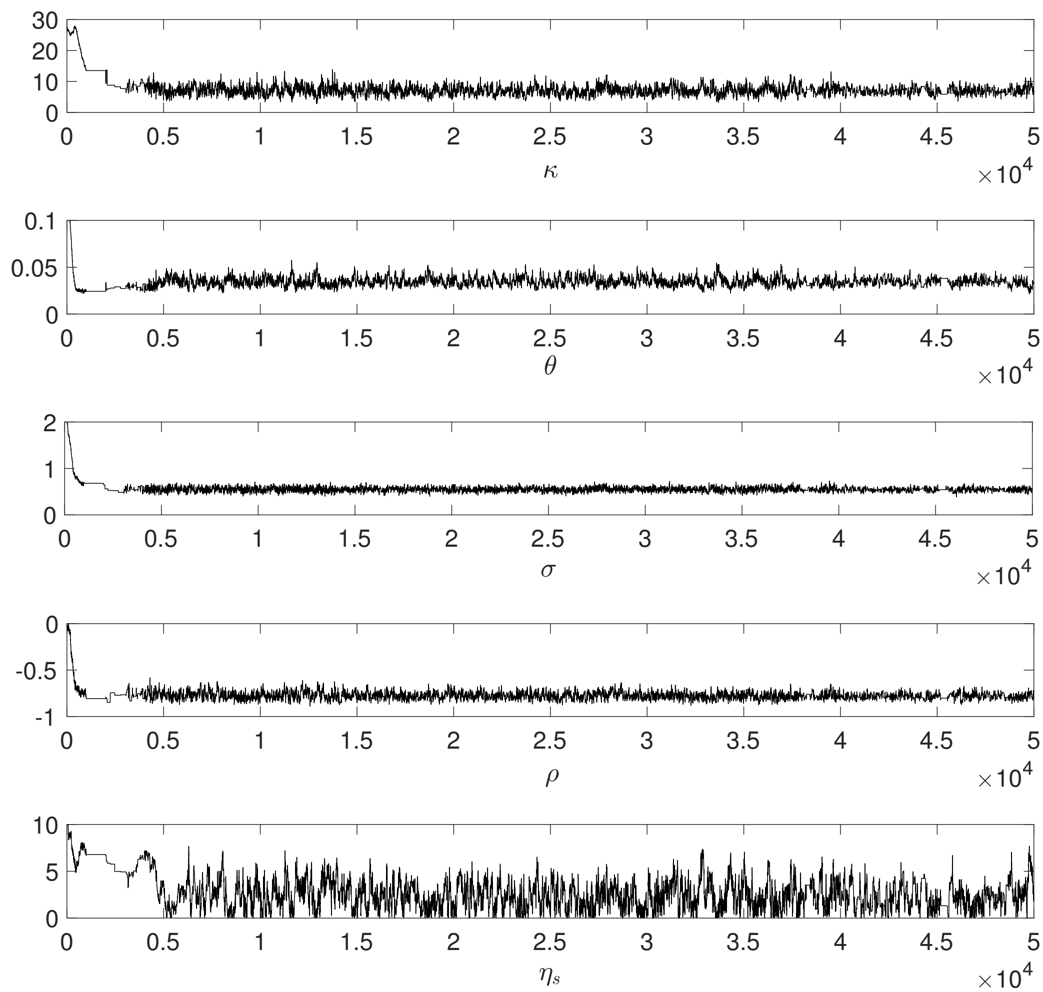
Notes: We show how the sum of squared pricing errors changes with the particle values in the case of the SV model. For each of the particle values (spot variances), we calculate the sum of squared pricing errors. We use option data for December 1, 2015 and the following parameter values: $\kappa = 3$,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

Figure 1.3: Log-Likelihoods Based on the ISV Particle Filtering Method



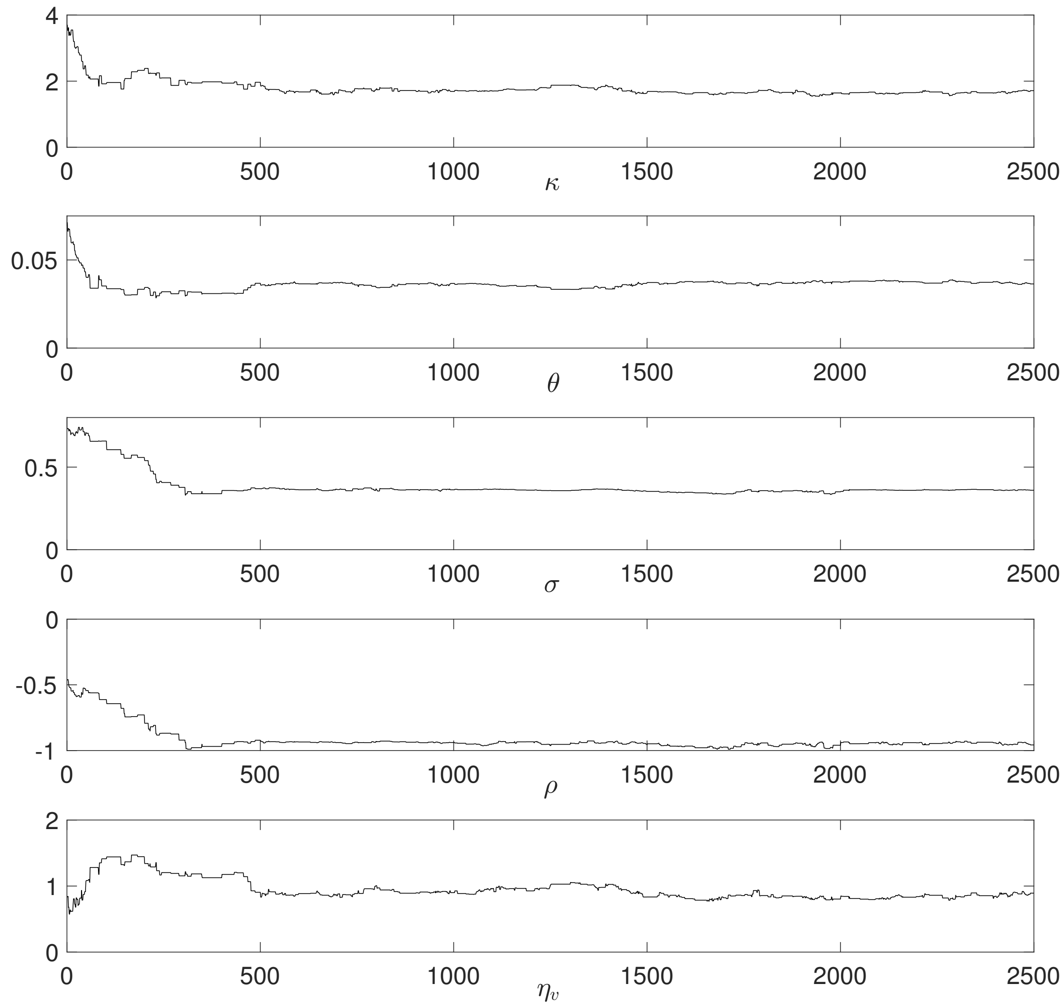
Notes: Using all option data in the sample for the year 2015, we plot the log likelihood based on particle filtering with the ISV method (ISV), and we compare it to the log likelihood using the conventional implementation of the particle filter (CPF).

Figure 1.4: Parameter Trace for SV Model Parameters. Return-Based Estimation



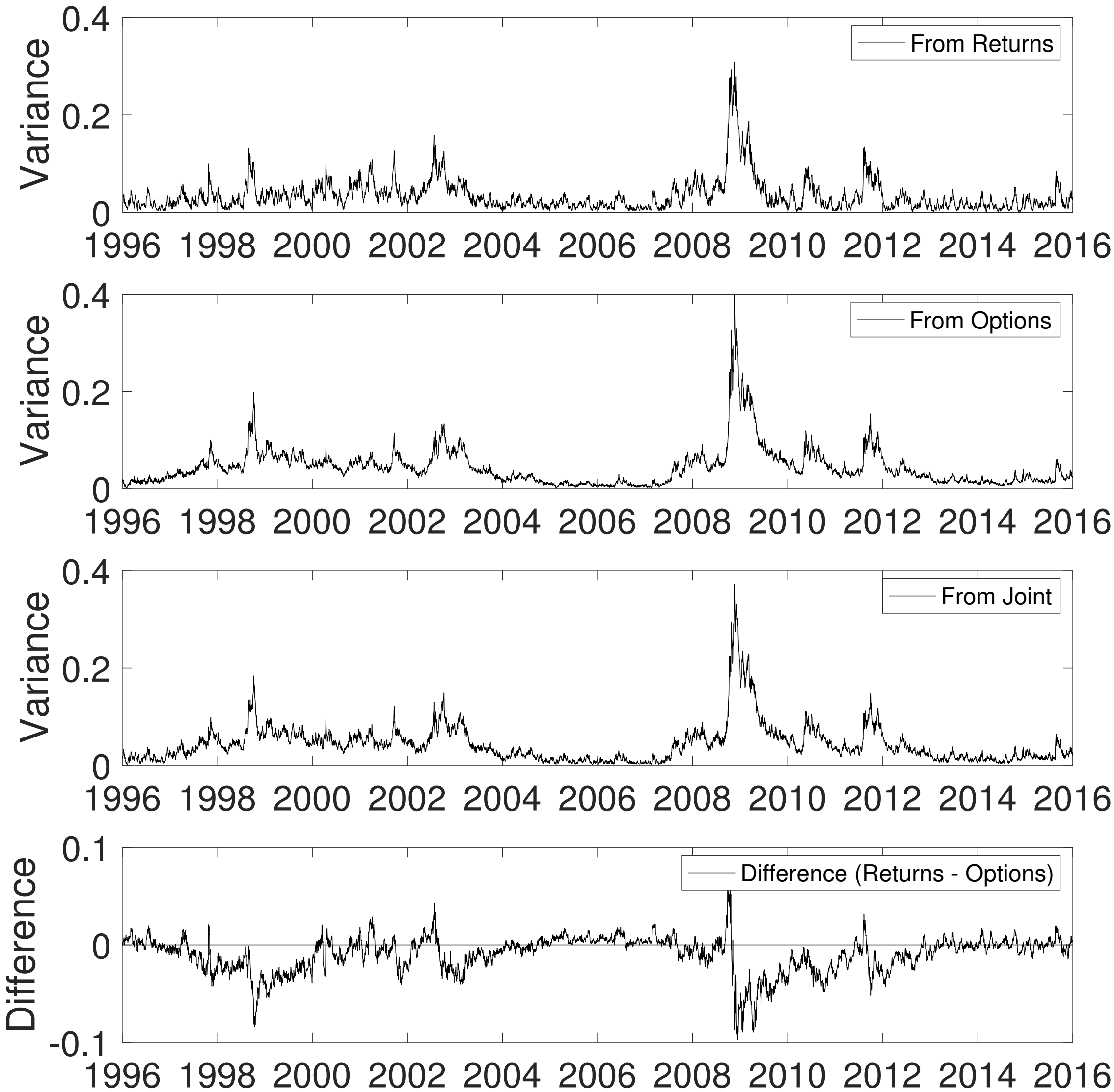
Note: We plot the full traces for each parameter in the SV model. We use 50,000 iterations. The first 1/4 of the iterations are treated as burn-in.

Figure 1.5: Parameter Trace for SV Model Parameters. Option-Based Estimation



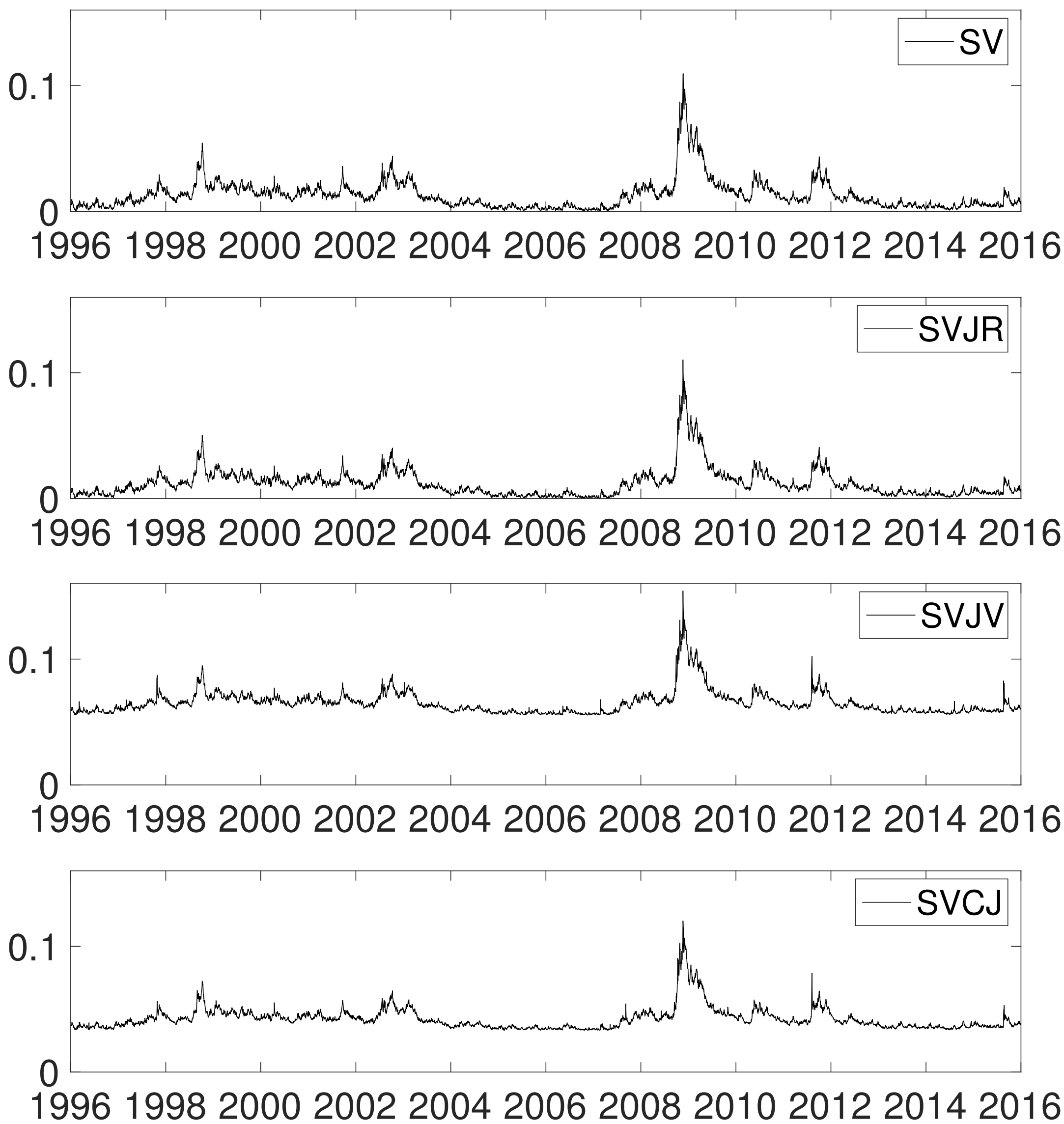
Note: We plot the full traces for each parameter in the SV model. We use 2500 iterations. The first 1/4 of the iterations are treated as burn-in.

Figure 1.6: Filtered Variance Paths. SV Model



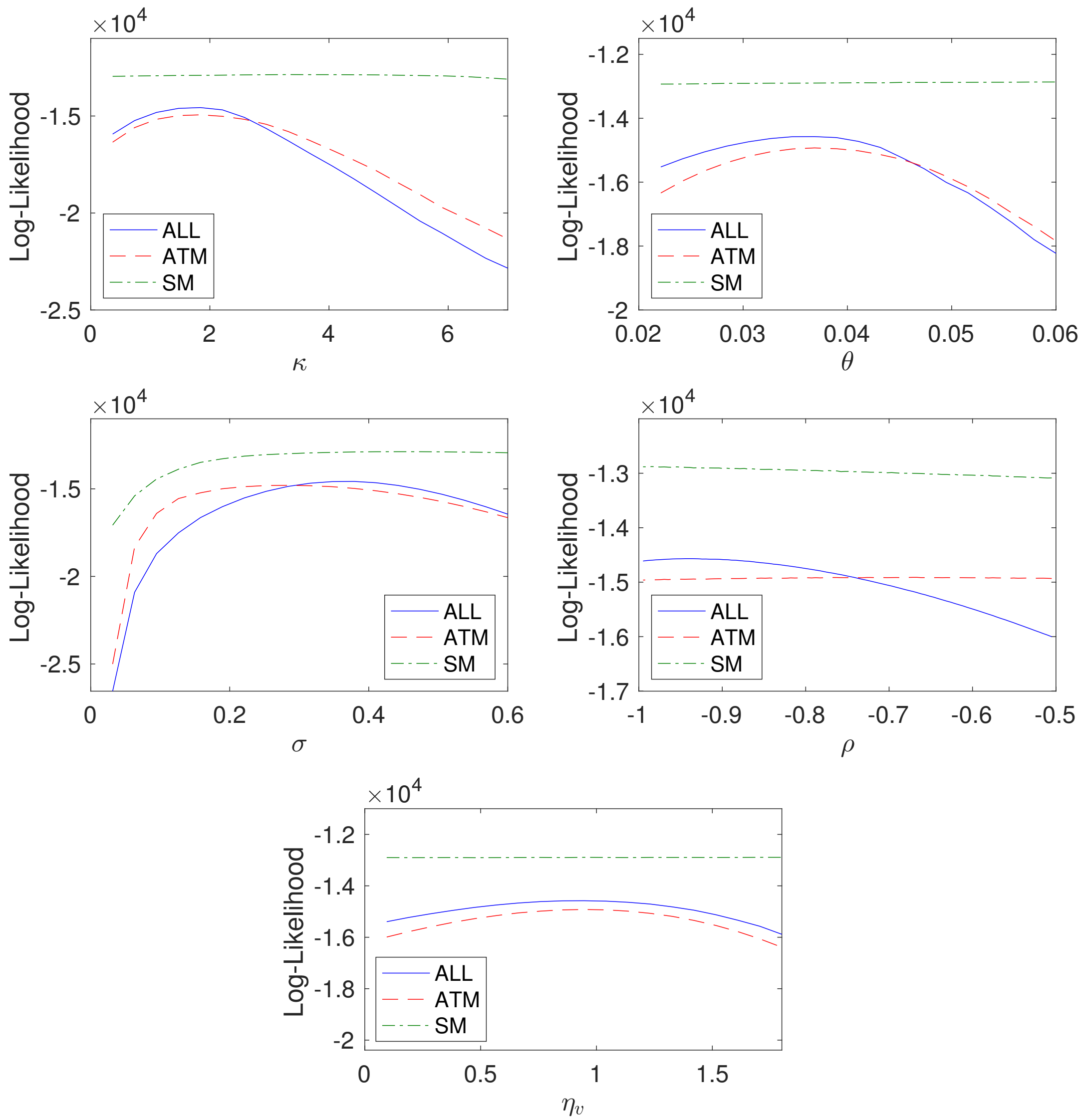
Notes: We plot the filtered variance path estimated from returns in the top panel, the variance estimated from options in the second panel, and the variance from joint estimation in the third panel. The bottom panel plots the difference between the variance estimated from returns and the variance estimated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

Figure 1.7: Conditional Variance of Variance. Various Models



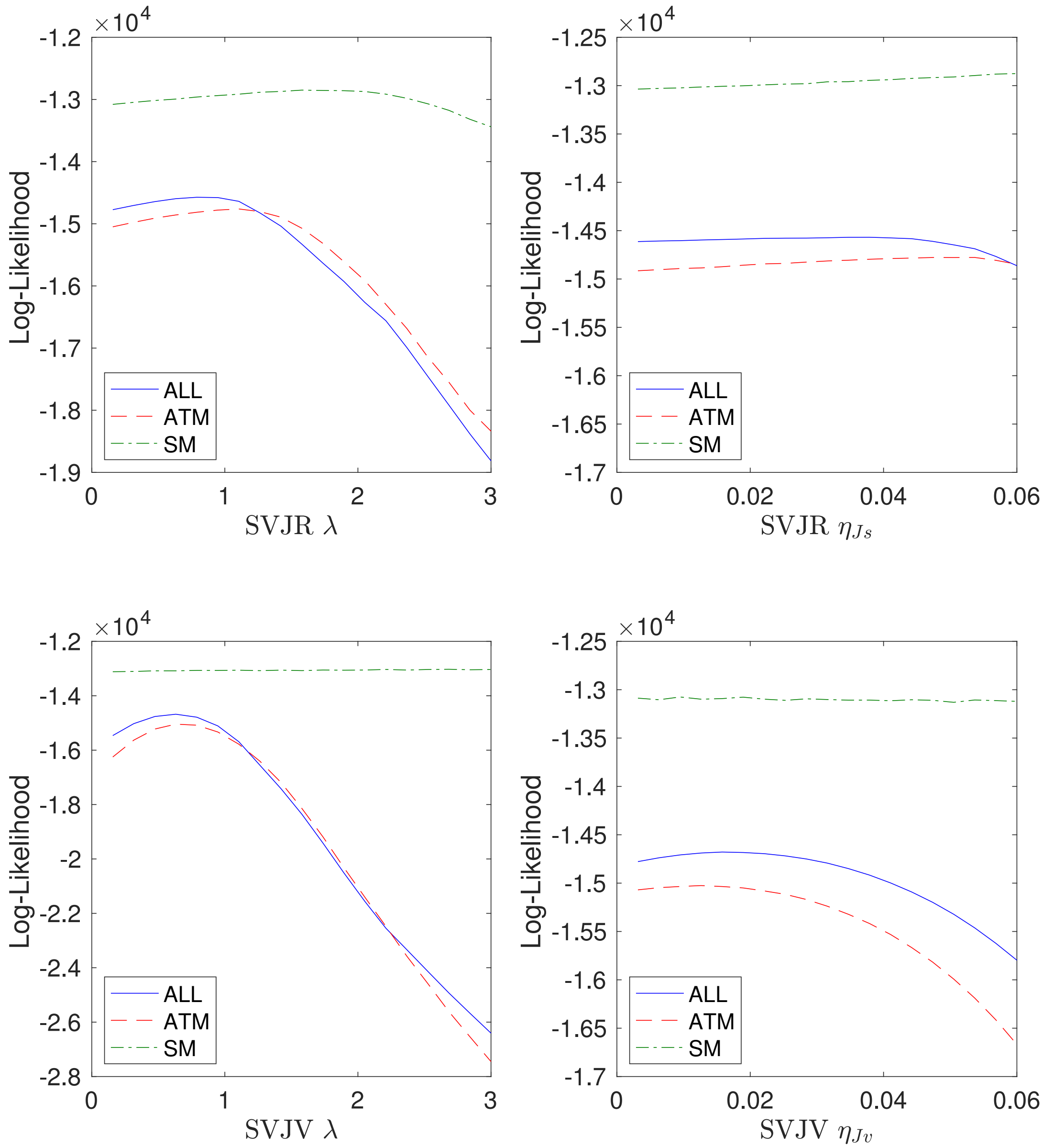
Notes: We plot the conditional variance of variance for four option pricing models.

Figure 1.8: Log-Likelihoods for Restricted Option Samples: The SV Model



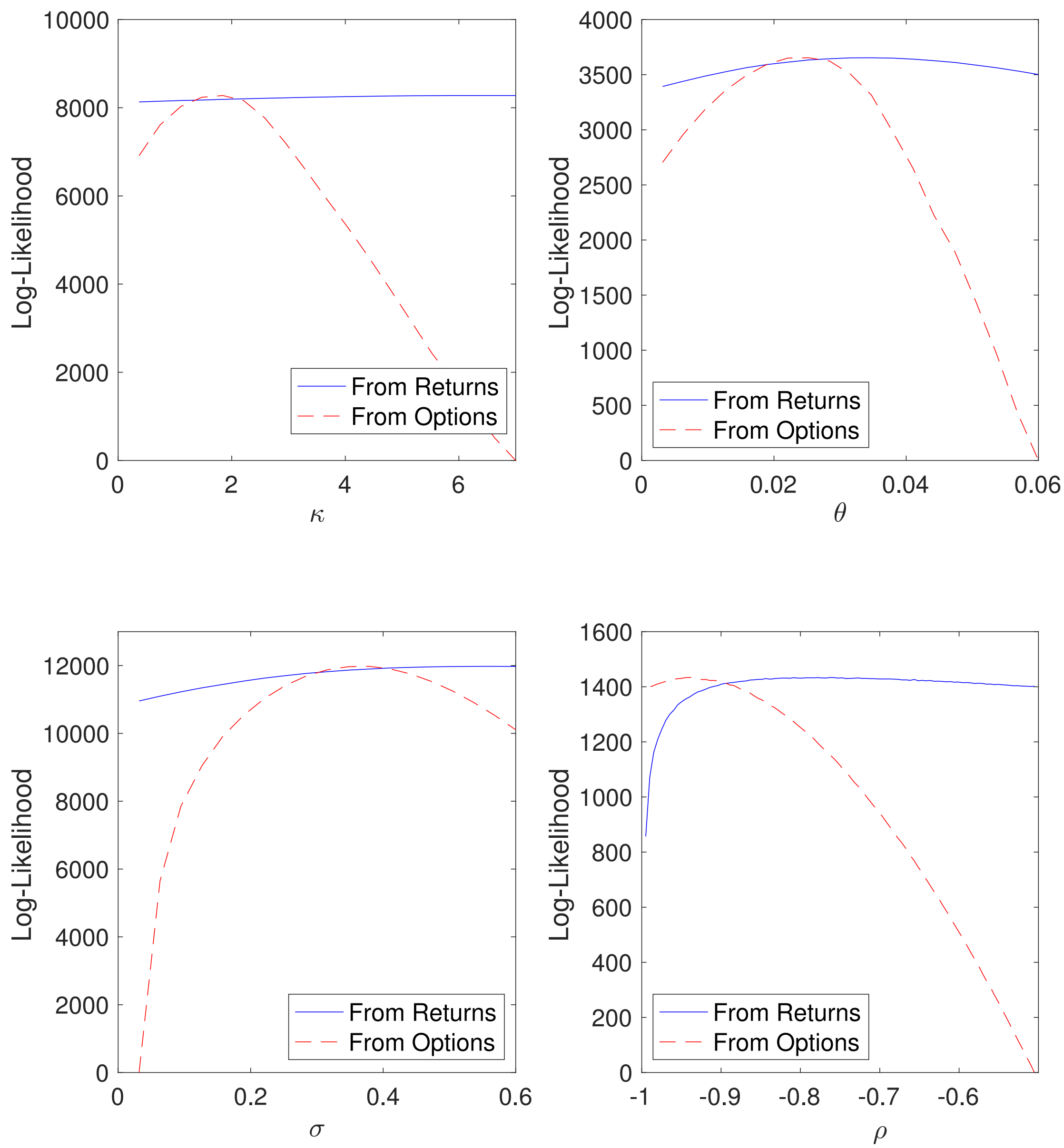
Notes: We plot the log-likelihood for the SV model as a function of the model parameters. We either use the option sample used in our estimation exercise or samples that are restricted based on maturity and/or moneyness. ALL represents the sample with all options. ATM represents the sample with ATM options but all maturities and SM represents the sample with short maturity options but all moneyness.

Figure 1.9: Log-Likelihoods for Restricted Option Samples: Jump Parameters



Notes: We plot the log-likelihood for the jump parameters in the SVJR and SVJV models. We either use the option sample used in our estimation exercise or samples that are restricted based on maturity and/or moneyness. ALL represents the sample with all options. ATM represents the sample with ATM options but all maturities and SM represents the sample with short maturity options but all moneyness.

Figure 1.10: Log-Likelihood for Index Returns and Options



Notes: We plot the log-likelihood for index returns vs index options as a function of model parameters. For ease of presentation, we shift the level of the likelihood to have the same value at the maximum. Note that we do not scale the gradient of the likelihood.

Table 1.1: Return and Option Data

Panel A: Return Data

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

Panel B: Option Data

	Maturity (days)					
Moneyness (K/S)	5-30	30-60	60-90	90-180	180-365	All
Number of Option Contracts						
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
Average Call Prices						
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
Average Implied Volatilities						
0.85-0.90	0.34	0.28	0.26	0.25	0.24	0.27
0.90-0.95	0.27	0.24	0.23	0.23	0.22	0.24
0.95-1.00	0.21	0.20	0.20	0.21	0.21	0.21
1.00-1.05	0.18	0.18	0.18	0.19	0.19	0.18
1.05-1.10	0.20	0.16	0.16	0.17	0.18	0.17
1.10-1.15	0.31	0.20	0.18	0.17	0.17	0.18
All	0.24	0.21	0.21	0.20	0.20	0.21

Panel C: Restricted Samples. Descriptive Statistics

	Total No. of Options	Avg.Maturity($days$)	Avg.Moneyness(K/S)
ALL	129182	112.02	0.9886
ATM	109987	94.61	1.0001
SM	28831	22.7978	0.9786

Notes: Panel A reports descriptive statistics for the sample of index returns. The mean and standard deviation are annualized. Panel B reports the number of contracts, average call price and average implied volatility in the option data set where we choose the most liquid (highest trading volume) option within each moneyness-maturity range. Moneyness is defined as K/S. Due to the fact that OTM options are generally more heavily traded, this data set mainly consists of OTM call and OTM put options. Panel C presents the sample size for the restricted samples used in Figures 1.8 and 1.9. ALL represents the sample with all options. ATM represents the sample with ATM options but all maturities and SM represents the sample with short maturity options but all moneyness.

