THREE ESSAYS ON OBSERVABLE COVARIATES IN OPTION PRICING

by

Yoontae Jeon

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of The Joseph L. Rotman School of Management
University of Toronto

Abstract

THREE ESSAYS ON OBSERVABLE COVARIATES IN OPTION PRICING

Yoontae Jeon

Doctor of Philosophy

Graduate Department of The Joseph L. Rotman School of Management

University of Toronto

2017

This dissertation contains three essays on observable covariates in option pricing. In the first essay, I propose firm-specific public news arrival from Factiva database as an observable covariate in equity options market and study how the public news arrival is priced. I first establish the empirical relationship between the firm-specific public news arrival and jumps in individual equity returns. Subsequently, I build a continuous-time stochastic volatility jump diffusion model where news arrivals driving the jump dynamics. When estimated on equity options data for 20 individual firms, the premia placed on jump frequency and size turn out to be consistent with the theories highlighting both positive and negative effects of public news arrival.

The second essay, based on a joint work with Peter Christoffersen, Bruno Feunou and Chayawat Ornthanalai, studies how the stock market illiquidity affects the market crash risk. Our empirical approach is to estimate a continuous-time model with stochastic volatility and dynamic crash probability where stock market illiquidity is used as an observable covariate driving the crash probability. While the crash probability is time-varying, its dynamic depends only weakly on return variance once we include market illiquidity as an economic variable in the model. This finding suggests that the relationship between variance and jump risk found in the literature is largely due to their common exposure to market illiquidity. Our study highlights the importance of equity market frictions in index return dynamics and explains why prior studies find that crash risk increases with market uncertainty level.

The third essay, based on a joint work with Peter Christoffersen and Bruno Feunou, proposes the realized jump variation measure constructed from the intraday S&P500 returns data as an observable covariate that helps pricing of index options. The volatility and jump intensity dynamics in the model are directly driven by model-free empirical measures of diffusive volatility and jump variation. Because the empirical measures are observed in discrete intervals, our option valuation model is cast in discrete time, allowing for straightforward filtering and estimation of the model. When estimated on S&P500 index options and returns the new model performs well compared with standard benchmarks.

Acknowledgements

I would like to first thank the members of my dissertation committee Peter Christoffersen, Tom McCurdy, and Jason Wei for their support and guidance throughout my entire Ph.D. study. I wish to express my special thanks to Peter Christoffersen, who was my supervisor, and Tom McCurdy for their continuous support and encouragement. I am also heavily indebted to Redouane Elkamhi and Chayawat Ornthanalai for their patience and effort on guiding my research. I was very lucky to be in the environment with wonderful faculty members of the finance department. In particular, I benefited a lot from Michael Hasler, Raymond Kan, and Liyan Yang. And of course, I would also like to thank my fellow Ph.D. students in finance who have been my best colleagues during my study. Especially, I benefited heavily from all conversations I had with my officemate Joon Woo Bae. I would like to acknowledge the financial support from the Rotman School of Management and Social Science and Humanities Research Council (SSHRC). All my work would have been not possible without their financial support. Last but most importantly, I wish to express great thanks to my parents, my wife, and my sister for supporting me with patience throughout my study.

Contents

List of Tables

List of Figures

Chapter 1

The Price of News Arrivals: Evidence from Equity Options

1.1 Introduction

One of the first empirical puzzles found in the financial literature is the leptokurtic distribution of stock returns. Earlier research tried to explain the puzzle by the so called Mixture of Distributions Hypothesis (MDH) (e.g., Clark (1973), Epps and Epps (1976), Tauchen and Pitts (1983)). MDH conjectures that trading activities are triggered by randomly spaced arrivals from a latent information process. However, the biggest problem with the empirical testing of MDH lies in the latent nature of the information process. The literature has thus largely focused on testing the implications of MDH, namely the volume-volatility relationship. Instead of relying on the implications of the hypothesis, in this paper, I directly test the role of specific information process, the firm-specific public news arrival, and how the uncertainty of its arrival is priced in the market.

To construct a measure of public news arrival, I use one of the most widely used database, Factiva. For each of the 20 firms in my sample, I construct two measures of public news arrival. The first is a simple count of the daily number of news articles appearing in the database while the second measures the news tone associated with individual article by applying the textual analysis technique developed in Loughran and McDonald (2011).

Large amounts of public news arrival coincides with large discrete movements in daily stock returns. For example, Microsoft's stock price dropped by 14.47% on Apr 3^{rd}, 2000 following a judge's ruling that Microsoft had violated antitrust laws. Of course, the day came with excessive amount of news articles,

348 news articles in my sample compared to the average news articles per day of 71, reporting the judge's ruling. On the theoretical side, Andersen (1996) proposes a modified Mixture of Distributions Hypothesis where the information arrival induces Poisson-type jumps in returns and finds strong empirical support. Along these lines, I conjecture that public news arrivals are related to the stock return jumps, rather than its continuous movements, and I find supporting empirical evidences.

Detecting jumps in stock returns has been one of the most active research area in financial econometrics in last decade.[1] I employ the method developed in Lee and Mykland (2008) to identify individual daily stock returns as jump or no jump days. Using the identified jump days, I find strong evidence linking public news count with the probability of jumps. In particular, the news tone that measures the tone of each news article, does not have a significant relationship to the occurrence of jumps, but it is strongly correlated with the size of a jump conditional on its occurrence. This evidence is consistent with Tetlock (2007) who has found the correlation between pessimistic tones extracted from the Wall Street Journal and large negative returns. I also find additional evidence from the equity options market using the slope of the implied volatility surface. Both news counts and news tones are shown to be related with the IV-SKEW that proxies for the embedded risk-adjusted jump risk.

I then use the findings from the daily returns as a guidance to build a continuous time stochastic-volatility jump-diffusion model of daily returns, with the goal to study the market price of risk associated with the public news arrival. The major innovation of the model is to feature a time-varying jump-intensity where its variation solely depends on the observed public news arrival. By fitting the model to the daily returns, news arrivals, and equity options prices, I find a significant positive risk premium associated with the public news innovation while the actual jump size due to the news carries a large negative risk premium. This suggests that a public news arrival is not viewed as redundant, but rather viewed as something investors prefer to have. Through this public news interpretation, I am able to both reconcile and explain the previously documented puzzling positive jump-timing risk premium.

There are mainly two channels for public news and its contents to cause stock returns to jump. The first channel is by influencing the beliefs of either noise or liquidity traders. The seminal paper by De Long, Shleifer, Summers, and Waldmann (1990) studies how noise trader risk can explain various empirical puzzles, while Campbell, Grossman, and Wang (1993) focuses on how sudden changes in liquidity traders can affect short-term returns. Tetlock (2007) provides empirical evidence consistent with these theoretical models. On the other hand, public news arrival can be viewed as a resolution of information asymmetry. If the public news articles do contain information that was only know to a group of privately informed investors, the arrival of such news instantaneously resolves the information

[1] For reference, see Huang and Tauchen (2005) and Gilder, Shackleton, and Taylor (2014) for concise summary.

asymmetry, thus causing the stock price to jump.

Furthermore, it is not clear how the public information should be priced in the market. If the noise trading effect is true, then public information is nothing more than exogenous shocks that cause short-term market movements that in turn quickly revert to the fundamental, hence should carry a negative risk premium. On the other hand, if the public information indeed resolves the information asymmetry, then it should carry a positive risk premium as shown in Easley and O'hara (2004). My findings suggest that, at least when using Factiva based news counts and contents as a proxy for public information, both theories co-exist in the market. Specifically, the positive jump-timing risk premium can be viewed as evidence of resolution of information asymmetry via a public news arrival story. On the other hand, the significantly negative jump-size premium can be viewed as a negative risk premium associated with risk aversion against the effect of noise trading induced price jumps. Thus, I conclude that public news arrival is an important economic factor that is strongly priced in the market, and the resulting evidences are consistent with both views from the previous theories.

This paper is perhaps most closely related to the work by Engle, Hansen, and Lunde (2012). They use the same news dataset from Factiva and study whether the news information can improve the forecasting power of daily realized volatility. In contrast, I emphasize the contemporaneous relationship between news arrival and stock return jumps. Also, I focus on using news arrivals as an exogenous observable to extract the risk premium associated with it instead of forecasting. In this respect, Lee (2012)'s work serves as good evidence why I focus on jumps. Lee (2012) finds that there is a higher chance of observing intra-day jumps in returns during scheduled firm-specific news announcement times. I follow the same intuition, with the notable difference that I study the impact of a daily time-series of news arrival instead of focusing on specific events. In other words, I am specifically interested in the role played by the unexpected and mostly unscheduled component of public news arrival.

There exists an extensive amount of literature studying the time-varying jump-intensity in returns and options prices.[2] The typical approach, mostly taken for its analytical tractability, is to assume an affine functional form for the jump-intensity in the latent variance process.[3] However, this approach does not allow for separate interpretations of the diffusive variance risk premium and jump-timing risk premium because two come from the same source. Particularly, two risk premiums are forced to have the same sign, which is not a required restriction. Perhaps due to this analytical complexity, the jump-timing premium has been mostly neglected and assumed away from in most studies. My model contributes to

[2]Too cite few, see Andersen, Benzoni, and Lund (2002), Eraker (2004), Broadie, Chernov, and Johannes (2007), Christoffersen, Jacobs, and Ornthanalai (2012), Ornthanalai (2014), and Andersen, Fusari, and Todorov (2015b).

[3]Maheu and McCurdy (2004), Santa-Clara and Yan (2010) and Maheu, McCurdy, and Zhao (2013) are notable exceptions where the jump-intensity process is modeled as a separate latent process.

this literature by proposing to bypass this issue by allowing a purely public news dependent process that enters the jump-intensity equation, thus allowing one to identify the two premiums separately.

In terms of methodology, my paper is also related to the literature on explaining derivatives prices using economic co-variates. Usage of stochastic co-variates has been a popular approach in the credit derivatives literature.[4] There are way fewer studies linking economic co-variates to the pricing of options, perhaps due to the different modeling approach and difficulties associated with assigning appropriate co-variates. I contribute to this literature by proposing an observed news process as a possible candidate for an economic co-variate.

The remainder of the paper is organized as follows. Section 2 describes the dataset used for the analysis and provides preliminary non-parametric evidence from both equity and options markets. In Section 3, I develop the structural model that builds on the findings from Section 2 and discuss the estimation strategy. Section 4 focuses on the resulting implications from the estimated parameters and its properties. Section 5 concludes.

1.2 Data and Non-Parametric Analysis

In this section, I first describe the data sets used in the paper. I then perform non-parametric analysis to look for the evidence that links firm-specific news arrivals to jumps in returns.[5] The result not only indicates that firm-specific news arrival is related to return jumps, but also provides good intuition on how the parametric models should be structured.

1.2.1 Data

The main variable of interest is firm-specific news arrival. The focus of this paper is placed on the role played by the daily arrival of firm-specific news instead of specific corporate events with large news flows, including both scheduled and unscheduled. Therefore, I require a comprehensive database that contains as many firm-specific news articles as possible. In this regard, I use Factiva database to search for comprehensive list of news articles.[6] Due to the technological advances such as Internet, the number of daily news articles have dramatically increased since the early 2000s. For this reason, I start my sample period on January, 2000 and ends on July, 2012 to avoid issues with obvious trend in news data.

Factiva database conveniently identifies each news article by its own ticker, which allows me to easily

[4] See Altman (1968), Shumway (2001), Duffie, Saita, and Wang (2007), Dionne, Gauthier, Hammami, Maurice, and Simonato (2011), etc.

[5] The dataset and most of non-parametric analysis in this section are taken from Jeon, McCurdy, and Zhao (2016).

[6] Bajgrowicz and Scaillet (2011) and Engle, Hansen, and Lunde (2012) also use Factiva database.

merge CRSP database with Factive news articles at daily level by its unique ticker. I identify 20 firms with the most amount of news articles for this study as the firms with smaller variation in its amount of daily news flow would not provide as much reliable conclusions as the firms with large amount of news flows. For each firm, I simply count the number of news articles with its ticker that appears in the Factiva database.[7] Table 1.1 shows descriptive statistics of daily news counts for 20 selected firms. The mean number of news articles observed each day is around 44 while median number is 32, indicating that significant amount of weight is placed on the large news counts outlier. Table 1.1 also shows that daily news counts are highly volatile. Standard deviation of each firm is very large, sometime being much bigger than the mean, indicating that there are significant variation associated with the amount of news flowing into the equity market.

Besides the absolute number of news articles, its individual content might be also relevant for investors. To quantify the individual content, I rely on the recent development in textual analysis to measure the tone of each individual news articles. Due to the limitation of computing power and resources, I only download the first paragraph of each article. Then, I count the number of positive and negative words used in the first paragraph using the words list provided by Loughran and McDonald (2011). The final measure of news tone of each individual article is simply the difference between the percentage of positive and negative words. To ensure that longer articles carry more weight, I value-weight them by the number of words in each article everyday. Table 1.2 shows descriptive statistics of daily news tones for 20 selected firms. News tone measure of 0 thus corresponds to the neutral news tone day that had equal amount of positive and negative wordings. The news tone is negative in average with an exception of IBM and Cisco. Similar to the news counts reported in Table 1.1, the news tones are also highly volatile.

1.2.2 Evidences from Daily Jump Detection

In this section, I show preliminary evidence on the relationship between firm-specific news arrivals and daily return jumps. In order to classify each day as jump day or no jump day, I rely on the non-parametric method developed in Lee and Mykland (2008). Their method first normalizes each return observation by the non-parametric spot variance estimator, and then it is being compared to the specific quantiles provided by the limiting distributions of interest.[8] To be conservative, I consider four different statistics that differ in its significance level and asymptotic distributions. Specifically, $J99$ and $J95$ denote the

[7] Ederington and Lee (1993), Mitchell and Mulherin (1994), and Berry and Howe (1994) show that simple count of number of news articles is a good measure of public information arrival.

[8] I have used the corrections pointed out by Gilder, Shackleton, and Taylor (2014) in deriving the quantiles of the asymptotic distribution.

jumps detected at 99% and 95% significance level using Gumbel distribution as in Lemma 1 of Lee and Mykland (2008), while $J_0 99$ and $J_0 95$ denote the jumps detected at 99% and 95% significance level using more relaxed normal distribution as in Theorem 1 of Lee and Mykland (2008).[9]

Having identified the jumps at daily level for individual securities, I first run the following pooled logit regression to test whether the jumps are more likely to occur on days with more news arrivals. NewsCount and NewsTone denote the daily count and tone of news articles described in the previous section, where I standardize them to have same mean and standard deviation across 20 firms.

$$\text{logit}(p_{it}) = a + b_1 \times \text{NewsCount}_{it} + b_2 \times \text{NewsTone}_{it} + b_3 \times \text{ret}_{i,t-1} + \epsilon_{it} \tag{1.1}$$

Table 1.3 reports the pooled logit regression result for four different threshold of detecting jumps. Same findings hold across all four cases: more news counts are associated with higher probability of having jump while news tone does not have statistically significant relationship to the probability of jump. The result shows that jumps are more likely to occur when excessive amount of information flows, regardless of its actual content.

Next, I further explore the role of news tone given that it does not affect the occurrence of jumps itself. Tetlock (2007) shows that the level of news pessimism extracted from Wall Street Journal is related to the downward pressure on market prices. In other words, the news tone is linked to the size of the market price movements, or jumps. Motivated by this, I run the following regression to test whether the news tone matters for the size of jumps conditional on having jumps.

$$r_{it}|\text{Jump} = a + b_1 \times \text{NewsCount}_{it} + b_2 \times \text{NewsTone}_{it} + \epsilon_{it} \tag{1.2}$$

For each of four jump detection statistics, I first take subsamples classified as jump days only. Then, I assume the entire daily return of those days are due to the jump component. Table 1.4 reports the OLS regression result. First panel provides the estimates of coefficients when entire subsample of jump days are used. Again, no qualitative differences are found among four different jump detection statistics. NewsTone variable shows statistically significant positive coefficient, meaning negative NewsTone comes with negative jump returns which is consistent with Tetlock (2007). Coefficients estimates for NewsCount shows much weaker statistical significance and even has a negative sign. The potential issue with NewsCount variable is that it can take only positive values while the dependent variable, size of jump

[9] Each of four statistics $\{J99, J95, J_0 99, J_0 95\}$ thus identifies the jump day if the absolute value of daily return is above $\{5.1024, 4.4881, 3.2283, 2.4565\}$ times of the daily spot volatility.

return, can be both positive and negative. Thus, I further divide the jump days into two subsamples, one with positive jump returns and other with negative jump returns.

Middle panel reports the result for positive jumps only subsample. First and most interestingly, the significance associated with NewsTone variable disappears and the sign becomes negative. In other words, the actual content of news matters less for positive jump returns and actually it even reduces the size of jumps, given that NewsTone will be in general positive on those days. Second, NewsCount is now strongly related with the size of jumps, carrying a statistically significant positive coefficient.

Bottom panel shows the coefficient estimates on negative jumps only subsample. In this case, all coefficients for both NewsTone and NewsCount are statistically significant at 1% level. The signs of estimated coefficients for NewsTone and NewsCount are positive and negative, respectively, indicating that pessimistic news tone and more number of negative news come with larger negative returns. Looking at the size of the coefficients, NewsTone dominates the NewsCount in its impact on the jump returns. Using the estimated coefficients from the $J99$ statistic in column (1), one standard deviation decrease in NewsTone decreases jump size by 1.47% while one standard deviation increase in NewsCount decreases jump size by 0.66%.

The subsample results are largely consistent with the findings by Chen and Ghysels (2011). Using intra-day returns as the sign of news, they find that moderately good news actually reduces the volatility while bad news and very good news increase volatility. News tone result implies the same conclusion that negative news tone increases volatility via having larger jump sizes while positive news actually reduces, or does not impact, volatility.

1.2.3 Evidences from Implied Volatility

Having established the linkage between news measures and jumps in the previous section, I now move on to find evidences from the options market. Options market reflects investor's risk-adjusted expectation, thus reveals forward-looking information. The most well-known pattern of implied volatility is perhaps around the scheduled earnings announcement (Rogers, Skinner, and Buskirk (2009)). Empirically, implied volatility spikes a day before the earnings announcement, then shows a slight drop on the announcement date, and large drops follow. One way to think about earnings announcement is to classify it as a day with large information arrival, as the number of news counts is excessively high around the earnings announcement date. However, there is a fundamental difference between earnings announcement date and other dates with large news flows. That is, the timing of earnings announcements is known in advance, while other large news flows come at surprise without fixed date in advance.

Figure 1.2 compares the behavior of implied volatility in [-5,+5] days window around scheduled earnings announcement date and unscheduled large news flow dates. The top panel plots the average one-month maturity at-the-money (ATM) implied volatility within 5 days window around the quarterly earnings announcement dates. Similarly, in the bottom panel, I plot the average ATM implied volatilities for four dates each year with largest news count that does not belong to within 5 days of quarterly earnings announcement dates. First, top panel reveals consistent pattern with what was reported in the literature, both peak a day before and a large drop afterwards, around the scheduled earnings announcement dates. On the other hand, bottom panel shows very different pattern. The level of average implied volatility now peaks on the day of large news flow instead of the day before. Also, there is no large drop in average implied volatility afterwards, but it rather persists. With this interesting differences between the impact of scheduled and unscheduled news in mind, I move on to study how the jump risk component in equity options is related to the public news arrival.

To measure perceived jump risk embedded in equity options prices, I choose the steepness of volatility smirk, or implied volatility skew (IV-SKEW), as the measure of investor's risk aversion to expected negative jumps. This measure was explored in Xing, Zhang, and Zhao (2010) where they show the predictability of cross-sectional returns using IV-SKEW. I closely follow their definition of IV-SKEW where I use Black-Scholes delta as the definition of moneyness where Xing, Zhang, and Zhao (2010) uses the ratio of the strike price to the stock price. To construct IV-SKEW measure, I first obtain end-of-day option prices and implied volatility as well as Black-Scholes delta from IVY OptionMetrics database for 20 firms in my sample. The sample period is again from Jan, 2000 to Jul, 2012. Then, for each day with options traded, I choose OTM puts with maturity being closest to 30 days and BS-delta value being closest to -0.25. Similarly, I choose ATM calls by looking for options having maturity closest to 30 days and BS-delta closest to 0.5. The final daily measure of IV-SKEW is computed as the difference between the implied volatilities of average OTM puts and ATM calls selected.

$$\text{IV-SKEW}_{it} = \text{IV}_{it}^{OTMP} - \text{IV}_{it}^{ATMC} \tag{1.3}$$

The famous volatility smirk puzzle basically translates to this measure of IV-SKEW being positive. I also find this in my sample where the average daily IV-SKEW is 3.86%. Larger IV-SKEW reflects larger risk-adjusted expected jump risk for investors. Hence, I next run the simple linear regression of news variables from the previous section on the IV-SKEW. To avoid noise associated with daily measure of IV-SKEW, I average them at monthly level as well as all explanatory variables. Following equation summarizes the regression model used.

$$\text{IV-SKEW}_{it} = a + b_1 \text{NewsCount}_{it} + b_2 \text{NewsTone}_{it} + b_3 \text{ATM IV} + \epsilon_{it} \tag{1.4}$$

Table 1.5 summarizes the result where column (1), (2), and (3) report different combinations while column (4) reports the full model result. The resulting estimates of coefficients are all statistically significant in all models and generally agrees with the results from Table 1.4. For instance, NewsCount is positively related with the IV-SKEW, meaning that more news comes with steeper smirk. Also note that sign for NewsTone variable is negative, indicating pessimistic news content also makes IV smirk steeper. In terms of its magnitude, one standard deviation increase in NewsCount will increase IV-SKEW by 0.17% while one standard deviation decrease in NewsTone will increase IV-SKEW by 0.61%. Given that IV-SKEW embeds both information about the occurrence and size of jumps, where it is hard to disentangle them non-parametrically under the risk-neutral measure, the results are largely consistent with the findings from Table 1.3 and Table 1.4.

Having preliminary evidences above established, I next move on to build a reduced-form model that features findings of this section. The goal of reduced-form study is to identify the prices associated with the inherent news arrival process, which is shown to be related to stock return jumps.

1.3 Reduced-Form Model of News and Jumps

In this section, I build a reduced-form model of stock price process that features the empirical findings of the previous section. Specifically, I embed the news process, to be filtered from the observed information from daily news counts and news tones, in the standard affine stochastic volatility jump-diffusion model. Then, I discuss the risk-neutralization of the process that delivers closed-form option pricing formula. Lastly, I outline the two-step filtering process used to estimate parameters and infer the latent states of the proposed dynamics from the observable data.

1.3.1 Reduced-Form Model of Stock Price Dynamics

I begin by specifying the process governing the log stock variance, spot variance, and news under the physical measure ($\mathbb{P}$). I use S_t, V_t, and I_t to denote stock price, spot variance, and spot news at time t. The following dynamics fully describe the process of three factors under the physical measure $\mathbb{P}$:

$$
\begin{aligned}
d\log(S_t) &= (\mu - \frac{1}{2}V_t - \xi\lambda_t)dt + \sqrt{V_t}(\sqrt{1-\rho^2}dW_t^1 + \rho dW_t^2) + q_t dN_t && (1.5)\\
dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^2 && (1.6)\\
dI_t &= \kappa_I(\theta_I - I_t)dt + \sigma_I\sqrt{I_t}dW_t^3 && (1.7)
\end{aligned}
$$

where μ denotes the return drift of individual equity. For simplicity, I treat μ as a constant and fix it at the sample average of daily returns throughout the paper. All Brownian motions $dW_t^i, i = 1$ to 3 are assumed to be independent to each other.

I assume standard square-root process for the variance and news process, V_t and I_t, as in Heston (1993). The log stock price $\log(S_t)$ also follows standard jump-diffusion process with $q_t dN_t$ representing the compound Poisson distributed jump process with time-varying intensity λ_t. Each individual jump is assumed to be independent and identically distributed normal distribution with mean jump size η and jump standard deviation δ. The jump compensation term ξ is set to be equal to $e^{(\eta+\frac{1}{2}\delta^2)} - 1$ to ensure log stock price is a martingale process.

What is new to the model is the specification of jump intensity, λ_t, dynamics. Standard assumptions made in the literature is to define it as either a constant or an affine function of spot variance V_t.[10] In this paper, I take different approach to use observed firm-specific news flow to anchor the jump intensity in contrast to using latent process V_t. The empirical findings of the previous section ensures the validity of this specification which I re-confirm in the reduced-form estimation later. To keep the model within affine class for the analytical tractability, I impose the following affine functional form of jump intensity:

$$\lambda_t = \gamma_0 + \gamma_1 I_t \tag{1.8}$$

where γ_0 is a constant term that captures the residual of jump-intensity not explained by the news process I_t. This specification belongs to the two-factor affine stochastic volatility jump-diffusion framework and thus the option pricing can be done analytically using the general result of Duffie, Pan, and Singleton (2000).

Note that I only model the dynamics of individual firm's returns, thus abstracting away from the potential factor structure in returns.[11] I do so because the paper focuses on firm-level dynamics and

[10] For example, see Pan (2002) and Bates (2006).

[11] The factor structure and pricing of idiosyncratic risk in equity options markets have been started to gain attention only recently. See Christoffersen, Fournier, and Jacobs (2015), Gourier (2016), and Bégin, Dorion, and Gauthier (2016) for the recent development in this subject.

risk premium only, instead of at portfolio level. Thus, the loss by not considering the potential factor structure is rather minimal while the gain from analytical tractability is huge.

1.3.2 Risk Neutralization

The model has three sources of diffusive risk represented by Brownian increments and one source of jump risk. I impose linear form of price of risk for three diffusive Brownian motions to preserve same square-root functional form under the risk-neutral measure as in Heston (1993). As discussed in Pan (2002), the pricing kernel for jump risk under the incomplete market can take virtually any arbitrary form by allowing it to change its entire distribution.

In this paper, I only consider two sources of jump risk premium, namely jump-timing and jump-size premium. Because the jump intensity λ_t is driven by the news process that is independent of the diffusive variance process V_t, the risk premium imposed on the Brownian motion dW_t^3, denoted by λ_I, effectively controls the jump-timing premium by allowing risk-neutral jump-intensity to differ from its physical counterpart. Lastly, jump-size premium is introduced by simply shifting the mean of normally distributed individual jumps by the amount of $\eta^{\mathbb{Q}} - \eta$. Below summarizes the change of measure where λ_V and λ_I denote the diffusive risk premium placed on variance and news, respectively.

$$
\begin{aligned}
dW_t^{2,\mathbb{Q}} &= dW_t^2 + \lambda_V \sigma \sqrt{V_t} dt && (1.9)\\
dW_t^{3,\mathbb{Q}} &= dW_t^3 + \lambda_I \sigma_I \sqrt{I_t} dt && (1.10)\\
\eta^{\mathbb{Q}} &= \eta + (\eta^{\mathbb{Q}} - \eta) && (1.11)
\end{aligned}
$$

I do not specify the risk premium associated with the Brownian motion dW_t^1 associated with log stock price as it has to be fixed to have risk-neutral drift equal to the risk-free rate r. Under this change of measure, the risk-neutral dynamics preserves the following same functional form:

$$
\begin{aligned}
d\log(S_t) &= (r - \frac{1}{2}V_t - \xi^{\mathbb{Q}}\lambda_t)dt + \sqrt{V_t}(\sqrt{1-\rho^2}dW_t^{1,\mathbb{Q}} + \rho dW_t^{2,\mathbb{Q}}) + q_t^{\mathbb{Q}} dN_t && (1.12)\\
dV_t &= \kappa^*(\theta^* - V_t)dt + \sigma\sqrt{V_t}dW_t^{2,\mathbb{Q}} && (1.13)\\
dI_t &= \kappa_I^*(\theta_I^* - I_t)dt + \sigma_I\sqrt{I_t}dW_t^{3,\mathbb{Q}} && (1.14)
\end{aligned}
$$

where the mapping between physical and risk-neutral parameters are given by:

$$\kappa^* = \kappa + \lambda_V \sigma \tag{1.15}$$

$$\kappa_I^* = \kappa_I + \lambda_I \sigma_I \tag{1.16}$$

$$\theta^* = \frac{\kappa\theta}{\kappa^*} \tag{1.17}$$

$$\theta_I^* = \frac{\kappa_I \theta_I}{\kappa_I^*} \tag{1.18}$$

1.3.3 Filtering and Estimation

As all latent state continuous models do, my model also needs to jointly estimate the parameters and filter the latent states. Given that my focus is on identifying the risk premiums associated with news process, I follow the approach from Christoffersen, Heston, and Jacobs (2013) and perform a sequential estimation.[12] Specifically, the estimation procedure is divided into two steps. The first step identifies all parameters and spot states under the physical measure only using the daily returns and observed news data. Then, I take the physical parameters and states as given in the second step and only estimates the risk premium parameters using equity options data. Pros of this approach is that I can avoid the difficulty of weighting the likelihood between physical and risk-neutral counterparts. Meanwhile, the obvious cons of this approach is that it is not statistically efficient as the joint estimation procedure. Since my focus is placed heavily on the qualitative outcome of resulting pricing kernel estimates rather than exactly quantifying the risks, I argue that sequential estimation procedure is better-suited for my model.

Estimated under the Physical Measure

I first define the state-space system by discretizing $\mathbb{P}$-measure equations (1.5), (1.6), and (1.7) using Euler scheme at daily interval. The discretized state-space equations are written as below:

$$r_{t+1} = (\mu - \frac{1}{2}V_t - \xi\lambda_t)\Delta t + \sqrt{\Delta t V_t}(\sqrt{1-\rho^2}\epsilon_{t+1}^1 + \rho\epsilon_{t+1}^2) + \sum_{j=0}^{N_{t+1}} y_{j,t+1} \tag{1.19}$$

$$V_{t+1} = V_t + \kappa(\theta - V_t)\Delta t + \sigma\sqrt{\Delta t V_t}\epsilon_{t+1}^2 \tag{1.20}$$

$$I_{t+1} = I_t + \kappa_I(\theta_I - I_t)\Delta t + \sigma_I\sqrt{\Delta t I_t}\epsilon_{t+1}^3 \tag{1.21}$$

$$N_{t+1} \sim \text{Poisson}(\gamma_0 + \gamma_1 I_t) \tag{1.22}$$

[12] Christoffersen, Fournier, and Jacobs (2015) and Andersen, Fusari, and Todorov (2015b) take the opposite approach by starting from the risk-neutral measure and sequentially estimate risk premium parameters by matching it to the physical measure.

where the innovation terms ϵ^i_{t+1} for $i = 1$ to 3 are i.i.d. standard normal random variables, the counting process N_{t+1} denotes the number of jumps between time t and $t+1 = t+\Delta t$, and individual jump terms $y_{j,t+1}$ are i.i.d. normally distributed random variables with mean η and standard deviation δ. I set the daily time interval to Δt to be 1/252 so that all parameters are expressed in annual terms.

Under the physical probability measure, I have three observables, namely daily returns, news count, and news tone. In order to simplify the filtering procedure while maintaining the empirical findings from the previous section, I first construct the tone-adjusted news count measure as follows:

$$\tilde{I}_t = \text{NewsCount} \times \exp(-\text{NewsTone}) \tag{1.23}$$

The intuition behind this measure is as follows. It was shown that negative news tones emphasize the size of jumps where positive news tones reduces the size of jumps (although statistically less significant) in Table 1.4. Since the size of individual jumps is fixed to be constant, η, in the model, I effectively embed the effect of news tone on the jump size into the news count measure by the above adjustment. The negative news tone thus results in higher tone-adjusted news count $\tilde{I}_t$ as $\exp(-\text{NewsTone})$ is greater than 1 when NewsTone is negative, and positive news tone will lower the tone-adjusted news count in the same fashion.