Table 1.2: Parameter Estimates Based on Joint Estimation Using Returns and Options

	SV	SVJR	SVJV	SVCJ
κ	2.1564	1.5531	1.0158	1.1248
	(0.0567)	(0.0404)	(0.0522)	(0.0489)
θ	0.0351	0.0359	0.0282	0.0241
	(0.0008)	(0.0008)	(0.0009)	(0.0008)
σ	0.4262	0.4152	0.3892	0.3450
	(0.0087)	(0.0085)	(0.0092)	(0.0074)
ρ	-0.9161	-0.9378	-0.9417	-0.9237
	(0.0129)	(0.0134)	(0.0118)	(0.0107)
η_s	2.5016	2.3513	2.7833	3.0401
	(0.2805)	(0.3081)	(0.3077)	(0.3486)
η_v	1.0836	0.5753	-0.1376	0.0498
	(0.0472)	(0.0454)	(0.0421)	(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)
Diffusive ERP	0.0878	0.0844	0.0785	0.0733
Jump ERP		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 posterior standard deviation for each parameter. Diffusive ERP and Jump ERP represent the equity risk premium due to the diffusive and jump components.

Table 1.3: Parameter Estimates in Existing Studies: The Heston SV Model.

Panel A: Based on Returns

Author	Period	κ	θ	σ	ρ
ABL	1953-1996	4.032	0.017	0.202	-0.380
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

Panel B: Based on Options

Author	Period	κ	θ	σ	ρ	η_v
BCC	1988-1991	1.150	0.040	0.390	-0.640	
Bates_2	1988-1993	1.490	0.067	0.742	-0.571	
	1988-1993	1.260	0.071	0.694	-0.587	2.28
BCJ	1987-2003	7.056	0.019	0.361	-0.397	
CJM	1996-2004	2.879	0.063	0.537	-0.704	

Panel 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
		-0.500*	-0.195*			
Eraker	1987-1990	4.788	0.049	0.554	-0.569	2.520
		2.268*	0.103*			
ASK	1990-2004	5.070	0.046	0.480	-0.767	
HLM	1990-2007	1.879	0.037	0.386	-0.741	1.9894
		-0.111*	-0.630*			

Notes: We report parameters for the SV model in existing studies. Estimates in Panel A are physical values. Estimates in Panel B are risk-neutral values. In Panel C, estimates with a star (*) indicate risk-neutral values and the rest are physical 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 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. The second row of Bates_2 presents the estimates with dynamic constraint on the spot variance; CJM: Christoffersen, Jacobs, and Mimouni (2010), based on the S&P 500; BCJ: Broadie, Chernov, and Johannes (2007), based on the S&P 500; Pan: Pan (2002), based on the S&P 500; ASK: Aït-Sahalia and Kimmel (2007), based on the S&P 500; HLM: Hurn, Lindsay, and McClelland (2015), based on the S&P 500.

Table 1.4: Conditional Moments

Panel A: Conditional Moments

	Conditional Variance(R)	Conditional Variance(V)	Conditional Covariance(R, V)
SV	V_t	$\sigma^2 V_t$	$\rho\sigma V_t$
SVJR	$V_t + \lambda(\mu_s^2 + \sigma_s^2)$	$\sigma^2 V_t$	$\rho\sigma V_t$
SVJV	V_t	$\sigma^2 V_t + \lambda\mu_v$	$\rho\sigma V_t$
SVCJ	$V_t + \lambda E(\xi)^2$	$\sigma^2 V_t + \lambda\mu_v$	$\rho\sigma V_t + \lambda\rho_J\mu_v^2$

Panel B: Conditional Moments. Sample Averages

	Conditional Variance(R)	Conditional Variance(V)	Conditional Covariance(R, V)
SV	0.0438	0.0080	-0.0171
SVJR	0.0455	0.0075	-0.0169
SVJV	0.0450	0.0610	-0.0165
SVCJ	0.0435	0.0416	-0.0143

Notes: Panel A presents closed-form expressions for three conditional moments for each of the four models we study. Panel B presents the sample averages for these moments. Note that $E(\xi) = \mu_s^2 + 2\mu_s\mu_v\rho_J + \rho_J^2\mu_v^2 + \sigma_s^2$.

Chapter 2

The Pricing of Volatility and Jump Risks in the Cross-Section of Index Option Returns

2.1 Introduction

One of the most enduring puzzles of asset pricing literature is that out-of-the-money (OTM) index put options are associated with large negative average returns (e.g., Jackwerth, 2000; Santa-Clara and Saretto, 2009; Bondarenko, 2014). While an index put option is a negative beta asset and thus is expected to have a negative rate of return, the magnitudes in the data seem too large to be consistent with standard models (Chambers, Foy, Liebner, and Lu, 2014). On the other hand, Bakshi, Madan, and Panayotov (2010) document that average returns of OTM index call options are also negative and declining with the strike price. This stylized fact is somewhat less known, but is perhaps even more puzzling because it contradicts the prediction from standard theories that expected call option returns should be positive and increase with the strike price (Coval and Shumway, 2001).[1]

[1] Related, Constantinides and Jackwerth (2009) and Constantinides et al. (2011) document widespread violations of stochastic dominance by OTM index and index futures options.

Previous research suggests several explanations for these large negative OTM option returns. For example, Broadie, Chernov, and Johannes (2009) conclude that the presence of a jump risk premium or estimation risk is consistent with historical put option returns. Bakshi, Madan, and Panayotov (2010) relate negative OTM call option returns to a U-shaped pricing kernel that arises in a model featuring short-selling and heterogeneity in investors' belief about return outcomes. Polkovnichenko and Zhao (2013) consider a rank-dependent utility model with a particular probability weighting function to explain the data. Negative OTM option returns can also be explained with theories of skewness/lottery preferences and leverage constraints (Barberis and Huang, 2008; Brunnermeier, Gollier, and Parker, 2007; Mitton and Vorkink, 2007; Frazzini and Pedersen, 2012). OTM options are often associated with substantial skewness and embedded leverage, which makes them particularly attractive for investors who have skewness preferences or face leverage constraints. The demand pressure will drive up prices and consequently lead to low returns in equilibrium (Gârleanu, Pedersen, and Poteshman, 2009).

This paper argues that the low returns on OTM index options are primarily due to the pricing of market volatility risk. Options are volatility-sensitive assets, and therefore their expected returns will critically depend on investor's attitudes towards volatility risk. We find that with a negative volatility risk premium, expected option returns implied from a stochastic volatility model match the average returns of call and put options across all strikes as well as the average returns of all the option portfolios that we consider. In particular, consistent with the data, not only does the pricing of volatility risk imply a steep relationship between expected put option returns and the strike price with OTM put options earning large negative rate of returns, it also implies a non-monotonic relation between expected call option returns and the strike price with OTM call options earning large negative expected returns. These results are robust to different parameterizations of the stochastic volatility process and also hold in the presence of a variance-dependent pricing kernel of Christoffersen, Heston, and Jacobs (2013).

Further corroborating the volatility risk premium hypothesis, we document that the volatility risk premium positively predicts future index option returns with OTM options and ATM straddles exhibit the strongest return predictability. Both the sign and patterns of index option return predictability are in line with the impact of the volatility risk premium on expected option returns in a stochastic volatility model. The index option return predictability is not due to the underlying return predictability by the volatility risk premium (Bollerslev, Tauchen, and Zhou, 2009). It is robust to different empirical implementations as well as to controlling for other predictors. We also show that the index option return predictability is economically significant and can be translated into large economic gains. We propose an option selling strategy that exploits option return predictability. The new strategy significantly outperforms a benchmark strategy that writes index options every month.

Lastly, we study the relationship between the pricing of price jump risk and expected option returns. We consider a stochastic volatility jump model (SVJ) that incorporates a compensation for price jump risk, reflecting investors' crash fear. Consistent with Broadie, Chernov, and Johannes (2009), we find that when jump risk is priced, the SVJ model yields large negative expected returns on OTM put options with magnitudes very close to the data. While the jump risk premium fits put option returns extremely well, it fails to match the average returns of OTM calls. In particular, the presence of a jump risk premium would imply an increasing relation between expected call option returns and the strike price with OTM calls earning large positive expected returns, which is contrary to the data. Our results about the jump risk premium are robust to different parameterizations. On the empirical side, we find that the jump risk premium significantly predicts future returns on OTM put options, but it does not predict call option and straddle returns.

Taken together, the evidence leads us to conclude that the negative average OTM index option returns are primarily due to the pricing of market volatility risk, although the jump risk premium also accounts for some portion of OTM put option returns.[2] Overall our results suggest that the pricing of volatility risk has a first-order effect on the cross-section of index option returns. The rest of the paper is organized as follows. Section 2 discusses related literature. Section 3 investigates relative roles of the volatility and jump risk premiums in fitting historical S&P 500 option returns. Section 4 studies the index option return predictability. Section 5 contains robustness results, and Section 6 concludes the paper.

2.2 Related Literature

The bulk of option literature focuses on the behavior of *option prices*. For example, it is well known that in the equity index options market implied volatilities from OTM put options are consistently higher than their ATM counterparts following 1987 market crash (Rubinstein, 1994). This stylized fact is often referred to as the implied volatility skew or volatility smirk, and it contradicts the prediction of the BSM model that implied volatility is constant across strikes.[3] The presence of a pronounced volatility skew has inspired many subsequent studies. For example, there is an extensive literature that demonstrates stochastic volatility and jumps are needed in order to fit the rich option price dynamics, although the empirical evidence is somewhat mixed regarding the relative importance of these additional factors as well as their pricing. For important contributions, see Bakshi, Cao, and Chen (1997), Bates (2000), Chernov and Ghysels (2000), Pan (2002), Jones (2003), Eraker (2004), Broadie, Chernov, and Johannes (2007) and Andersen, Fusari, and Todorov (2015a).[4] Our paper

[2]OTM call options are priced with a significant premium during crisis periods. For example, buying 1-month 5% OTM call in February 2009 and holding it to maturity generates a return of -54%, despite the fact that the S&P 500 actually went up by 6.2% over the month. By contrast, a similar episode during normal times would be associated with a large positive return. For example, the rate of return for a same 5% OTM call is 841% from October 2004 to November 2004 during which the S&P 500 index went up by 6.3%. This premium reflects investors' concerns about both tails (variance), not just the crash risk.

[3]Closely related, Aït-Sahalia and Lo (1998), Jackwerth and Rubinstein (1996) and Jackwerth (2000) document that the risk-neutral distribution inferred from option prices is not log-normal and systematically skewed more to the left.

[4]Another strand of literature addresses the implied volatility skew relying on equilibrium models (e.g., Benzoni, Collin-Dufresne, and Goldstein, 2011; Du, 2011; Seo and Wachter, 2017). Existing studies also suggest that the implied volatility skew might be related to demand pressure (Bollen and Whaley, 2004;

differs from this literature in that we model *option returns* rather than *option prices*. Our analysis suggests that the pricing of volatility risk has a first-order effect on the cross-section of index option returns.

Our paper is closely related to the expanding literature that investigates index option returns. Previous studies usually focus on put options and they find surprisingly low returns for OTM index put options (e.g., Bondarenko, 2014; Jackwerth, 2000). Broadie, Chernov, and Johannes (2009) and Chambers et al. (2014) formally compare historical put option returns to option pricing models, and their results suggest that index put options are likely embedded with a jump risk premium.[5] We extend their analysis to include call options. Understanding index call option returns is important for two reasons. First, Bakshi, Madan, and Panayotov (2010) document that average returns of OTM index call options are negative and decreasing with the strike price. Their findings about call options are puzzling because under general economic conditions, Coval and Shumway (2001) show that expected call option returns should be positive and increase with the strike price. Therefore any theory that tries to explain why OTM put options have large negative returns should also explain the puzzling returns patterns observed on OTM call options. Studying call options is also important because call options, which are claims on the upside, are critical for disentangling the volatility risk premium from the jump risk premium. As we will show in the next section, the volatility and jump risk premiums have drastically different predictions on expected OTM call option returns. Consistent with Broadie, Chernov, and Johannes (2009) and Chambers et al. (2014), we find that expected put option returns computed with the jump risk premium are consistent with the observed data. However, the jump risk premium would also imply that expected OTM call option returns are positive and increasing with the strike price, which is contrary to the data. In contrast, we show that expected option returns computed with

Gârleanu, Pedersen, and Poteshman, 2009), aversion to model uncertainty (Liu, Pan, and Wang, 2005), or investor sentiment (Han, 2007).

[5]The two papers have different conclusions about whether index put option returns are consistent with standard option pricing models with only equity risk premium (e.g., volatility and jump risks are not priced). Our analysis confirms the results in Chambers et al. (2014) that the hypothesis of no additional risk premiums can be rejected in general.

the volatility risk premium match the low OTM call and put option returns simultaneously. Related, several papers use factor-based approaches to gain a better understanding of index option returns. Examples include Jones (2006), Cao and Huang (2007), and Constantinides, Jackwerth, and Savov (2013). Israelov and Kelly (2017) propose a method for constructing conditional distribution for index option returns. Driessen and Maenhout (2007) and Faias and Santa-Clara (2017) study index option returns from portfolio allocation perspective. Santa-Clara and Saretto (2009) investigate the impact of margin requirements on option trading strategies.

There is a large body of literature on the pricing of volatility and jump risks in the financial markets. The pricing of aggregate volatility and jump risks has been studied extensively in the cross-section of stock returns.[6] See, among others, Ang, Hodrick, Xing, and Zhang (2006), Adrian and Rosenberg (2008) and Cremers, Halling, and Weinbaum (2015). Our paper is more related to studies that focus on the pricing of aggregate volatility and jump risks in the equity index options market. Index options market, where stochastic volatility and jump risks play a prominent role, contains rich economic information about the pricing of these risk factors. For example, Coval and Shumway (2001) report that zero-beta at-the-money straddle positions produce large losses and they interpret it as evidence that systematic stochastic volatility is priced in option returns. Bakshi and Kapadia (2003) find that delta-hedged option portfolios have negative average returns which indicates the volatility risk premium is negative. Our results are consistent with the findings of Coval and Shumway (2001) and Bakshi and Kapadia (2003) that the volatility risk premium is negative in the index options market. The key difference between the above studies and this paper is that their emphasis is on using option portfolios to infer the existence and sign of the volatility risk premium, while this paper aims to quantify the impact of the volatility risk

[6] A number of studies examine the volatility risk premium based on variance swaps. See, among others, Egloff, Leippold, and Wu (2010), Aït-Sahalia, Karaman, and Mancini (2015b), and Dew-Becker et al. (2017). Another strand of literature provides additional evidence on the market volatility risk premium by comparing option implied volatility with realized volatility. See, among others, Lamoureux and Lastrapes (1993), Fleming, Ostdiek, and Whaley (1995) and Christensen and Prabhala (1998).

premiums on the cross-section of *unhedged* index option returns. We also document that the volatility risk premium predicts future option returns and characterize the effect of the jump risk premium on expected option returns.

We also contribute to the growing literature on the variance risk premium (e.g., Bollerslev, Tauchen, and Zhou, 2009; Carr and Wu, 2009; Drechsler and Yaron, 2011; Eraker, 2012). Existing studies find that the volatility risk premium is a strong predictor of short term U.S. stock index returns (e.g., Bollerslev et al., 2009), and the predictability is also seen in the international data (Bollerslev et al., 2014). This is in contrast to many traditional predictors (e.g., dividend yield) which often operate over long horizons. Related, Bali and Hovakimian (2009), Goyal and Saretto (2009) and Della Corte, Ramadorai, and Sarno (2016) investigate the role of the volatility risk premium in predicting the cross-section of asset returns. Our paper expands the existing evidence on the predictive power of the volatility risk premium by documenting a significant time-series index return predictability by the volatility risk premium in the S&P 500 index options market. These results are new in the literature. It is important to note that the index option return predictability cannot be attributed to the underlying stock return predictability by the volatility risk premium. Instead, we show the option return predictability is likely due to the time-varying volatility risk premiums embedded in index options.

2.3 The Volatility Risk Premium, the Jump Risk Premium and Expected Option Returns

In this section, we begin by reviewing historical returns of S&P 500 index options across a wide range of strikes as well as returns of a number of option portfolios. We then compare these average returns in the data to expected option returns implied by option pricing models. We investigate whether index option returns are consistent with the pricing of volatility risk, or the pricing of price jump risk or both.

2.3.1 Historical S&P 500 Index Option Returns

This paper focuses on historical returns from holding S&P 500 index options. We download S&P 500 index options (SPX) data from OptionMetrics through WRDS. The sample period for our analysis is from March 1998 to August 2015.[7] In particular, on the first trading day after monthly option expiration date, we collect SPX options that will expire over the next month. These options are the most frequently traded options in the marketplace and they have maturities ranging from 25 to 33 calendar days. Prior to February 2015, the expiration day for index options is the Saturday immediately following the third Friday of the expiration month. Starting in February 2015, the option expiration day is the third Friday of a month.[8] We also apply standard filters to option data and relegate details to Appendix.

Following the existing literature, we construct time-series of monthly holding-to-maturity returns to S&P 500 index options for fixed moneyness, ranging from 0.96 to 1.08 for calls, 0.92 to 1.04 for puts with an increment of 2%.[9] Moneyness is defined as the strike price over the underlying index: K/S. We do not investigate options that are beyond 8% OTM or 4% ITM because of potential data issues (e.g., low price or low trading volume or missing observations). We also compute returns on a number of option portfolios including at-the-money straddles (ATMS), put spreads (PSP), crash-neutral spreads (CNS) and call spreads (CSP). As pointed out by Broadie, Chernov, and Johannes (2009), returns on option portfolios are more informative than individual option returns and therefore they provide more powerful tests. ATMS involves the simultaneous purchase of a call option and a put option with $K/S = 1$. PSP consists of a short position in a 6% OTM put and a long

[7]OptionMetrics data starts from January 1996. However, the settlement values (SET) for SPX options required to compute holding-to-maturity returns are only available from April 1998. As a result, we start sampling options in March 1998. The settlement values for S&P 500 index options are calculated using the opening sales price in the primary market of each component security on the expiration date and are obtained from the CBOE.

[8]This means we usually select options on Mondays. If Monday is an exchange holiday (e.g., Martin Luther King Day or President's Day), we use Tuesday data.

[9]The bid-ask spreads in the option market are usually large, and monthly holding-to-maturity option returns mitigate this problem because they only incur the trading cost at initiation. Holding-to-maturity returns are also easy to analyze analytically and avoid a number of theoretical and statistical issues associated with high frequency option returns (Broadie et al., 2009).

position in an ATM put. CNS consists of a long position in an ATM straddle and a short position in a 6% OTM put. Finally, CSP combines a long position in an ATM call with a short position in a 6% OTM call. When computing option returns, we use the mid-point of bid-ask quotes as a proxy for option price, and we calculate option payoff at maturity based on the index settlement values. Notice that Broadie, Chernov, and Johannes (2009) and a subsequent study by Chambers et al. (2014) focus on index put options and several option portfolios. We extend their analysis to include index call options. Call options are claims on the upside, which will be critical for differentiating the volatility risk premium from the jump risk premium.

Table 2.1 reports average monthly returns for the cross-section of index options with different strikes as well as option portfolios. Panel A of Table 2.1 shows that average returns of OTM index call options are negative and declining with the strike price. For example, the average returns from buying a 4% OTM call and a 6% OTM call are -1.47% and -18.12% per month, respectively. Our results confirm the findings of Bakshi, Madan, and Panayotov (2010) which study 1%, 3% and 5% OTM S&P 500 call options from 1988 to 2007.

Panel B of Table 2.1 presents another stylized fact in the equity index option market that put options, especially OTM put options, have very large negative returns. For example, over our sample period, buying a 6% OTM put option would lose about 45% per month on average. These estimates are largely consistent with the existing literature.

Panel C reports average returns on option portfolios. Coval and Shumway (2001) find that zero-beta straddles have negative average returns. We do not investigate zero-beta straddles as in Coval and Shumway (2001) because constructing a zero-beta straddle would require a model to determine the portfolio weights. Nevertheless, confirming their results, we find that simple ATM straddles on average lose 8.47% per month over our sample period. Also notice that the average return for call spreads is 13.56% per month. Call spreads earn high returns because both the long position in ATM call and the short position in 6% OTM call generate positive returns as shown in Panel A.

Conducting statistical inference on option returns reported in Table 2.1 is in general difficult because option returns are highly non-normal. For example, Figure 2.1 and 2.2 plot the time series of OTM index call and put option returns. As can be seen from these figures, option returns are often associated with extreme observations, which makes the standard CAPM-type of linear models inappropriate. Broadie, Chernov, and Johannes (2009) propose to evaluate average option returns relative to what would have been obtained in an option pricing model. Their methodology not only automatically takes into account the leverage and kinked payoffs of options, but also anchors hypothesis tests at appropriate null values. Following Broadie, Chernov, and Johannes (2009), we compare historical option returns with those generated by various option pricing models. We are particularly interested in understanding to what extent these negative OTM index option returns are due to the pricing of stochastic volatility risk, or the pricing of price jump risk or both.

2.3.2 Analytical Framework

To assess the relative roles of the volatility and jump risk premiums in explaining the cross section of index option returns, we consider a standard affine jump diffusion framework with mean-reverting stochastic volatility and Poisson-driven jumps in stock price. The model is commonly referred to as the SVJ model (Bates, 1996) and nests the Black-Scholes-Merton model (Black and Scholes, 1973; Merton, 1973), the Heston stochastic volatility model (Heston, 1993) and the Merton jump diffusion model (Merton, 1976) as special cases. The SVJ model says the index level (S_t) and its spot variance (V_t) have the following dynamics under the physical measure ($\mathbb{P}$):

$$
\begin{aligned}
dS_t &= (\mu + r - d)S_t dt + S_t\sqrt{V_t}dW_1 + (e^Z - 1)S_t dN_t - \lambda\overline{\mu}S_t dt \\
dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_2
\end{aligned}
$$

where μ is the equity risk premium, r is the risk-free rate, d is the dividend yield, N_t is a $\mathbb{P}$-measure Poisson process with a constant intensity λ, $Z \sim N(\mu_z, \sigma_z^2)$, $\overline{\mu}$ is the mean jump size with $\overline{\mu} = \exp(\mu_z + \frac{1}{2}\sigma_z^2) - 1$, θ is the long-run mean of variance, κ is the rate of mean reversion, σ is volatility of volatility, and W_1 and W_2 are two correlated Brownian motions with $\mathbb{E}[dW_1 dW_2] = \rho dt$. The dynamics under the risk-neutral measure ($\mathbb{Q}$) are:

$$
\begin{aligned}
dS_t &= (r-d)S_t + S_t\sqrt{V_t}dW_1^{\mathbb{Q}} + (e^{Z^{\mathbb{Q}}} - 1)S_t dN_t^{\mathbb{Q}} - \lambda^{\mathbb{Q}}\overline{\mu}^{\mathbb{Q}} S_t dt \\
dV_t &= [\kappa(\theta - V_t) - \eta V_t]dt + \sigma\sqrt{V_t}dW_2^{\mathbb{Q}}
\end{aligned}
$$

where η is the price of volatility risk, $N_t^{\mathbb{Q}} \sim \text{Poisson}(\lambda^{\mathbb{Q}} t)$, $Z^{\mathbb{Q}} \sim N(\mu_z^{\mathbb{Q}}, (\sigma_z^{\mathbb{Q}})^2)$ and $\overline{\mu}^{\mathbb{Q}} = \exp(\mu_z^{\mathbb{Q}} + \frac{1}{2}(\sigma_z^{\mathbb{Q}})^2) - 1$. Throughout the paper, risk neutral quantities will be denoted with $\mathbb{Q}$ and all other quantities are taken under the physical measure. Note that there are three types of risk premiums in the SVJ model: the equity risk premium (μ), the volatility risk premium (ηV_t) and the jump risk premium (price jump has different distributions under $\mathbb{P}$ and $\mathbb{Q}$ probability measures).

Broadie, Chernov, and Johannes (2009) point out that expected option returns can be computed analytically within the above framework. Their insight is particularly useful as it allows one to quantitatively analyze the impact of different risk premiums. To better understand relative effects of the volatility and jump risk premiums on expected option returns, we will focus on three versions of the SVJ model: a benchmark BSM model in which neither volatility nor jump risk is priced (BSM), the Heston stochastic volatility model with a volatility risk premium (SV), and finally a SVJ model in which only jump risk is priced, but stochastic volatility risk is not (SVJ).

Computing model-implied expected option returns requires the knowledge of parameter values of each model. Our approach to infer model parameters is very similar to Broadie, Chernov, and Johannes (2009). In particular, we calibrate the equity risk premium, the risk-free rate and the dividend yield based on those realized over our sample period. The remaining $\mathbb{P}$-measure parameters are estimated from the time-series of index returns.[10] Our

[10] As argued by Broadie, Chernov, and Johannes (2009), while one can estimate these parameters using

estimation is based on particle filtering and Appendix B describes the details. In the robustness analysis, we show our results are robust to different parameterizations of stochastic volatility and jumps.

Moreover, we obtain estimates of the volatility and jump risk premiums by observing that in a standard power utility environment (e.g., Bakshi and Kapadia, 2003; Broadie, Chernov, and Johannes, 2009; Christoffersen, Heston, and Jacobs, 2013; Naik and Lee, 1990), the risk adjustment for volatility risk is given by:

$$\eta V_t = Cov(\gamma \frac{dS_t}{S_t}, dV_t) \Longrightarrow \eta = \gamma\sigma\rho \tag{2.1}$$

and the risk adjustment for price jump risk is given by:

$$\begin{aligned} \lambda^{\mathbb{Q}} &= \lambda \exp(-\mu_z \gamma + \frac{1}{2}\gamma^2\sigma_z^2) \\ \mu_Z^{\mathbb{Q}} &= \mu_z - \gamma\sigma_z^2. \end{aligned} \tag{2.2}$$

where γ is relative risk aversion of the agent. For our benchmark analysis, we follow Broadie, Chernov, and Johannes (2009) and assume a risk aversion of 10. We also perform an extensive sensitivity analysis with respect to risk aversion and those results are contained in Section 2.5.1.

Table 2.2 reports (annualized) parameter values that we use to compute expected option returns for different models. For the BSM model, the constant volatility parameter is set equal to the square root of the long run mean of stock variance (θ) in the SV model ($\sigma_{BSM} =$ 19.05%). For the SV model, given a risk aversion of 10 and a negative ρ, equation (2.1) indicates that the volatility risk premium parameter η must be negative and is equal to -4.347. Pan (2002) finds that the magnitudes of the volatility risk premium needed to reconcile time-series and option-based spot volatility measures imply explosive risk-neutral volatility dynamics ($\kappa + \eta < 0$). In contrast, our calibration does not have this issue: volatility process

option data, this approach might be problematic for our purpose because we would be explaining option returns using information extracted from option prices in the first place.

under the risk neutral measure remains mean-reverting ($\kappa + \eta > 0$). For the SVJ model, we set the volatility risk premium to zero ($\eta = 0$) so that we can focus exclusively on the jump risk premium. Consistent with the notion that investors fear large adverse price jumps, the risk corrections in equation (2.2) indicate that price jumps occur more frequently and more severely under the risk-neutral measure. Our estimates imply about 1.50 jumps per year on average ($\lambda^{\mathbb{Q}} = 1.4969$) and a mean jump size of -6.67% ($\mu_z^{\mathbb{Q}} = -0.0667$) under $\mathbb{Q}$ probability measure, and about 0.97 jumps per year on average ($\lambda = 0.9685$) with a mean jump size of -2.09% ($\mu_z = -0.0209$) under $\mathbb{P}$ probability measure.

Based on parameter values reported in Table 2.2, we compute expected option returns implied from the BSM, SV and SVJ models and compare them to realized average option returns in the data. Following Broadie, Chernov, and Johannes (2009), we also simulate each model to form a finite-sample distribution of average option returns, which allows one to test whether realized option returns are significant relative to a model. Specifically, we simulate 25000 sample paths of the index, with each sample path having 210 months (the sample length of our data). For each sample path, we compute time series average option returns. The p-values are then calculated as the percentile of realized option returns relative to the 25000 simulated options returns. If the percentile is higher than 0.5, we report the p-value as 1 minus the percentile.

2.3.3 Results

Table 2.3 contains results for the BSM model. Expected option returns computed analytically are labeled by "Model". We also report the average simulated option returns, denoted by "Simulation". Not surprisingly, those two are very close to each other. Historical option returns taken from Table 2.1 are denoted by "Data". Table 2.3 shows that the BSM model cannot account for the empirical option return patterns. First, confirming the results of Chambers, Foy, Liebner, and Lu (2014), we find that the BSM model is rejected by OTM put option returns and ATM straddle returns. For example, according to the BSM model, an

ATM straddle should earn 0.71 percent per month. In the data, the monthly average return for ATM straddles is -8.47 percent with a p-value of 0.03. The p-value of 0.03 means that only 3% of the 25000 simulated average straddle returns are less than the -8.47 percent realized return in the data. Furthermore, we find that the BSM model is also unable to explain OTM call option returns. In particular, the model predicts that expected call option return is an increasing function of the strike price with OTM call options earning large positive return, which is contrary to the data.[11] The difference between the BSM model implied returns and the data is statistically significant for 6% and 8% OTM calls as indicated by p-values.

Table 2.4 shows that the stochastic volatility model is able to quantitatively match average returns of call and put options across all strikes as well as average returns of all option portfolios. In particular, consistent with the data, the model implies that expected returns of OTM call options are negative and decreasing with the strike price. The expected monthly return decreases monotonically from 1.89 percent for ATM calls to -25.05 percent for 8% OTM calls. This property of the stochastic volatility model sharply contrasts with the Black-Scholes-Merton model where the expected call option return is monotonically increasing function of the strike price. In regards to put options, the stochastic volatility model predicts more negative expected returns on OTM put options relative to the BSM model, which again is consistent with the data. The p-values suggest that realized average option returns are not significantly different from those generated by the stochastic volatility model. Finally, Panel C shows that the SV model is also consistent with returns on all option portfolio that we consider. The fact that a simple stochastic volatility model describes average option returns remarkably well is somewhat surprising given it is a clearly misspecified model for option prices.[12] As discussed in Section 2, there is an extensive literature

[11]In the BSM model, a call option is effectively a levered position in the underlying asset. Moreover, the embedded leverage is an increasing function of the strike price. Because the underlying asset (i.e., the index) typically has a positive expected return that is higher than the risk-free rate, the expected call option return should be greater than the expected return of the underlying, and it is increasing with the strike price.

[12]Empirical shortcomings of the SV model are of course well-documented. For example, the estimated SV model often generates a steeply upward-sloping term structure of implied volatility, which is incompatible with the observed term structure. Moreover, the model implies the instantaneous change in volatility is Gaussian and homoskedasticity, which is unable to capture the sudden and abrupt moves in the observed

that demonstrates several factors are required to fit the rich dynamics of option prices. Our results somewhat echo the findings of Cochrane and Piazzesi (2005) in the bond market that although multiple factors are needed to describe empirical patterns in bond prices, a single factor summarizes nearly all information about risk premium.

Table 2.5 contains results for the SVJ model. Confirming the findings of Broadie, Chernov, and Johannes (2009) and Chambers et al. (2014), Panel B shows that the presence of a jump risk premium is consistent with observed put option returns; the SVJ model predicts very large negative expected returns on OTM put options. While the jump risk premium matches put returns almost perfectly, it fails to explain the large negative average returns of OTM calls. For example, Panel A shows that the SVJ model actually predicts an increasing relation between expected call option returns and the strike price with OTM call options earning large positive returns, which is contrary to the data. It is important to note that this result about the SVJ model is not due to our parameterization. We also use SVJ parameters reported in Broadie, Chernov, and Johannes (2009) and Chambers et al. (2014) and find very similar results. Also see Section 2.5.1 for additional discussion. Interestingly, despite the return difference between the data and the model is quite large for OTM call options, the p-values actually indicate that the SVJ model is not rejected. The reason is that OTM call option returns have very large standard deviations, which makes them less informative as compared to portfolio returns. Portfolio-based evidence in Panel C indeed shows that the SVJ model is rejected by call spreads. The monthly average return of call spreads is 13.56 percent, much higher than the model-implied return which is only 1.46 percent. The difference is statistically significant with a p-value of 0.04.

To summarize our findings, we plot expected option returns computed in Tables 2.3 to 2.5 against the strike price in Figure 2.3: Panel A for call options and Panel B for put options. Panel A shows that the volatility and jump risk premiums have drastically different predictions on OTM call options. The SVJ model implies that expected returns of OTM

volatility dynamics. For detailed discussions, see Bates (2003), Broadie, Chernov, and Johannes (2007), and Christoffersen, Jacobs, and Mimouni (2010).

call options are positive and increasing with the strike price, which is qualitatively similar to the BSM model. In contrast, the SV model predicts a decreasing relationship between expected returns and the strike price with OTM calls earning large negative returns. Panel B shows that the three models yield similar predictions on put options in that expected put returns should be negative. The SVJ model yields the most negative estimates, followed by the SV model.

2.3.4 Discussion

The above section shows that when volatility risk is priced, a stochastic volatility model is able to match the average returns of both OTM call and put options. On the other hand, when the jump risk is priced, a SVJ model is able to match the average returns of OTM put options. It is important to note that the presence of the volatility and jump risk premiums is critical for the SV and SVJ models to fit the data. In unreported results, we find that when volatility and jump risks are not priced, expected option returns implied from the SV and SVJ models are similar to those in the BSM model, and both models will be rejected.

The volatility and jump risk premiums affect expected option returns because they induce changes in the risk-neutral index return distribution under which option prices are determined. A negative volatility risk premium will add probability mass to both tails and thus increase the value of both OTM call and put options, which leads to lower expected returns. On the other hand, the jump risk premium has two effects on the risk-neutral distribution. First, the presence of a jump risk premium will result in a more negatively-skewed risk neutral distribution. This in turn will increase the value of OTM put options and decrease the value of OTM call options. Second, the jump risk premium also tends to fatten both tails because it also introduces a wedge between risk neutral and physical variance. For OTM put options, both effects will lead to a higher valuation and this is the reason why the SVJ model yields most negative OTM put option returns. On the other hand, for OTM call options, the skewness effect tends to dominate and therefore a presence of the jump risk premium will result in higher expected returns for OTM call options.[13]

We also analyze how expected option returns vary with respect to changes in the volatility risk premium in the SV model. Based on the same parameter values reported in Table 2.2, we plot in Figure 2.4 expected returns on call options, put options and straddles in the SV model against risk aversion γ for different moneyness. A higher γ means a larger volatility risk premium (more negative). Figure 2.4 reveals several interesting results. First, as γ increases (meaning the volatility risk premium becomes more negative), expected returns on calls, puts and straddles monotonically decrease regardless of moneyness. Again, this is because a negative volatility risk premium makes options more valuable. However, the magnitude of the effect of the volatility risk premium on expected returns crucially depends on the moneyness of an option. In particular, the relation between the volatility risk premium and the expected option return is much stronger for OTM call and put options with the steepest slope. As options move towards the in-the-money direction, expected returns become less sensitive to the volatility risk premium and the slope flattens out. On the other hand, straddles have their own unique pattern. ATM straddle returns are more sensitive to changes in the volatility risk premium as compared to their ITM and OTM counterparts. In Section 4, we show these theoretical results are consistent with our findings on index option return predictability.

Figure 2.4 also helps understand why the volatility risk premium fits option return data well. As discussed, a negative volatility risk premium increases option value which then leads to a lower expected return. Moreover, this effect is disproportionately stronger for out-of-the-money options. As a result, a negative volatility risk premium is able to generate not only a steeper relation between expected put option returns and the strike price with OTM put options earning large negative returns, but also a decreasing relation between expected OTM call option returns and the strike price.

[13] If one ignores the equilibrium restrictions and allows the variance of jump size to take different values under the physical and risk neutral measures, expected option returns become even more complicated. See Branger, Hansis, and Schlag (2010) for a related discussion.

2.4 Predicting Index Option Returns

In this section, we investigate whether the volatility and jump risk premiums can forecast future index option returns. If options were embedded with time-varying volatility and jump risk premiums, then one might observe some predictability of option returns by these risk premiums. Indeed, we find that the volatility risk premium predicts subsequent index option returns. The index option return predictability is both statistically and economically significant, and is not due to the underlying return predictability afforded by the volatility risk premium. We also find that the jump risk premium predicts future OTM put option returns.

2.4.1 Predicting Option Returns: the Volatility Risk Premium

Following the definition of the equity risk premium, we define the volatility risk premium as the difference between physical and risk neutral expectations of future realized volatility:

$$\text{VRP}_t = \mathbb{E}_t(RV_{t,t+1}) - \mathbb{E}_t^{\mathbb{Q}}(RV_{t,t+1}).$$

The volatility risk premium is constructed each month on the option selection date and will be used to forecast option returns over the following month. For the baseline results, we follow Bollerslev, Tauchen, and Zhou (2009) and measure the volatility risk premium as the difference between realized volatility and the VIX index:

$$\text{VRP}_t = RV_{t-1,t} - VIX_t$$

where realized volatility is computed as the square-root of the sum of squared 5-min log returns on S&P 500 futures over the past 30 days.[14] The VIX index is published by the

[14] The use of intra-day data is motivated by the realized volatility literature which demonstrates the critical role of high frequency data in volatility measuring and modeling. See, among others, Andersen et al. (2001), Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2002). Following the literature, we focus on 5-

Chicago Board Options Exchange (CBOE), and it tracks 30-day risk neutral expectation of future realized volatility inferred from option prices.[15] In the robustness analysis, we show that our empirical results are not sensitive to how we measure the volatility risk premium. For example, instead of using realized volatility, we also estimate expected volatility and find similar results. Figure 2.5 plots realized volatility, the VIX and the volatility risk premium throughout our sample period. Consistent with findings in Todorov (2010) and Carr and Wu (2009), Figure 2.5 shows that the volatility risk premium, like volatility itself, fluctuates substantially over time.

To investigate whether the volatility risk premium predicts future index option returns, we run the following time-series predictive regressions at monthly frequency:

$$option\_ret^{i}_{t,\, t+1} = \alpha^i + \beta^i \mathrm{VRP}_t + \epsilon, \quad i \in \{call, put, straddle\} \tag{2.3}$$

where the dependent variable *option_ret* is returns from holding call options, put options and straddles from month t to month $t+1$. The analysis in Section 2.3.4 shows that options with different moneyness have different sensitivities with respect to the volatility risk premium and therefore we run predictive regressions separately for different moneyness groups. In particular, for call options, we consider the following three groups: $0.96 <= K/S < 1.00$, $1.00 <= K/S < 1.04$ and $1.04 <= K/S < 1.08$. For put options, we consider $0.92 <= K/S < 0.96$, $0.96 <= K/S < 1.00$ and $1.00 <= K/S < 1.04$. Again we do not investigate options that are beyond 8% OTM or 4% ITM because of potential data issues. For straddles, we consider the following three moneyness groups: $0.94 <= K/S < 0.98$, $0.98 <= K/S < 1.02$ and $1.02 <= K/S < 1.06$.

min returns to avoid potential microstructure effects. Liu et al. (2015) argue that it is difficult to outperform 5-minute realized variance even with more sophisticated sampling techniques. We also treat overnight and weekend returns as an additional 5-minute interval.

[15] The CBOE developed the first-ever volatility index in 1993, which then was based on the average implied volatilities of at-the-money options on S&P 100. In 2003, the CBOE started publishing a new index that is calculated using S&P 500 index option prices in a model-free approach. We uses the new VIX index. For more details on the model-free approach, see for example Dupire (1994), Neuberger (1994), Britten-Jones and Neuberger (2000), and Jiang and Tian (2005).

Panel A of Table 2.6 reports predictive regression results for call options. The volatility risk premium significantly predicts future returns with a positive and statistically significant coefficient for call options that are between 4% and 8% OTM. The t-statistic is 2.11 and the adjusted R^2 of the regression is 0.99%. Throughout the paper, t-statistics are computed using Newey-West standard errors with four lags (Newey and West, 1987, 1994).[16] Interestingly, the forecasting power of the volatility risk premium, in terms of the slope estimates, statistical significance and R^2, decreases monotonically as call options move towards the in-the-money direction.

Panel B shows that the volatility risk premium also positively predicts future index put option returns. Similar to calls, return predictability however becomes increasingly weak as put options move towards the in-the-money direction, judging by the predictive coefficient, statistical significance as well as R^2. Specifically, among put options that are between 4% OTM and 8% OTM, the volatility risk premium significantly predicts subsequent returns, with a Newey-West t-statistic of 2.21 and an adjusted R^2 of 2.04%. The predictability, however, becomes insignificant for other groups.

Panel C shows that straddle has its own distinct pattern. In particular, the volatility risk premium exhibits the most significant forecasting power over at-the-money straddles. As straddles move away from the money, the predictive power of the volatility risk premium drops substantially and is only marginally significant.

To sum up, we document that the volatility risk premium positively forecasts future option returns. Moreover, this index option return predictability exhibits an interesting dependence on the moneyness, with OTM options and ATM straddles having the strongest predictability.

[16] Results with OLS standard errors are stronger. To be conservative, we report only Newey-West t-stats.

2.4.2 The Economic Significance of Index Option Return Predictability

To assess the economic significance of the index option return predictability that we document above, this section proposes a trading strategy that exploits option return predictability in the context of selling index options. Writing index options is popular because historically it tends to yield higher returns by collecting the volatility risk premium. Since the volatility risk premium is positively associated with future option returns, the simplest strategy would be to sell options only in months when the volatility risk premium is negative. This strategy relies only on an ex-ante market signal and does not require investors to estimate any model. Moreover, since return predictability is significant among out-of-the-money options and at-the-money straddles, we will test the performance of the new trading strategy in the context of selling a 4% OTM call, a 6% OTM call, a 4% OTM put, a 6% OTM put and an ATM straddle. As a benchmark, we consider a strategy that writes options in every month of the sample. The new strategy is called "VRP $<$ 0", and the benchmark strategy is called "Always". The performance of the S&P 500 over the same period is also included for comparison.

Table 2.7 reports the results. First, by comparing the benchmark strategy "Always" to the S&P 500 index, it appears that writing index options is more profitable than investing in the S&P 500. For example, over our sample period, average returns and Sharpe ratios from selling put options and straddles are higher than those obtained with the S&P 500. Selling OTM call options is also more profitable than just buying the index, but it has very large standard deviation which actually makes the Sharpe ratio lower than the S&P 500.

More importantly, Table 2.7 shows that our new strategy outperforms the benchmark strategy. Taking ATM straddles as an example, following the new strategy, one would obtain a monthly average return of 0.106 with a Sharpe ratio of 0.151. In contrast, the average return and Sharpe ratio for the benchmark strategy are 0.085 and 0.115, respectively. Note that with the new trading strategy, one sells options less often. The last column of Table 2.7 indicates the number of months in which options are shorted. We also report skewness of different trading strategies. Despite the improvements in the Sharpe ratio, the new strategy that we propose has a similar or even lower skewness relative to the benchmark strategy. Finally, it should be emphasized that Sharpe ratio is a poor performance measure of derivatives trading strategies which often yield highly non-normal payoffs (Goetzmann et al., 2004). The strategy proposed in this paper is only suitable for institutional investors with deep pockets and long investment horizon.

Table 2.7 also shows that overall writing OTM put options tends to be more profitable than writing OTM call options. As our analysis suggests, one potential explanation is that selling OTM put options earns both the volatility and jump risk premiums. In contrast, by selling OTM call options, one collects mainly the volatility risk premium. The divergence between selling calls and puts might also be related to institutional frictions and order flow. For example, it is in general easy to sell calls via covered calls, but difficult to sell naked puts. Moreover, OTM put options can be used as portfolio insurance and therefore attract much more demand than OTM calls.

2.4.3 Option Return Predictability and Index Return Predictability

Bollerslev, Tauchen, and Zhou (2009), among others, document that the volatility risk premium predicts future index returns at short horizons. Therefore a natural interpretation of option return predictability is that it is merely a manifestation of the underlying index return predictability afforded by the volatility risk premium. While this explanation appears plausible, it can be ruled out based on the fact that the volatility risk premium forecasts future option returns all in the same direction: a more negative volatility risk premium this month is associated with lower option returns in the subsequent month. If option return predictability were caused by stock return predictability, then one would observe opposite signs on the predictive coefficients for calls and puts because the expected stock return has differential impacts on call and put returns. In particular, the expected call (put) option return increases (decreases) with the expected stock return.

On the other hand, both the sign of predictive coefficients and the predictability patterns are consistent with the impact of the volatility risk premium on expected option returns in a stochastic volatility model. Section 2.3.4 shows that as the volatility risk premium becomes more negative, expected option returns decrease, which is consistent with a positive predictive coefficient. Furthermore, as also discussed in Section 2.3.4, OTM options and ATM straddles are most sensitive to changes in the volatility risk premium. These predictions are in line with our empirical findings that OTM options and ATM straddles exhibit strongest return predictability. This suggests that the economic source of option return predictability is likely due to the time varying volatility risk premium embedded in index options.

2.4.4 Predicting Option Returns: the Jump Risk Premium

We also investigate whether the jump risk premium can predict future option returns. Dennis and Mayhew (2002) and Bakshi and Kapadia (2003) establish the links between risk neutral skewness and the slope of the implied volatility curve, and therefore we use the difference in average implied volatilities between OTM and ATM put options as a proxy for the jump risk premium:

$$\text{JUMP}_t = IVOL_{OTM,t} - IVOL_{ATM,t} \tag{2.4}$$

where OTM and ATM refer to put options with 0.90 $=<$ K/S $<=$ 0.94 and 0.98 $=<$ K/S $<=$ 1.02, respectively.

We also report the predictability regression results with the jump risk premium. The jump risk premium does not contain predictive information on future call option returns. It is insignificant across all moneyness groups and R^2s are close to zero. The jump risk premium negatively forecasts future OTM put option returns, with a highly significant Newey-West t-statistic of -2.94 and an adjusted R^2 of 1.45%. The predictability, however, becomes insignificant for other moneyness groups. Lastly, Panel C shows that the jump risk premium does not predict straddle returns.

Overall, we conclude that the jump risk premium significantly predicts future OTM put option returns, but it does not seem to be a significant predictor for other options.

2.5 Robustness

This section includes several robustness checks. We study how different parameterizations might affect expected option returns in the SV and SVJ models. We also investigates the robustness of option return predictability results to a number of implementation choices.

2.5.1 Parameters

Our main analysis shows that the presence of the volatility risk premium implies that both OTM call and put options should earn large negative expected returns, which is consistent with the data. On the other hand, the jump risk premium implies that OTM put options should have large negative expected returns, whereas OTM call options should have large positive expected returns. In this section, we assess how these results might be affected by different parameterizations with respect to both physical measure parameters and risk aversion parameter. We first discuss the volatility risk premium and subsequently the jump risk premium.

Table 2.8 recalculates expected option returns in the SV model by increasing/decreasing each $\mathbb{P}$-measure parameter by one standard deviation. We continue to assume a risk aversion of 10 when computing expected option returns. The results suggest that expected option returns are not particularly sensitive to changes in $\mathbb{P}$-measure parameters, and overall expected returns are very close to those obtained with our baseline parameterization. For example, it is well-known that the volatility mean reverting parameter κ is notoriously difficult to pin down precisely. However, its impact on expected option returns turns out to be small: decreasing or increasing it by one standard deviation produces very similar expected returns.

Table 2.9 reports expected option returns in the SV model with different values of risk aversion ranging from 0 to 20. For $\mathbb{P}$-measure parameters, we use our baseline estimates reported in Table 2.2. Table 2.9 shows that risk aversion has a much larger effect on expected option returns, especially for OTM options. When risk aversion is equal to zero (e.g., volatility risk is not priced), return patterns in the SV model are actually similar to the BSM model: both expected call and put option returns increase with the strike price. In this case, confirming the results in Chambers et al. (2014), we find the SV model is rejected. As risk aversion increases, namely the volatility risk premium becomes more negative, expected option returns decrease. Notice that in order for the SV model to match the data quantitatively, the value of risk aversion should be around 8-12.

We also compute expected option returns using the variance-dependent pricing kernel of Christoffersen, Heston, and Jacobs (2013). Their variance-dependent pricing kernel, when projected onto the stock return, is U-shaped and able to explain a range of option anomalies.[17] With this particular pricing kernel, Christoffersen, Heston, and Jacobs (2013) show that the volatility risk premium is given by:

$$\eta = \gamma\sigma\rho - \xi\sigma^2 \tag{2.5}$$

where the first term is related to risk aversion as before, and the second term originates from variance preferences ξ which, according to Christoffersen, Heston, and Jacobs (2013), should be positive. We assume $\xi = 10$. Table 2.10 repeats the same exercise in Table 2.9 using the above new specification of the volatility risk premium. With the variance-dependent pricing

[17] Many papers find that the pricing kernel is not a monotonically decreasing function of index return. See, among others, Aït-Sahalia and Lo (1998), Jackwerth (2000), Rosenberg and Engle (2002) and Chaudhuri and Schroder (2015). On the other hand, Linn, Shive, and Shumway (2018) point out potential biases in the existing estimates of the pricing kernel. After properly accounting for the conditioning information, they show the pricing kernel is monotonically decreasing with index returns.

kernel, expected option returns are lower than those in Table 2.9 because the volatility risk premium now has an extra component resulting from variance preferences ξ. Also notice that the variance-dependent pricing kernel implies that the risk aversion value needed for the SV model to fit option returns is small.

The Online Appendix also report the corresponding results for the SVJ model. We investigate if our results about the jump risk premium are sensitive to our characterization of jump process under the physical measure by increasing or decreasing each of the jump parameters by one standard deviation. We only focus on jump-related parameters since expected option returns do not vary much with parameters associated with stochastic volatility. Overall, we find the return patterns are similar to our benchmark case. Specifically, the jump risk premium implies very large negative expected returns for OTM put options which is consistent with the data. However, it also implies that expected OTM call option returns are positive and increasing with the strike price, which is inconsistent with the data.

The Online Appendix reports the effect of risk aversion on expected option returns in the SVJ model. An increase in risk aversion leads to a larger jump risk premium, meaning that price jumps occur more frequently and more severely under the risk neutral measure. While the jump risk premium is able to match put option returns easily, its implications on call options are in general inconsistent with the data. For example, across a wide range of risk aversion values, expected returns on OTM calls are positive and increasing with the strike price. If risk aversion is high enough (e.g., 20), it is possible for the SVJ model to yield negative expected returns on OTM calls. However, a very large risk aversion also implies that ATM and ITM calls should earn negative expected returns which is inconsistent with the data.

2.5.2 The Measurement of the Volatility Risk Premium

In the main analysis, we measure the volatility risk premium as the difference between realized volatility and the VIX. In other words, we use realized volatility as a proxy for the expected future volatility under the physical measure. To ensure our empirical results are not driven by this assumption, we also estimate expected physical volatility using the heterogeneous autoregressive model (the HAR model) proposed by Corsi (2009). In particular, we obtain conditional forecasts of future volatility by projecting realized volatility onto lagged realized volatilities computed over difference frequencies:

$$\log RV_{t,t+1} = \delta_0 + \delta_1 \log RV_{t-1,t} + \delta_2 \log RV_t^W + \delta_3 \log RV_t^D + \epsilon$$

where $RV_{t-1,t}$ is realized volatility over the past month, and RV_t^W and RV_t^D denote realized volatilities over the past week and day, respectively. Because realized volatilities are approximately log-normally distributed (Andersen et al., 2001), it is more appropriate to forecast logarithmic of realized volatilities with linear models. The log specification also ensures that we will always obtain positive volatility forecasts. We estimate the above model based on the full sample and take the fitted values as expectations of future realized volatility:

$$\mathbb{E}_t(RV_{t,t+1}) = \exp(\delta_0 + \delta_1 \log RV_{t-1,t} + \delta_2 \log RV_t^W + \delta_3 \log RV_t^D + \frac{1}{2}\sigma_\epsilon^2).$$

Finally we compute the difference between $\mathbb{E}_t(RV_{t,t+1})$ and the VIX to obtain the volatility risk premium.

Table 2.11 contains results of predictive regressions based on this new measure of the volatility risk premium. Consistent with our benchmark results, the volatility risk premium positively predicts futures option returns, with OTM calls, OTM puts and ATM straddles exhibiting the strongest return predictability. We also estimate the volatility risk premium simply as the difference between 30-day historical volatility based on daily returns and the average option implied volatility and find very similar results.

2.5.3 Controlling for Other Predictors

So far we have only focused on univariate predictive regression. In this section, we evaluate the performance of the volatility risk premium in predicting future option returns controlling for other factors including the jump risk premium and the level of volatility. Given return predictability is concentrated among OTM options and ATM straddles, we will focus on these options only.

The results of multivariate predictive regressions are summarized in Table 2.12. Specification (1) considers both the volatility risk premium and the jump risk premium as predictors. After including the jump risk premium as a control, we find the predictive power of the volatility risk premium remains statistically significant. We also find that the volatility risk premium does not subsume the jump risk premium. While the jump risk premium does not predict returns on calls and straddles, it remains a significant predictor over future OTM put option return. This suggests that both the volatility and jump risk premiums are important for OTM put options.

Specification (2) of Table 2.12 controls for the level of volatility. Including volatility as a control does not change our results. The volatility risk premium remains significant in all cases. Note that volatility itself has some forecasting power over future option returns. Specifically, volatility negatively predicts future straddle and call option returns and positively predicts future put option returns, but the predictability is not always statistically significant. These results are broadly consistent with the analysis in Hu and Jacobs (2017).

Finally, specification (3) includes all three variables into the predictive regression. Including both volatility and the jump risk premium as controls does not affect the predictive power of the volatility risk premium. The volatility risk premium remains significant in forecasting future option returns. Finally, we also find that the predictive power of the volatility risk premium is robust to controlling for option betas (e.g., loadings on price, volatility and jump risks), and these results are summarized in the Online Appendix.

2.5.4 Holding-Period Option Returns

In the main analysis, we study the predictive relation between the volatility risk premium and holding-to-maturity index option returns. We also examine if our empirical findings persist to other holding periods. In particular, instead of holding options to maturity, we consider a holding period of 15 calendar days. When option liquidation dates land on a holiday (e.g., the New Year and the Fourth of July), we use the option price information the day before and we assume options trade at the mid-point of bid-ask quotes. Overall, we find similar results when using the volatility risk premium to predict holding-period option returns.

2.6 Conclusion

Both out-of-the-money S&P 500 index call and put options are associated with large negative average returns. This paper investigates how these negative returns are related to the pricing of stochastic volatility and jump risks. We show that the low returns on OTM option are primarily due to the pricing of market volatility risk. A stochastic volatility model in which volatility risk is negatively priced is able to match average returns of call and put options across all strikes as well as returns of a number of option portfolios. Further corroborating the volatility risk premium hypothesis, we document a statistically and economically significant index option return predictability by the volatility risk premium. Overall, our results suggest that the pricing of volatility risk has a first-order effect on the cross-section of index option returns. On the jump risk premium side, our analysis suggests that the pricing of jump risk is also important and some portion of OTM put option returns are related to the jump risk premium.

This paper can be extended in several ways. First, in our theoretical analysis we assume there is only one factor that drives time-varying stochastic volatility. In the data volatility dynamics are much more complex and our analysis can be extended to take this into account.[18] Second, we have focused on the volatility risk premium and extensions to investigating the impact of higher moment (e.g., skewness and kurtosis) risk premiums on expected option returns would be useful. Third, we consider the predictability of option returns in the U.S. market, and it may prove interesting to extend our analysis to international data. We plan to address these in future research.

[18]Existing studies find that at least two factors are needed in order to characterize the volatility dynamics. See, among others, Alizadeh, Brandt, and Diebold (2002), Engle and Rangel (2008) and Christoffersen, Heston, and Jacobs (2009). Another strand of literature emphasizes the importance of incorporating jumps into the volatility dynamics. See, among others, Broadie, Chernov, and Johannes (2007) and Eraker, Johannes, and Polson (2003).

Table 2.1: Average Monthly Returns of S&P 500 Index Options

Panel A: Call Option							
K/S	0.96	0.98	1	1.02	1.04	1.06	1.08
Ret	6.36	6.89	6.50	1.89	-1.47	-18.12	-25.05
Panel B: Put Option							
K/S	0.92	0.94	0.96	0.98	1	1.02	1.04
Ret	-52.07	-45.02	-37.86	-27.76	-22.36	-15.76	-13.15
Panel C: Option Portfolio							
	ATMS	PSP	CNS	CSP			
Ret	-8.47	-18.54	-3.93	13.56			

Notes: This table reports average monthly returns of S&P 500 call and put options for different moneyness (defined as the strike price over the index: K/S), as well as average monthly returns of several option portfolios. For option portfolios, we consider an at-the-money straddle (ATMS), a put spread (PSP) that consists of a short position in a 6% OTM put and a long position in an ATM put, a crash neutral spread (CNS) that consists of a long position in an ATM straddle and a short position in a 6% OTM put, and a call spread (CSP) that consists of a long position in an ATM call and a short position in a 6% OTM call. Returns are reported in percent per month. The sample period is March 1998 to August 2015.