After the adjustment, I end up with two observables under the physical probability measure, daily log-returns r_{t+1} and tone-adjusted news count $\tilde{I}_t$. They are linked to the state equation by the simple measurement relationship that the daily log-returns are observed without an error and the news process I_t is observed with normally distributed measurement error ($\tilde{I}_t = I_t + \epsilon^m_t$). Then, I estimate the physical parameters and filter the latent states at the same time by maximizing the likelihood of observing daily log-returns and tone-adjusted news count via Particle Filtering (PF) algorithm.

Pricing Kernel Parameter Estimation

Given the estimated physical parameters and latent states from the previous section, I next estimate parameters associated with pricing kernel where I treat all else being fixed. The end-of-day options prices for 20 firms in the sample are obtained from OptionMetrics database. I follow literature and pick only Wednesday prices in order to avoid potential issues using daily data.[13] Commonly-used option data filters, such as strictly positive volume and in-violation of put-call parity, are applied to raw data. For each day, I pick options with maturity between 15 and 250 calendar days to ensure only liquid options

[13] I use the previous business day if Wednesday turns out to be holiday. See Dumas, Fleming, and Whaley (1998) for more detailed description of advantage using Wednesday options data.

are considered. Lastly, I pick six strike prices with the highest trading volume for each fixed maturity every Wednesday.

Since I am dealing with the individual equity options that are American, I follow Broadie, Chernov, and Johannes (2007) and convert them into the corresponding European options prices.[14] All put options are converted into corresponding call options via put-call parity for the ease of implementation later. This leaves us with a total of 191,625 options for 20 individual firms.

The pricing kernel equations defined in (1.9), (1.10), and (1.11) mean that there are only three extra parameters to be estimated once physical parameters and states are fixed. The estimation is performed by minimizing vega-weighted root mean squared error (VWRMSE) proposed by Trolle and Schwartz (2009). It is based on simplifying assumption that vega-weighted option errors are i.i.d. normally distributed. Thus, I estimate three parameters $\Theta^{\mathbb{Q}} = \{\lambda_V, \lambda_I, \eta^{\mathbb{Q}} - \eta\}$ by minimizing the following VWRMSE-based likelihood:

$$\tilde{\Theta}^{\mathbb{Q}} = \arg\min_{\Theta} -\frac{1}{2}\sum_{i=1}^{N}[\log(VWRMSE^2) + e_i^2/VWRMSE^2] \tag{1.24}$$

$$\text{VWRMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} e_j^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}((C_i^{Mkt} - C_i^{Mod}(\Theta))/BSV_i^{Mkt})^2} \tag{1.25}$$

where C_i^{Mkt}, C_i^{Mod}, and BSV_iMkt denote market price of call option, model-implied price of call option, and market-implied Black-Scholes Vega, respectively.

The option pricing formula is available in closed-form up to the Fourier transform, as the model falls into the class of affine stochastic-volatility jump-diffusion model. The following proposition summarizes the characteristic function of the log-spot stock price under the physical measure. Since the model preserves identical functional form under the risk-neutral measure, the same formula is applied with the appropriate parameter mappings.

Proposition 1 *Denote the risk-neutral characteristic function of log-spot price by* $E_t[\exp(iu\log(S_{t+\tau})])] = S_t^{iu} f(u, \tau, V_t, I_t)$. *Then function f is given by*

$$f(u, \tau, V_t, L_t) = \exp(A(u, \tau) + B_1(u, \tau)V_t + B_2(u, \tau)I_t)$$

A, B_1, *and* B_2 *are given as the solution to the following Ricatti ODE with the initial conditions* $A(0) =$

[14] OptionMetrics provides implied volatility computed using CRR binomial-tree model, zero-rates, and ex-post divided rates that are sufficient for this conversion.

$B_1(0) = B_2(0) = 0$.

$$\frac{dA}{d\tau} = (r - \xi\gamma_0)iu + \gamma_0\theta_u + \kappa\theta B_1 + \kappa_I\theta_I B_2 \tag{1.26}$$

$$\frac{dB_1}{d\tau} = -\frac{1}{2}u(i+u) - (\kappa - \rho\sigma iu)B_1 + \frac{1}{2}\sigma^2 B_1^2 \tag{1.27}$$

$$\frac{dB_2}{d\tau} = \gamma_1\theta_u - \gamma_1\xi iu - \kappa_I B_2 + \frac{1}{2}\sigma_I^2 B_2^2 \tag{1.28}$$

where $\theta_u = \exp(\theta iu - \frac{1}{2}\delta^2 u^2) - 1$. *All three ODEs have closed-form analytical solution similar to the Heston (1993)'s expression.*

Proof. *Direct application of Duffie, Pan, and Singleton (2000) result.* ■

Once the characteristic function is available in the closed-form, European call options can be valued using the following formula following Heston (1993).

$$C_t = S_t P_1 - Ke^{-r\tau}P_2 \tag{1.29}$$

where the P_1 and P_2 probabilities are computed using Fourier inversion:

$$P_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re[\frac{e^{iu\log(\frac{S_t}{K})}f(u+1,\tau,V_t,I_t)}{iuS_t e^{r\tau}}]du \tag{1.30}$$

$$P_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re[\frac{e^{iu\log(\frac{S_t}{K})}f(u,\tau,V_t,I_t)}{iu}]du \tag{1.31}$$

The integrands in the above expression vanish quickly and can be computed effectively using a numerical integration scheme such as quadrature.

1.4 Estimation Results

1.4.1 Physical Parameter Estimates

Table 1.6 reports parameter estimates for 20 firms in the sample. For brevity, I omit the parameter estimates associated with the information process I_t.

The speed of mean-reversion parameter κ for the diffusive variance V_t has average of 1.90 in average. Cisco has the slowest mean-reversion speed having estimated κ equal to 0.10 while Pfizer mean-reverts the fastest with κ being equal to 3.99. The magnitudes are in general consistent with those reported in the prior literature including Dubinsky and Johannes (2006) and Christoffersen, Fournier, and Jacobs

(2015). The long-run mean level of diffusive variance θ has average estimate of 0.084, or 28.98% of annual volatility. Overall, estimates for diffusive variance process are mostly consistent with previous studies.

The estimates of individual jump-distribution parameters are summarized by its mean η and standard deviation σ. Average estimate of η is 0.3% where it varies from -5.2% of Merck & Co. to 8.1% of Amazon. In average, it can be seen as an evidence that positive and negative jumps are equally likely for individual equity returns. Note that my model does not feature separate positive and negative sized jumps, hence the estimated average jump-size is close to 0. Along the same intuition, the expected standard deviation of individual jumps must be large. This is indeed the case, the average estimated δ is 8.2%, enormously larger than the mean.

The parameters of focus in this paper are γ_1 that measures the relationship between the news process I_t and jump-intensity λ_t. Estimated parameter γ_1 is positive in all 20 firms which is consistent with the previous non-parametric findings that more information comes with higher probability of jumps. In terms of magnitude, the average γ_1 is 0.072. Given that average I_t in the entire sample is 47.45, this roughly translates to 3.4 jumps per year explained by the news process I_t. The average total number of jumps per year is then given by γ_0, which is 1.11 in average, plus news induced jumps. Thus, news process carries the first-order importance in explaining jumps in which roughly 75% of the time-varying jump-intensity is captured by the news.

Overall, physical parameter estimates emphasize the benefit of having news process, which is filtered from observable public news arrival data, in capturing time-varying jump-intensity just using physical observables. Having established the estimates and states, I next discuss the pricing kernel estimates, which is the central findings of this paper.

1.4.2 Pricing Kernel Parameter Estimates

Table 1.7 reports estimates of three pricing kernel parameters defined in equation (1.9), (1.10), and (1.11) for 20 firms in the sample. Three parameters λ_V, λ_I, and $\eta^{\mathbb{Q}} - \eta$ each represents the diffusive variance risk premium, news risk premium, and jump-size risk premium.

The variance risk premium (VRP) is arguably one of the most actively researched topic in recent finance literature. The significantly negative variance risk premium, often measured by the difference between physical realized volatility and risk-neutral volatility such as VIX, is found in index options market. However, relatively little is known about the VRP at individual firm levels. Existing studies such as Carr and Wu (2009) and Drissen, Maenhout, and Vilkov (2009) have found much smaller amount

of VRP in individual firm levels that are sometimes indistinguishable from being zero. In this paper, rather than trying to pin down the exact mechanisms behind why individual variance risk premiums are smaller, I focus on extracting risk premium components involving news process and studies its further implications.

The estimated diffusive variance risk premium parameter λ_V is in mostly negative with an average value of -0.119. This value is much smaller than what was estimated for index options market in the prior literature.[15] It is also consistent with the prior non-parametric findings documenting much smaller magnitude of variance risk premium in individual equity options. For instance, Carr and Wu (2009) finds only small portion of their sample firms exhibit statistically significant variance risk premium estimate. Drissen, Maenhout, and Vilkov (2009) also finds that large portion of their sample cannot reject the null of zero variance risk premium. Overall, the estimated diffusive variance risk premium is consistent with the prior findings.

Recall that there are two distinct risk premiums associated with the jump component, namely jump-timing risk premium λ_I and jump-size premium $\eta^{\mathbb{Q}} - \eta$. The unique feature of my model is that estimated pricing kernel jointly identifies these two parameters. The estimated jump-timing risk premium λ_I is mostly positive with a single exception of Wal Mart, averaging to the value of 1.951 across 20 firms. What this means is that instead of having a negative risk premium on the jump-timing, there is a large positive risk premium associated with the jump-timing. In other words, risk-neutral world has higher probability weight on the state of the world with smaller number of jumps. This is highly counter-intuitive, because risk-averse investors do not like jumps. Instead, estimated result implies that risk-averse investors favor having more jumps.

This puzzling finding was acknowledged in the prior literature that studied index options. For instance, in her seminal paper, Pan (2002) (Section 5.2) found that jump-intensity estimates become smaller when it was allowed to vary. Aït-Sahalia, Karaman, and Mancini (2015) (Section 5.3), using OTC variance swap data, has also found this positive jump-timing premium and concluded it as an evidence of limited ability of estimating flexible change of measure. My result, although estimated using individual equity options, is consistent with their findings. In particular, I used observed news process as an exogenous identifier of jump-intensity in order to circumvent the problem of limited ability in estimating general pricing kernel.

In order to explain the positive jump-timing premium, I rely on two arguments. First, my pricing kernel jointly identifies the jump-timing and jump-size premium. Looking at the estimated jump-size

[15] For example, Christoffersen, Fournier, and Jacobs (2015) reports estimated λ_V to be -1.48 without jumps in the index returns process.

risk premium parameter $\eta^{\mathbb{Q}} - \eta$, it is found to be largely negative with a single exception of Cisco. The average jump-size premium is very large being -5.4% where the average jump-size under the physical measure was found to be only 0.3% in Table 1.6. Thus, aggregated jump-risk premium still remains negative once both timing and size premiums are considered. Therefore, I interpret the result as a decomposition of individual equity jump risk premium, rather than an evidence of positive jump risk premium. Second, recall that the source of jump-timing, or jump-intensity, in my model solely comes from the news process I_t. Thus, resulting estimates of the jump premiums have a direct interpretation in terms of how investors view news uncertainty. The fact that jump-timing being positive implies that investors prefers the state of the world with more news arrivals. On the other hand, the negative jump-size premium implies that investors are really afraid of having large negative jumps in returns due to the news arrival. Putting these together, I conclude that news is viewed as preferable to investors once the negative impact is accounted for.

To engage economic interpretation of the findings, I consider two channels for public news to cause jumps in equity returns. The first channel argues that sudden arrival of massive amount of public information triggers the rapid increase in the noise or liquidity trading activity via distorting their belief. Other possible channel is that the arrival of public information comes together with the resolution of information asymmetry, thus resulting in sudden movements in the equity price. Both channels have same implications that public news arrival is related with return jumps, but have opposite interpretation in terms of risk premium. If the public information merely serves as a channel to increase noise trading activity, it should be negatively priced as it only increases potential jump risk faced by investors. On the other hand, if it indeed resolves the information asymmetry between privately informed and uninformed investors, it should be positively priced. Indeed, Easley and O'hara (2004)'s noisy rational expectations equilibrium model implies that firms facing higher information asymmetry requires higher return. Empirically, Zhao (Forthcoming) shows this is true by measuring firm's information intensity by its form 8-K filing frequency.

My findings are consistent with both theories. The somewhat puzzling positive jump-timing risk premium associated with the positive estimates of λ_I can be explained by the resolution of information asymmetry story. Investors seek to have more public information, although it can cause prices to jump, because it resolves the potential information asymmetry. On the other hand, they do not like public information to increase the noise trading activity and cause returns to jump, especially negatively signed jumps, thus placing more subjective probability weight on the state of the world with large negative jumps in returns. Therefore, the estimated jump-size risk premium parameter $\eta^{\mathbb{Q}} - \eta$ is largely negative. In recent article, Han, Tang, and Yang (2016) theoretically studies conflicting role of public information

and shows that public information improves market liquidity but at the same time can harm price efficiency. Thus, their result is perhaps most closely related to my results that document both positive and negative implications of public information arrival.

Overall, the findings of this section highlights the importance of separating the precise source of jump risk premium. In future research, it would be interesting to study the quantitative implications of the estimated parameters in the context of noisy rational expectations equilibrium model featuring both positive and negative effect of public news arrival.

1.5 Conclusion

I first study the role of firm-specific public news arrival on jumps in equity returns. Using comprehensive news data from Factiva database, I find news to be strongly related with jumps in both physical and risk-neutral measures. I then estimate a continuous-time model with stochastic volatility and news driven jump-intensity. In particular, the variation in probability of jump is driven by the observable news process instead of latent state variables. The model is estimated in sequential fashion to ensure the identification of risk premium parameters associated with news. Resulting estimates reveal an important finding: jump-timing risk is positively priced while jump-size risk is significantly negatively priced. I interpret this result as investor's preference for having more public news arrivals while disliking the potential large negative returns news can induce. Thus, public news is not redundant and it carries significant risk premium. The question of exact source of news arrival risk premium is left unresolved and is left as a venue for future research.

Table 1.1: **Summary Statistics of News Counts. 2000-2012**

	Summary Statistics			
Company Name	Total	Mean	Median	Std. Dev.
Microsoft	325,150	71.0	92	53.4
GE	265,302	57.9	72	82.9
IBM	165,213	36.3	45	31.2
Chevron	106,356	23.7	29	21.8
UTC	53,809	12.6	14	13.1
Pfizer	111,363	24.7	29	26.5
Johnson & Johnson	103,271	23.3	26	26.5
Merck & Co.	51,170	12.3	12	26.2
Disney	160,245	34.9	41	15.5
JP Morgan	232,971	51.2	63	48.2
WalMart	165,220	36.0	43	100.4
American Express	54,216	12.3	13	83.6
Intel	171,146	37.8	41	17.5
Bank of America	202,898	44.9	45	86.0
Verizon	159,291	35.8	42	113.6
AT&T	139,631	31.0	37	99.3
Cisco	113,818	25.4	29	78.4
Yahoo	85,902	19.6	20	30.9
Amazon	60,519	13.6	14	17.5
Ebay	73,672	16.5	19	18.3
Total	3,303,317	44.3	32	67.9

This table reports summary statistics of daily news counts downloaded from the Factiva database. The first column reports the total number of news articles for each firm during the sample period. The last three columns report the daily mean, median, and standard deviation of news counts for each firm. The sample period is from January 2000 to July 2012.

Table 1.2: **Summary Statistics of Daily News Tones. 2000-2012**

	Summary Statistics				
Company Name	Mean	Std. Dev.	25 Prctile	Median	75 Prctile
Microsoft	-5.33	13.67	-11.79	-3.06	3.35
GE	-2.32	11.32	-7.78	-1.72	3.39
IBM	4.32	12.98	-2.18	5.01	11.68
Chevron	-7.21	17.58	-15.68	-6.96	0.23
UTC	-2.93	18.91	-11.08	-0.78	6.28
Pfizer	-6.96	16.99	-15.27	-5.90	1.74
Johnson & Johnson	-3.32	16.26	-10.03	-1.61	4.36
Merck & Co.	-4.53	23.91	-12.75	-0.44	5.47
Disney	-1.46	13.02	-7.84	-0.49	5.65
JP Morgan	-9.39	12.79	-16.11	-7.40	-1.39
WalMart	-10.30	13.61	-17.31	-8.99	-1.84
Americal Express	-2.19	21.48	-12.42	-0.37	8.20
Intel	-0.16	15.08	-7.48	1.10	8.58
Bank of America	-10.67	15.07	-17.78	-9.51	-1.71
Verizon	-6.16	15.57	-13.15	-4.70	2.33
AT&T	-5.64	14.88	-12.73	-4.39	2.59
Cisco	3.71	16.29	-5.01	4.34	12.78
Yahoo	-2.52	18.21	-11.48	-1.28	7.77
Amazon	-0.76	19.65	-9.93	0.00	9.68
Ebay	-4.85	20.06	-13.53	-3.27	4.88
Total	-3.92	17.13	-12.05	-2.67	4.93

This table reports the summary statistics of daily news tones (in percentage). The daily news tone variable is constructed by analyzing the first paragraph of each news article. I search for the percentage of positive and negative words using the list from Loughran and McDonald (2011). Then, tones from each individual articles are aggregated to the daily level using the total number of words in each article as a weight. The sample period is from January 2000 to July 2012.

Table 1.3: **Effect of News Counts on the Probability of a Daily Jump. 2000-2012**

	(1) $J99$	(2) $J95$	(3) J_099	(4) J_095
Intercept	-4.9743^{***}	-4.5062^{***}	-3.3069^{***}	-2.4112^{***}
	(0.0847)	(0.0596)	(0.0282)	(0.0131)
NewsCount	0.1944^{***}	0.1885^{***}	0.1747^{***}	0.1579^{***}
	(0.0452)	(0.0443)	(0.0405)	(0.0328)
NewsTone	-0.0710	-0.0693	-0.0062	-0.0061
	(0.0495)	(0.0427)	(0.0233)	(0.0164)
Ret_{t-1}	-1.8709	-2.6763	-2.7729^{**}	-3.6652^{***}
	(2.6206)	(1.8715)	(1.1521)	(0.8004)

This table reports the coefficients from the pooled logit regression of daily news count, news tone, and lagged return on the daily jump indicator defined using Lee and Mykland (2008). The explanatory variables are the total number of news reported on Factiva database each day and its news tone, standardized to have same mean and standard deviation across firms, and lagged daily returns. News tone measure is constructed first at each individual article level by counting the number of positive and negative words from Loughran and McDonald (2011), they are then aggregated by a value-weighting scheme using the total number of words in the article. The daily return jump indicator is identified using 4 different statistics. $J99$ and $J95$ indicators use Lee and Mykland (2008)'s Lemma 1 statistic at 99% and 95% significance, respectively. I use the correction term from Gilder, Shackleton, and Taylor (2014). The J_099 and J_095 indicators use looser bound from the normal distribution as in Theorem 1 of Lee and Mykland (2008). Each of four statistics $\{J99, J95, J_099, J_095\}$ thus identifies the jump day if the absolute value of daily return is above $\{5.1024, 4.4881, 3.2283, 2.4565\}$ times of the daily spot volatility. The sample period is from January 2000 to July 2012. Statistical significance levels of 1%, 5%, and 10% are indicated with ***, **, and *, respectively. Standard errors clustered at individual firm levels are reported in parentheses.

Table 1.4: **Effect of News Counts and Tones on Daily Jump Size. 2000-2012**

	(1) $J99$	(2) $J95$	(3) J_099	(4) J_095
Intercept	0.0148***	0.0117***	0.0100***	0.0063***
	(0.0047)	(0.0034)	(0.0014)	(0.0009)
NewsCount	-0.0006	-0.0007	-0.0008	-0.0008
	(0.0021)	(0.0017)	(0.0010)	(0.1074)
NewsTone	1.8468***	1.4412***	0.8202***	0.5045***
	(0.2355)	(0.1721)	(0.0786)	(0.0432)
N	452	713	2267	5243
R^2	12.65%	9.33%	4.72%	2.61%
Positive Jumps Only	(1) $J99$	(2) $J95$	(3) J_099	(4) J_095
Intercept	0.0609***	0.0583***	0.0485***	0.0414***
	(0.0044)	(0.0032)	(0.0013)	(0.0007)
NewsCount	0.0168***	0.0127***	0.0086***	0.0053***
	(0.0025)	(0.0019)	(0.0011)	(0.0006)
NewsTone	-0.2758	-0.1853	-0.0793	−0.0899**
	(0.2418)	(0.1823)	(0.0794)	(0.0429)
N	241	374	1256	2816
R^2	16.96%	10.42%	4.75%	2.57%
Negative Jumps Only	(1) $J99$	(2) $J95$	(3) J_099	(4) J_095
Intercept	−0.0553***	−0.0514***	−0.0429***	−0.0371***
	(0.0037)	(0.0026)	(0.0012)	(0.0007)
NewsCount	−0.0066***	−0.0073***	−0.0064***	−0.0046***
	(0.0013)	(0.0011)	(0.0007)	(0.0004)
NewsTone	0.8593***	0.7561***	0.5142***	0.3807***
	(0.1736)	(0.1264)	(0.0607)	(0.0342)
N	211	339	1011	2427
R^2	25.85%	24.20%	15.55%	9.75%

This table reports the coefficients from the linear regression of daily news counts and news tones on the daily jump size. I assume the entire daily return is due to the jump component on the jump days detected using Lee and Mykland (2008). NewsCount measures the absolute number of news articles appeared in the Factiva database per each day. The NewsTone measure is constructed first at each individual article level by counting the number of positive and negative words from Loughran and McDonald (2011), then they are aggregated by a value-weighting scheme using total number of words in the article. The daily return jump indicator is identified using 4 different statistics. $J99$ and $J95$ indicator uses Lee and Mykland (2008)'s Lemma 1 statistic at 99% and 95% significance, respectively. We use the correction term from Gilder, Shackleton, and Taylor (2014). J_099 and J_095 indicator uses looser bound from the normal distribution as in Theorem 1 of Lee and Mykland (2008). Each of the four statistics $\{J99, J95, J_099, J_095\}$ thus identifies the jump day if the absolute value of daily return is above $\{5.1024, 4.4881, 3.2283, 2.4565\}$ times the daily spot volatility. The sample period is from January 2000 to July 2012. Statistical significance levels of 1%, 5%, and 10% are indicated with ***, **, and *, respectively. Standard errors are reported in parentheses.

Table 1.5: **Effect of News Counts and Tones on IV-SKEW. 2000-2012**

	(1)	(2)	(3)	(4)
Intercept	-0.0065^{***}	0.0070^{***}	-0.0067^{***}	-0.0073^{***}
	(0.0011)	(0.0011)	(0.0011)	(0.0011)
NewsCount			0.0018^{**}	0.0017^{**}
			(0.0008)	(0.0008)
NewsTone		-0.3563^{***}		-0.3547^{***}
		(0.0548)		(0.0548)
ATM IV	0.2035^{***}	0.1996^{***}	0.2047^{***}	0.2008^{***}
	(0.0044)	(0.0044)	(0.0044)	(0.0045)
N	3020	3020	3020	3020
R^2	41.29%	42.10%	41.38%	42.18%

This table reports the coefficients from the linear regression of monthly average news counts and news tones on the monthly average IV-SKEW. IV-SKEW is defined as the difference between the implied volatility of the call option having Black-Scholes delta closest to 0.5 and put option having delta closest to -0.25. Both options are chosen to have maturity as close as possible to 30 days. NewsCount measures the absolute number of news articles that appeared in Factiva database during each month. The NewsTone measure is constructed first at each individual article level by counting the number of positive and negative words from Loughran and McDonald (2011), then they are aggregated by value-weighting scheme using total number of words in the article. Statistical significance levels of 1%, 5%, and 10% are indicated with ***, **, and *, respectively. Standard errors are reported in parentheses.

Table 1.6: **Model Parameter Estimates under the Physical Measure**

	Estimated Parameters							
Company Name	κ	θ	σ	η	δ	γ_0	γ_1	ρ
Microsoft	2.46	0.045	0.53	0.015	0.105	0.42	0.035	-0.43
GE	1.36	0.074	0.45	0.004	0.075	0.89	0.036	-0.49
IBM	2.56	0.057	0.56	0.012	0.066	1.00	0.037	-0.50
Chevron	2.51	0.074	0.31	0.005	0.074	1.00	0.012	-0.50
UTC	2.60	0.091	0.31	0.007	0.050	0.95	0.091	-0.49
Pfizer	3.99	0.049	0.44	-0.011	0.065	1.07	0.064	-0.45
Johnson & Johnson	3.20	0.032	0.45	0.004	0.046	1.14	0.130	-0.42
Merck & Co.	0.45	0.049	0.53	-0.052	0.087	1.02	0.173	-0.45
Disney	0.60	0.058	0.48	0.014	0.053	1.15	0.056	-0.60
JP Morgan	1.25	0.274	0.90	-0.002	0.086	0.70	0.047	-0.38
Wal Mart	2.53	0.050	0.45	0.015	0.043	1.18	0.043	-0.38
Americal Express	1.40	0.114	0.60	0.003	0.085	0.96	0.178	-0.30
Intel	1.61	0.138	0.48	-0.039	0.084	0.96	0.058	-0.35
Bank of America	0.64	0.048	0.55	-0.010	0.147	0.51	0.024	-0.50
Verizon	1.87	0.047	0.49	0.002	0.051	1.25	0.049	-0.50
AT&T	1.69	0.075	0.43	0.006	0.067	1.01	0.054	-0.48
Cisco	0.10	0.161	0.56	0.002	0.082	1.03	0.056	-0.49
Yahoo	1.93	0.064	0.64	-0.003	0.184	3.56	0.007	-0.26
Amazon	1.97	0.130	0.65	0.081	0.165	1.00	0.214	-0.50
Ebay	3.19	0.050	0.52	-0.001	0.032	1.49	0.073	-0.47
Average	1.90	0.084	0.52	0.003	0.082	1.11	0.072	-0.45

This table reports the estimated model parameters under the physical measure using daily returns and news counts from Jan, 2000 to Jul, 2012 for 20 individual equities. A Particle Filtering (PF) algorithm was used to estimate the parameters by maximizing the likelihood of observing daily returns and tone-adjusted news counts. The tone-adjusted news counts is defined as below:

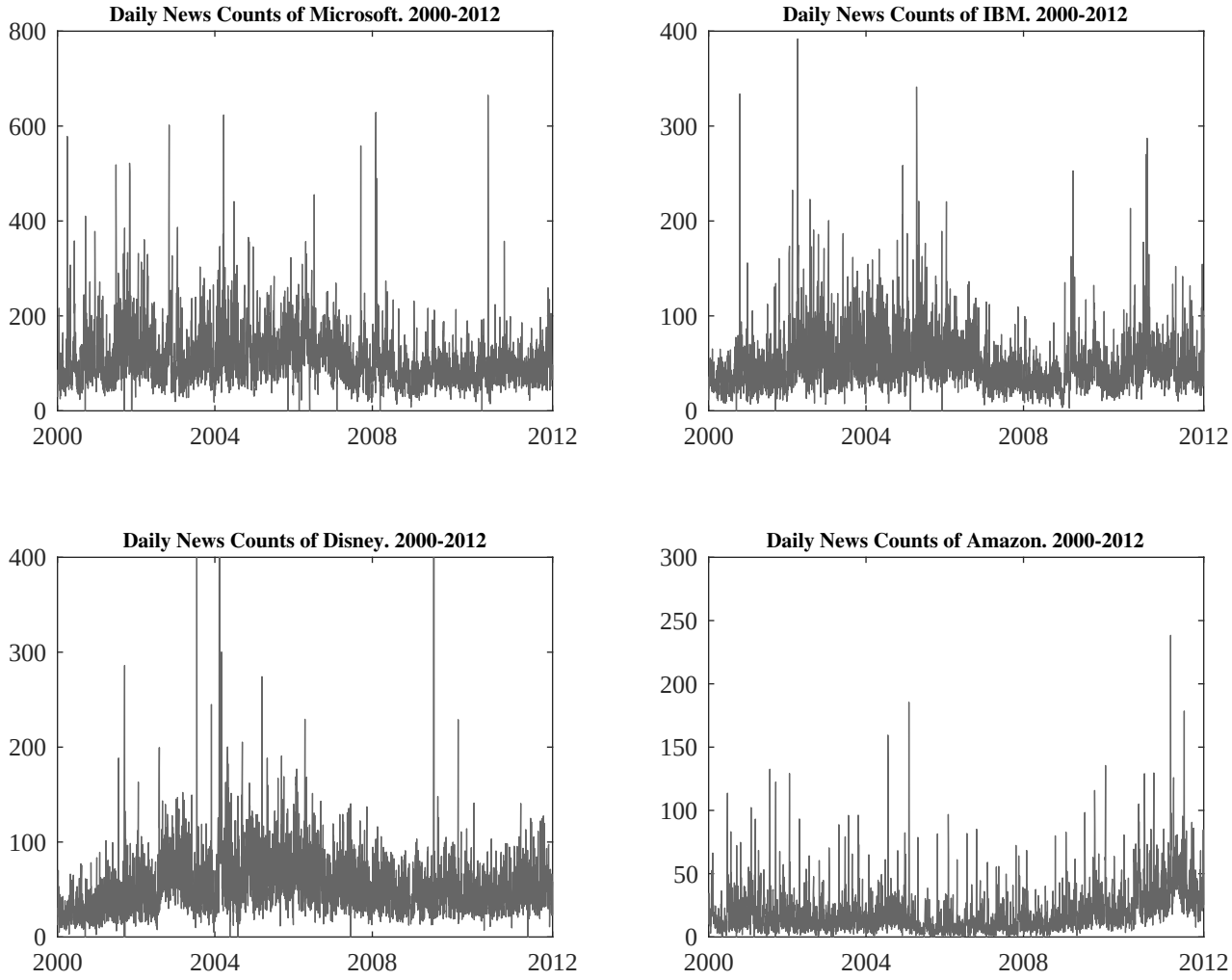
$$\tilde{I}_t = \text{NewsCount} \times \exp(-\text{NewsTone})$$

Table 1.7: **Pricing Kernel Parameter Estimates**

	Estimated Parameters		
Company Name	λ_V	λ_I	$\eta^{\mathbb{Q}} - \eta$
Microsoft	-0.888	1.629	-0.088
GE	-0.623	1.626	-0.074
IBM	-0.509	0.167	-0.051
Chevron	-0.353	1.799	-0.139
UTC	-1.842	1.008	-0.028
Pfizer	0.060	0.033	-0.059
Johnson & Johnson	0.007	0.007	-0.035
Merck & Co.	0.000	0.000	-0.009
Disney	0.002	1.398	-0.032
JP Morgan	0.497	1.027	-0.021
Wal Mart	-0.042	-0.006	-0.026
Americal Express	-0.384	2.647	-0.099
Intel	0.448	6.528	-0.079
Bank of America	0.202	0.787	-0.057
Verizon	-0.191	0.358	-0.041
AT&T	-0.073	0.096	-0.033
Cisco	0.087	0.064	0.007
Yahoo	1.282	18.222	-0.025
Amazon	-0.015	1.609	-0.147
Ebay	-0.039	0.017	-0.041
Average	-0.119	1.951	-0.054