Table 2.2: Parameters

	BSM	SV	SVJ
μ	0.0506	0.0506	0.0506
r	0.0201	0.0201	0.0201
d	0.0174	0.0174	0.0174
σ_{BSM}	0.1905		
κ		6.4130	5.9859
		(0.923)	(0.909)
θ		0.0363	0.0358
		(0.004)	(0.004)
σ		0.5472	0.5423
		(0.033)	(0.035)
ρ		-0.7944	-0.8015
		(0.026)	(0.028)
λ			0.9658
			(0.114)
μ_z			-0.0209
			(0.007)
σ_z			0.0677
			(0.009)
η		-4.3470	0.0000
$\lambda^{\mathbb{Q}}$			1.4969
$\mu_z^{\mathbb{Q}}$			-0.0667

Notes: This table reports parameter values that we use to compute expected option returns for different models. The equity risk premium (μ), risk-free rate (r) and dividend yield (d) are calibrated to match those observed in our sample. For the BSM model, the constant volatility parameter (σ_{BSM}) is equal to the square root of the long-run variance (θ) in the SV model. For the SV and SVJ models, we use particle filtering to estimate the remaining $\mathbb{P}$-measure parameters and report standard errors of those estimates in the parentheses. The volatility risk premium (η) and $\mathbb{Q}$-measure jump parameters ($\lambda^{\mathbb{Q}}$ and $\mu_z^{\mathbb{Q}}$) are obtained based on equations (2.1) and (2.2) with a risk aversion of 10. All parameters are reported in annual terms.

Table 2.3: Expected Option Returns: the Black-Scholes-Merton Model

			Panel A: Call Option				
K/S	0.96	0.98	1	1.02	1.04	1.06	1.08
Data	6.36	6.89	6.50	1.89	-1.47	-18.12	-25.05
Model	7.25	8.67	10.30	12.13	14.12	16.27	18.54
Simulation	7.18	8.58	10.19	12.00	14.01	16.18	18.53
p-value	0.45	0.42	0.37	0.23	0.19	0.07	0.08
			Panel B: Put Option				
K/S	0.92	0.94	0.96	0.98	1	1.02	1.04
Data	-52.07	-45.02	-37.86	-27.76	-22.36	-15.76	-13.15
Model	-16.01	-14.07	-12.24	-10.54	-8.99	-7.61	-6.40
Simulation	-15.92	-13.99	-12.19	-10.49	-8.94	-7.56	-6.35
p-value	0.10	0.05	0.04	0.06	0.06	0.12	0.12
			Panel C: Option Portfolio				
Portfolio	ATMS	PSP	CNS	CSP			
Data	-8.47	-18.54	-3.93	13.56			
Model	0.71	-8.03	1.97	8.88			
Simulation	0.67	-7.99	1.93	8.77			
p-value	0.03	0.09	0.11	0.29			

Notes: This table compares average option returns in the data reported in Table 2.1 with expected option returns implied from the Black-Scholes-Merton model (BSM). "Model" represents expected option returns computed analytically using BSM parameters reported in Table 2.2. We also simulate 25000 sample paths of the index from which we report the average simulated option returns (denoted by "Simulation") and p-values. The p-values are calculated as the percentage of the 25000 simulated option returns that is less than realized option returns. Sample paths are simulated based on the same parameters used for computing expected option returns.

Table 2.4: The Volatility Risk Premium and Expected Option Returns

Panel A: Call Option							
K/S	0.96	0.98	1	1.02	1.04	1.06	1.08
Data	6.36	6.89	6.50	1.89	-1.47	-18.12	-25.05
Model	4.82	5.10	5.14	4.50	1.12	-8.53	-22.54
Simulation	4.90	5.24	5.38	4.94	1.78	-7.99	-21.34
p-value	0.40	0.40	0.44	0.41	0.48	0.50	0.30
Panel B: Put Option							
K/S	0.92	0.94	0.96	0.98	1	1.02	1.04
Data	-52.07	-45.02	-37.86	-27.76	-22.36	-15.76	-13.15
Model	-30.43	-26.69	-22.96	-19.29	-15.74	-12.39	-9.40
Simulation	-30.54	-26.77	-23.03	-19.33	-15.76	-12.42	-9.42
p-value	0.22	0.19	0.18	0.26	0.26	0.34	0.28
Panel C: Option Portfolio							
Portfolio	ATMS	PSP	CNS	CSP			
Data	-8.47	-18.54	-3.93	13.56			
Model	-5.24	-12.33	-2.52	6.74			
Simulation	-5.13	-12.37	-2.41	6.94			
p-value	0.24	0.25	0.36	0.20			

Notes: This table compares average option returns in the data reported in Table 2.1 with expected option returns implied from the Heston stochastic volatility model (SV) in which volatility risk is priced. "Model" represents expected option returns computed analytically using SV parameters reported in Table 2.2. We also simulate 25000 sample paths of the index from which we report the average simulated option returns (denoted by "Simulation") and p-values. The p-values are calculated as the percentage of the 25000 simulated option returns that is less than realized option returns. Sample paths are simulated based on the same parameters used for computing expected option returns.

Table 2.5: The Jump Risk Premium and Expected Option Returns

Panel A: Call Option							
K/S	0.96	0.98	1	1.02	1.04	1.06	1.08
Data	6.36	6.89	6.50	1.89	-1.47	-18.12	-25.05
Model	2.96	2.70	2.32	2.34	7.31	29.39	64.09
Simulation	3.01	2.74	2.35	2.39	7.63	29.65	64.05
p-value	0.26	0.26	0.30	0.50	0.40	0.25	0.29
Panel B: Put Option							
K/S	0.92	0.94	0.96	0.98	1	1.02	1.04
Data	-52.07	-45.02	-37.86	-27.76	-22.36	-15.76	-13.15
Model	-41.93	-35.61	-29.10	-22.71	-16.85	-11.95	-8.38
Simulation	-42.22	-35.93	-29.40	-22.98	-17.05	-12.09	-8.46
p-value	0.32	0.31	0.29	0.35	0.30	0.33	0.24
Panel C: Option Portfolio							
Portfolio	ATMS	PSP	CNS	CSP			
Data	-8.47	-18.54	-3.93	13.56			
Model	-7.21	-9.30	-2.45	1.46			
Simulation	-7.30	-9.47	-2.50	1.49			
p-value	0.41	0.16	0.38	0.04			

Notes: This table compares average option returns in the data reported in Table 2.1 with expected option returns implied from the SVJ model in which only jump risk is priced, but volatility risk is not. “Model” represents expected option returns computed analytically using SVJ parameters reported in Table 2.2. We also simulate 25000 sample paths of the index from which we report the average simulated option returns (denoted by “Simulation”) and p-values. The p-values are calculated as the percentage of the 25000 simulated option returns that is less than realized option returns. Sample paths are simulated based on the same parameters used for computing expected option returns..

Table 2.6: Predicting Option Returns: the Volatility Risk Premium

	Panel A: Call Option		
	$0.96 <= K/S < 1.00$	$1.00 <= K/S < 1.04$	$1.04 <= K/S < 1.08$
Intercept	0.11	0.27	1.28
	(1.66)	(1.57)	(1.50)
VRP	-0.06	4.35	24.44
	(-0.04)	(1.70)	(2.11)
Adj. R^2	-0.06%	0.37%	0.99%
	Panel B: Put Option		
	$0.92 <= K/S < 0.96$	$0.96 <= K/S < 1.00$	$1.00 <= K/S < 1.04$
Intercept	-0.23	-0.12	-0.09
	(-1.21)	(-0.83)	(-0.72)
VRP	8.37	4.38	2.70
	(2.21)	(1.53)	(1.21)
Adj. R^2	2.04%	0.64%	0.55%
	Panel C: Straddle		
	$0.94 <= K/S < 0.98$	$0.98 <= K/S < 1.02$	$1.02 <= K/S < 1.06$
Intercept	0.08	0.04	-0.04
	(1.91)	(0.83)	(-0.50)
VRP	1.47	2.63	2.39
	(1.84)	(2.83)	(1.76)
Adj. R^2	0.87%	1.59%	1.41%

Notes: This table reports results of the following option return predictability regression:

$$option\_ret^i_{t,\,t+1} = \alpha^i + \beta^i \mathrm{VRP}_t + \epsilon, \quad i \in \{call, put, straddle\}$$

where *option_ret* is monthly holding-to-maturity returns on call options (Panel A), put options (Panel B) and straddles (Panel C). The predictor VRP_t is computed as the difference between realized volatility and the VIX. Realized volatility is constructed based on 5-min log returns on S&P 500 futures over past 30 calendar days. We run predictive regressions for different moneyness groups as indicated by different columns. Newey-West t-statistics with 4 lags are reported in the parentheses. The sample period is March 1998 to August 2015.

Table 2.7: Descriptive Statistics of Option Trading Strategies

	mean	std	sr	skew	holding-period
Panel A: Index					
	mean	std	sr	skew	holding-period
S&P 500	0.004	0.045	0.082	-0.639	210
Panel B: 4% OTM Call					
	mean	std	sr	skew	holding-period
Always	-0.015	3.672	-0.004	6.803	209
VRP< 0	-0.157	3.272	-0.048	8.146	187
Panel C: 6% OTM Call					
	mean	std	sr	skew	holding-period
Always	-0.181	6.179	-0.029	12.695	206
VRP< 0	-0.581	1.873	-0.310	5.605	184
Panel D: 4% OTM Put					
	mean	std	sr	skew	holding-period
Always	-0.379	2.164	-0.175	4.250	207
VRP< 0	-0.470	1.973	-0.238	4.792	185
Panel E: 6% OTM Put					
	mean	std	sr	skew	holding-period
Always	-0.450	2.468	-0.182	5.219	206
VRP< 0	-0.575	2.216	-0.259	6.221	185
Panel F: ATM Straddle					
	mean	std	sr	skew	holding-period
Always	-0.085	0.739	-0.115	1.430	209
VRP< 0	-0.106	0.704	-0.151	1.462	188

Notes: This table reports mean, standard deviation (std), Sharpe ratio (sr) and skewness (skew) of returns of several trading strategies. Panel A reports on the S&P 500. Panels B to F report the performance of writing a 4% OTM call, a 6% OTM call, a 4% OTM put, a 6% OTM put and an ATM straddle. We consider two option selling strategies: "Always" and "VRP <0". "Always" shorts index options in every month. "VRP <0" shorts index options only in months when the observed market volatility risk premium is negative. Returns are reported from the perspective of a long investor. The sample period is March 1998 to August 2015.

Table 2.8: Robustness: Stochastic Volatility Parameters

Panel A: Call Option							
K/S	0.96	0.98	1	1.02	1.04	1.06	1.08
Baseline	4.82	5.10	5.14	4.50	1.12	-8.53	-22.54
$\kappa+$	4.81	5.08	5.03	4.29	1.07	-6.99	-20.07
$\kappa-$	4.97	5.46	5.94	6.12	1.85	-10.64	-26.04
$\theta+$	4.42	4.49	4.23	3.21	0.03	-7.94	-20.06
$\theta-$	5.27	5.85	6.42	6.67	2.81	-9.64	-25.68
$\sigma+$	4.61	4.85	4.86	4.18	0.01	-11.26	-26.08
$\sigma-$	4.96	5.30	5.43	4.99	2.31	-5.89	-18.75
$\rho+$	5.00	5.39	5.58	5.15	1.94	-7.13	-20.14
$\rho-$	4.70	4.91	4.84	4.07	0.50	-10.01	-25.32
Panel B: Put Option							
K/S	0.92	0.94	0.96	0.98	1	1.02	1.04
Baseline	-30.43	-26.69	-22.96	-19.29	-15.74	-12.39	-9.40
$\kappa+$	-30.18	-26.43	-22.69	-19.03	-15.51	-12.22	-9.31
$\kappa-$	-31.00	-27.26	-23.52	-19.81	-16.19	-12.72	-9.53
$\theta+$	-29.56	-25.88	-22.25	-18.70	-15.29	-12.11	-9.29
$\theta-$	-31.42	-27.61	-23.79	-20.01	-16.30	-12.76	-9.52
$\sigma+$	-30.88	-27.13	-23.39	-19.69	-16.09	-12.67	-9.56
$\sigma-$	-29.85	-26.13	-22.44	-18.83	-15.36	-12.11	-9.23
$\rho+$	-30.21	-26.50	-22.80	-19.16	-15.64	-12.31	-9.33
$\rho-$	-30.77	-26.98	-23.21	-19.50	-15.90	-12.52	-9.48
Panel C: Option Portfolio							
	ATMS	PSP	CNS	CSP			
Baseline	-5.24	-12.33	-2.52	6.74			
$\kappa+$	-5.28	-12.12	-2.57	6.57			
$\kappa-$	-5.05	-12.87	-2.39	7.55			
$\theta+$	-5.48	-11.68	-2.66	6.06			
$\theta-$	-4.87	-13.13	-2.27	7.80			
$\sigma+$	-5.55	-12.50	-2.74	6.65			
$\sigma-$	-4.90	-12.10	-2.27	6.90			
$\rho+$	-4.96	-12.37	-2.32	7.08			
$\rho-$	-5.47	-12.38	-2.69	6.50			

Notes: This table reports expected option returns for the SV model by increasing (+) and decreasing (-) each $\mathbb{P}$-measure parameter by one standard deviation. Expected option returns based on our baseline parameterization are also included for comparison. Returns are in percent per month.

Table 2.9: The Impact of Risk Aversion: A Sensitivity Analysis

Panel A: Call Option							
$\gamma \backslash K/S$	0.96	0.98	1	1.02	1.04	1.06	1.08
0	7.49	9.74	13.36	19.77	30.47	42.42	52.86
2	6.99	8.86	11.76	16.66	24.06	30.17	32.46
4	6.55	8.08	10.36	14.03	18.66	19.58	15.45
6	5.98	7.09	8.60	10.75	12.36	8.94	0.35
8	5.49	6.24	7.11	8.00	7.06	-0.14	-12.30
10	4.82	5.10	5.14	4.50	1.12	-8.53	-22.54
12	4.25	4.15	3.54	1.75	-3.73	-15.87	-31.44
14	3.80	3.38	2.18	-0.72	-8.24	-22.77	-39.74
16	3.01	2.14	0.22	-3.85	-12.88	-28.33	-45.40
18	2.40	1.15	-1.40	-6.53	-17.08	-33.73	-51.12
20	1.87	0.25	-2.90	-9.04	-21.06	-38.81	-56.33
Panel B: Put Option							
$\gamma \backslash K/S$	0.92	0.94	0.96	0.98	1	1.02	1.04
0	-11.04	-10.62	-10.15	-9.61	-8.96	-8.15	-7.06
2	-15.34	-14.12	-12.88	-11.63	-10.35	-9.00	-7.52
4	-19.58	-17.61	-15.66	-13.73	-11.83	-9.93	-8.02
6	-23.37	-20.74	-18.15	-15.60	-13.12	-10.72	-8.45
8	-27.16	-23.93	-20.73	-17.58	-14.52	-11.62	-8.94
10	-30.43	-26.69	-22.96	-19.29	-15.74	-12.39	-9.40
12	-33.77	-29.55	-25.32	-21.14	-17.09	-13.28	-9.90
14	-37.16	-32.49	-27.78	-23.09	-18.51	-14.20	-10.42
16	-39.85	-34.85	-29.77	-24.69	-19.73	-15.05	-10.97
18	-42.71	-37.39	-31.94	-26.45	-21.05	-15.96	-11.52
20	-45.56	-39.94	-34.14	-28.24	-22.41	-16.88	-12.06
Panel C: Option Portfolio							
$\gamma \backslash$ Portfolio	ATMS	PSP	CNS	CSP			
0	2.27	-8.51	3.68	11.99			
2	0.78	-9.30	2.45	10.97			
4	-0.66	-10.22	1.25	10.14			
6	-2.19	-10.91	0.00	8.97			
8	-3.64	-11.76	-1.21	8.05			
10	-5.24	-12.33	-2.52	6.74			
12	-6.71	-13.09	-3.74	5.78			
14	-8.10	-14.01	-4.92	4.96			
16	-9.70	-14.53	-6.22	3.80			
18	-11.17	-15.27	-7.46	2.84			
20	-12.60	-16.10	-8.66	1.95			

Notes: This table reports expected option returns for the SV model using different values of risk aversion (γ) ranging from 0 to 20. The remaining parameters are based on Table 2.2. Returns are in percent per month.

Table 2.10: The Impact of Risk Aversion: A Variance-Dependent Pricing Kernel

Panel A: Call Option							
$\gamma\backslash K/S$	0.96	0.98	1	1.02	1.04	1.06	1.08
0	5.69	6.59	7.72	9.12	9.45	4.58	-5.38
2	5.11	5.60	6.01	6.08	3.95	-4.12	-16.75
4	4.57	4.68	4.42	3.23	-1.09	-11.84	-26.63
6	4.11	3.90	3.04	0.76	-5.68	-19.05	-35.54
8	3.47	2.84	1.30	-2.18	-10.39	-25.32	-42.44
10	2.95	1.98	-0.16	-4.71	-14.71	-31.32	-49.06
12	2.34	0.98	-1.79	-7.38	-18.86	-36.51	-54.26
14	1.49	-0.31	-3.75	-10.28	-22.70	-40.52	-57.83
16	0.85	-1.32	-5.35	-12.81	-26.36	-44.80	-61.92
18	0.29	-2.23	-6.83	-15.20	-29.89	-48.90	-65.78
20	-0.50	-3.42	-8.59	-17.73	-33.06	-52.10	-68.51
Panel B: Put Option							
$\gamma\backslash K/S$	0.92	0.94	0.96	0.98	1	1.02	1.04
0	-24.90	-22.01	-19.16	-16.36	-13.64	-11.05	-8.64
2	-28.44	-24.99	-21.57	-18.21	-14.96	-11.89	-9.11
4	-31.91	-27.95	-24.00	-20.10	-16.33	-12.78	-9.62
6	-35.38	-30.95	-26.49	-22.06	-17.76	-13.70	-10.13
8	-38.38	-33.56	-28.68	-23.80	-19.05	-14.57	-10.65
10	-41.45	-36.27	-30.97	-25.65	-20.42	-15.48	-11.17
12	-44.24	-38.75	-33.10	-27.38	-21.73	-16.37	-11.72
14	-46.56	-40.85	-34.94	-28.92	-22.95	-17.28	-12.34
16	-49.10	-43.16	-36.96	-30.60	-24.26	-18.21	-12.93
18	-51.64	-45.50	-39.02	-32.33	-25.59	-19.13	-13.50
20	-53.76	-47.48	-40.81	-33.87	-26.86	-20.09	-14.16
Panel C: Option Portfolio							
$\gamma\backslash$ Portfolio	ATMS	PSP	CNS	CSP			
0	-2.90	-11.15	-0.58	8.36			
2	-4.41	-11.87	-1.83	7.27			
4	-5.89	-12.66	-3.06	6.29			
6	-7.29	-13.57	-4.25	5.47			
8	-8.81	-14.24	-5.50	4.40			
10	-10.23	-15.09	-6.70	3.52			
12	-11.70	-15.82	-7.93	2.56			
14	-13.29	-16.32	-9.23	1.51			
16	-14.75	-17.05	-10.46	0.59			
18	-16.16	-17.88	-11.65	-0.28			
20	-17.68	-18.48	-12.92	-1.21			

Notes: This table reports expected option returns for the SV model using different values of risk aversion (γ) ranging from 0 to 20. The volatility risk premium is computed based on the variance-dependent pricing kernel of Christoffersen, Heston, and Jacobs (2013). The remaining parameters are based on Table 2.2. Returns are in percent per month.

Table 2.11: Robustness: Alternative Measures of the Volatility Risk Premium

		Panel A: Call Option	
	$0.96 <= K/S < 1.00$	$1.00 <= K/S < 1.04$	$1.04 <= K/S < 1.08$
Intercept	0.10	0.22	1.04
	(1.67)	(1.40)	(1.39)
VRP	-0.37	3.32	20.56
	(-0.31)	(1.41)	(2.05)
Adj. R^2	-0.04%	0.20%	0.73%
		Panel B: Put Option	
	$0.92 <= K/S < 0.96$	$0.96 <= K/S < 1.00$	$1.00 <= K/S < 1.04$
Intercept	-0.28	-0.15	-0.11
	(-1.65)	(-1.13)	(-0.98)
VRP	8.03	4.14	2.50
	(2.27)	(1.49)	(1.15)
Adj. R^2	1.98%	0.60%	0.49%
		Panel C: Straddle	
	$0.94 <= K/S < 0.98$	$0.98 <= K/S < 1.02$	$1.02 <= K/S < 1.06$
Intercept	0.07	0.02	-0.07
	(1.68)	(0.38)	(-0.89)
VRP	1.16	2.27	2.02
	(1.52)	(2.48)	(1.52)
Adj. R^2	0.56%	1.24%	1.01%

Notes: This table reports results of the following option return predictability regression:

$$option\_ret^i_{t,\, t+1} = \alpha^i + \beta^i \mathrm{VRP}_t + \epsilon, \quad i \in \{call, put, straddle\}$$

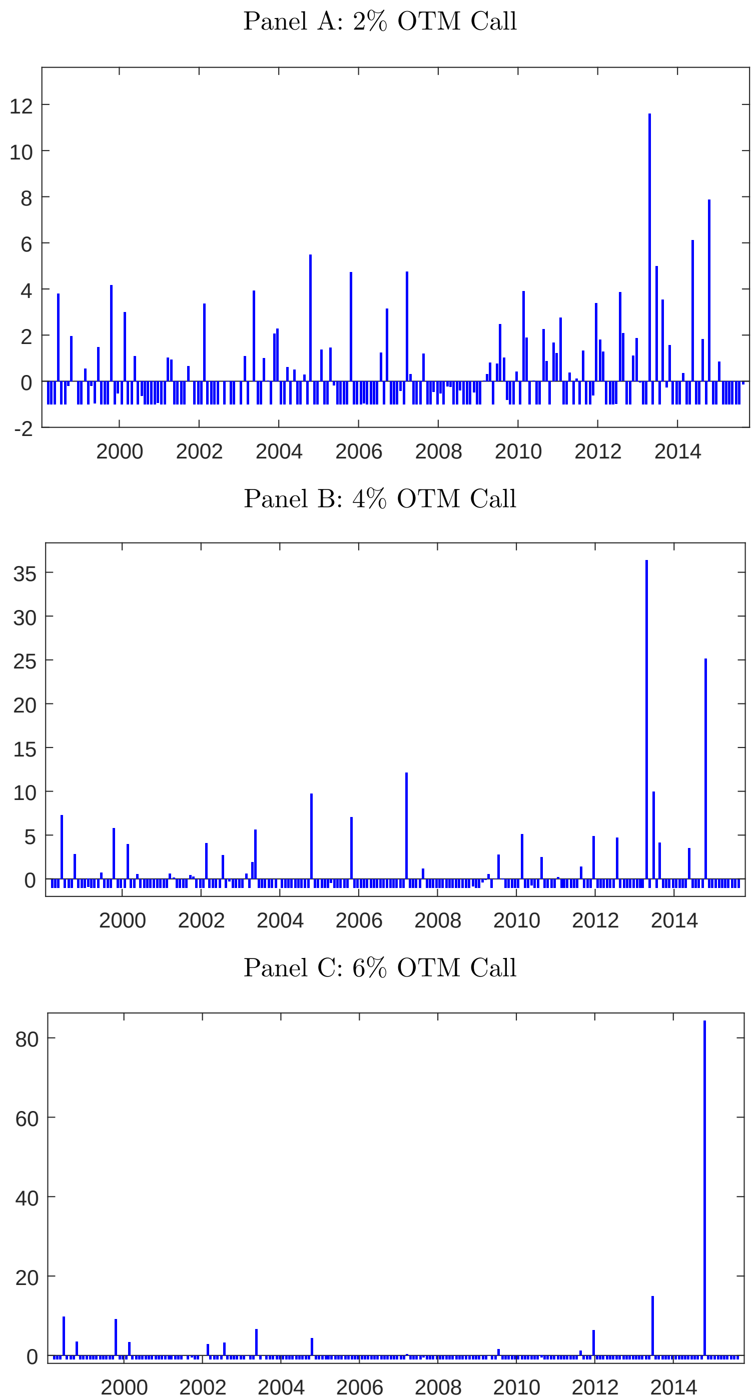
where *option_ret* is monthly holding-to-maturity returns on call options (Panel A), put options (Panel B) and straddles (Panel C). The predictor VRP_t is computed as the difference between expected future realized volatility and the VIX. Expected future realized volatility is estimated using the Heterogeneous Autoregressive Model (the HAR model) of Corsi (2009). We run predictive regressions for different moneyness groups as indicated by different columns. Newey-West t-statistics with 4 lags are reported in the parentheses. The sample period is March 1998 to August 2015.

Table 2.12: Robustness: Controlling for Other Predictors

	OTM Call			OTM Put			ATM Straddle		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Intercept	-0.05	1.48	0.02	1.13	-0.82	0.31	0.09	0.05	0.11
	(-0.08)	(1.44)	(0.03)	(1.83)	(-5.06)	(0.55)	(0.44)	(0.70)	(0.55)
VRP	24.82	25.31	24.98	7.99	5.71	5.98	2.62	2.67	2.69
	(2.09)	(2.09)	(2.10)	(2.19)	(1.99)	(2.01)	(2.85)	(3.14)	(3.13)
JUMP	18.50		18.12	-18.47		-13.68	-0.63		-0.77
	(1.32)		(1.27)	(-2.89)		(-2.26)	(-0.25)		(-0.31)
RV		-1.02	-0.21		3.16	2.51		-0.04	-0.08
		(-0.53)	(-0.10)		(2.82)	(2.31)		(-0.12)	(-0.22)
Adj. R^2	0.99%	0.93%	0.92%	3.27%	3.48%	4.07%	1.55%	1.53%	1.49%

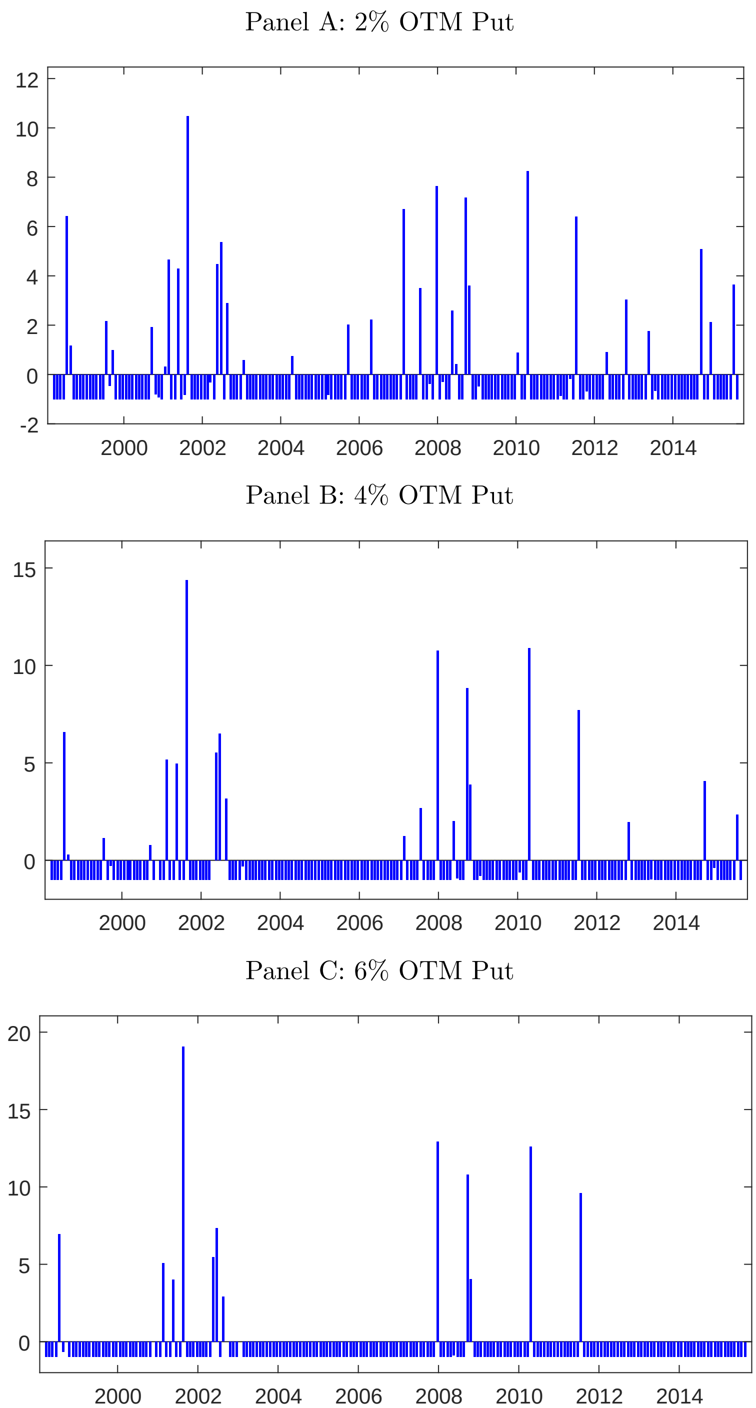
Notes: This table reports results of multivariate predictive regressions. We use realized volatility (RV), the volatility risk premium (VRP) and the jump risk premium (JUMP) to predict future returns on OTM calls ($1.04 <= K/S < 1.08$), OTM puts ($0.92 <= K/S < 0.96$) and ATM straddles ($0.98 <= K/S < 1.02$). RV is constructed based on 5-min log returns on S&P 500 futures over past 30 calendar days. VRP is computed as the difference between RV and the VIX. JUMP is computed as the difference between the average implied volatility from OTM put options and that from ATM put options. Newey-West t-statistics with 4 lags are reported in the parentheses. The sample period is March 1998 to August 2015.