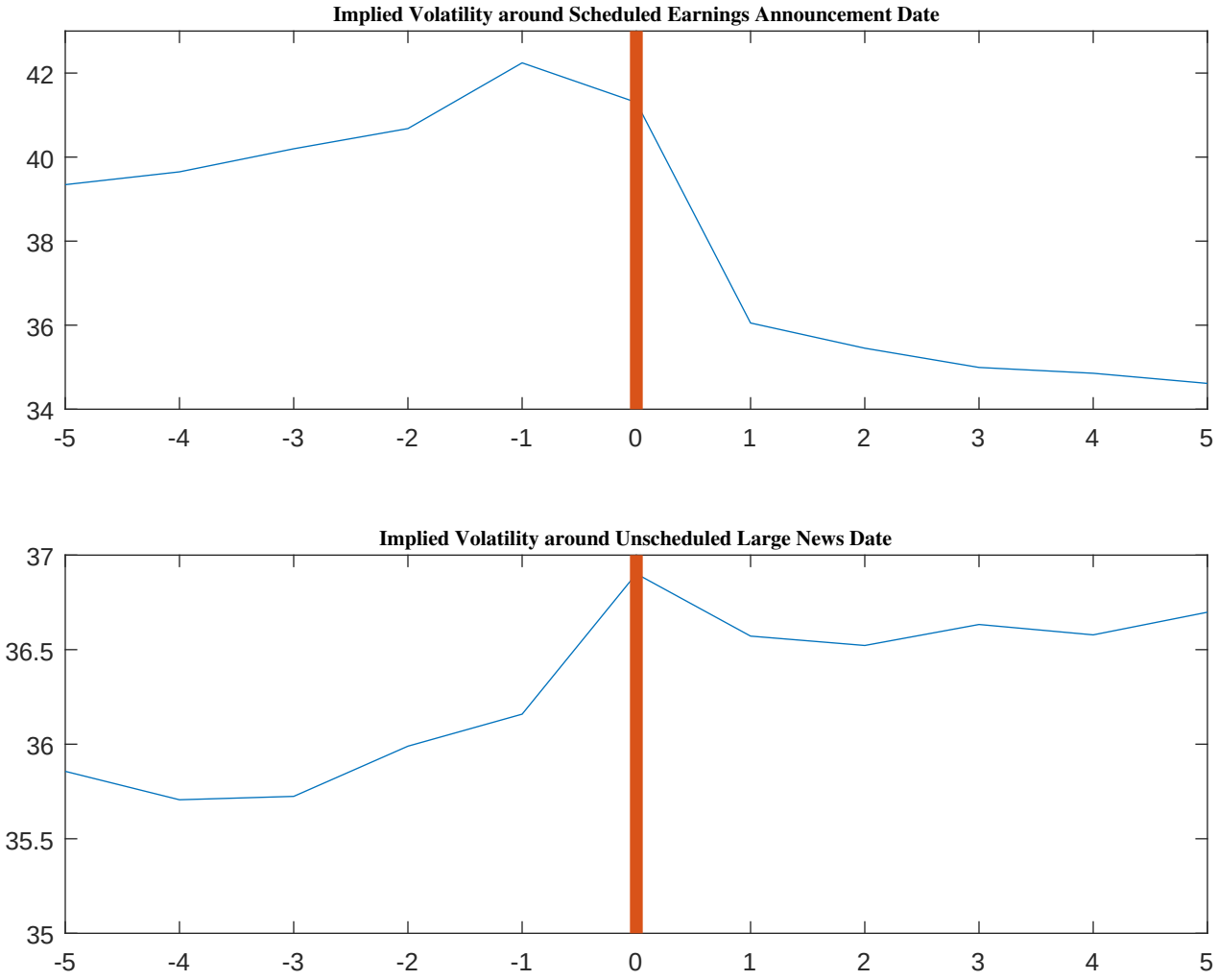
This table reports the pricing kernel parameters estimated by minimizing Vega-weighted root mean squared error (VWRMSE). Estimation was performed first by fixing the physical dynamics parameters and spot variances filtered from Table 1.6, then only allowing the pricing kernel parameters to vary. Three pricing kernel parameters λ_V, λ_I, and $\eta^{\mathbb{Q}} - \eta$ each represents the diffusive variance risk premium, news risk premium, and jump size risk premium, respectively.

Figure 1.1: **Time-series Plot of Daily News Counts for Selected Firms. 2000-2012**



This figure plots the time-series of daily news counts of four selected firms. The sample begins in Jan, 2000 and ends Jul, 2012. Y-axis represents absolute counts of news articles that appear in Factiva database each day.

Figure 1.2: **Implied Volatility Around Scheduled vs. Unscheduled Dates**



This figure plots the behavior of the average implied volatility around scheduled vs. unscheduled news dates. The top panel plots the average implied volatility in the [-5,5]-day window around the scheduled quarterly earnings announcement dates. The bottom panel plots the average implied volatility in the [-5,5]-day window around the dates in top 2% range of news counts that are not within 5 days from the earnings announcement date. Both panels plot the average implied volatility of all 20 firms in the sample from Jan, 2000 to Jul, 2012.

Appendix

Appendix 1A: Benchmark Model

In this section, I provide a result comparing the performance of my model to the benchmark model. Benchmark model considered is the plain stochastic-volatility jump-diffusion (SVJ) model with constant jump-intensity, thus serves as a special case of my model. Specifically, it corresponds to the case $\gamma_1 = 0$.

Table 1A.1 reports the estimated parameter for the benchmark model. The estimated parameters are mostly similar to the full news model estimates reported in Table 1.6. To compare the performance of two models, Table 1A.2 provides the negative log-likelihood of observing daily returns as well as model-implied number of expected jumps per year. Improvement in return likelihood is observed in virtually all 20 firms in-sample, with varying degree of the differences between two. News model implies mostly more expected number of jumps, average being 3.72 jumps per year compared to 3.17 of the benchmark model case.

Appendix 1B: Variance Risk Premium

So far, I have focused on the implications of point estimates of each parameters, and thus have not quantified model-implied risk premiums for individual firms. Given the parameter estimates and the affine structure of the model, it is straightforward to extract the relevant measures.

First, the model-implied variance risk premium is computed as the difference between the unconditional variance of log-returns under the risk-neutral and physical probability measures. As the model features two sources of risks, diffusive and jump, the resulting functional form for variance risk premium has also two component stemming from each of two. The following proposition provides an expression for the variance risk premium.

Proposition 2 *The unconditional variance risk premium is given by*

$$VRP = \underbrace{(\theta^* - \theta)}_{Diffusive} + \underbrace{[(\gamma_0 + \gamma_1\theta_I^*)((\eta^{\mathbb{Q}})^2 + \delta^2) - (\gamma_0 + \gamma_1\theta_I)(\eta^2 + \delta^2)]}_{Jump} \tag{1.32}$$

Table 1A.3 reports the model-implied variance risk premiums. The first three columns report diffusive, jump, and total variance risk premiums, respectively. Consistent with the previous literature, the resulting variance risk premium is very small. Also, diffusive and jump components have similar magnitude of contribution on average to total variance risk premium. To convert the numbers into more conventional definition of variance risk premium that uses the difference between annualized volatility, the last column reports the following expression.

$$\text{VRP in Vol.} = \sqrt{\theta^* + (\gamma_0 + \gamma_1\theta_I^*)((\eta^{\mathbb{Q}})^2 + \delta^2)} - \sqrt{\theta + (\gamma_0 + \gamma_1\theta_I)(\eta^2 + \delta^2)} \tag{1.33}$$

Again, the average variance risk premium is very small, being only -0.36% in annualized volatility terms.

Table 1A.1: **Benchmark Model Parameter Estimates under the Physical Measure**

	Estimated Parameters						
Company Name	κ	θ	σ	η	δ	λ	ρ
Microsoft	2.51	0.057	0.45	0.013	0.084	2.74	-0.48
GE	2.36	0.052	0.39	0.007	0.061	2.73	-0.18
IBM	2.56	0.057	0.56	0.012	0.066	2.70	-0.24
Chevron	2.54	0.075	0.30	0.005	0.073	1.42	-0.52
UTC	2.57	0.089	0.31	0.007	0.050	2.62	-0.51
Pfizer	2.59	0.077	0.32	-0.013	0.058	2.48	-0.49
Johnson & Johnson	2.62	0.045	0.31	0.005	0.041	3.82	-0.49
Merck & Co.	0.44	0.050	0.51	-0.051	0.085	4.78	-0.09
Disney	2.61	0.113	0.33	0.012	0.076	2.50	-0.49
JP Morgan	0.49	0.100	0.59	0.010	0.163	4.55	-0.50
Wal Mart	2.58	0.067	0.32	0.011	0.042	2.95	-0.49
Americal Express	1.64	0.074	0.50	0.003	0.081	2.78	-0.17
Intel	2.58	0.071	0.45	-0.029	0.082	2.19	-0.41
Bank of America	0.64	0.047	0.53	-0.010	0.149	2.70	-0.50
Verizon	1.29	0.076	0.39	0.002	0.065	2.49	-0.47
AT&T	2.09	0.056	0.39	0.006	0.089	2.00	-0.40
Cisco	0.10	0.156	0.55	0.002	0.079	4.83	-0.43
Yahoo	1.93	0.064	0.64	-0.003	0.176	3.56	-0.26
Amazon	1.97	0.129	0.62	0.081	0.165	6.55	-0.37
Ebay	1.50	0.060	0.50	-0.001	0.036	3.00	-0.41
Average	1.88	0.076	0.45	0.003	0.086	3.17	-0.40

This table reports the estimated benchmark model parameters under the physical measure using daily returns from Jan, 2000 to Jul, 2012 for 20 individual firms. Particle Filtering (PF) algorithm was used to estimate the parameters by maximizing the likelihood of observing daily returns.

Table 1A.2: **Comparison between Benchmark Model and News Model**

Company Name	Benchmark Likelihood	News Model Likelihood	Benchmark λ	News Model $E[\lambda]$
Microsoft	-8402.44	-8418.78	2.74	4.30
GE	-8481.90	-8488.88	2.73	3.95
IBM	-8920.27	-8921.26	2.70	2.88
Chevron	-8822.36	-8824.40	1.42	1.44
UTC	-8595.54	-8602.05	2.62	2.54
Pfizer	-8705.53	-8730.83	2.48	3.50
Johnson & Johnson	-9887.53	-9931.67	3.82	5.56
Merck & Co.	-8583.35	-8593.89	4.78	4.01
Disney	-8179.39	-8231.86	2.50	4.05
JP Morgan	-7830.30	-7845.75	4.55	4.61
Wal Mart	-9080.98	-9114.28	2.95	3.71
Americal Express	-8044.31	-8048.48	2.78	4.13
Intel	-7595.94	-7629.32	2.19	4.18
Bank of America	-8010.00	-8024.50	2.70	2.32
Verizon	-8949.93	-8952.91	2.49	3.95
AT&T	-8822.86	-8826.19	2.00	3.55
Cisco	-7459.82	-7459.40	4.83	3.33
Yahoo	-6744.76	-6745.92	3.56	3.76
Amazon	-6574.41	-6604.12	6.55	5.22
Ebay	-8749.92	-8750.22	3.00	3.31
Average	-8322.08	-8337.23	3.17	3.72

This table compares the return likelihood and estimated unconditional number of jumps per year between the benchmark model and the news model. In the benchmark model, parameter λ represents unconditional number of jumps per year. In the news model, the annual jump-intensity is equal to $\gamma_0 + \gamma_1 I_t$, thus $E[\lambda] = \gamma_0 + \gamma_1 E[I_t]$, where $E[I_t]$ is computed as the in-sample average of filtered sates I_t.

Table 1A.3: **Variance Risk Premium implied by Model Parameters**

Company Name	Diffusive VRP	Jump VRP	Total VRP	VRP in Vol.
Microsoft	0.0107	-0.0125	-0.0019	-0.31%
GE	0.0191	0.0006	0.0198	3.04%
IBM	0.0071	0.0030	0.0102	1.86%
Chevron	0.0034	0.0206	0.0240	3.92%
UTC	0.0255	-0.0004	0.0251	3.79%
Pfizer	-0.0003	0.0167	0.0164	3.05%
Johnson & Johnson	0.0000	0.0049	0.0049	1.13%
Merck & Co.	0.0000	0.0040	0.0040	0.67%
Disney	-0.0001	-0.0034	-0.0035	-0.66%
JP Morgan	-0.0720	-0.0174	-0.0894	-8.72%
Wal Mart	0.0004	-0.0004	0.0000	-0.01%
Americal Express	0.0224	0.0049	0.0273	3.44%
Intel	-0.0163	-0.0062	-0.0225	-2.79%
Bank of America	-0.0070	-0.0032	-0.0102	-1.69%
Verizon	0.0025	0.0032	0.0057	1.16%
AT&T	0.0014	0.0006	0.0020	0.32%
Cisco	-0.0514	-0.0013	-0.0528	-6.64%
Yahoo	-0.0191	-0.0036	-0.0226	-2.67%
Amazon	0.0006	-0.0775	-0.0768	-7.46%
Ebay	0.0003	0.0059	0.0062	1.32%
Average	-0.0036	-0.0031	-0.0067	-0.36%

This table reports the variance risk premium implied by the estimated model parameters. The first three columns report diffusive, jump, and total variance risk premium, respectively, computed using the expression given in the equation (1.32). The last column (VRP in Vol.) reports the total variance risk premium computed as the difference between annualized volatility under the risk-neutral and physical measure. The exact expression is given in the equation (1.33), also shown below.

$$\text{VRP in Vol.} = \sqrt{\theta^* + (\gamma_0 + \gamma_1\theta_I^*)((\eta^{\mathbb{Q}})^2 + \delta^2)} - \sqrt{\theta + (\gamma_0 + \gamma_1\theta_I)(\eta^2 + \delta^2)}$$

Chapter 2

Time-Varying Crash Risk: The Role of Stock Market Liquidity

2.1 Introduction

What is the impact of market liquidity on the volatility and crash probability of the aggregate stock market? To answer this question, we estimate a continuous-time model with stochastic volatility and dynamic crash probability. The innovation of our method is the introduction of market illiquidity as an economic factor driving the dynamics of volatility and jump intensity. We measure daily market illiquidity (i.e., lack of market liquidity) as the average effective bid-ask spreads of securities constituting the S&P 500 index estimated from high-frequency trades.[1] It can be thought of as the average cost of a round-trip trade for stocks in the index. We estimate the model during 2004–2012 using daily S&P 500 index options, realized spot variance, and market illiquidity, and find that 64% of the time-varying crash probability is explained by the stock market's exposure to market illiquidity.

Market liquidity, defined as the ease with which securities can be bought or sold without significant price impact, has become an increasing concern in financial markets. This is evidenced, for example, by the "flash crash" of May 2010, when major US stock indices fell by almost 10% before recovering quickly. Similarly, market-wide trading halts on August 24, 2015 generated spikes in asset price volatility across financial markets. These two incidents were quickly identified as symptoms of market illiquidity because they occurred in the absence of major news about fundamentals. As documented in Chung and Chuwonganant (2014), regulatory changes in the US markets have increased the role of non-bank traders in liquidity provision, which has intensified the relationship between volatility and market liquidity. Thus, the influence of market liquidity on the economy appears to be increasing in importance.

[1] This measure is motivated by Aït-Sahalia and Yu (2009) and Goyenko, Holden, and Trzcinka (2009) who find strong empirical support for using effective bid-ask as a measure for market illiquidity.

Market crashes refer to large, unexpected drops in asset prices. Crashes can occur in the presence of news about fundamentals, as well in their absence. In the latter case, market illiquidity is often the culprit. Huang and Wang (2009) show in an equilibrium framework that when trading is costly, potential traders are deterred from participating in the market continuously. They will enter the market only when large trading needs arise, i.e., when hit by sufficiently large idiosyncratic shocks, and importantly more on the selling side.[2] This is because idiosyncratic shocks push investors away from their optimal positions, making them more risk averse and less willing to hold the asset. The increase in risk aversion exacerbates the selling-need for potential sellers, and dampens the demand for potential buyers. This, in turn, leads to order imbalances in the form of excess supply, and therefore price decreases in response. As also shown in Lo, Mamaysky, and Wang (2004), the asymmetry in desire to trade between traders with offsetting shocks arises when trading becomes more costly.

While there exists some empirical evidence suggesting that crashes in the stock market are often driven by market illiquidity, they are typically anecdotal (e.g., "flash crash") or limited to individual stocks.[3] Relatedly, there is an extensive literature on index return models which unanimously agrees that index prices "jump."[4] In these models, crashes are large negative jumps in index returns that cannot be explained by the current level of the index's volatility. More recently, several studies have advocated that the probability of observing crashes is time-varying. The common approach is to let the jump arrival rate increase with the level of the stock return variance.[5] Although this modeling framework is parsimonious, it is inconsistent with the notion that crashes are sudden price drops *unexplainable* by the current volatility level.[6] Therefore, the relationship between stock market crash risk and market return volatility remains an open discussion. Importantly, while recent studies agree that crash risk is time-varying, they are silent on the economic variables driving its dynamic. Our study contributes by providing economic underpinnings to models with time-varying crash risk, and showing that much of the variation in jump intensity is driven by market trading frictions.

To motivate our subsequent modeling framework, we apply a predictive regression analysis linking our market illiquidity measure to a non-parametrically estimated realized jump measure for daily S&P 500 index returns (e.g., Andersen, Bollerslev, and Diebold, 2007; Huang and Tauchen, 2005). Realized

[2]Gennotte and Leland (1990) develop a rational expectation model explaining why a large price drop can occur when there is a relatively small amount of selling in the market.

[3]For instance, Jiang and Yao (2013) find that illiquid stocks have higher daily jump returns in the cross-section. Brogaard, Li, and Xia (2016) show that enhanced stock liquidity decreases the firm's default risk. For an alternative view on individual stock liquidity and crash risk, see Chang, Chen, and Zolotoy (2016).

[4]This literature is too large to cite in full; see Pan (2002), Maheu and McCurdy (2004), Eraker, Johannes, and Polson (2003), Eraker (2004), Broadie, Chernov, and Johannes (2007), and Bakshi, Carr, and Wu (2008).

[5]For examples, see Pan (2002), Andersen, Benzoni, and Lund (2002), and Bates (2006, 2012).

[6]Santa-Clara and Yan (2010) is a notable exception, as they model jump intensity as a quadratic function of state variables. In an affine framework, Andersen, Benzoni, and Lund (2002) and Andersen, Fusari, and Todorov (2016) find a statistically weak relationship between crash intensity and spot variance.

jump variation measures the portion of daily return variance that is due to stock price jumps, and we find that it significantly increases with the level of market illiquidity on the previous day. Importantly, the effect of market illiquidity crowds out the predictive ability of realized variance on realized jump variation, suggesting that market illiquidity is the more robust predictor of crash probability for the stock market index. We confirm this finding by running daily time-series regressions on changes in risk-neutral skewness estimated from index option prices, as well as on changes in realized skewness estimated from high-frequency index returns. In either case, we obtain similar conclusions confirming the robust linkage between market illiquidity and crash risk.

Armed with the evidence above, we estimate a continuous-time model similar to the stochastic volatility with jump model (SVJ) studied by Pan (2002) and Bates (2006), among others. In this model, the jump arrival rate is affine in return variance. We extend this framework by letting the time-varying jump intensity dynamic be a function of return variance, market illiquidity, and a latent state variable. We estimate the model by extracting information embedded in index options and high-frequency intraday trades. We use the unscented Kalman filter (UKF) to extract daily state variables. This filtering method allows for sequential learning in the dynamics of latent jump intensity, variance, and illiquidity processes.

We refer to the most general model that we study as the stochastic jump with variance and illiquidity (SJVI). In this model, the jump intensity dynamic is stochastic and affine in the spot variance, the market illiquidity level, and the latent jump-intensity-specific variable designed to capture the omitted risk factor. For comparisons, we estimate two other benchmark models with jump intensity dynamics that are unrelated to the market illiquidity level. In all specifications, we model the spot variance as a two-factor square-root process, with market illiquidity being one of the factors. Our estimation results show a strong contemporaneous relationship between market illiquidity and spot variance. On average, a one-standard-deviation increase in the level of market illiquidity increases the spot variance by about 12%. This finding lends support to previous studies that have documented a positive relationship between return volatility and trading activity (e.g., Schwert, 1989; Lamoureux and Lastrapes, 1990; Chae, 2005).

We find that the average jump probability is between 2 and 3 per year. When a jump occurs, its average size is between -3.7% and -5.9% in daily return units, with a standard deviation between 3.1% and 4.7%. Therefore, the jump dynamic that we estimate represents a large drop of daily index price, a "crash," and not a market surge. We find strong evidence that during our sample period, crash risk in the S&P 500 index mostly reflects investors' fear of market illiquidity. We arrive at this conclusion by examining the contribution of market illiquidity to the jump intensity dynamic in the new model and find a contemporaneous positive relationship with a strong statistical significance. On the other hand, the contribution of market spot variance is modest and statistically weak, which supports our preliminary

evidence found using regression analysis. Collectively, these findings suggest the reason previous studies find that jump intensity increases with the level of spot variance is due to the strong positive relationship between variance and market illiquidity.

In terms of economic magnitudes, we find that market illiquidity explains more than half of the S&P 500 index's crash probability level during our sample (64% on average). On the contrary, the contribution of market spot variance to the jump intensity dynamic is only about 12%, with the remaining 24% coming from the latent jump-intensity-specific factor. However, during the six-month period after the Lehman Brothers' collapse in 2008, we find that market spot variance dominates other factors in explaining the time-varying crash probability, with the contribution as high as 70%. This finding suggests that investors' fear of crash risk during the subprime crisis reflects uncertainty about the market's fundamentals, while outside the crisis period, crash risk mostly reflects investors' fear of market illiquidity.

We emphasize that the relationship between market illiquidity and time-varying volatility and crash risks is not due to market microstructure noise. The market illiquidity proxy that we use is derived from effective spreads of 500 firms constituting the S&P 500 index and *not* from trades on index funds nor index futures. We believe this measure is indicative of the transaction cost of replicating the index, which directly affects the ability of authorized participants to create and redeem shares of S&P 500 ETFs in order to keep their level at a fair value. We confirm that our main conclusions hold using various robustness checks. For instance, we show that our estimation results are qualitatively similar before and after implementation of the "circuit breaker" in 2010. We re-estimate our models using alternative market illiquidity measures including Amihud (2002), and reach the same conclusions.

In summary, our findings illustrate the importance of market illiquidity in explaining time-varying volatility and crash risks, which is largely missing from prior empirical studies examining index return dynamics. That said, our results do not speak to what gives rise to the initial need for liquidity.[7] Our objective is to establish the empirical relationship between market illiquidity and stock market crash risk, and quantify its economic magnitude.

The remaining parts of this paper proceed as follows. Section 2.2 describes the data and sample selection, and reports preliminary evidence found using regression analyses. Section 2.3 describes the model and estimation procedure. Section 2.4 discusses estimation results and interpretation of our findings. Section 2.5 demonstrates the robustness of our findings. Finally, Section 2.6 concludes.

[7]The lack of liquidity can arise due to various reasons, which can lead to "crashes." For instance, heterogeneous future liquidity needs (Allen and Gale, 1994) and adverse selection costs (Grossman and Miller, 1988) facing liquidity providers can limit the supply of liquidity. As argued in Easley, López de Prado, and O'Hara (2010), many high-frequency trading firms are in the business of liquidity provision and their ability to vanquish quickly from the market when faced with toxic order flows portends episodes of sudden illiquidity, and crashes. In a more recent study, Cespa and Foucault (2014) show that when liquidity providers learn information about an asset from prices of other assets, illiquidity contagion can occur and is a source of market fragility.

2.2 Data and Preliminary Evidence

The sample period goes from January 1, 2004 through December 31, 2012. We focus on the recent period because the global financial market has gone through a drastic transformation, e.g., new banking regulations, proliferation of algorithmic trading and exchanged-traded funds.

2.2.1 Market Illiquidity

We construct a time-series measure of market liquidity at the daily level. We focus on the trading friction associated with the cost of participating in the stock market. We measure this using effective bid-ask spreads following Goyenko, Holden, and Trzcinka (2009), who find strong empirical supports for using intraday bid-ask spreads as the measure of market illiquidity.

We obtain all transactions recorded on securities constituting the S&P 500 index from the TAQ database. Then, for each stock i on day t, we calculate the effective spread of its k^{th} trade as

$$ILQ^i_{t,k} = \frac{2|S^{i,P}_{t,k} - S^{i,M}_{t,k}|}{S^{i,M}_{t,k}}, \tag{2.1}$$

where $S^{i,P}_{t,k}$ is the price of the k^{th} trade of stock i on day t, and $S^{i,M}_{t,k}$ is the midpoint of the best prevailing bid and ask at the time of the k^{th} trade. The daily effective spread of stock i on day t is then computed as the dollar-volume weighted average effective spreads over all trades during the day[8]

$$ILQ^i_t = \frac{\sum_{k=1}^{K} DolVol^i_{t,k} ILQ^i_{t,k}}{\sum_{k=1}^{K} DolVol^i_{t,k}}. \tag{2.2}$$

Lastly, we aggregate the effective spreads of firms constituting the S&P 500 index on each day by equally weighting their daily illiquidity measures:

$$ILQ_t = \frac{1}{N} \sum_{i=1}^{N} ILQ^i_t. \tag{2.3}$$

Daily S&P 500 cash index returns are plotted in the top-left first panel of Figure 2.1, while the top-right panel plots the daily time series of market illiquidity, ILQ. We plot the annualized market illiquidity measure by multiplying their daily levels by 252. The mean annualized ILQ measure is 16.85%, which translates to a 0.067% trading cost at the daily level. The standard deviation of the annualized market

[8]Figure 2A.1 in the Appendix shows percentiles of daily dollar effective spread distribution for the S&P 500 constituents. For the majority of firms, their trading cost measured by the dollar effective spread is well above one cent, which is the minimum tick size set by the exchanges. This finding suggests that the effective spread measure that we use is minimally affected by the minimum tick-size rule.

illiquidity is 5.77% with an interquartile range of 13.90% and 18.26%. We see that the market illiquidity measure rises significantly during the financial crisis period but stays relatively stable in other periods, with occasional few spikes. Figure 2.1 shows a sharp spike on May 6, 2010, which is associated with the "flash crash" incident.

2.2.2 Realized Variance and Jump Variation

We construct daily realized variance and jump variation measures using intraday S&P 500 cash index returns obtained from TickData. Using the latest observation at each minute, we construct a grid of one-minute intraday returns starting from 9:30 a.m. and ending at 4:00 p.m.

Calculations of realized variance and realized jump variation have been studied extensively in the recent literature.[9] We calculate the daily realized variance measure, RV, as the sum of squared one-minute log returns: $RV_t^N = \sum_{i=1}^{N} r_{i,t}^2$. This method measures the total quadratic variation in returns. We measure the variation in daily index returns that is due to the diffusive component using the jump-robust realized variance $MinRV$ of Andersen, Dobrev, and Schaumburg (2012). It is calculated as follows:

$$MinRV_t^N = \frac{\pi}{\pi - 2}\left(\frac{N}{N-1}\right)\sum_{i=1}^{N-1}\min(|r_{i,t}|, |r_{i+1,t}|)^2. \tag{2.4}$$

Following Barndorff-Nielsen and Shephard (2004), we define daily realized jump variation, RJV, as the component in total realized variance RV that is not explained by $MinRV$. On each day, it is calculated as:

$$RJV_t = \max(RV_t - MinRV_t,\ 0). \tag{2.5}$$

We can think of RJV as the proxy for jump risk in daily index returns. The bottom-left and bottom-right panels in Figure 2.1 plot the annualized daily time series of $MinRV$ and RJV, respectively. The mean and standard deviation of the $MinRV$ measure over this period are 2.36% and 6.55%, respectively. For the RJV measure, the mean and standard deviation are 0.33% and 0.94%. The daily index return variance is thus mostly composed of the continuous component of stock price change.[10]

[9] See Huang and Tauchen (2005) for a concise summary.

[10] This finding is consistent with Huang and Tauchen (2005) who find that jumps account for 7% of stock market price variance.

2.2.3 Predicting Realized Jumps

This section provides preliminary evidence on the economic relationship between market illiquidity and jump risk. We estimate a predictive regression model on the realized jump variation measure RJV. The objective is to identify the economic variables that robustly predict the occurrence and magnitudes of jumps the next day.

We examine three variables of interest and their various combinations, namely, the market illiquidity measure ILQ, the diffusive quadratic variation measure $MinRV$, and the option-implied volatility index VIX. The most general predictive regression specification is

$$RJV_{t+1} = \beta_0 + \beta_1 MinRV_t + \beta_2 ILQ_t + \beta_3 VIX_t + \beta_3 Ret_t + \beta_3 RJV_t + \epsilon_{t+1}. \quad (2.6)$$

We obtain daily VIX levels from the Chicago Board of Options Exchange (CBOE) that represent the market's fear index calculated as the 30-day implied volatility level of S&P 500 index options. We include lagged log return of the S&P 500 (Ret) and the autoregressive term for RJV as control variables. Year and day-of-the-week fixed effects are present in all regression specifications, but their estimates are not reported here to save space.

Table 2.1 summarizes the regression results for six specifications based on the general model described in equation (2.6). We report the heteroskedasticity-consistent t-statistic (White, 1980) in parentheses below each parameter estimate. Columns (1)–(3) show that when each of the three variables enters in the regression model, it appears statistically significant in predicting the realized jump variation the next day. We find that $MinRV_t$, ILQ_t, and VIX_t are positive and highly significant at the 95% level or higher. These positive coefficients confirm the intuition that jumps are more likely to occur following a day of more volatile and illiquid market conditions. Looking at the size of the coefficients, we find that market illiquidity, ILQ, is the dominant variable. A one-standard-deviation increase in market illiquidity today would increase realized jump variation by 9.1% the next day.

Columns (4)–(5) report results for combining market illiquidity and the two volatility measures. The results clearly show that ILQ_t is the dominant variable in predicting daily realized jump variation. When the market illiquidity measure ILQ_t is added to the regression, the coefficients on $MinRV_t$ and VIX_t variables lose statistical significance. Column (6) reports results for the most general specification where all independent variables are included. We find that the coefficient on ILQ_t decreases by half due to influences of the two market variance measures. Nevertheless, Column 6 shows that ILQ_t is the only variable that remains statistically significant, confirming that it is the leading predictor of realized jump variation in the stock market index.

We emphasize that the market illiquidity measure that we use is calculated from effective spreads of 500 firms constituting the S&P 500 index while the dependent variable, *RJV*, is constructed from one-minute log returns on the S&P 500 cash index. Market illiquidity and realized return jump variations are thus not related. Further, our finding that market illiquidity dominates return variance in explaining the time-varying market jump risk is confirmed using both $MinRV_t$ and VIX_t.

Overall, the results in Table 2.1 provide preliminary evidence for the importance of market illiquidity in explaining time-varying jumps in index returns. Further, it shows that omission of the market illiquidity variable can lead to a different conclusion regarding the role of market return variance in explaining jumps on the stock market index.

2.2.4 Market Illiquidity and Crash Risk

The realized jump variation *RJV* measure we used in the previous section captures the magnitude of positive and negative jumps in index returns and therefore does not identify a stock market crash from a stock market surge. We provide further evidence linking the role of market illiquidity to crash risk by estimating the impact of *ILQ* on daily skewness of the S&P 500 index. Because crashes are large sudden drops in asset prices, a more negative skewness measure would signal a higher probability of crash risk (Chen, Hong, and Stein, 2001). Therefore, if market illiquidity is strongly linked to the stock market crash, we expect that ILQ_t would be negatively related to the stock market skewness measure.