Figure 2.1: The Time Series of OTM Call Option Returns



Notes: This figure plots the time series of monthly holding-to-maturity returns of 2% OTM call (Panel A), 4% OTM call (Panel B) and 6% OTM call (Panel C).

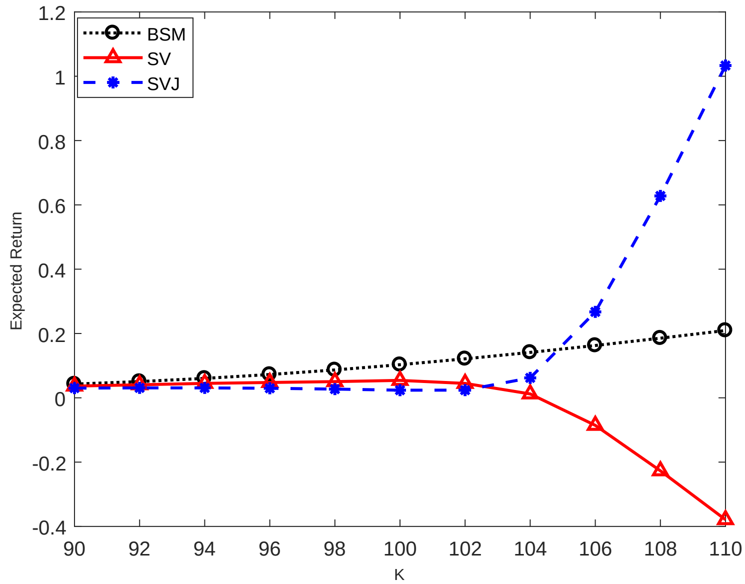
Figure 2.2: The Time Series of OTM Put Option Returns



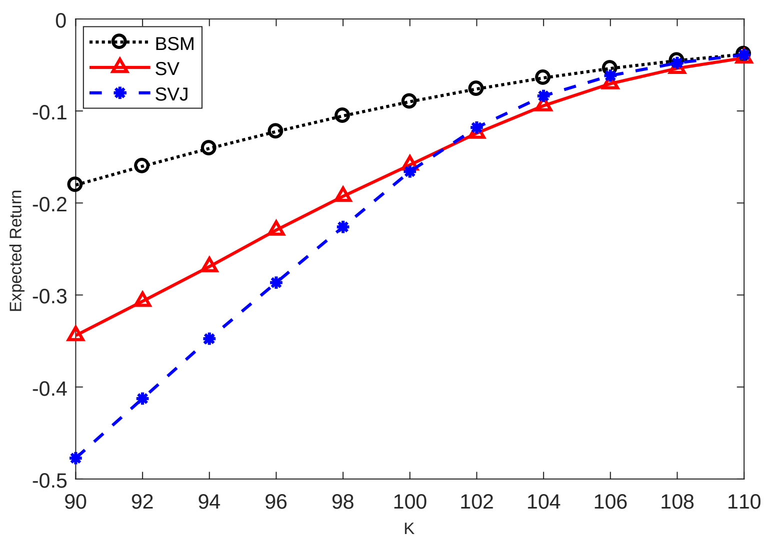
Notes: This figure plots the time series of monthly holding-to-maturity returns of 2% OTM put (Panel A), 4% OTM put (Panel B) and 6% OTM put (Panel C).

Figure 2.3: Moneyness and Expected Option Returns

Panel A: Expected Call Option Returns

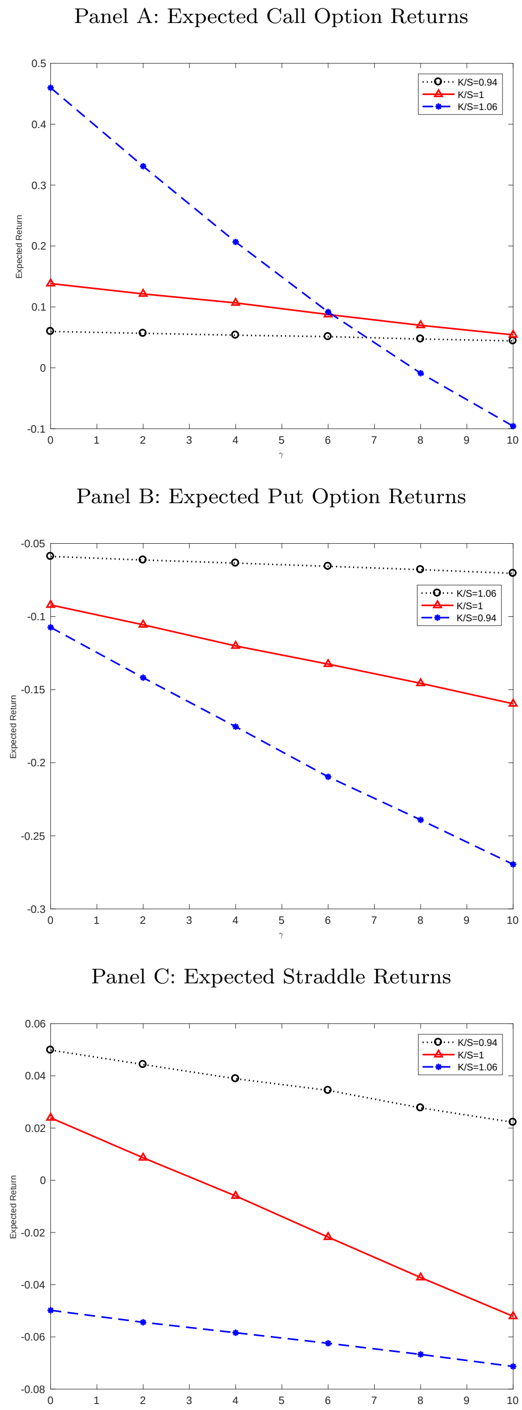


Panel B: Expected Put Option Returns



Notes: This figure plots expected option returns against the strike price for a benchmark BSM model, a SV model in which volatility risk is priced and a SVJ model in which jump risk is priced, but volatility risk is not. Panel A is for call options and Panel B for put options. Expected option returns are computed analytically based on parameters reported in Table 2.2.

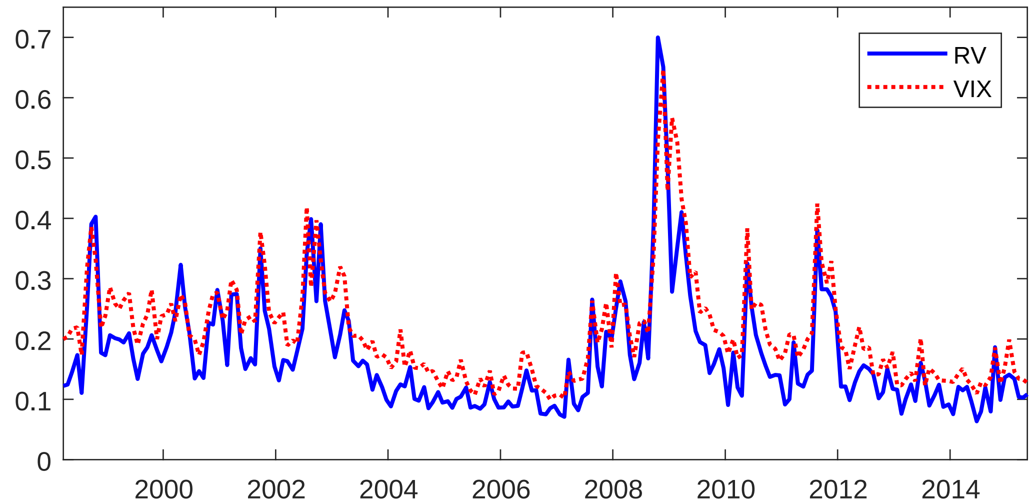
Figure 2.4: The Volatility Risk Premium and Expected Option Returns



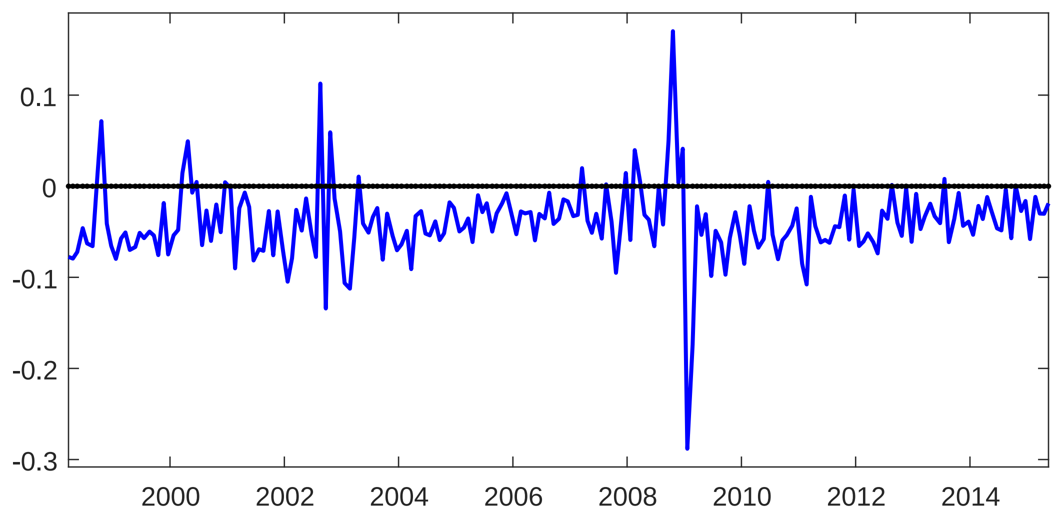
Notes: This figure plots expected option returns against risk aversion coefficient (γ) in the SV model: Panel A for calls, Panel B for puts and Panel C for straddles. A higher γ corresponds to a more negative volatility risk premium. The remaining parameters required for computing expected returns are based on Table 2.2.

Figure 2.5: Realized Volatility, the VIX and the Volatility Risk Premium

Panel A: Realized Volatility and the VIX



Panel B: Volatility Risk Premium



Notes: This figure plots the time series of monthly realized volatility (RV) and the VIX (Panel A), as well as their difference which is the volatility risk premium (Panel B). The sample period is March 1998 to August 2015.

Chapter 3

The Time-Series Information in Options: Implications for the Index Option Return Puzzle

3.1 Introduction

The option pricing models require the identification of three sets of unknowns: physical parameters, risk-neutral parameters[1] and the (physical) latent spot variances.[2] We can extract the physical information, which includes physical parameters and the latent spot variances, from underlying returns and risk-neutral parameters from options. We therefore usually equate the underlying with physical information and options with risk-neutral information. An option contract, a derivative on the underlying, is also a tradeable asset, and the its price changes in the time-series dimension, namely option returns, contain all relevant physical information about the underlying. At each time, option prices are evaluated under the risk-neutral measure. Whereas option price changes are also driven by the dynamics of underlying

[1] Risk premia can be inferred as the differences between physical and risk-neutral parameters.

[2] There are no risk-neutral spot variances, or the risk-neutral spot variances should be identical to physical spot variances.

price and the spot variance under the physical measure.

Our main contribution in this paper is to identify the physical information from daily options, without using the underlying returns. We use the spot variance dynamic to filter daily spot variances, and write the option return as a function of option partial derivatives and physical parameters. Therefore, all the physical parameters and the spot variances can be identified from options. We show that both the physical parameters and the spot variances can be more precisely estimated using options than using underlying returns. Moreover, parameters estimated exclusively from options improve the fit of monthly hold-to-maturity option returns.

Our method differs from the existing literature because we identify the physical information from options, rather than from underlying returns. The estimation of physical information from underlying returns is relatively straightforward. A large number of studies are available on this estimation using various methods.[3] The literature on the estimation of options is relatively more limited, partly because of the computational burden on option evaluation. More importantly, existing literature effectively focuses on the cross-sectional information and partly or entirely ignores the time-series information in options. Some existing studies fix both the physical parameters and spot variances identified from underlying returns and then estimate only risk-neutral parameters from large cross-section of option prices (see Broadie et al., 2007, among others). Others make partial use of option time-series to filter the spot variances but do not identify all the physical parameters (see, for example Bates, 2000; Andersen et al., 2015a).[4] Joint estimation using both underlying returns and option prices simultaneously is possible, but usually only a relatively small number of options in the cross-sectional dimension (sometimes only at-the-money short maturity options) are included (see, among others, Pan, 2002; Eraker, 2004; Aït-Sahalia and Kimmel, 2007; Hurn et al., 2015). In this joint estimation, options help to identify the spot variances, but the

[3]See for example Chernov and Ghysels (2000), Andersen et al. (2002), Chacko and Viceira (2003), Chernov et al. (2003), Eraker et al. (2003), Jones (2003), Bates (2006), Christoffersen et al. (2010)

[4]Note that Andersen et al. (2015a) impose additional consistency restriction with realized variances from high-frequency data, but the spot variances information can be thought of as mainly comes from options.

physical parameters information still comes from underlying returns.

Our paper is also closely related to the Recovery Theorem (see Ross, 2015) as we also recover the physical probability from derivatives (options). Ross (2015) recovers the physical probability from the cross-sectional option prices. In our paper, we estimate the physical probability from the time-series option prices. We show that options contain the risk-neutral information in the cross-sectional prices and the physical information in the time-series prices.

We compare our estimation results based on $S\&P$ 500 index option prices and index option returns with estimates obtained using index returns only or using index returns and option prices. Most physical parameters are identified with much smaller standard deviations from index options than from index returns. This improvement in parameter identification is due to our large option panels that help to pin down the spot variances. The more precisely identified spot variances in turn reduce the standard deviations of physical parameters that either describe or depend on the spot variances. Therefore, once option prices are included, using more than one option return does not help much to identify the parameters related to the index return[5], because all the option returns are driven by the same realization of the index return process and thus are perfectly correlated. We show that the equity risk premium parameter estimated using options is consistently different from that estimated using index returns, including all models from the Black-Scholes Model (Black and Scholes, 1973) to the Double Jump Stochastic Volatility Model (Duffie et al., 2000). Based on the posterior distribution, we cannot conclude that the two markets are inconsistent, but we show that the posterior mean from index returns is indeed inconsistent with options market. In other words, completely ignoring parameter uncertainty and simply assuming the posterior mean from index returns to be the true parameter value for options market may lead to biased estimation result.

Using the physical information extracted from options, we re-examine the index option

[5]For example, the equity risk premium parameter and mean jump size parameter. These parameters are not much more precisely identified from index options than from index returns and index option prices jointly.

return puzzle. The index option return puzzle refers to the existence of extremely negative out-of-the-money (OTM) put and call returns, and abnormal returns of option strategies, such as at-the-money (ATM) straddle. The early literature maintained that options are mispriced and realized option returns are inconsistent with theoretical returns predicted by classic asset pricing models.[6] More recently, Broadie et al. (2009) argue that if estimation risk is properly taken into account, there is no puzzle.[7] Branger et al. (2010) show that models with stochastic volatility and jumps can generate almost any structure of option returns by changing various risk premia, especially the jump-related ones. However, Chambers et al. (2014) re-examine the puzzle based on the Broadie et al. (2009) approach using 1987-2012 data and conclude that option returns are still puzzling. These existing studies usually estimate and fix the physical parameters from index returns and then infer risk-neutral parameters to investigate the puzzle. Researchers agree that risk premia and estimation risk should play important roles, but there is no consistency on how to take them into account.

Tests based on estimates from index returns constitutes a joint hypothesis test in the sense that when rejected, we do not know whether the options are mispriced, or the parameters estimated from index returns are not consistent with index options. In this paper, we test the index option return puzzle by decomposing the test into an option mispricing test and index inconsistency test. The rejection in either of the above two can lead to the option return puzzle. In the option mispricing test, we find that as long as jumps, either in return or variance, or both, are introduced, there is no option mispricing. The rejection of the Black-Scholes and Heston Stochastic Volatility model is likely due to model misspecification. In the index inconsistency test, index returns could not reject any model we test. Our parameters estimated from options provide a direction to take into account risk premia and estimation risk for the traditional approach. If all the parameters are adjusted to our option

[6] See, for example, Jackwerth (2000), Coval and Shumway (2001), Bondarenko (2014), Santa-Clara and Saretto (2004), Jones (2006).

[7] To capture the impact of estimation risk, one may adjust each parameter by a certain amount (say, one standard deviation), but it is not clear whether we should increase or decrease a certain parameter in order to resolve the puzzle.

based values, most models would be consistent with both index returns and option returns.

The rest of this paper is structured as follows: Section 3.2 introduces the models and presents the dynamics of option returns. We also explain how we calculate the expected option returns to test the index option return puzzle. Section 3.3 discusses the estimation method based on option prices and option returns. Section 3.4 presents the model estimation results. Section 3.5 re-examines the index option return puzzle using these estimates. Section 3.6 concludes.

3.2 The Model

We first discuss the model dynamics for both index returns and option returns. Then we discuss the expected option return in closed-form, which will be needed in testing the index option return puzzle. We consider the stochastic volatility model with contemporaneous jumps in return and variance (SVCJ) model[8] and all its nested models. The reason we choose this classic model class that we have rich existing results to compare with. Note that the method proposed in this paper also applies to other model classes, such as the multi-factor model. Suppose the spot price and variance follow the continuous dynamic as below:

$$\frac{dS_t}{S_t} = (r_t - \delta_t + \gamma_t - \lambda\bar{\mu}_s)dt + \sqrt{V_t}dZ_t + (e^{J_t^s} - 1)dN_t \tag{3.1}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t + J_t^v dN_t \tag{3.2}$$

where S_t stands for the index level, r_t is the risk-free rate, δ_t is the dividend yield, γ_t is the total equity risk premium, κ denotes the speed of mean reversion, θ is the unconditional long run 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 related to return and variance, with correlation ρ_J. We

[8]Proposed by Duffie et al. (2000).

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:

$$\bar{\mu}_s = \frac{e^{(\mu_s + \sigma_s^2/2)}}{1 - \rho_J \mu_v} - 1 \tag{3.3}$$

We assume that the risk-neutral dynamic follows the same structure:

$$\frac{dS_t}{S_t} = (r_t - \delta_t - \lambda\bar{\mu}_s^{\mathbb{Q}})dt + \sqrt{V_t}dZ_t^{\mathbb{Q}} + (e^{J_t^{s\mathbb{Q}}} - 1)dN_t^{\mathbb{Q}} \tag{3.4}$$

$$dV_t = \kappa^{\mathbb{Q}}(\theta^{\mathbb{Q}} - V_t)dt + \sigma\sqrt{V_t}dW_t^{\mathbb{Q}} + J_t^{v\mathbb{Q}}dN_t^{\mathbb{Q}} \tag{3.5}$$

where we assume that the diffusive variance risk premium is equal to $\eta_v V_t$, and thus $\kappa^{\mathbb{Q}} = \kappa - \eta_v$ and $\theta^{\mathbb{Q}} = (\kappa\theta)/\kappa^{\mathbb{Q}}$. We assume that the jump risk premia are entirely attributable to the mean jump size in return and mean jump size in variance: $\eta_{J^s} = \mu_s - \mu_s^{\mathbb{Q}}$ and $\eta_{J^v} = \mu_v - \mu_v^{\mathbb{Q}}$. Therefore, $J_t^{v\mathbb{Q}} \sim Exp(\mu_v^{\mathbb{Q}})$ and $J_t^{s\mathbb{Q}} | J_t^{v\mathbb{Q}} \sim N(\mu_s^{\mathbb{Q}} + \rho_J J_t^{v\mathbb{Q}}, \sigma_s^2)$. The jump intensity λ and the return jump standard deviation σ_s do not change across measures. Therefore, the total equity risk premium can be decomposed as $\gamma_t = \eta_s V_t + \lambda(\bar{\mu}_s - \bar{\mu}_s^{\mathbb{Q}})$.

The above specification nests many of the widely-used models in existing literature. If we shut down jumps in return, it reduces to the stochastic volatility with jumps in variance (SVJV) model of Barndorff-Nielsen and Shephard (2002), while it becomes the stochastic volatility with jumps in return (SVJR) model of Bates (1996) if we shut down jumps in variance. By shutting down all jumps, we arrive at the stochastic volatility (SV) model of Heston (1993), and by further restricting V_t to be constant over time, we have the Black-Scholes (BS) model developed by Black and Scholes (1973).

Given the dynamics above, the option price C_t is a function of S_t and V_t.[9] Applying Ito's lemma, we write the instantaneous option price dynamic under physical measure as:[10]

[9]We use C to denote an option which can be either a call or a put. Whenever we need to distinguish between the types of option, we use Call and Put.

[10]The dynamic under risk-neutral measure can be derived similarly, but we do not need it in this paper.

$$
\begin{aligned}
\frac{dC_t}{C_t} &= \left[r_t + \frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}\eta_s V_t - \frac{\partial C_t}{\partial V_t}\frac{1}{C_t}\eta_v V_t - \lambda\frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}\bar{\mu}_s^{\mathbb{Q}} - \lambda\frac{\partial C_t}{\partial V_t}\frac{1}{C_t}\bar{\mu}_v^{\mathbb{Q}} \right] dt \\
&\quad + \frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}\sqrt{V_t}dZ_t + \frac{\partial C_t}{\partial V_t}\frac{1}{C_t}\sigma\sqrt{V_t}dW_t + \frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}(e^{J_t^s} - 1)dN_t + \frac{\partial C_t}{\partial V_t}\frac{1}{C_t}J_t^v dN_t
\end{aligned}
\tag{3.6}
$$

where C_t represents the price of an option, either a call or a put, at time t. Taking expectations on both sides of the equation, we obtain an expression of the instantaneous expected option return:

$$
\frac{1}{dt}\mathbb{E}_t\left[\frac{dC_t}{C_t}\right] = r_t + \frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}\eta_s V_t - \frac{\partial C_t}{\partial V_t}\frac{1}{C_t}\eta_v V_t + \lambda\frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}(\bar{\mu}_s - \bar{\mu}_s^{\mathbb{Q}}) + \lambda\frac{\partial C_t}{\partial V_t}\frac{1}{C_t}(\bar{\mu}_v - \bar{\mu}_v^{\mathbb{Q}})
\tag{3.7}
$$

The expected option return is equal to the risk-free rate plus four types of risk premiums: the underlying diffusive risk premium, the variance diffusive risk premium, the underlying jump risk premium and the variance jump risk premium. Note that all the physical parameters that enter the underlying return dynamic also show up in the option return dynamic. We thus need only one return dynamic, either from underlying return or from option return, together with the variance dynamic, to identify all the model parameters. If the information across underlying and options markets is entirely consistent, estimation for both markets should yield identical results. However, the existing literature has found that index market and options market may not be entirely consistent. Therefore, by assuming consistency across two markets, any error will be reflected in the parameter estimates and misleading conclusion might be drawn when we use one market to explain the other.

These model dynamics are affine, and therefore, quasi-closed-form solutions are available not only for option price but also for its derivatives (Option Greeks). According to the Fourier-Cosine Series Expansion method introduced by Fang and Oosterlee (2008), the price for an option is given by:

$$C_t = e^{-r\tau} \sum_{j=0}^{' N-1} Re\{\phi(\frac{j\pi}{b-a}; V_t) e^{-ij\pi\frac{a}{b-a}}\} U_j \tag{3.8}$$

$$U_j = \frac{2}{b-a} \int_a^b v(y, T) cos(j\pi \frac{y-a}{b-a}) dy \tag{3.9}$$

where $\phi(\omega; V_t)$ is the characteristic function of underlying log-return y conditional on V_t. U_j is the Fourier-Cosine coefficient of option payoff function $v(y, T)$. Fang and Oosterlee (2008) further show that U_j can be computed analytically for vanilla European options.

An important advantage of the Fourier-Cosine method is that several option Greeks can be obtained along with the evaluation of option price. In this paper, we focus on the option partial derivatives to underlying price (Delta) and spot variance (Vega). For the stochastic volatility model, Vega can be defined in different ways, here we mean the derivative of the option price with respect to the spot variance V_t.[11] The Delta is given by:

$$\frac{\partial C_t}{\partial S_t} = e^{-r\tau} \sum_{j=0}^{' N-1} Re\{\phi(\frac{j\pi}{b-a}; V_t) e^{ij\pi\frac{y-a}{b-a}} \frac{ij\pi}{b-a}\} \frac{U_j}{S_t} \tag{3.10}$$

while the Vega is given by:

$$\frac{\partial C_t}{\partial V_t} = e^{-r\tau} \sum_{j=0}^{' N-1} Re\{\frac{\partial \phi(\frac{j\pi}{b-a}; V_t)}{\partial V_t} e^{ij\pi\frac{y-a}{b-a}}\} U_j \tag{3.11}$$

3.2.1 Hold-to-Maturity Option Returns

Existing studies of the option return puzzle focus on monthly hold-to-maturity returns. One important motivation is that hold-to-maturity option return is available in closed-form.[12] In line with the existing literature, we therefore also use monthly hold-to-maturity option returns to analyze the index option return puzzle, even though our model estimation is based

[11] Alternative definitions are the derivative with respect to volatility $\sqrt{V_t}$ or the long run mean variance θ or even $\sqrt{\theta}$.

[12] Note that the instantaneous option return we discussed in the previous subsection is valid only locally. For option return over a holding period, it generally requires integration, which is much harder to derive.

on daily option returns.

3.2.1.1 Model-Implied Option Returns

As discussed in Broadie et al. (2009), for a hold-to-maturity option, the expected return has closed-form expression. We denote a call return from t to $t+\tau$ as $R^{call}_{t,\tau}$[13]. Its expected return is then given by:

$$\mathbb{E}_t(R^{call}_{t,\tau}) = \frac{\mathbb{E}_t[(S_{t+\tau} - K)^+]}{\mathbb{E}^{\mathbb{Q}}_t[e^{-r\tau}(S_{t+\tau} - K)^+]} - 1 = \frac{\mathbb{E}_t[(R^g_{t,\tau} - k)^+]}{\mathbb{E}^{\mathbb{Q}}_t[e^{-r\tau}(R^g_{t,\tau} - k)^+]} - 1 \tag{3.12}$$

where r and $R^g_{t,\tau}$ are the risk-free rate and the gross index return from t to $t+\tau$ and k is the moneyness defined as K/S. The expected call option return depends only on the moneyness, the maturity, the risk-free rate and the expected index gross returns under the physical and the risk-neutral measures. The risk-free rate is given and moneyness and maturity are option characteristics that do not change with parameter values. The option return thus depends only on the pricing kernel that differentiates the physical and the risk-neutral underlying return distribution. However, Equation (3.12) does not provide a closed-form expression itself, and we need to make use of closed-form option price and rewrite the equation as:

$$\mathbb{E}_t(R^{call}_{t,\tau}) = \frac{\mathbb{E}_t[(S_{t+\tau} - K)^+]}{\mathbb{E}^{\mathbb{Q}}_t[e^{-r\tau}(S_{t+\tau} - K)^+]} - 1 = \frac{e^{(r+\gamma)\tau} Call^{\mathbb{P}}(t,\tau)}{Call(t,\tau)} - 1 \tag{3.13}$$

where γ is the equity risk premium from t to $t+\tau$.[14] $Call(t,\tau)$ is the option price evaluated under risk-neutral measure with closed-form solution as discussed in the previous subsection. $Call^{\mathbb{P}}(t,\tau)$ is the present value of option payoff with the analytical formula similar to $Call(t,\tau)$ but replacing all the risk-neutral parameters with the physical parameters.

Similarly, the expected put option return can be written as:

[13] Throughout this paper, we use R to denote the index return and R^c to denote the index option return which can be either a call return or a put return. Whenever distinguishing call and put is necessary, we use R^{call} and R^{put}.

[14] Here r and γ are the risk-free rate and equity risk premium from t to $t+\tau$, which are different from the instantaneous risk free rate and equity risk premium r_t and γ_t

$$\mathbb{E}_t(R_{t,\tau}^{put}) = \frac{\mathbb{E}_t[(K - S_{t+\tau})^+]}{\mathbb{E}_t^{\mathbb{Q}}[e^{-r\tau}(K - S_{t+\tau})^+]} - 1 = \frac{e^{(r+\gamma)\tau} Put^{\mathbb{P}}(t,\tau)}{Put(t,\tau)} - 1 \tag{3.14}$$

Therefore, for any model dynamic, if the option price has a closed-form solution, the option return also has a closed-form expression. Option strategies that consist of a combination of call or/and put options can be expressed accordingly, such as a straddle:

$$\begin{aligned} \mathbb{E}_t(R_{t,\tau}^{straddle}) &= \frac{\mathbb{E}_t(|S_{t+\tau} - K|)}{\mathbb{E}_t^{\mathbb{Q}}[e^{-r\tau}(S_{t+\tau} - K)^+] + \mathbb{E}_t^{\mathbb{Q}}[e^{-r\tau}(K - S_{t+\tau})^+]} - 1 \\ &= \frac{e^{(r+\gamma)\tau}[Call^{\mathbb{P}}(t,\tau) + Put^{\mathbb{P}}(t,\tau)]}{Call(t,\tau) + Put(t,\tau)} - 1 \end{aligned} \tag{3.15}$$

We examine several other option strategies, including call spreads (long an ATM call and short a 6% OTM call), put spreads (long an ATM put and short a 6% OTM put), crash neutral spreads (long an ATM straddle and short a 6% OTM put), lottery neutral spreads (long an ATM straddle and short a 6% OTM call) and crash lottery neutral spreads (long an ATM straddle and short a 6% OTM call and put).[15] The expected returns for these option strategies can be derived as a function of the expressions for puts, calls and straddles.