We follow the nonparametric method developed by Bakshi and Madan (2000) and calculate daily skewness from S&P 500 index options. We refer to this measure as risk-neutral skewness, *RNSkew*. We obtained end-of-day S&P 500 index option prices from OptionMetrics. *RNSkew* is calculated from option contracts with approximately one month to maturity. Therefore, the measure $RNSkew_t$ on day t that we use represents investors' forward-looking risk-neutral expectation of the stock market crash risk from the end of day t to day $t+30$. We discuss details on the skewness measure in the Appendix 2.6.

We examine the impact of market illiquidity on the stock market's risk-neutral skewness using the following regression model:

$$\Delta RNSkew_t = \beta_1 \Delta MinRV_t + \beta_2 \Delta ILQ_t + \sum_{i=1}^{p} \alpha_i \Delta RNSkew_{t-i} + \sum_{j=1}^{q} \delta_j \epsilon_{t-j} + \epsilon_t, \qquad (2.7)$$

where Δ indicates that we are examining the change in daily variables, and ϵ_t is a normally distributed error term. We estimate the model on change in daily skewness and not on its level because $RNSkew_t$ is highly persistent with an autocorrelation of 0.98. As a result, the explanatory variables we use are changes in daily *MinRV* and *ILQ*. The regression shown in equation (2.7) is an autoregressive-moving-

average model (ARMA) with p-order lag in the autoregressive term and q-order lag in the error term. Although not shown in equation (2.7), we include *Ret*, and day-of-the-week fixed effects as control variables in the regression model. We estimate the model using maximum likelihood.

Table 2.2 reports estimation results for four regression specifications based on the general model in equation (2.7). We choose the ARMA model with lags of $p = 2$ in the autoregressive term and $q = 1$ in the error term. The number of lags is determined based the Ljung-Box test corresponding to the most parsimonious model that sufficiently removes autocorrelations in the residuals. Column (1) provides the baseline regression results. It shows that the change in risk-neutral skewness is negatively related to the index return and is strongly explained by its autoregressive terms. Columns (2) and (3) show that when $\Delta MinRV$ and ΔILQ are separately included in the regression, they load negatively on the change in risk-neutral skewness. However, ΔILQ is the only statistically significant variable. The negative and highly significant coefficient on ΔILQ suggests that when the average trading cost in the stock market increases, investors' expectation of the market crash risk also increases.

Column 4 reports results for the full regression model. We observe two striking findings. First, the coefficient on $\Delta MinRV$ switches sign from negative to positive and is significant at the 10% level. Second, the coefficient on ΔILQ remains negative and significant but approximately doubles in terms of magnitude. These results suggest that both $MinRV$ and ILQ are important determinants of daily stock market skewness, i.e., crash risk. The positive coefficient on $\Delta MinRV$, however, shows that an increasing market variance is related to a less negatively skewed risk-neutral distribution of daily index returns. This finding is intuitive because as the variance level increases, the index return distribution would become more fat-tailed on both positive and negative sides, indicating an increasing likelihood of market surges as well as market crashes. Consequently, the index return distribution appears more symmetric, i.e., less negatively skewed, conditional on an increasing variance level. As a robustness check, we verify this result by replacing $MinRV$ with the total quadratic variation (RV) in Table 2.2 and obtain the same conclusion.

For comprehensiveness, we verify our results using realized skewness constructed from high-frequency trades on the S&P 500 constituents. Our method follows that in Amaya, Christoffersen, Jacobs, and Vasquez (2015). We obtain the same conclusion when using daily realized skewness. To save space, the results are reported in Appendix 2.6, which shows that aggregate market illiquidity is a leading determinant of stock market crash risk and that its influence dominates the impact of market return variance. Motivated by this non-parametric evidence, we develop a continuous-time model that allows market illiquidity to act as an economic covariate in explaining the time-varying volatility and crash risks.

2.3 Model and Estimation

2.3.1 The SJVI Model

We begin by specifying the processes governing the log index price, spot variance, spot illiquidity, and latent component of jump intensity dynamic under the risk-neutral measure ($\mathbb{Q}$). We use the notation S_t and V_t to denote index price and spot variance at time t. We let L_t represent the spot market illiquidity, which measures the average cost of trading in the stock market at time t, with a higher value indicating a more illiquid market. We include a stochastic process Ψ_t that is designed to capture the latent time-varying jump intensity in index returns. Thus, the model consists of four factors that fully describe the return dynamics under $\mathbb{Q}$:

$$
\begin{aligned}
d\log(S_t) &= (r - \frac{1}{2}V_t - \xi\lambda_t)dt + \sqrt{V_t}(\sqrt{1-\rho^2}dW_t^1 + \rho dW_t^2) + q_t dN_t && (2.8)\\
dV_t &= \kappa_V(\theta_V - V_t)dt + \gamma dL_t + \xi_V\sqrt{V_t}dW_t^2 && (2.9)\\
dL_t &= \kappa_L(\theta_L - L_t)dt + \xi_L\sqrt{L_t}dW_t^3 && (2.10)\\
d\Psi_t &= \kappa_\Psi(\theta_\Psi - \Psi_t)dt + \xi_\Psi\sqrt{\Psi_t}dW_t^4, && (2.11)
\end{aligned}
$$

where r denotes the risk-free rate and all Brownian motions dW_t^i, for $i = 1$ to 4, are independent of each other.

We assume the market illiquidity process L_t, and the latent jump intensity process Ψ_t, in equations (2.10) and (2.11) follow the standard square-root model with long-run mean levels of θ_L and θ_Ψ, respectively. The variance dynamic in equation (2.9) follows Heston's (1993) square-root process with an additional term γdL_t. We discuss our specification choice for the variance dynamic later in this subsection.[11]

The log index price dynamic described in equation (2.8) follows a standard jump-diffusion process where $q_t dN_t$ denotes the jump component. Following the extant literature on index return models, we assume that jumps follow a compound Poisson process with intensity λ_t and each individual jump is independent and identically distributed (i.i.d.) normal with the jump mean size θ and the jump size standard deviation δ. To ensure the discounted stock price is a martingale, we include the jump compensation term $\xi = e^{(\theta + \frac{\delta^2}{2})} - 1$ in equation (2.8). Lastly, to complete the model, we specify the

[11] Our main conclusions are unaffected when estimating a simpler model without illiquidity feedback in the spot variance, i.e., $\gamma = 0$ in equation (2.9).

dynamic of the time-varying jump intensity λ_t as follows:

$$\text{SJVI model:} \qquad \lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t. \tag{2.12}$$

This jump intensity specification is motivated by numerical tractability and ease of economic interpretation. Equation (2.12) shows that the time-varying jump arrival rate is determined jointly by the levels of spot variance V_t, spot market illiquidity L_t, and state variable Ψ_t. The latent state variable Ψ_t is designed to capture the portion of jump intensity dynamic not explained by the covariates V_t and L_t. For the remaining parts of this paper, we refer to this general specification as the SJVI model.

Equation (2.9) shows that the evolution of spot variance depends on its own mean-reverting drift, the diffusive component, and the market illiquidity process L_t. This specification allows for changes in the market spot variance V_t and market illiquidity L_t to be contemporaneously related, which is supported by Lamoureux and Lastrapes (1990) who find that daily trading volume significantly explains daily return variance.

We choose a parsimonious modeling framework that lets dL_t enter the dynamic of dV_t, and not vice versa. Besides parsimony, this choice is motivated by their joint time-series estimates, which show that the change in market illiquidity leads the change in market spot variance. We show this by estimating the Vector Autoregressive Moving-Average (VARMA) model below:

$$\begin{pmatrix} \Delta MinRV_t \\ \Delta ILQ_t \end{pmatrix} = \delta + \Phi \begin{pmatrix} \Delta MinRV_{t-1} \\ \Delta ILQ_{t-1} \end{pmatrix} - \Theta u_{t-1} + u_t, \tag{2.13}$$

where δ is a 2×1 vector of coefficients, and u_t is a 2×1 vector of normally distributed residuals. Φ and Θ are 2×2 matrices of VARMA model coefficients. The variables $MinRV$ and ILQ are defined in the previous section. We estimate the model above and find that

$$\hat{\Phi} = \begin{pmatrix} -0.149^{***} & 0.479^{***} \\ (-3.59) & (4.50) \\ -0.023 & 0.237^{***} \\ (-1.47) & (5.45) \end{pmatrix} \quad \text{and} \quad \hat{\Theta} = \begin{pmatrix} 0.717^{***} & -0.031 \\ (19.97) & (-0.39) \\ -0.014 & 0.730^{***} \\ (-1.18) & (22.09) \end{pmatrix}, \tag{2.14}$$

where the t-statistic is reported in parentheses below each parameter estimate.

The diagonal elements in Φ measure the impact of autoregressive terms for $\Delta MinRV_t$ and ΔILQ_t, which, as we expected, are statistically significant. The off-diagonal elements in Φ provide insight on the

cross-impacts between $\Delta MinRV_t$ and ΔILQ_t. We find the coefficient estimate measuring the impact of ΔILQ_{t-1} on $\Delta MinRV_t$ is 0.479 with a t-statistic of 4.50. This shows that a change in ILQ_{t-1} on the previous day has a positive and statistically significant impact on a change in $MinRV_t$ today. On the other hand, we do not find statistically significant evidence that a change in ILQ_t today is driven by a change in $MinRV_{t-1}$ on the previous day; the coefficient estimate is -0.023 with the t-statistic of -1.47. As a robustness check, we re-estimate the VAR model by replacing $MinRV$ with RV in equation (2.13) and obtain the same conclusion. Further, looking at the coefficient estimates in Θ, we find that only the diagonal elements are statistically significant. This shows that the residual terms in the market spot variance L_t and the market illiquidity level ILQ_t do not affect each other. This result supports our modeling assumption in equations (2.9) and (2.10), where the Brownian shocks dW_t^2 and dW_t^3 are independent.

Finally, we note that the variance dynamic that we consider in equation (2.9) falls under the class of two-factor stochastic volatility models, which have been shown to effectively explain the term structure of index option prices.[12] Our model differs from the existing two-factor volatility literature in that we allow the expected future variance to depend on the level of spot variance, V_t, and spot market illiquidity, L_t, as shown below:

$$E_t[V_T] = \theta_V + (V_t - \theta_V)e^{-\kappa_V(T-t)} + [(L_t - \theta_L)\frac{\gamma\kappa_L}{\kappa_V - \kappa_L}](e^{-\kappa_V(T-t)} - e^{-\kappa_L(T-t)}). \tag{2.15}$$

This equation shows that the long-run mean of the spot variance is θ_V, and the mean-reversion speed to the long-run variance is denoted by κ_V. It also shows that the current level of market illiquidity positively affects the shape of the expected term structure of variance. Its impact, however, dissipates as the time horizon increases. This is seen from the third term on the right-hand side, which converges to 0 as time T goes to infinity.

2.3.2 Benchmark Models

We consider two nested specifications of the SJVI model. In the first specification, we shut off the illiquidity channel in the time-varying jump intensity dynamic by setting $\gamma_L = 0$ in equation (2.12). As the result, the probability of observing jumps depends on the level of spot variance and the latent state component as follows:

$$\text{SJV model:} \qquad \lambda_t = \Psi_t + \gamma_V V_t. \tag{2.16}$$

[12] See for examples, Egloff, Leippold, and Wu (2010), and Andersen, Fusari, and Todorov (2015b).

We refer to this as the stochastic jump intensity with variance (SJV) model. When Ψ_t is constant, it nests the affine jump intensity dynamic, $\lambda_t = \gamma_0 + \gamma_V V_t$, commonly adopted in time-varying jump studies (e.g., Pan, 2002; Bates, 2006).

The second nested specification that we study shuts off the impact of both market illiquidity and the spot variance on the jump probability. That is, we set γ_V and γ_L equal to zero in equation (2.12). This yields

$$\text{SJ model:} \qquad \lambda_t = \Psi_t, \tag{2.17}$$

We refer to this as the stochastic jump intensity model (SJ).

We keep all other aspects of the three models that we study identical. This allows us to focus solely on the role of market illiquidity and spot variance in assessing time-varying jump risk.

2.3.3 Filtering

Each of the three models we study contains three latent state variables: V_t, L_t, and Ψ_t. We extract the latent state variables using the square-root unscented Kalman filter (UKF) of Van der Merwe and Wan (2001). We apply the UKF method because the option prices data that we fit the models to are non-linear in the state variables.[13]

The state variables in the filtering equations evolve under the physical probability ($\mathbb{P}$) measure. We therefore need to define the state variables' dynamic under the physical measure. We do not impose risk premiums on the L_t and Ψ_t processes for simplicity and also because the literature has not yet provided clear guidance on how to model their risk premiums. As a result, there is no change to these two processes from $\mathbb{Q}$ to $\mathbb{P}$. We apply the commonly used functional form of the variance price of risk to the spot variance process, which is given by $\nu_V\sqrt{V_t}$ as in Heston (1993). This price of risk specification shifts the Brownian shock in equation (2.9) by $dW_t^{2,\mathbb{P}} = dW_t^2 - \nu_V\sqrt{V_t}dt$, where the superscript $\mathbb{P}$ denotes that it is evaluated under the physical probability measure. Applying this transformation, the resulting variance process under $\mathbb{P}$ can be written as

$$dV_t = \kappa_V^{\mathbb{P}}(\theta_V^{\mathbb{P}} - V_t)dt + \gamma dL_t + \xi_V\sqrt{V_t}dW_t^{2,\mathbb{P}}, \tag{2.18}$$

where we have the following parameter mappings $\kappa_V^{\mathbb{P}} = \kappa_V - \nu_V\xi_V$ and $\theta_V^{\mathbb{P}} = \theta_V\kappa_V/\kappa_V^{\mathbb{P}}$.

[13] For recent papers using UKF as the filtering method, see Bakshi, Carr, and Wu (2008) and Filipović, Gourier, and Mancini (2016). We refer to Christoffersen, Dorion, Jacobs, and Karoui (2014) for technical details and comparison between different filtering methods.

We discretize the $\mathbb{P}$-measure state dynamics using the conventional Euler scheme at the daily interval. The discretized state-space system can be written as follows:

$$V_{t+1} = V_t + \kappa_V^{\mathbb{P}}(\theta_V^{\mathbb{P}} - V_t)\Delta t + \gamma\kappa_L(\theta_L - L_t)\Delta t + \xi_V\sqrt{\Delta t V_t}\epsilon_{t+1}^1 + \gamma\xi_L\sqrt{\Delta t L_t}\epsilon_{t+1}^2 \quad (2.19)$$

$$L_{t+1} = L_t + \kappa_L(\theta_L - L_t)\Delta t + \xi_L\sqrt{\Delta t L_t}\epsilon_{t+1}^2 \quad (2.20)$$

$$\Psi_{t+1} = \Psi_t + \kappa_\Psi(\theta_\Psi - \Psi_t)\Delta t + \xi_\Psi\sqrt{\Delta t \Psi_t}\epsilon_{t+1}^3, \quad (2.21)$$

where the error terms ϵ_{t+1}^i, for $i = 1$ to 3, are i.i.d. standard normal. In the above state-space system, we set the time step $\Delta t = 1/252$ to reflect the daily discretization interval. To keep notation to a minimum, we apply the superscript $\mathbb{P}$ only to parameters under the physical measure that differ in values from their corresponding risk-neutral parameters.

We next describe the functional relationships linking the latent state variables to the observed data used in the estimation. The first observable is the illiquidity measure denoted by ILQ_t, which we introduced earlier in Section 2.2. The other observables that we use are daily at-the-money (ATM) and out-of-the-money (OTM) S&P 500 index options. These three sets of observables are used in the measurement equations in the UKF procedure. We write the system of measurement equations as follows:

$$\log(\text{ILQ}_{t+1}) = \log(E_t[\int_t^{t+1} L_s ds]) + u_{t+1}^1 \quad (2.22)$$

$$ATM_{t+1}^O = ATM_{t+1}^M(V_{t+1}, L_{t+1}, \Psi_{t+1}) + u_{t+1}^2 \quad (2.23)$$

$$OTM_{t+1}^O = OTM_{t+1}^M(V_{t+1}, L_{t+1}, \Psi_{t+1}) + u_{t+1}^3, \quad (2.24)$$

where measurement errors u_{t+1}^i, for $i = 1$ to 3, are independent normal random variables with constant variances. The above filtering equations are applied to all trading days from January 2, 2004 to December 31, 2012, resulting in 2,262 observation days.