In order to compute the unconditional expected option return, we also need the distribution of spot variances. If there is no jump in variance, the unconditional variance follows a gamma distribution with the probability density of:

$$f(V_\infty) = \frac{\omega^\upsilon}{\Gamma(\upsilon)} V_\infty exp(-\omega V_\infty) \tag{3.16}$$

where $\omega = 2\kappa/\sigma^2$ and $\upsilon = 2\kappa\theta/\sigma^2$. For the models with jumps in variance, we rely on simulation to approximate the spot variance distribution.

[15] All these option strategies are model free and do not require hedging based on any specific model.

3.2.1.2 Realized Option Returns

To compute option returns from the data, we use options with exactly one month maturity from 1996 to 2015 and the monthly hold-to-maturity return is simply defined as the realized payoff divided by its price minus 1.[16] For a call option:

$$R_{t,\tau}^{call} = \frac{(S_{t+\tau} - K)^+}{call(t,\tau)} - 1 \tag{3.17}$$

where $(S_{t+\tau} - K)^+$ is the realized payoff at maturity and $call(t,\tau)$ is the observed market call price at time t with time-to-maturity $\tau = 1$ month.

Similarly, for a put option and a straddle:

$$R_{t,\tau}^{put} = \frac{(K - S_{t+\tau})^+}{put(t,\tau)} - 1 \tag{3.18}$$

$$R_{t,\tau}^{straddle} = \frac{|K - S_{t+\tau}|}{call(t,\tau) + put(t,\tau)} - 1 \tag{3.19}$$

Other option strategy returns can be defined similarly.

3.2.1.3 The Finite Sample Distribution of Option Returns

We need the statistical distribution to compare the model-implied option return and realized option return. This statistical distribution can be obtained from the realized monthly option return data. However, Broadie et al. (2009) argue that the relatively short sample period, as well as the extreme skewness and non-normality of option returns can make the finite-sample distribution approximation extremely inaccurate. Therefore, they suggest simulating option return distributions using model parameters. We follow Broadie et al. (2009) and simulate the average option returns (for puts, calls and option strategies) over our sample period.

[16] Weekly option is also available, but since weekly option started to trade only after 2005, we have to give up some information from the data by either reducing the sample length or holding a monthly option for only a week.

Within each simulation, we simulate 240 months of index prices and spot variances, with ten steps per day (which amounts to 240*21*10 steps in total), and calculate the average returns for every moneyness over the 240 months. For a call with strike price K:

$$R_j^{call} = \frac{1}{240} \sum_{i=1}^{240} \left[\frac{(K - S_{t_i+\tau})^+}{Call(t_i, \tau)} - 1 \right] \tag{3.20}$$

where R_j^{Call} is the average call option return for the j-th simulation run, t_i represents the starting point for each of the $N = 240$ months. With 10,000 simulations, $\{R_j^{call}\}$ can be used to approximate the finite sample distribution of call option returns. Similarly, we may calculate the distributions for calls and puts, as well as option strategies.

3.3 Estimation

In this section, we first briefly explain the data we use. Then, we discretize the continuous time model and present the dynamics for both observables and unobservables. The spot variances are latent while index returns, option prices and option returns are observed at each period. We derive the likelihood functions for index returns, option prices and option returns. Subsequently, we combine these likelihoods and explain how to weigh them.

3.3.1 Option Data

We use $S\&P$ 500 index returns and option prices for the period from January 1, 1996 to December 31, 2015, a total of 5031 trading days or 240 months. The $S\&P$ 500 index returns and risk-free rates are obtained from CRSP. Index option prices, zero coupon yields, and dividend yields are downloaded from OptionMetrics. We use both put and call options and impose the following standard filters on the option data:

1. Discard options with fewer than five 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 five 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 a negative put-call parity implied price.

5. Discard options with price less than 50 cents.

After imposing these filters, the dataset contains 945,110 options in total. The option data set is not balanced over time as we have almost eight times more options in 2015 than 1996. We then create a more balanced panel by including only the most liquid option from each moneyness-maturity bin. We use six moneyness and five maturity bins as shown in Panel B of Table 3.1. The balanced data set thus has thirty options per day unless no option is available for a certain bin. This procedure yields a data set with 129,182 options. Panel B of Table 3.1 provides sample sizes, option prices and implied volatilities for these moneyness-maturity bins. Note that the data are still not perfectly balanced because in the early years of the sample we do not have many observations with short maturity and moneyness greater than 1.10. Due to the computational constraint in the filtering process, we use the balanced dataset for parameter estimation and the full dataset to construct the monthly realized option returns.

3.3.2 Model Discretization

By applying Ito's lemma and discretizing equation (3.1), (3.2) and (3.6), we have:

$$R_{t+1} = log(\frac{S_{t+1}}{S_t}) = r_t - \delta_t - \frac{1}{2}V_t + \gamma_t - \lambda\bar{\mu}_s + \sqrt{V_t}z_{t+1} + J^s_{t+1}B_{t+1} \tag{3.21}$$

$$V_{t+1} - V_t = \kappa(\theta - V_t) + \sigma\sqrt{V_t}w_{t+1} + J^v_{t+1}B_{t+1} \tag{3.22}$$

$$
\begin{aligned}
R_{t+1}^{c} = \frac{C_{t+1}^{*} - C_t}{C_t} &= r_t + \frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}\eta_s V_t - \frac{\partial C_t}{\partial V_t}\frac{1}{C_t}\eta_v V_t - \lambda\frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}\bar{\mu}_s^{\mathbb{Q}} - \lambda\frac{\partial C_t}{\partial V_t}\frac{1}{C_t}\bar{\mu}_v^{\mathbb{Q}} \\
&+ \frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}\sqrt{V_t}z_{t+1} + \frac{\partial C_t}{\partial V_t}\frac{1}{C_t}\sigma\sqrt{V_t}w_{t+1} + \frac{\partial C_t}{\partial S_t}\frac{S_t}{C_t}(e^{J_t^s} - 1)B_{t+1} \\
&+ \frac{\partial C_t}{\partial V_t}\frac{1}{C_t}J_t^v B_{t+1}
\end{aligned}
\tag{3.23}
$$

where $\gamma_t = \eta_s V_t + \lambda(\bar{\mu}_s - \bar{\mu}_s^{\mathbb{Q}})$ and z_{t+1} and w_{t+1} are distributed standard normal. The discrete jump frequency B_{t+1} follows the Bernoulli distribution and thus, at each time point, there is either no jump or one jump. C_{t+1}^{*} is the time-to-maturity adjusted option price at time $t+1$.[17] We estimate this discretized model using daily data, but we report annualized parameter estimates through our analysis.

We do not directly observe the spot variances, but the observables are generated by the spot variances. The whole dynamic system can be described as Equation (3.24). The spot variance is unobservable and evolves according to Equation (3.22). At each time, the spot variance generates all the observables: one index return according to Equation (3.21), one set of option prices according to the risk-neutral dynamics Equation (3.4)-(3.5), and one set of option returns according to Equation (3.23).

$$
\begin{array}{ccccccc}
(V_t) & \rightarrow & (V_{t+1}) & \rightarrow & (V_{t+2}) & \rightarrow & .... \\
\downarrow & & \downarrow & & \downarrow & & \\
R_{t+1}, \{C_t\}, \{R_{t+1}^c\} & & R_{t+2}, \{C_{t+1}\}, \{R_{t+2}^c\} & & R_{t+3}, \{C_{t+2}\}, \{R_{t+3}^c\} & & ....
\end{array}
\tag{3.24}
$$

The spot variance can be filtered from any one of the three observables. The main focus of this paper is to utilize option prices and option returns to identify model dynamics. We compare this estimation approach with two benchmark estimation approaches that have been well developed in the literature: one is the estimation using index returns only; the

[17] We need to be very careful about the option price change from t to $t+1$. Both calls and puts have negative option Theta, meaning that the option prices decrease with time, even though everything else remains the same. We have to adjust back the part of option price change that is related solely to changes in maturity. Rather than directly using the observed option price C_{t+1} at $t+1$, we therefore calculate C_{t+1}^{*} by assuming the time-to-maturity is the same as when we observe it at time t.

other is the joint estimation using both index returns and option prices.

The estimation of parameters and latent factors can be implemented in several ways, such as EMM, MCMC, Particle Filtering and so on. Our estimation method is based on Jacobs and Liu (2017), and we combine the Adaptive Metropolis-Hastings MCMC (AMH-MCMC) (see Metropolis et al., 1953; Hastings, 1970; Johannes and Polson, 2009) and Particle Filter (see Christoffersen et al., 2010; Johannes et al., 2009, among others). The benefit of this method is that we can incorporate relatively large option panels, which are important to identify model parameters as well as to reduce measurement error in option returns. Given a set of parameters, we first filter the spot variance using the Particle Filtering and then search for the optimal parameters based on AMH-MCMC. The whole idea is to maximize the likelihood of generating the observed data, no matter what the observables are. We provide a detailed description of the implementation in Appendix. We now characterize the likelihood function implied from different observables.

3.3.3 The Likelihood Based on Index Returns

The likelihood of observing the index return conditional on the spot variance V_t^i is given by Equation (3.21):

$$f(R_{t+1}|V_t^i) = \frac{1}{\sqrt{2\pi V_t^i}} \exp\left\{ -\frac{1}{2} \frac{\left[R_{t+1} - (r_t - \delta_t - \frac{1}{2}V_t^i + \gamma_t - \lambda\bar{\mu}_s + J_{t+1}^{s,i} B_{t+1}^i)\right]^2}{V_t^i} \right\} \quad (3.25)$$

where V_t^i denotes the i-th particle for the spot variance at time t. According to Pitt (2002), the total likelihood of Particle Filtering given a set of parameters can be expressed as:

$$f(R_{1:T}|\Theta) = \prod_{t=1}^{T} \left\{ 1/N \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi V_t^i}} \exp\left\{ -\frac{1}{2} \frac{\left[R_{t+1} - (r_t - \delta_t - \frac{1}{2}V_t^i + \gamma_t - \lambda\bar{\mu}_s + J_{t+1}^{s,i} B_{t+1}^i)\right]^2}{V_t^i} \right\} \right\} \quad (3.26)$$

Our implementation of Particle Filter uses 10,000 particles and 50,000 iterations in AMH-MCMC. 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.

3.3.4 The Likelihood Based on Option Prices

If we observe the index option prices rather than index returns, the spot variances can still be inferred. The option price has two essential characteristics: 1) at each time point, a whole cross-section of option prices are observed; 2) unlike index prices, options prices are observed with errors.

We use lower case $c_{t,h}$ to represent market option price which is equal to the model price plus some error:

$$c_{t,h} = C_{t,h}(V_t|\Theta) + \epsilon_{t,h} \tag{3.27}$$

where $c_{t,h}$ is the h-th option at time t, which can be either a call or a put. $h = 1, 2, .., H_t$ and H_t is the total number of options at time t. We assume $\epsilon_{t,h}$ is normally distributed and $\epsilon_{t,h} \sim N(0, (\sigma_c)^2)$. Thus the likelihood of option prices conditional on V_t^i is:

$$\begin{aligned} f(\{c_t\}|V_t^i) &= \left(\prod_{h=1}^{H_t} f(c_{t,h}|C_{t,h}(V_t^i|\Theta))\right)^{1/H_t} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma_c}\right) exp\left(-\frac{\sum_{h=1}^{H_t}(c_{t,h} - C_{t,h}(V_t^i|\Theta))^2}{2(\sigma_c)^2 H_t}\right) \end{aligned} \tag{3.28}$$

We take the H_t-th root in order to normalize the likelihood with respect to the different number of options on each day. The total likelihood for the entire sample conditional on a given set of parameters is:

$$f(\{c_{1:T}\}|\Theta) = \prod_{t=1}^{T}\left\{1/N \sum_{i=1}^{N}\left[\left(\frac{1}{\sqrt{2\pi}\sigma_c}\right) exp\left(-\frac{\sum_{h=1}^{H_t}(c_{t,h} - C_{t,h}(V_t^i|\Theta))^2}{2(\sigma_c)^2 H_t}\right)\right]\right\} \tag{3.29}$$

Quasi-closed-form solutions for the option price are available in the affine models we consider, and each option price takes less than 0.01 seconds to evaluate However, for each function evaluation we have to evaluate option prices along three dimensions: for each option in a day, for each particle, and for each day, which is computationally infeasible. Therefore, rather than directly computing the pricing error for each option conditional on each particle, we follow Jacobs and Liu (2017) to implement the Implied Spot Variance method. The total likelihood can thus be re-written as a function of the implied spot variance:

$$f(\{c_{1:T}\}|\Theta) = \prod_{t=1}^{T} \left\{ 1/N \sum_{i=1}^{N} \left[\left(\frac{1}{\sqrt{2\pi}\sigma_c} \right) exp\left(-\frac{SSE_{ISV}(\hat{b}_t * (V_t^i - ISV_t)^2 + 1)}{2(\sigma_c)^2 H_t} \right) \right] \right\} \tag{3.30}$$

Rather than computing option prices for $N = 10,000$ particles, we now only need to compute a few grids in the ISV search step and then for each particle we calculate its likelihood based on its distance to ISV.

3.3.5 The Likelihood Based on Option Returns

The likelihood for option return is given by Equation (3.23). Since option prices are observed with errors for both the numerator and denominator, the ratio (option return) should theoretically follow Cauchy distribution[18], which is relatively hard to model. Therefore, rather than modeling the option return, we model the option price change instead. Multiplying both sides of Equation (3.23) by C_t:

$$\begin{aligned} C_{t+1}^* - C_t = r_t C_t + \frac{\partial C_t}{\partial S_t} S_t \eta_s V_t - \frac{\partial C_t}{\partial V_t} \eta_v V_t - \lambda \frac{\partial C_t}{\partial S_t} S_t \bar{\mu}_s^{\mathbb{Q}} - \lambda \frac{\partial C_t}{\partial V_t} \bar{\mu}_v^{\mathbb{Q}} + \frac{\partial C_t}{\partial S_t} S_t \sqrt{V_t} z_{t+1} \\ + \frac{\partial C_t}{\partial V_t} \sigma \sqrt{V_t} w_{t+1} + \frac{\partial C_t}{\partial S_t} S_t (e^{J_t^s} - 1) B_{t+1} + \frac{\partial C_t}{\partial V_t} J_t^v B_{t+1} \end{aligned} \tag{3.31}$$

[18]Under certain conditions, this Cauchy distribution can be approximated by a normal distribution, see Marsaglia (1965, 2006).

The diffusion part is $\frac{\partial C_t}{\partial S_t} S_t \sqrt{V_t} z_{t+1} + \frac{\partial C_t}{\partial V_t} \sigma \sqrt{V_t} w_{t+1}$. The correlation between z_{t+1} and w_{t+1} is equal to ρ and the option pricing error is σ_c. The total diffusive variance is:

$$\sigma_{cr} = \sqrt{(\frac{\partial C_t}{\partial S_t} S_t)^2 V_t + (\frac{\partial C_t}{\partial V_t} \sigma)^2 V_t + \rho \frac{\partial C_t}{\partial S_t} S_t \frac{\partial C_t}{\partial V_t} \sigma V_t + 2(\sigma_c)^2} \tag{3.32}$$

The expected change in option price is:

$$\mu_{cr} = r_t C_t + \frac{\partial C_t}{\partial S_t} S_t \eta_s V_t - \frac{\partial C_t}{\partial V_t} \eta_v V_t - \lambda \frac{\partial C_t}{\partial S_t} S_t \bar{\mu}_s^{\mathbb{Q}} - \lambda \frac{\partial C_t}{\partial V_t} \bar{\mu}_v^{\mathbb{Q}} + \frac{\partial C_t}{\partial S_t} S_t (e^{J_t^s} - 1) B_{t+1} + \frac{\partial C_t}{\partial V_t} J_t^v B_{t+1} \tag{3.33}$$

Therefore, the likelihood of the h-th option return conditional on V_t^i is given by:

$$f(c^*_{t+1,h} - c_{t,h} | V_t^i) = \frac{1}{\sqrt{2\pi} \sigma^i_{cr,h}} \exp\left\{ -\frac{1}{2} \frac{(c^*_{t+1,h} - c_{t,h} - \mu^i_{cr,h})^2}{(\sigma^i_{cr,h})^2} \right\} \tag{3.34}$$

The total likelihood for the whole option panels is:

$$f(\{c^*_{2:T+1} - c_{1:T}\} | \Theta) = \prod_{t=1}^{T} \left\{ 1/N \sum_{i=1}^{N} \left[\left(\frac{1}{\sqrt{2\pi} \sigma^i_{cr,h}} \right) exp \left(- \frac{\sum_{h=1}^{H_t} (c^*_{t+1,h} - c_{t,h} - \mu^i_{cr,h})^2}{2(\sigma^i_{cr,h})^2 H_t} \right) \right] \right\} \tag{3.35}$$

Remember that all these partial derivatives (option Greeks) are in fact byproducts of the option price evaluation. We may thus calculate this option return (option price change) likelihood along with option price likelihood using the ISV method without significant additional cost.

3.3.6 Combining the Likelihoods

We have discussed the likelihood function for each of the observables: index returns, option prices, and option returns. When combining different observables, the weights of each likelihood becomes a question. In this paper, we choose to 'equally' weigh the information from the three types of observables, no matter how many data samples we have at each period. Note that, every day, we have up to 30 option prices and option returns,[19] but only one index return. We take the H_t-th root for option prices and option returns so that the option information will not dominate index return information simply because more data points are available for options. Note that the H_t-th root in fact equally weighs not only on the different types of observations but also on the information from different time periods, and it guarantees a date with more available options will not significantly outweigh a date with fewer options.

3.4 Parameter Estimates

Table 3.2 presents the parameter estimates from option prices and option returns. Parameters for the BS model are calculated by matching the options data: θ is simply the square of average implied volatility for all options and η_s is the average index return implied from all option returns divided by θ.[20] For the other models, We estimate the parameters by maximizing the likelihood from option prices and option returns using the Particle Filter together with the Adaptive Metropolis-Hastings MCMC method. Parameters are reported under the physical measure. The risk-neutral parameters can be inferred by subtracting risk premia from the corresponding physical parameters. We report the posterior mean and the posterior standard deviation after burn-in in parenthesis.

We compare our parameter estimates with those from index returns only (Table 3.3) and those from index returns and option prices jointly (Table 3.4).[21] To provide additional insight, Figure 3.1 plots the parameter traces for the SV model using index returns as an

[19] Option return data points are a little bit fewer than option price data points because a valid option return needs an option to have valid prices for two consecutive days.

[20] We calculate η_s using option returns, rather than index returns.

[21] Admittedly, the joint estimation using index returns, option prices and option returns all together is also an interesting application. However in our paper, we would like to provide some ideas about the potential differences between index and option markets without imposing consistency restrictions.

example. The parameters converge to their average values with less than 5000 iterations and then fluctuate around the averages. We pick 'crazy' initial values for each parameter that are far away from conventional estimates in order to confirm the robustness of the method.

We find that the long-term mean variances estimated from options and those estimated from index returns are not very different in values. Note that κ and σ from index options are smaller than those from index returns, implying a more consistent and smooth spot variance dynamic, which can be confirmed by the plot of the spot variance path in Figure 3.2. ρ is more negative from index options than from index returns. One possible explanation could be that, due to the consistency restriction across the physical and risk-neutral measures, ρ is mainly driven by risk-neutral skewness rather than physical skewness.[22] The mean jump size in returns μ_s is more negative and the mean variance jump size μ_v is smaller when estimated from options.

We also find that the posterior standard deviations for most of the parameters are much smaller when estimated from options than from index returns. We argue that since options are available across moneyness and maturity, this information helps to identify the level of spot variances[23] and thus reduce the standard deviations for all the parameters that describe or depend on the spot variances.[24] The improved identification of parameters is evident from the likelihood curve for each parameter plotted in Figure 3.3. We shift the level of the likelihood to have the same value at maximum but we keep the scale, and thus the 'gradient' remains unchanged. The standard deviation of each parameter is closely related to the 'gradient' at the optimal point and the steeper the curve, the smaller the standard deviation. As can be seen from the figure, most parameters have much steeper likelihoods from index options than from index returns.

[22] It is commonly found in existing literature that ρ estimated from index options is more negative than that from index returns, (Eraker, 2004; Aït-Sahalia and Kimmel, 2007; Andersen et al., 2015a; Hurn et al., 2015)

[23] For example, if we only observe a single maturity of options, when the prices are high, we do not know whether it is due to a high spot variance or a high long run mean variance. But if we have multiple-maturity options, both the spot variance and long run mean variance can be well identified.

[24] A detail discussion of why options covering large moneyness and maturity range improve parameter identification can be found in Jacobs and Liu (2017).

We also find that η_s from options is consistently smaller than that from index returns for all models, which may indicate different required returns from the participants in the underlying and options markets. The next question is: is the difference in η_s statistically significant and does it imply inconsistencies between these two markets? The top panel in Figure 3.4 plots the posterior distribution between η_s from index returns and η_s from index options. The probability of index returns based η_s being larger than index options based η_s is 0.6625, meaning that we cannot reject the hypothesis that η_s from the two different sources of information are the same. However, if we plot the posterior distribution of η_s from index options and compare it against the mean value of $\eta_s = 2.5374$ from index returns (panel (b) in Figure 3.4), we clearly reject the hypothesis about consistency. In conclusion, there is no inconsistency for η_s, given the estimation risk or parameter uncertainty from both markets; However, if we use the posterior mean of η_s from index returns to explain the options market, while entirely ignoring the estimation risk, the inconsistency does exist.

Note that the posterior standard deviation for η_s from options is much smaller than that from index returns, but not smaller than that from index returns and option prices jointly. We conclude that this improved identification of η_s is not due to the inclusion of multiple option returns, but due to the improved identification of the spot variances from option prices. Even though more than one option return are available at each period, they are all driven by the same realization of the index process and thus are perfectly correlated.[25] At last, we find that the mean jump size in returns μ_s is more negative and the mean jump size in variances μ_v is much smaller from options than those from index returns.

[25]The effective sample size is still one even if more than one option is used.

3.5 The Index Option Return Puzzle

3.5.1 Realized Option Returns

The index option return puzzle refers to the extremely negative returns on OTM puts, OTM calls, and option strategies, such as straddles. However, there is little consensus about whether, from statistic point of view, index option return is puzzling or not. Jackwerth (2000) documents that monthly put writing strategies deliver abnormal high returns, in both absolute and risk-adjusted levels, and he concludes that options are likely to be mispriced. Coval and Shumway (2001) find that put returns are too negative to be consistent with a single factor model based on weekly options and straddles. Bondarenko (2014) analyzes monthly index futures options based on equilibrium models and finds significantly negative put returns that are inconsistent with single-factor equilibrium model. Santa-Clara and Saretto (2004) compute returns for a number of index option portfolios and argue that the returns are implausibly large, although these returns might be difficult to achieve due to transaction costs. Jones (2006) analyzes daily option returns based on a nonlinear multi-factor model and finds that deep OTM put options have significant alphas. More recently, Broadie et al. (2009) argue that the traditional statistical test based on finite sample distribution might be inaccurate due to relatively short sample period and extreme skewness of option returns. They propose Monte Carlo simulation to determine the finite sample distribution of the average option returns. By reviewing some of the existing models with no risk premia, they find that it is hard to reject these models using monthly hold-to-maturity put returns, and with proper correction of estimation risk, there is no puzzle for option portfolios including ATM straddle. Branger et al. (2010) show that models with stochastic volatility and jumps can generate almost any structure of option returns by changing various risk premia, especially the jump-related ones. In contrast, Chambers et al. (2014) re-examine the puzzle based on Broadie et al. (2009) method using data from 1987-2012 and conclude that option returns are still puzzling.

The traditional approach used to test the index option return puzzle in the existing literature, no matter what the conclusion is, typically consists of 3 steps: 1) estimate the physical parameters from index returns; 2) infer the risk-neutral parameters without using option prices; 3) calculate model-implied option returns and compare with market realized counterparts.

The second step is implemented differently in different studies. Some studies simply assume all risk premia to be zero, and thus the risk-neutral parameters are exactly the same as the physical parameters. Some studies derive risk premium parameters from the general equilibrium models developed by Bates (1988) and Naik and Lee (1990).[26] Other studies back-engineer the risk-neutral parameters by calibrating risk premia to fit market option returns. These tests based on estimates from index returns are therefore joint hypothesis tests, and when rejected, we do not know whether options are mispriced, or the parameters estimated from index returns are inconsistent with index options. To assess the impact of these potential inconsistencies or estimation risk, one can, for example, increase or decrease each parameter by a certain amount (say, one standard deviation). However, since we do not know how exactly investors calculate and price estimation risk, the adjustment is still ad hoc.

Based on our parameter estimates from options, we re-evaluate the index option return puzzle differently by decomposing the test into two parts. One is an option mispricing test: do the market realized option returns reject a model with parameters extracted exclusively from options? The other is an index inconsistency test: do the index returns reject a model with parameters extracted from options? If a model, estimated by fitting the options, is rejected by the market realized option returns, the model is likely to be misspecified; if a model is not rejected by option returns, but rejected by index returns, it is likely that the information from index returns is inconsistent with that from index options.[27] Rejecting

[26] Risk aversion level must be assumed, and the choice of this value is usually arbitrary.

[27] The estimation risk from options is much smaller than the estimation risk from index returns.

in either hypothesis can lead to the option return puzzle. The benefit of going from index options to index returns, rather than the other way around as in the existing literature, is that parameters estimated from options generally have much smaller estimation risk.

We first document the index option return puzzle using parameters estimated from index returns. Due to the noise in option trading, we follow the literature and focus on monthly hold-to-maturity option returns. The final payoff of an option contract is simply $(S_T - K)^+$ for a call and $(K - S_T)^+$ for a put. For each expiration date, we count back 30 days to record the corresponding option trading price. If the target date is a holiday, we use the nearest trading day instead. The option return is then computed as the final payoff divided by its original cost. Figures 3.5 shows some examples for option returns at each month. Option returns can be extremely skewed for the deep OTM options, where most of the returns are -100%, and only in few months huge returns are realized.

Table 3.5 presents statistics for market realized option returns. Put option returns display an approximately monotonic decreasing trend as we move from ITM to OTM options. Call option returns do not exhibit a clear trend but become increasingly negative for OTM options. Both deep OTM calls and puts yield very negative returns. The skewness and kurtosis also increase sharply for deep OTM options. The majority of realized returns for OTM options are -100% while at the same time, due to the low price of these options, returns can be huge once a positive payoff is realized. Option strategy returns display less skewness and kurtosis compared to call and put returns.

3.5.2 The Option Return Puzzle Implied by Parameters from Index Returns

We first calculate model-implied option returns using parameters estimated from index returns. In order to calculate option prices, we need the risk-neutral parameters. Since finding the right risk premia to solve the option return puzzle is not our objective, we follow Broadie et al. (2009) and assume no risk premia. Thus, the risk-neutral parameters are the same as the physical parameters.

Table 3.6 reports the model-implied returns as well as the market realized option returns. P-values are reported in parenthesis. The p-values are highly correlated as they come from the same simulation. To be conservative, we reject a model as long as at least one p-value is smaller than 1% (rather than 5% used in some existing studies) for any moneyness or option strategy.[28] All models are rejected by the data for some moneyness or option strategy. For call options, the BS, SV and SVJR model can only generate monotonically increasing option returns, which are inconsistent with the data especially for the OTM calls. Stochastic volatility models with jumps in variance are not rejected as they can generate decreasing returns for deep OTM call options. For put options, consistent with the existing literature, almost all models have low p-values for the moneyness ranging between 0.94 and 1.[29] With huge standard error, the very deep OTM put option return does not reject any model even though the difference in the average option returns is large. Note that a model in line with call returns may still be rejected by put returns (and vice versa).[30] Option strategies are often considered more informative as they either reduce the exposure to the index returns or dampen the effect of rare events. All models are rejected by ATM straddles and some are also rejected by other option strategies. In conclusion, all models are rejected by at least one option or option strategy return, which is the index option return puzzle documented in the existing literature. Of course, the existence of the puzzle may simply due to the zero risk premium assumption or parameter estimation risk. With an alternative assumption that variance risk premium $\eta_v = 4$, the option return puzzle does disappear even for the SV model. Besides, as argued in Broadie et al. (2009), by increasing or decreasing parameters and spot variances in specific ways, the puzzle also disappears. However, including the risk premia

[28] This choice of significance level is related to multiple comparisons problem: when individual comparisons are not perfectly dependent, as the number of comparisons increases, it becomes more likely that the groups being compared will appear to differ in terms of at least one attribute due to random sampling error alone. Therefore, unless perfectly dependent, multiple comparisons generally require a stricter significance threshold for individual comparisons, so as to compensate for the number of inferences being made.

[29] These moneyness could have been rejected at significance level of 5% (See Broadie et al., 2009; Chambers et al., 2014, et. al).

[30] Put-Call Parity holds only for option prices, not for option returns.

and adjusting parameters require much effort to justify a specific direction or amount and may potentially introduce additional contradictions. For example, a variance risk premium of 4 would imply a risk-neutral long-run mean variance three times as high as the physical variance, which makes it impossible to match the term structure of option implied volatility.

As another benchmark, we also report option returns using parameters estimated from index returns and option prices jointly. All models are rejected based on at least ATM straddle.

3.5.3 The Option Return Puzzle Implied by Parameters from Options

Based on the parameters estimated from option prices and option returns, we break down the index option return puzzle into an option mispricing test and an index inconsistency test. One may argue that since we estimate parameters by matching option prices as well as option returns, by definition these parameters should be able to generate market option returns. However, this is not necessarily the case. If a model is misspecified, we can estimate parameters by fitting the data as well as possible, but the model may still have low power explaining the data. In other words, if realized option returns reject a model with parameters estimated from the options prices and option returns, then unless we accept a worse fit for option prices, these returns would reject the model with any other set of parameters. For such model, no matter how we adjust the parameters, we will not be able to fit option prices and option returns at the same time.

3.5.3.1 Option Mispricing

Table 3.7 reports the option returns and p-values using parameters estimated from options. The rejection of the BS and SV models implies that the options are indeed mispriced for these two models, implying that these models might be misspecified. These results contribute to the existing literature by reducing the possibility that the index option return puzzle can be resolved based on the risk premia and estimation risk. Risk premia in our test have been set to match option prices and option returns. Moreover, parameters estimated from index options are subject to much less estimation error than those from index returns. Therefore, the rejections of these two models are even stronger than in the existing literature. For the other three models, we do not reject the hypothesis that the model-implied option returns are consistent with realized option returns. Introducing jumps in either the underlying returns or/and the variances leads to large improvement in the modeling of option returns.

In the above analysis, we focus on whether an estimated model is rejected by the data using p-values. However, in order to compare different models or different estimation methods, we use the Option Return Root Mean Squared Error (OR-RMSE). We define the OR-RMSE as the normalized average distance between the model and data:[31]

$$OR - RMSE = \sqrt{\frac{1}{18}\sum_{\{k\}}\left[\left(\frac{\mathbb{E}_t^{\mathbb{P}}(R_{t,\tau}^{call}(k)) - R_{t,\tau}^{call}(k)}{\mathbb{E}_t^{\mathbb{P}}(std(R_{t,\tau}^{call}(k)))}\right)^2 + \left(\frac{\mathbb{E}_t^{\mathbb{P}}(R_{t,\tau}^{put}(k)) - R_{t,\tau}^{put}(k)}{\mathbb{E}_t^{\mathbb{P}}(std(R_{t,\tau}^{put}(k)))}\right)^2\right]} \tag{3.36}$$

where $\{k\}$ represents the moneyness set we would like to match and $\{k\}$ is from 0.92 to 1.08 with step size of 0.02. The rationale behind this measure is simple: according to the Central Limit Theorem, the average option returns at each moneyness are normally distributed and the standard deviations can be used to normalize the distance between the model and data.

Table 3.8 reports the OR-RMSE for each model using different estimation methods. For example, using parameters estimated from index returns, the BS model-implied option returns are, on average, 1.24 standard deviation away from realized option returns, while the SVCJ model implied returns, using parameters estimated from options, are only 0.62 standard deviation away from realized returns. OR-RMSE cannot be directly considered as a measure to reject a model or not, but it indicates how close on average the model predicted

[31]The idea is the same as Z-score. For the moneyness with large standard deviation, we can accept a larger difference between the model return and the realized return.

return is to market realized return given the standard deviation at each moneyness. Overall more complex models tend to fit option returns better and the estimation using option prices and option returns provides the best fit among the three methods.

3.5.3.2 Index Inconsistency

For the SVJR, SVJV, and SVCJ models that can not be rejected under the option mispricing test, we use their corresponding physical parameters to test whether index returns are puzzling. If these physical parameters are rejected based on index returns, it means the index is mispriced with respect to options, and this inconsistency may lead to the index option return puzzle. On the other hand, if the parameters are not rejected based on index returns, it means that the puzzle for these models in the existing literature, if there is any, was mainly due to parameter estimation risk. Once the uncertainty in parameter estimation is incorporated correctly, the puzzle does not exist anymore. To check whether the physical parameters estimated from options are consistent with index returns, we compare the first and second moments of index returns. We first simulate 10,000 index price and variance paths, and then calculate the p-values of realized average index return and variance.[32] Figure 3.6 plots the histograms of simulated average returns and variances for the SVJR, SVJV, and SVCJ models.

The index returns do not reject any of the three models, even though the parameters estimated from index returns and those estimated from index options seem to be different in values. In other words, for all the models we test, the parameters estimated from options are not inconsistent with index returns. This is not surprising, given the identification difficulties based on index returns. The advantage of such analysis is that we provide a direction of how to take into account estimation risk and risk premia, without causing additional contradictions. If we start from parameters from index returns, once we adjust all the

[32] Alternatively, we can directly compare the parameters estimated from index returns and those from index options to see whether they are significantly different. However, the comparison of parameters requires more assumptions on the standard deviation and also joint hypothesis test on a group of parameters.

parameters to exactly match the values estimated from options, the option return puzzle would disappear and more importantly, no additional contradiction to neither index returns nor option prices is generated.

3.6 Conclusion

We propose a new method to extract physical parameters and the latent spot variances from option prices and option returns. We show that using large panels of options helps to identify the physical parameters. Because the option data cover a wide moneyness and maturity range, the likelihood curve becomes much more informative for parameters, and as a result, the posterior standard deviations for most of the parameters are much smaller when based on index options. We cannot reject the consistency hypothesis between the index market and options market, but we show that the traditional way of using the parameter point estimates from index returns, while entirely ignoring parameter uncertainty, to explain option market may lead to biased results. We examine the monthly option return fit using different estimates and find that the parameters estimated exclusively from options provide the best fit.

With the information extracted from options, we re-examine the index option return puzzle by decomposing it into an option mispricing test and an index inconsistency test. Provided that jumps, in either returns or variances, or both, are introduced, options are not mispriced. We conclude that the option mispricing for the Black-Scholes and the Heston Stochastic Volatility models is likely due to model misspecification. Our parameters estimated from options provide a direction for the traditional approach to take into account risk premia and estimation risk.

Table 3.1: Return and Option Data

Panel A: Return Data

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

Panel B: Option Data

	Maturity (days)					
Moneyness (K/S)	5-30	30-60	60-90	90-180	180-365	All
Number of Option Contracts						
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
Average Call Prices						
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
Average Implied Volatilities						
0.85-0.90	0.34	0.28	0.26	0.25	0.24	0.27
0.90-0.95	0.27	0.24	0.23	0.23	0.22	0.24
0.95-1.00	0.21	0.20	0.20	0.21	0.21	0.21
1.00-1.05	0.18	0.18	0.18	0.19	0.19	0.18
1.05-1.10	0.20	0.16	0.16	0.17	0.18	0.17
1.10-1.15	0.31	0.20	0.18	0.17	0.17	0.18
All	0.24	0.21	0.21	0.20	0.20	0.21

Notes: Panel A reports descriptive statistics for the sample of index returns. The mean and standard deviation are annualized. Panel B reports the number of contracts, average call price and average implied volatility in the option data set where we choose the most liquid (highest trading volume) option within each moneyness-maturity range. Moneyness is defined as K/S. Because OTM options are generally more heavily traded, this data set mainly consists of OTM calls and OTM puts.

Table 3.2: Parameter Estimates from Option Prices and Option Returns

	BS	SV	SVJR	SVJV	SVCJ
κ	0.0406	1.6830	1.3561	1.2853	0.9183
	(-)	(0.0219)	(0.0478)	(0.0581)	(0.0312)
θ		0.0363	0.0348	0.0274	0.0264
		(0.0007)	(0.0011)	(0.0012)	(0.0015)
σ		0.3705	0.3807	0.3686	0.3324
		(0.0062)	(0.0081)	(0.0120)	(0.0131)
ρ		-0.9256	-0.9305	-0.9015	-0.9118
		(0.0108)	(0.0156)	(0.0171)	(0.0244)
η_s	1.6432	1.6030	1.5177	1.8899	2.0291
	(-)	(0.2486)	(0.2205)	(0.2774)	(0.2520)
η_v		0.8918	0.6751	0.0486	-0.0643
		(0.0261)	(0.0431)	(0.0334)	(0.0455)
λ			0.7673	0.8276	0.6719
			(0.0324)	(0.0534)	(0.0458)
μ_s			-0.0136		-0.0366
			(0.0030)		(0.0038)
σ_s			0.0581		0.0396
			(0.0024)		(0.0015)
η_{J^s}			0.0330		0.0112
			(0.0023)		(0.0034)
μ_v				0.0246	0.0022
				(0.0051)	(0.0021)
η_{J^v}				-0.0412	-0.0576
				(0.0043)	(0.0055)
ρ_J					-0.5567
					(0.0184)

Notes: We report parameters estimated using option prices and option returns from 1996 to 2015 for the BS, SV, SVJR, SVJV and SVCJ models. Parameters are all annualized and reported in the physical measure. The risk-neutral parameters can be inferred by subtracting risk premia from the corresponding physical parameters. Parameters for the BS model are calculated by matching the option sample: θ is simply the square of average implied volatility for all options and η_s is the average index return implied from all option returns divided by θ. For the other models, we report the posterior mean and the posterior standard deviation in parenthesis.

Table 3.3: Parameter Estimates from Index Returns

	BS	SV	SVJR	SVJV	SVCJ
κ		6.9691	6.3308	8.0033	7.5660
		(0.9034)	(1.0469)	(1.1433)	(1.1116)
θ	0.0384	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	2.0763	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)

Notes: We report parameters estimated using index returns only from 1996 to 2015 for the BS, SV, SVJR, SVJV and SVCJ models. Parameters are annualized. Parameters for the BS model are calculated by matching the sample average return and variance of the index. For the other models, we report the posterior mean and the posterior standard deviation after burn-in in parenthesis.

Table 3.4: Parameter Estimates from Index Returns and Option Prices

	BS	SV	SVJR	SVJV	SVCJ
κ	0.0406	2.1564	1.5531	1.0158	1.1248
		(0.0567)	(0.0404)	(0.0522)	(0.0489)
θ		0.0351	0.0359	0.0282	0.0241
		(0.0008)	(0.0008)	(0.0009)	(0.0008)
σ		0.4262	0.4152	0.3892	0.3450
		(0.0087)	(0.0085)	(0.0092)	(0.0074)
ρ		-0.9161	-0.9378	-0.9417	-0.9237
		(0.0129)	(0.0134)	(0.0118)	(0.0107)
η_s	1.9638	2.5016	2.3513	2.7833	3.0401
		(0.2805)	(0.3081)	(0.3077)	(0.3486)
η_v		1.0836	0.5753	-0.1376	0.0498
		(0.0472)	(0.0354)	(0.0421)	(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)

Notes: We report parameters estimated using index returns and option prices jointly from 1996 to 2015 for the BS, SV, SVJR, SVJV and SVCJ models. Parameters are all annualized and reported in the physical measure. The risk-neutral parameters can be inferred by subtracting risk premia from the corresponding physical parameters. Parameters for the BS model are calculated by matching the data: θ is simply the square of average implied volatility for all options and η_s is the average index return divided by θ. For the other models, we report the posterior mean and the posterior standard deviation in parenthesis.

Table 3.5: Realized Option Returns: Descriptive Statistics

Panel A: Call Returns

K/S	0.92	0.94	0.96	0.98	1.00	1.02	1.04	1.06	1.08
Mean	9.24	6.21	7.35	4.54	4.44	7.43	-11.33	-15.13	-52.07
Std	2.86	3.61	4.55	5.87	7.90	11.98	18.58	26.82	21.64
Skew	-0.49	-0.16	0.22	0.50	1.07	2.26	4.81	6.84	7.82
Kurt	2.87	2.42	2.53	2.36	3.51	8.77	28.39	55.52	63.56
2.5%	0.87	-2.24	-2.22	-7.10	-11.24	-14.78	-44.41	-65.64	-96.37
97.5%	17.63	14.87	17.01	16.30	20.04	32.03	28.09	48.51	12.20

Panel B: Put Returns

K/S	0.92	0.94	0.96	0.98	1.00	1.02	1.04	1.06	1.08
Mean	-56.80	-54.75	-47.93	-36.26	-27.10	-18.45	-12.84	-5.84	-1.53
Std	17.13	15.35	13.02	11.73	9.32	7.61	6.05	5.99	5.18
Skew	7.05	6.44	5.22	3.90	2.74	2.06	1.40	1.43	1.21
Kurt	54.55	47.14	33.35	20.32	11.53	8.06	5.07	5.27	4.62
2.5%	-87.23	-81.41	-71.38	-57.45	-44.54	-32.98	-26.49	-23.00	-20.71
97.5%	-18.65	-22.15	-20.36	-11.69	-8.09	-2.83	1.50	12.38	18.48

Panel C: Option Strategy Returns

Strategy	AS	CS	PS	CNS	LNS	CLNS
Mean	-11.24	5.72	-18.45	-5.51	-10.04	-3.94
Std	4.70	7.53	8.93	4.20	4.86	3.88
Skew	1.66	0.64	1.66	0.64	2.00	0.21
Kurt	7.36	1.98	4.69	3.02	9.21	2.25
2.5%	-20.43	-10.84	-34.95	-13.72	-20.07	-12.33
97.5%	-1.74	22.34	-0.53	3.04	0.68	4.56

Notes: We report average returns of calls, puts and several option strategies in our sample (240 months), with standard errors, skewness, kurtosis, and percentile statistics. AS: ATM Straddle, long an ATM call and an ATM put. CS: Call Spread, long an ATM call and short a 6% OTM call. PS: Put Spread, long an ATM put and short a 6% OTM put. CNS: Crash Neutral Spread, long an ATM straddle and short a 6% OTM put. LNS: Lottery Neutral Spread, long an ATM straddle and short a 6% OTM call. CLNS: Crash Lottery Neutral Spread, long an ATM straddle and short 6% OTM call and put.

Table 3.6: Model-Implied Option Returns Implied by Parameters from Index Returns

Panel A: Call Returns

K/S	0.92	0.94	0.96	0.98	1.00	1.02	1.04	1.06	1.08
Market	9.24	6.21	7.35	4.54	4.44	7.43	-11.33	-15.13	-52.07
BS	7.85	9.37	11.23	13.43	15.96	18.81	21.95	25.35	28.99*
	(0.376)	(0.278)	(0.279)	(0.145)	(0.137)	(0.208)	(0.026)	(0.038)	(0.002)
SV	8.82	10.77	13.55	17.73	24.50	36.62	58.91*	88.80	117.48
	(0.470)	(0.156)	(0.134)	(0.035)	(0.020)	(0.023)	(0.002)	(0.033)	(0.041)
SVJR	8.04	9.81	12.36	16.22	22.61	34.55	56.23*	78.87	96.31
	(0.379)	(0.215)	(0.187)	(0.053)	(0.030)	(0.033)	(0.005)	(0.049)	(0.055)
SVJV	6.72	8.18	10.24	13.26	17.94	25.44	33.82	35.02	32.60
	(0.173)	(0.421)	(0.401)	(0.179)	(0.134)	(0.156)	(0.042)	(0.175)	(0.119)
SVCJ	8.13	9.89	12.37	16.03	21.85	32.03	49.46*	61.31	63.14
	(0.273)	(0.338)	(0.319)	(0.097)	(0.065)	(0.071)	(0.008)	(0.109)	(0.078)

Panel B: Put Returns

K/S	0.92	0.94	0.96	0.98	1.00	1.02	1.04	1.06	1.08
Market	-56.80	-54.75	-47.93	-36.26	-27.10	-18.45	-12.84	-5.84	-1.53
BS	-23.36	-20.71	-18.19	-15.81	-13.62	-11.65	-9.90	-8.40	-7.13
	(0.097)	(0.028)	(0.013)	(0.028)	(0.058)	(0.168)	(0.308)	(0.296)	(0.085)
SV	-19.39	-18.68	-17.88	-16.95	-15.84	-14.46	-12.67	-10.60	-8.72
	(0.112)	(0.051)	(0.035)	(0.074)	(0.144)	(0.330)	(0.496)	(0.173)	(0.048)
SVJR	-18.01	-17.35	-16.60	-15.72	-14.68	-13.37	-11.63	-9.64	-7.90
	(0.108)	(0.045)	(0.035)	(0.061)	(0.112)	(0.271)	(0.427)	(0.230)	(0.068)
SVJV	-14.72	-14.36	-13.83	-13.11	-12.20	-11.04	-9.57	-7.96	-6.54
	(0.092)	(0.042)	(0.023)	(0.035)	(0.064)	(0.134)	(0.239)	(0.430)	(0.164)
SVCJ	-15.83	-15.77	-15.50	-14.98	-14.18	-13.05	-11.52	-9.70	-8.03
	(0.082)	(0.030)	(0.022)	(0.040)	(0.077)	(0.192)	(0.324)	(0.339)	(0.106)

Panel C: Option Strategy Returns

Strategy	AS	CS	PS	CNS	LNS	CLNS
Market	-11.24	5.72	-18.45	-5.51	-10.04	-3.94
BS	1.24*	13.57	-12.19	3.25	-1.50	0.47
	(0.005)	(0.186)	(0.217)	(0.033)	(0.021)	(0.111)
SV	4.45*	21.71	-15.08	6.94*	2.33*	4.66
	(0.000)	(0.035)	(0.368)	(0.003)	(0.004)	(0.015)
SVJR	4.08*	20.22	-13.97	6.34*	2.21*	4.33
	(0.001)	(0.054)	(0.311)	(0.005)	(0.004)	(0.018)
SVJV	2.96*	16.08	-11.59	4.84	1.35	3.13
	(0.005)	(0.168)	(0.169)	(0.019)	(0.010)	(0.035)
SVCJ	3.94*	19.22	-13.61	6.33*	1.89*	4.13
	(0.001)	(0.102)	(0.242)	(0.007)	(0.005)	(0.026)

Notes: We report the average returns for calls, puts and option strategies in our sample (240 months), with parameters estimated from index returns. In parenthesis, we report the p-values for each model to generate the market realized returns at the corresponding moneyness or option strategy. * represents significance at the 1% level.

Table 3.7: Model-Implied Option Returns Implied by Parameters from Option Prices and Option Returns

Panel A: Call Returns

K/S	0.92	0.94	0.96	0.98	1.00	1.02	1.04	1.06	1.08
Market	9.24	6.21	7.35	4.54	4.44	7.43	-11.33	-15.13	-52.07
BS	6.53	7.75	9.23	10.95	12.92	15.11	17.51	20.09	22.83*
	(0.265)	(0.389)	(0.389)	(0.221)	(0.211)	(0.285)	(0.046)	(0.057)	(0.002)
SV	5.53	6.73	8.53	11.59	18.79	75.71	110.38	115.71	83.95
	(0.159)	(0.434)	(0.401)	(0.167)	(0.096)	(0.059)	(0.197)	(0.468)	(0.308)
SVJR	4.62	5.42	6.60	8.67	14.20	48.87	87.24	127.78	172.32
	(0.100)	(0.474)	(0.485)	(0.253)	(0.163)	(0.256)	(0.241)	(0.326)	(0.211)
SVJV	4.87	5.84	7.19	9.19	12.61	16.10	6.65	-7.16	-23.19
	(0.180)	(0.412)	(0.373)	(0.168)	(0.074)	(0.464)	(0.491)	(0.331)	(0.305)
SVCJ	4.02	4.55	5.29	6.42	8.47	14.64	20.32	19.44	13.28
	(0.164)	(0.370)	(0.350)	(0.143)	(0.052)	(0.119)	(0.447)	(0.363)	(0.394)

Panel B: Put Returns

K/S	0.92	0.94	0.96	0.98	1.00	1.02	1.04	1.06	1.08
Market	-56.80	-54.75	-47.93	-36.26	-27.10	-18.45	-12.84	-5.84	-1.53
BS	-19.13	-16.94	-14.86*	-12.92	-11.15	-9.54	-8.13	-6.91	-5.88
	(0.060)	(0.012)	(0.005)	(0.012)	(0.026)	(0.099)	(0.207)	(0.420)	(0.151)
SV	-20.28	-18.74	-17.19	-15.61	-14.01	-12.18	-9.23	-7.06	-5.59
	(0.206)	(0.107)	(0.072)	(0.089)	(0.120)	(0.226)	(0.291)	(0.379)	(0.153)
SVJR	-34.33	-29.52	-24.54	-19.61	-15.12	-11.45	-8.09	-6.14	-4.90
	(0.246)	(0.144)	(0.113)	(0.144)	(0.159)	(0.213)	(0.237)	(0.431)	(0.184)
SVJV	-21.55	-18.06	-15.27	-13.06	-11.29	-9.77	-7.90	-6.28	-5.04
	(0.231)	(0.143)	(0.100)	(0.126)	(0.159)	(0.277)	(0.317)	(0.375)	(0.160)
SVCJ	-42.38	-33.97	-26.20	-19.60	-14.51	-10.91	-8.27	-6.41	-5.12
	(0.367)	(0.467)	(0.405)	(0.436)	(0.443)	(0.461)	(0.490)	(0.224)	(0.079)

Panel C: Option Strategy Returns

Strategy	AS	CS	PS	CNS	LNS	CLNS
Market	-11.24	5.72	-18.45	-5.51	-10.04	-3.94
BS	0.94*	10.96	-9.89	2.69	-1.37	0.35
	(0.007)	(0.280)	(0.132)	(0.047)	(0.018)	(0.118)
SV	2.54*	18.15	-12.94	4.20	1.75*	3.42
	(0.002)	(0.115)	(0.289)	(0.019)	(0.006)	(0.036)
SVJR	-0.32	12.64	-11.92	2.65	-1.60	1.37
	(0.032)	(0.214)	(0.271)	(0.038)	(0.036)	(0.089)
SVJV	0.74	12.69	-9.94	2.55	0.22	2.05
	(0.024)	(0.221)	(0.024)	(0.041)	(0.021)	(0.054)
SVCJ	-2.95	7.70	-8.58	1.19	-3.85	0.37
	(0.087)	(0.360)	(0.023)	(0.054)	(0.133)	(0.127)

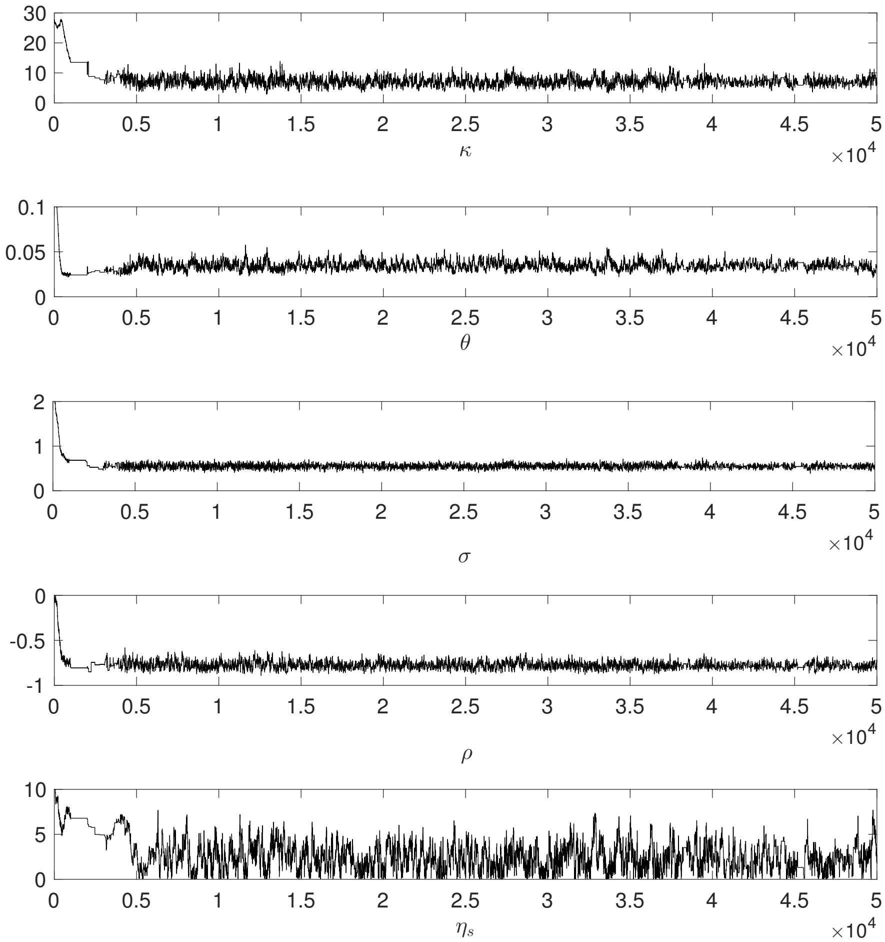
Notes: We report the average returns for calls, puts and option strategies in our sample (240 months), with parameters estimated from index option prices and option returns. In parenthesis, we report the p-values for each model to generate the market realized returns at the corresponding moneyness or option strategy. * represents significance at the 1% level.

Table 3.8: Option Return Root Mean Squared Error

	Index Returns	Index Returns & Option Prices	Option Returns & Option Prices
BS	1.24	1.27	1.23
SV	1.19	0.98	0.85
SVJR	1.16	0.88	0.76
SVJV	1.08	0.82	0.75
SVCJ	1.12	0.73	0.62

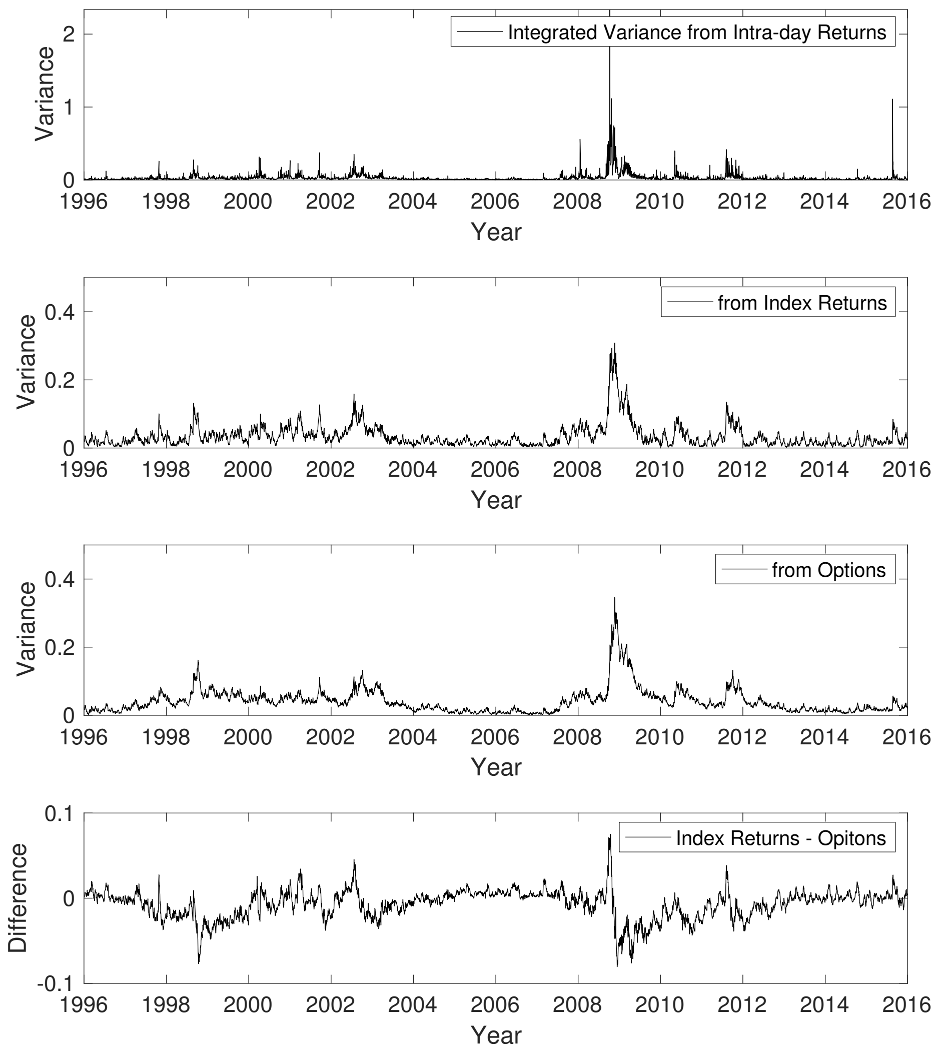
Notes: For each model, we report the Option Return Root Mean Squared Error (OR-RMSE) across all moneyness for calls and puts. Values in the table represent the average distances between model implied option returns and market realized option returns. OR-RMSE can not be directly considered as a measure to reject a model, but it indicates how close on average the model predicted returns are to market realized returns given the standard deviation at each moneyness. Overall more complicated models tend to fit option return better and the estimation using option prices and option returns provides the best fit among the three methods.

Figure 3.1: Parameter Trace for the SV Model Parameters. Estimation Based on Index Returns



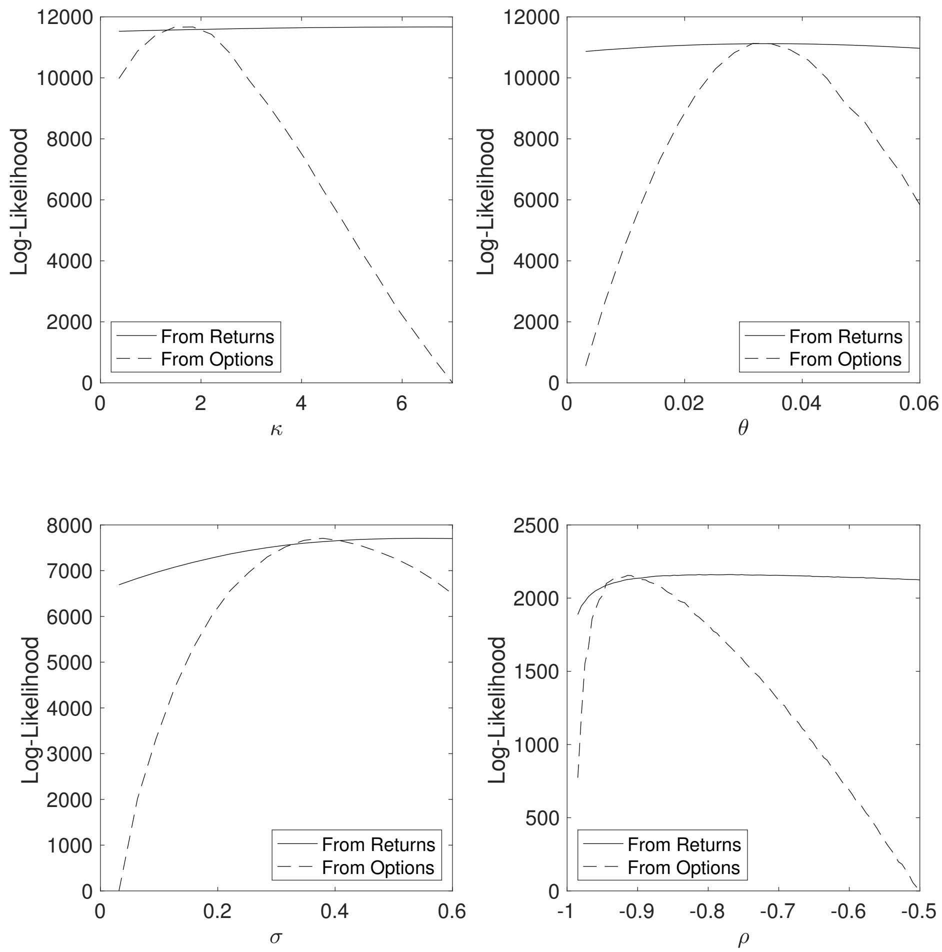
Note: We plot the full traces for each parameter in the SV model. We use 50,000 iterations. The first 1/4 of the iterations are treated as burn-in. We choose 'crazy' initial point for each parameter in order to confirm the robustness of the method. We report the average after burn-in as parameter value and the standard deviation after burn-in as parameter standard deviation.

Figure 3.2: Filtered Variance for the SV Model



Notes: We plot the realized variance integrated from intra-day 5-min index returns in the first panel. In the middle two panels, we plot the filtered variances for the SV model using either index returns or index options (option prices and option returns). We also plot the difference in filtered variances in the bottom panel.

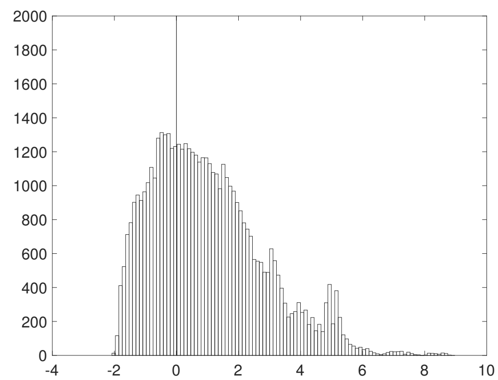
Figure 3.3: Log-Likelihood Curve Index Returns vs Index Options



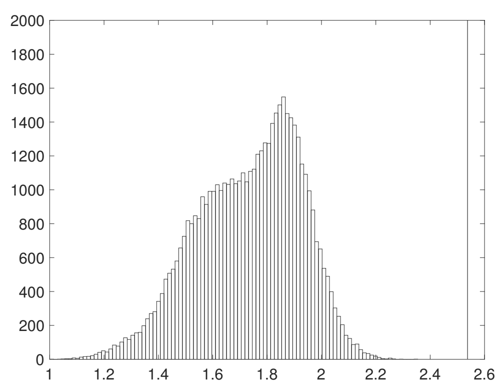
Notes: We plot the Log-Likelihood curve for index returns vs index options for different parameters. We shift the level of the likelihood to have the same value at maximum but we keep the scale and thus the 'gradient' remains the same.

Figure 3.4: Parameter Difference Posterior Distributions

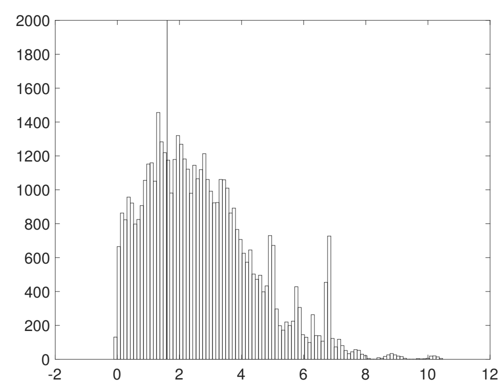
(a) Difference between η_s Estimated from Index Returns and η_s Estimated from Index Options



(b) Difference between η_s Estimated from Index Options and 2.5374



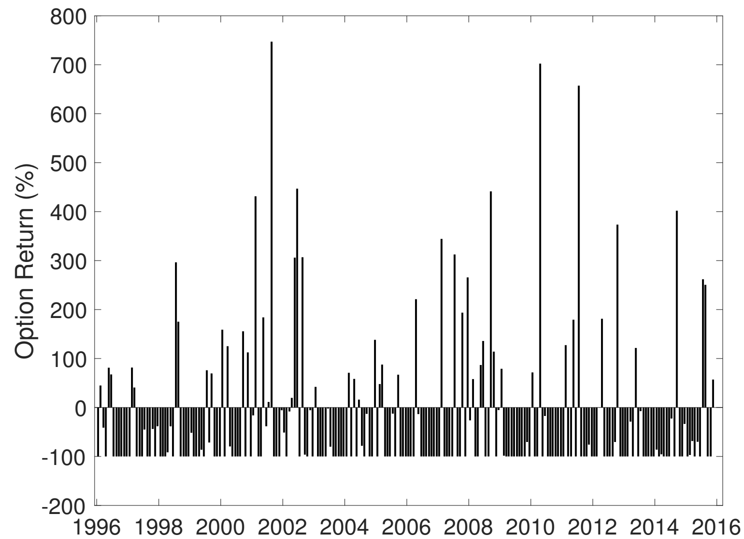
(c) Difference between η_s Estimated from Index Returns and 1.6030



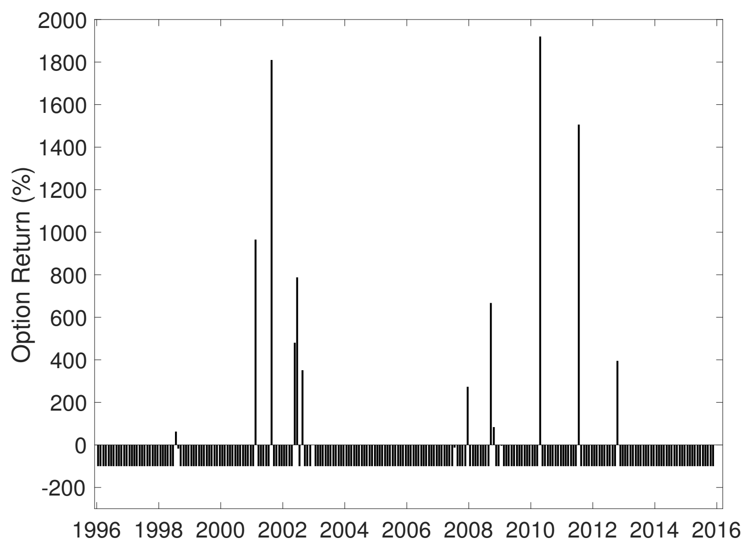
Notes: We histogram the posterior distribution for the difference between η_s estimated from index returns and η_s estimated from index options in panel (a). The probability of index returns based η_s being larger than index options based η_s is 0.6625 and we do not reject the hypothesis that η_s from the two different sources of information are the same. We histogram the posterior distribution of η_s estimated from index options and the posterior mean value of 2.5374 estimated from index returns in panel (b). We clearly reject the hypothesis that η_s estimated from index options is the same as the posterior mean value estimated from index returns. We histogram the posterior distribution of η_s estimated from index returns and the posterior mean value of 1.6030 estimated from index returns in panel (c). We do not reject the hypothesis that η_s estimated from index returns is the same as the posterior mean value estimated from index options.

Figure 3.5: Realized Monthly Option Returns

(a) ATM Call Option Returns



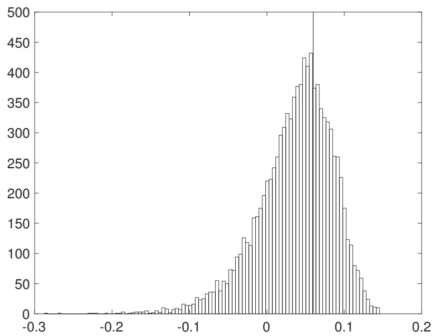
(b) 6% OTM Put Option Returns



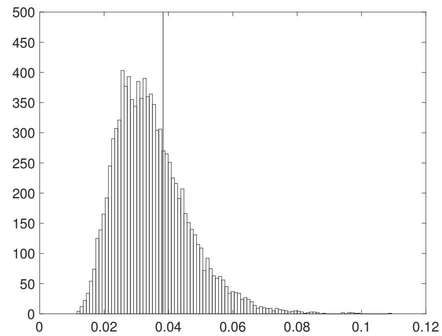
Notes: We plot option returns of ATM calls and 6% OTM puts for each month from 1996 to 2015. Option returns are reported in percentage. The maximum losses of ATM calls are -100% and the gains can be as high as 800%. The majority of realized returns for 6% OTM puts are -100% while at the same time, due to the low prices of these options, the returns can be very large if positive payoffs are realized.

Figure 3.6: Simulated Index Returns and Variances

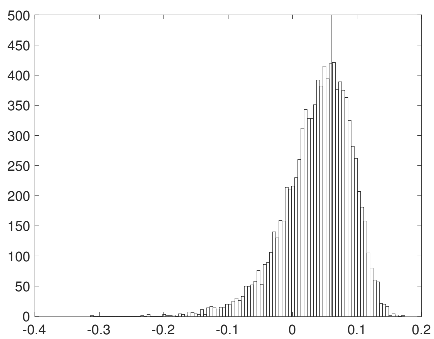
(a) Index Returns: SV



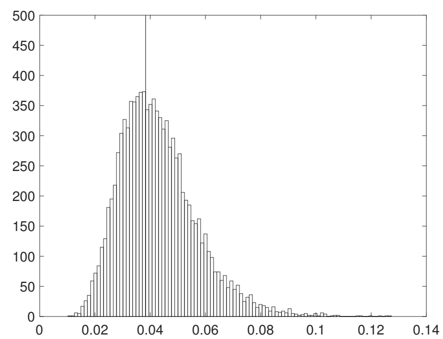
(b) Index Variances: SV



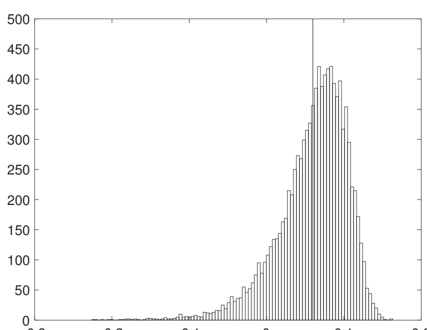
(c) Index Returns: SVJR



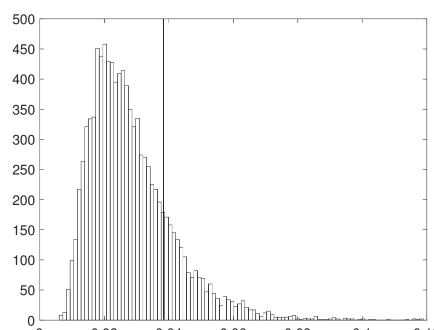
(d) Index Variances: SVJR



(e) Index Returns: SVCJ



(f) Index Variances: SVCJ



Notes: We histogram the simulated average index returns and variances for the SVJR, SVJV and SVCJ models. The vertical lines represent the realized average return and variance from market data.

Bibliography

Adrian, Tobias, and Joshua Rosenberg, 2008, Stock Returns and Volatility: Pricing the Long-Run and Short-Run Components of Market Risk, *Journal of Finance* 63, 2997–3030.

Aït-Sahalia, Yacine, Julio Cacho-Diaz, and Roger Laeven, 2015a, Modeling Financial Contagion Using Mutually Exciting Jump Processes, *Journal of Financial Economics* 117, 585–606.

Aït-Sahalia, Yacine, Mustafa Karaman, and Loriano Mancini, 2015b, The Term Structure of Variance Swaps and Risk Premia, *Working Paper, Princeton University* .

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

Aït-Sahalia, Yacine, and Andrew W. Lo, 1998, Nonparametric Estimation of State-Price Densitites Implicit in Financial Asset Prices, *Journal of Finance* 53, 499–547.

Alizadeh, Sassan, Michael W. Brandt, and Francis X. Diebold, 2002, Range-Based Estimation of Stochastic Volatility Models, *Journal of Finance* 57, 1047–1091.

Andersen, Torben G., Luca Benzoni, and Jesper Lund, 2002, An Empirical Investigation of Continuous-Time Equity Return Models, *Journal of Finance* 57, 193–223.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens, 2001, The Distribution of Realized Stock Return Volatility, *Journal of Financial Economics* 61, 43–76.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2003, Modeling and Forecasting Realized Volatility, *Econometrica* 71, 579–625.

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.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2017, Short-Term Market Risks Implied by Weekly Options, *Journal of Finance* 72, 1335–1386.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The Cross-Section of Volatility and Expected Returns, *Journal of Finance* 61, 259–299.

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

Bakshi, Gurdip, and Nikunj Kapadia, 2003, Delta-Hedged Gains and the Negative Market Volatility Risk Premium, *Review of Financial Studies* 16, 527–566.

Bakshi, Gurdip, Dilip Madan, and George Panayotov, 2010, Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels, *Journal of Financial Economics* 97, 130–154.

Bali, Turan G., and Armen Hovakimian, 2009, Volatility Spreads and Expected Stock Returns, *Management Science* 55, 1797–1812.

Bandi, F. M., and R. Reno, 2016, Price and Volatility Co-Jumps, *Journal of Financial Economics* 119, 107–146.

Barberis, Nicholas, and Ming Huang, 2008, Stocks as Lotteries: The Implication of Probability Weighting for Security Prices, *American Economic Review* 98, 2066–2100.

Barndorff-Nielsen, Ole E., and Neil Shephard, 2002, Econometric Analysis of Realised Volatility and Its Use in Estimating Stochastic Volatility Models, *Journal of the Royal Statistical Society* 64, 253–280.

Bates, David S, 1988, Pricing Options on Jump-Diffusion Process, *Working Paper, University of Pennsylvania* .

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

Bates, David S., 2000, Post-'87 Crash Fears in the S&P 500 Futures Option Market, *Journal of Econometrics* 94, 181–238.

Bates, David S., 2003, Empirical Option Pricing: A Retrospection, *Journal of Econometrics* 116, 387–404.

Bates, David S., 2006, Maximum Likelihood Estimation of Latent Affine Processes, *Review of Financial Studies* 19, 909–965.

Bates, David S., 2012, U.S. Stock Market Crash Risk, 1926-2010, *Journal of Financial Economics* 105, 229–259.

Bates, David S., 2018, How Crashes Develop: Intradaily Volatility and Crash Evolution, Forthcoming, Journal of Finance .

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2011, Explaining Asset Pricing Puzzles Associated with the 1987 Market Crash, *Journal of Financial Economics* 101, 552–573.

Black, Fischer, and Myron Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637–654.

Bollen, Nicolas P.B., and Robert E. Whaley, 2004, Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?, *Journal of Finance* 59, 711–753.

Bollerslev, Tim, James Marrone, Lai Xu, and Hao Zhou, 2014, Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence, *Journal of Financial and Quantitative Analysis* 49, 633–661.

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

Bondarenko, Oleg, 2014, Why are Put Options So Expensive?, *Quarterly Journal of Finance* 4, 1450015.

Branger, Nicole, Alexandra Hansis, and Christian Schlag, 2010, Expected Option Returns and the Structure of Jump Risk Premia, *Working Paper, University of Muenster* .

Britten-Jones, Mark, and Anthony Neuberger, 2000, Option Prices, Implied Price Processes, and Stochastic Volatility, *Journal of Finance* 55, 839–866.

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

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2009, Understanding Index Option Returns, *Review of Financial Studies* 22, 4493–4529.

Brunnermeier, Markus K., Christian Gollier, and Jonathan A. Parker, 2007, Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns, *American Economic Review* 97, 159–165.

Cao, Charles, and Jing-Zhi Huang, 2007, Determinants of S&P 500 Index Option Returns, *Review of Derivatives Research* 10, 1–38.

Carr, Peter, and Dilip B. Madan, 1998, Option Valuation Using the Fast Fourier Transform, *Journal of Computational Finance* 2, 61–73.

Carr, Peter, and Liuren Wu, 2009, Variance Risk Premiums, *Review of Financial Studies* 22, 1311–1341.

Chacko, George, and Luis M. Viceira, 2003, Spectral GMM Estimation of Continuous-Time Processes, *Journal of Econometrics* 116, 259–292.

Chambers, Donald R., Matthew Foy, Jeffrey Liebner, and Qin Lu, 2014, Index Option Returns: Still Puzzling, *Review of Financial Studies* 27, 1915–1928.

Chaudhuri, Ranadeb, and Mark Schroder, 2015, Monotonicity of the Stochastic Discount Factor and Expected Option Returns, *Review of Financial Studies* 28, 1462–1505.

Chernov, Mikhail, A. Ronald Gallant, Eric Ghysels, and George Tauchen, 2003, Alternative Models for Stock Price Dynamics, *Journal of Econometrics* 116, 225–257.

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.

Christensen, Bent J., and Nagpurnanand R. Prabhala, 1998, The Relation Between Implied and Realized Volatility, *Journal of Financial Economics* 50, 125–150.

Christoffersen, Peter, Steven Heston, and Kris 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, Peter, Steven Heston, and Kris Jacobs, 2013, Capturing Option Anomalies with a Variance-Dependent Pricing Kernel, *Review of Financial Studies* 26, 1962–2006.

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.

Cochrane, John H., and Monika Piazzesi, 2005, Bond Risk Premia, *American Economic Review* 95, 138–160.

Constantinides, George M., Michal Czerwonko, Jens Carsten Jackwerth, and Stylianos Perrakis, 2011, Are Options on Index Futures Profitable for Risk-Averse Investors? Empirical Evidence, *Journal of Finance* 66, 1407–1437.

Constantinides, George M., and Jens Carsten Jackwerth, 2009, Mispricing of S&P 500 Index Options, *Review of Financial Studies* 22, 1247–1277.

Constantinides, George M., Jens Carsten Jackwerth, and Alexi Savov, 2013, The Puzzle of Index Option Returns, *Review of Asset Pricing Studies* 3, 229–257.

Corsi, Fulvio, 2009, A Simple Approximate Long-Memory Model of Realized Volatility, *Journal of Financial Econometrics* 7, 174–196.

Coval, Joshua D., and Tyler Shumway, 2001, Expected Option Returns, *Journal of Finance* 56, 983–1009.

Cremers, Martijn, Michael Halling, and David Weinbaum, 2015, Aggregate Jump and Volatility Risk in the Cross-Section of Stock Returns, *Journal of Finance* 70, 577–614.

Della Corte, Pasquale, Tarun Ramadorai, and Lucio Sarno, 2016, Volatility Risk Premia and Exchange Rate Predictability, *Journal of Financial Economics* 120, 21–40.

Dennis, Patrick, and Stewart Mayhew, 2002, Risk-Neutral Skewness: Evidence from Stock Options, *Journal of Financial and Quantitative Analysis* 37, 471–493.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, 2017, The Price of Variance Risk, *Journal of Financial Economics* 123, 225–250.

Drechsler, Itamar, and Amir Yaron, 2011, What's Vol Got to Do with It, *Review of Financial Studies* 24, 1–45.

Driessen, Joost, and Pascal Maenhout, 2007, An Empirical Portfolio Perspective on Option Pricing Anomalies, *Review of Finance* 11, 561–603.

Du, Du, 2011, General Equilibrium Pricing of Options with Habit Formation and Event Risks, *Journal of Financial Economics* 99, 400–426.

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

Dupire, Bruno, 1994, Pricing with a Smile, *Risk* 7, 1–10.

Egloff, Daniel, Markus Leippold, and Liuren Wu, 2010, The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments, *Journal of Financial and Quantitative Analysis* 45, 1279–1310.

Engle, Robert F., and Jose Gonzalo Rangel, 2008, The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes, *Review of Financial Studies* 21, 1187–1222.

Eraker, Bjørn, 2004, Do Stock Prices And Volatility Jump? Reconciling Evidence From Spot And Option Prices, *Journal of Finance* 59, 1–37.

Eraker, Bjørn, 2012, The Volatility Premium, *Working Paper, University of Wisconsin* .

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

Faias, José Afonso, and Pedro Santa-Clara, 2017, Optimal Option Portfolio Strategies: Deepening the Puzzle of Index Option Mispricing, *Journal of Financial and Quantitative Analysis* 52, 277–303.

Fang, F., and C.W. Oosterlee, 2008, A Novel Option Pricing Method for European Options Based on Fourier-Cosine Series Expansions, *SIAM Journal on Scientific Computing* 31, 826–848.

Feunou, Bruno, and Cédric Okou, 2017, Risk-Neutral Moment-Based Estimation of Affine Option Pricing Models, *Working Paper, Bank of Canada* .

Fleming, Jeff, Barbara Ostdiek, and Robert E. Whaley, 1995, Predicting Stock Market Volatility: A New Measure, *Journal of Futures Markets* 15, 265–302.

Frazzini, Andrea, and Lasse H. Pedersen, 2012, Embedded Leverage, *Working Paper, New York University* .

Fulop, Andras, Junye Li, and Jun Yu, 2015, Self-Exciting Jumps, Learning, and Asset Pricing Implications, *Review of Financial Studies* 28, 876–912.

Gârleanu, Nicolae, Lasse Heje Pedersen, and Allen M. Poteshman, 2009, Demand-Based Option Pricing, *Review of Financial Studies* 22, 4259–4299.

Goetzmann, William, Jonathan Ingersoll, Matthew Spiegel, and Ivo Welch, 2004, Sharpening Sharpe Ratios, *Working Paper, Yale University* .

Goyal, Amit, and Alessio Saretto, 2009, Cross-Section of Option Returns and Volatility, *Journal of Financial Economics* 94, 310–326.

Han, Bing, 2007, Investor Sentiment and Option Prices, *Review of Financial Studies* 21, 387–414.

Hastings, W K, 1970, Monte Carlo Sampling Methods Using Markov Chains and Their Applications, *Biometrika* 57, 97–109.

Heston, Steven 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.

Hu, Guanglian, and Kris Jacobs, 2017, Volatility and Expected Option Returns, *Working Paper, University of Houston* .

Hurn, A. S., K. A. Lindsay, and A. J. McClelland, 2015, Estimating the Parameters of Stochastic Volatility Models Using Option Price Data, *Journal of Business and Economic Statistics* 33, 579–594.

Israelov, Roni, and Bryan Kelly, 2017, Forecasting the Distribution of Option Returns, *Working Paper, University of Chicago* .

Jackwerth, Jens Carsten, 2000, Recovering Risk Aversion from Option Prices and Realized Returns, *Review of Financial Studies* 13, 433–451.

Jackwerth, Jens Carsten, and Mark Rubinstein, 1996, Recovering Probability Distributions from Contemporaneous Security Prices, *Journal of Finance* 51, 1611–1631.

Jacobs, Kris, and Yuguo Liu, 2017, The Information in Index Returns and the Cross-Section of Index Options, *Working Paper, University of Houston* .

Jiang, George J., and Yisong S. Tian, 2005, The Model-Free Implied Volatility and Its Information Content, *Review of Financial Studies* 18, 1305–1342.

Johannes, Michael, and Nicholas Polson, 2009, MCMC Methods for Continuous-Time Financial Econometrics, in *Handbook of Financial Econometrics, Vol 2*, 1–72.

Johannes, Michael, Nicholas Polson, and Jonathan Stroud, 2009, Optimal Filtering of Jump Diffusions: Extracting Latent States from Asset Prices, *Review of Financial Studies* 22, 2759–2799.

Jones, Christopher S., 2003, The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, *Journal of Econometrics* 116, 181–224.

Jones, Christopher S., 2006, A Nonlinear Factor Analysis of S&P 500 Index Option Returns, *Journal of Finance* 61, 2325–2363.

Lamoureux, Christopher G., and William D. Lastrapes, 1993, Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities, *Review of Financial Studies* 6, 293–326.

Linn, Matthew, Sophie Shive, and Tyler Shumway, 2018, Pricing Kernel Monotonicity and Conditional Information, *Review of Financial Studies* 31, 493–531.

Liu, Jun, Jun Pan, and Tan Wang, 2005, An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks, *Review of Financial Studies* 18, 131–164.

Liu, Lily Y., Andrew J. Patton, and Kevin Sheppard, 2015, Does Anything Beat 5-Minute RV? A Comparison of Realized Measures across Multiple Asset Classes, *Journal of Econometrics* 187, 293–311.

Malik, Sheheryar, and Michael K. Pitt, 2011, Particle Filters for Continuous Likelihood Evaluation and Maximisation, *Journal of Econometrics* 165, 190–209.

Marsaglia, George, 1965, Ratios of Normal Variables and Ratios of Sums of Uniform Variables, *Journal of the American Statistical Association* 60, 193–204.

Marsaglia, George, 2006, Ratios of Normal Variables, *Journal of Statistical Software* 16, 1–10.

Merton, Robert C., 1973, Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* 4, 141–183.

Merton, Robert C., 1976, Option Pricing when Underlying Stock Returns are Discontinuous, *Journal of Financial Economics* 3, 125–144.

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

Mitton, Todd, and Keith Vorkink, 2007, Equilibrium Underdiversification and the Preference for Skewness, *Review of Financial Studies* 20, 1255–1288.

Naik, Vasanttilak, and Moon Lee, 1990, General Equilibrium Pricing of Options on the Market Portfolio with Discontinuous Returns, *Review of Financial Studies* 3, 493–521.

Neuberger, Anthony, 1994, The Log Contract, *Journal of Portfolio Management* 20, 74–80.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–727.

Newey, Whitney K., and Kenneth D. West, 1994, Automatic Lag Selection in Covariance Matrix Estimation, *Review of Economic Studies* 61, 631–653.

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

Pitt, Michael K., 2002, Smooth Particle Filters for Likelihood Evaluation and Maximisation Smooth particle filters for likelihood evaluation and maximisation, *Working Paper, University of Warwick* .

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

Polkovnichenko, Valery, and Feng Zhao, 2013, Probability Weighting Functions Implied in Options Prices, *Journal of Financial Economics* 107, 580–609.

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

Rosenberg, Joshua V., and Robert F. Engle, 2002, Empirical pricing kernels, *Journal of Financial Economics* 64, 341–372.

Ross, Steve, 2015, The Recovery Theorem, *Journal of Finance* 70, 615–648.

Rubinstein, Mark, 1994, Implied Binomial Trees, *Journal of Finance* 49, 771–818.

Santa-Clara, Pedro, and Alessio Saretto, 2004, Option Strategies: Good Deals and Margin Calls, *Journal of Financial Markets* 12, 391–417.

Santa-Clara, Pedro, and Alessio Saretto, 2009, Option Strategies: Good Deals and Margin Calls, *Journal of Financial Markets* 12, 391–417.

Seo, Sang Byung, and Jessica A. Wachter, 2017, Option Prices in a Model with Stochastic Disaster Risk, *Working Paper, University of Pennsylvania* .

Singleton, Kenneth J., 2006, *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment, Princeton University Press*.

Todorov, Viktor, 2010, Variance Risk-Premium Dynamics: The Role of Jumps, *Review of Financial Studies* 23, 345–383.