The latent spot illiquidity process in the state-space dynamic describes the instantaneous level of illiquidity at each moment and not at the aggregated daily level. To filter L_t from the daily observed market illiquidity measure, we integrate the spot illiquidity process over the day as shown in equation (2.22). Because the spot illiquidity measure is assumed to follow a square-root process, its daily integrated value is available in closed form. We use the log effective spread in the measurement equation because the empirical distribution of effective spreads is close to log-normal.

Following Pan (2002), we collect two time series of closing mid-price of options quotes that we label ATM and OTM. We let ATM denote at-the-money call option that has moneyness, defined as the ratio of forward-to-strike price, closest to 1. Similarly, OTM refers to out-of-the-money put option that has

moneyness closest to 0.95. For both ATM and OTM options, we retain contracts that have time-to-maturity closest to 30 calendar days. Figure 2.2 plots daily Black-Scholes option-implied volatilities calculated from the ATM and OTM contracts that we use in our study. As argued by Pan (2002), we use OTM options in the measurement equation as it provides the richest information on investors' expectation of crash probability in the stock market.

We follow Trolle and Schwartz (2009) and use Black-Scholes vega-weighted price as the functional form in the measurement equations for options fitting in equations (2.23)–(2.24). This method scales the value of options across time making their prices more comparable, which in turn facilitates the assumption of the normally distributed errors in the measure equations. Therefore, ATM_{t+1}^{O} and OTM_{t+1}^{O} in equations (2.23)–(2.24) represent the scaled ATM and OTM option prices observed at the end of day t. Similarly, the variables ATM_{t+1}^{M} and OTM_{t+1}^{M} denote the model-implied option prices scaled by their market Black-Scholes vega.

The models that we study fall within the affine jump-diffusion framework. Therefore, the conditional characteristic function of log stock price is available in exponential affine form. Following Duffie, Pan, and Singleton (2000), we derive the log affine functional form of the characteristic function in Appendix 2.6. The coefficients in the characteristic function are not all available in terms of elementary functions, thus, we solve for them numerically in the Ricatti system of equations. We use the fast Fourier transform (FFT) method first developed by Carr and Madan (1999) to numerically evaluate option prices.

Lastly, we note that at this stage, we do not need to specify the risk premiums associated with the first Brownian motion, dW_t^1, and the compound Poisson jumps, qdN_t, because they only alter the drift term of returns dynamics that is not part of the estimation. We discuss the specification of equity and jump risk premiums in a later section, where they are estimated using a time series of daily index returns.

2.3.4 Estimation

We estimate the models by maximizing the log-likelihood function resulting from the UKF step. We assume the measurement errors are conditionally normal, therefore, the time t conditional log-likelihood takes the following form:

$$l_t(\Theta) = -\frac{3}{2}\log(2\pi) - \frac{1}{2}\log(\det|\Omega_t|) - \frac{1}{2}(Y_t - \bar{Y_t})^T(\Omega_t)^{-1}(Y_t - \bar{Y_t}), \tag{2.25}$$

where $\bar{Y_t}$ and Ω_t denote the ex ante forecasts of the mean and covariance matrix conditional on time $t-1$ information on observables Y_t. We let Θ denote the set of all parameters to be estimated.

In addition to the log-likelihood resulting from the measurement error equations, we follow Andersen,

Fusari, and Todorov (2015a) and add a penalizing term that compares the filtered spot variance component, V_t, to the model-free estimate of spot variance calculated from high-frequency data. Incorporating this penalizing term, the conditional log-likelihood function that we estimate at time t is

$$L_t(\Theta) = l_t(\Theta) + \omega \log(\ (\sqrt{V_t^n} - \sqrt{V_t})^2\ ), \tag{2.26}$$

where $l_t(\Theta)$ is given in equation (2.25), V_t^n is the realized spot variance computed using one-minute grid returns from the S&P 500 index and V_t is the filtered spot variance from the UKF procedure. We describe the construction of the realized spot variance measure in more detail in Appendix 2.6. Daily time-series dynamic of the realized spot volatility, $\sqrt{V_t^n}$, is shown in the bottom panel of Figure 2.2.

The tuning parameter ω in equation (2.26) is set equal to 0.05 following Andersen, Fusari, and Todorov (2015a).[14] The model parameters are then estimated by maximizing the sum of conditional log-likelihoods over the sample period from January 2, 2004 to December 31, 2012.

2.4 Results

2.4.1 Maximum likelihood estimates

Table 2.3 reports parameter estimates for the three models. The first, second, and third columns report results for the SJ, SJV and SJVI models, respectively. We report log-likelihood values of the three models in the bottom row.

We find that parameters governing the square-root dynamic of spot variance are well estimated. Their parameter estimates are fairly consistent across the models. The correlation estimates of the two Brownian shocks in return and spot variance, ρ, are about -35%, confirming the asymmetric return-variance relationship found in the extant literature. We find that the spot market illiquidity level, L_t, significantly impacts the level of spot variance, V_t. This is seen from the estimates of γ. We find that across the three models, the estimates γ are about 0.12. This suggests that a one-standard-deviation increase in the spot market illiquidity, L_t, would increase the spot variance level by about 12% after controlling for the persistence dynamic of the variance process.

The strong relationship we find between market illiquidity and return variance lends support to previous studies examining the relationship between return volatility and market trading activity. In particular, motivated by the mixture of distribution hypothesis (MDH), which assumes that volatility and volume simultaneously depend on a latent information process, past research efforts have been

[14] We verify that our main conclusions are unaffected to a reasonably large range of values for ω.

devoted to studying the relationship between stock return volatility and trading volume (e.g., Clark, 1973; Epps and Epps, 1976; Tauchen and Pitts, 1983). Nevertheless, the findings in this literature have been mixed and understanding the relationship between information flows and trading activity has been an active research area. For instance, Lamoureux and Lastrapes (1990) estimate a GARCH volatility model and find that trading volume is the main driver of stock return volatility and that past stock return innovations became insignificant once trading volume is included in the model.[15] While we find that market illiquidity significantly drives the dynamic of spot variance, its effect does not eliminate the strong persistence in the variance dynamic. Further, the recent literature agrees that trading volume is an inadequate measure of market liquidity.[16] Given the recent availability of intraday trading data, we can more precisely measure market illiquidity by calculating the cost of participating in the stock market (i.e., transaction cost). Our results estimated using a continuous-time model documenting a strong relationship between market illiquidity and return variance therefore contribute to this stream of literature.

Estimates of the jump-size mean, θ, and the jump-size standard deviation, δ, in Table 2.3 indicate that the jump dynamic that we estimate corresponds to crash risk in the stock market. The estimates of θ are negative and highly significant. The average jump mean size in daily index return is between -3.7% (SJV model) and -5.9% (SJ model). Therefore, the jump dynamic that we identify corresponds to large drops in daily S&P 500 index returns.

Table 2.3 shows that the SJ model has the largest magnitudes of θ and δ. This implies that crashes in the SJ model are larger and more dispersed in magnitude relative to the other two models. We next examine parameter estimates governing the time-varying jump intensity. First, we look at the dynamic of the latent jump-intensity-specific factor, Ψ_t. The magnitude of parameters driving the Ψ_t dynamic in the SJ model differs significantly from those in the other two models. For instance, the long-run mean θ_Ψ, the mean-reversion speed κ_Ψ, and the volatility ξ_Ψ of the jump-intensity-specific factor are significantly larger for the SJ model. These findings are expected because in the SJ model, jump intensity dynamic solely depends on the latent state variable Ψ_t. Further, these results confirm that the dynamic of jump intensity is time-varying and follows a mean-reverting process.

Table 2.3 shows that when we add covariates to the jump intensity dynamic in the SJV and SJVI models, the log-likelihood value increases substantially. The improvement is large with an increase of about 5% relative to the SJ model. We therefore find strong support for modeling jump intensity as a function of economic covariates. Looking at the SJV model, we find the impact of spot variance on jump

[15] In contrast, several studies find evidence conflicting with the MDH specification. These studies include Hiemstra and Jones (1994), Lamoureux and Lastrapes (1994), Richardson and Smith (1994) and Andersen (1996).

[16] See for examples, Lee, Mucklow, and Ready (1993), Jones (2002) and Fleming (2003).

intensity, γ_V, is positive and statistically significant at the 5% level (t-statistic is 2.17). This finding is consistent with Pan (2002), Bates (2006), and Andersen, Fusari, and Todorov (2015b).

For the SJVI model, we find that when we add the market illiquidity measure to the jump intensity specification, the estimate of γ_V substantially decreases in magnitude and its statistical significance diminishes (t-statistic is 1.61). On the other hand, the impact of spot market illiquidity loads very strongly (t-statistic is 13.87). This finding shows that the inclusion of market illiquidity as an economic covariate significantly weakens the relationship between jump intensity and spot variance. This finding is consistent with our conclusions from Table 2.1, which we obtained using regression analyses.

2.4.2 Time-Varying Volatility and Crash Risks

This section examines the time-series dynamics of market spot volatility and jump intensity. Table 2.4 reports descriptive statistics of daily jump intensity λ_t, spot variance V_t, and spot illiquidity L_t levels that we obtained using the UKF from 2004–2012. We find that the sample moments of daily spot illiquidity are almost identical across the three models. This suggests that its dynamic is well identified when we extract their information from the daily market illiquidity measure ILQ_t calculated using effective bid-ask spreads.

Figure 2.3 plots the daily annualized jump intensity for the three models. The jump intensity dynamic of the SJ model is very volatile relative to the other two models. For instance, looking at the time-series statistics of λ_t in Table 2.4, we find the average expected number of jumps implied by this model is 1.19 per year, but with a median of 0.44 and a standard deviation of 5.95. This shows that the distribution of jump intensities filtered from the SJ model is highly skewed and dispersed. The average jump intensity implied by the SJ model is about half of the other models. However, the rarer nature of jumps observed in this model is compensated by its larger jump mean size of $\theta = -5.9\%$ per each jump as shown in Table 2.3. Figure 2.3 shows the expected number of jumps in the SJ model increases dramatically during the 2008–2009 crisis period, while it is small outside the crisis period.

We find the jump intensity dynamic estimated from the SJV and SJVI models have comparable distributions with means of 2.2 and 2.9 jumps per year, respectively. In these two models, the levels of jump intensity are relatively stable before mid-2007, but rise after and peak in the fall of 2008. We believe the relatively more stable jump intensity dynamics observed in the SJV and SJVI models are due to improved identification resulting from the use of covariates in the jump intensity specification. This argument is supported by looking at the models' log-likelihood performance, which is substantially worse under the SJ model.

We next examine the economic contribution of the spot variance and market illiquidity to the jump intensity dynamic. Figure 2.4 plots the decomposition of daily jump intensity levels. Here, we decompose daily jump intensities filtered from the SJV model (top panel) and from the SJVI model (bottom panel) into their respective components.

For the SJV model, the top panel of Figure 2.4 shows that the market's spot variance is the main component driving jump intensity dynamic. We find that on average, 61% of the jump intensity level is explained by its covariation with the market's spot variance. The time-series average of its contribution is about 61%. We find the jump-intensity-specific factor Ψ_t explains about 39%. This finding shows that a non-trivially large portion of jump intensity cannot be explained by the dynamic of market's spot variance.

The bottom panel of Figure 2.4 shows the decomposition of daily jump intensities estimated from the SJVI model. Here, we find that the jump intensity dynamic is heavily dominated by its co-movement with equity market illiquidity. We plot daily percentage contributions of each jump intensity component in Figure 2.5. The results shown are largely consistent with the findings in the bottom panel of Figure 2.4. We find that, on average, the market illiquidity factor explains about 64% of the jump probability in the SJVI model. In contrast to our findings for the SVJ model, we find the market's spot variance explains, on average, only 12%, with the remaining 24% contribution coming from the jump intensity factor, Ψ_t. Once we control for market illiquidity as an economic variable driving time-varying crash risk, the relative contribution of spot variance significantly diminishes.

The above findings offer important insight into the existing literature on index return models that has increasingly documented the importance of time-varying crash risk (e.g., Bates, 2006, 2012; Maheu, McCurdy, and Zhao, 2013). The common practice is to let jump intensity be an affine function of spot variance. This modeling approach is appealing because it is parsimonious. It identifies time-varying jump intensity as a constant multiple of the spot variance, thereby eliminating the need to introduce an additional state variable to the model. We find that our estimation results for the SVJ model provide some support for this modeling approach. However, we emphasize that the key economic variable that matters most from our results for modeling time-varying crash probability is not the market spot variance, but the market illiquidity factor. Lastly, our findings suggest the reason previous studies find a positive relationship between the stock market's time-varying crash risk and spot variance is because of their common exposure to market illiquidity.

2.4.3 Impulse Response Function

We examine the impact of market illiquidity on the current and future crash probability using impulse response functions (IRFs). The IRF tells how much current and future values of crash intensity $\lambda_{t+\tau}$ respond to a one-standard-deviation increase in either the spot variance V_t, the spot market illiquidity L_t, or the level of latent state variable Ψ_t. The SJVI model that we propose yields an analytically tractable IRF for the jump intensity. We report the IRF formula in Appendix 2.6.

In Figure 2.6, we consider two dates where the levels of spot volatility are relatively high or low. The left-hand-side panels plot the IRF on March 11, 2009, with high spot volatility. The right-hand-side panels plot the IRF on January 8, 2004, which corresponds to the day with low spot volatility.

Looking at the IRF plots on the day with high volatility (left panels), we find that the impact of spot volatility dominates. A one-standard-deviation increase in the spot volatility V_t translates to an increase of 0.11 in jump intensity on the same day. Importantly, the impact of a shock to spot volatility is very persistent with a half life of about two months. A one-standard-deviation increase in the market spot illiquidity L_t increases the jump intensity by about 0.07. Looking at the impact of a shock to the latent factor Ψ_t, we find that it is trivially small at both the short- and long-run horizons.

We next look at the IRF plots on the day with low spot volatility. The state variables on this day are significantly less volatile and therefore the levels of IRF are much lower. On this low-volatility day, Figure 2.6 shows that a shock to market illiquidity dominates in term of magnitude as well as its lasting impact on the jump intensity. Similar to the day with high spot volatility, we find that the impact of a one-standard-deviation shock to the latent factor Ψ_t is small. Overall, Figure 2.6 shows that shocks to the spot volatility V_t and the market illiquidity L_t are highly persistent and drive most of the current and future increase in the probability of crash risk. However, the relative importance of V_t and L_t depends on the level of uncertainty in the market.

2.4.4 Forecast Error Variance Decomposition

We perform a forecast error variance decomposition (FEVD) on the jump intensity λ_τ dynamic for the SJVI model. This method helps determine the amount of information each variable contributes in explaining changes to the current and future crash probability.

The error from forecasting the jump intensity $\lambda_{t+\tau}$ with τ-period horizon conditional on day t is defined as

$$\hat{\epsilon}_{\lambda,t+\tau} = \lambda_{t+\tau} - E_t[\lambda_{t+\tau}]. \tag{2.27}$$

The idea behind FEVD is to find how much of the variation in $\hat{\epsilon}_{\lambda,t+\tau}$, i.e., $\text{Var}_t\left[\hat{\epsilon}_{\lambda,t+\tau}\right]$, can be explained by shocks to each state variable driving the jump intensity dynamic. In other words, the FEVD asks how much of the unexpected change in the jump intensity is explained by shocks to V_t, L_t, and Ψ_t. Derivation of the FEVD is tedious. For brevity, we report the expression in Appendix 2.6.

In Figure 2.7, we plot the proportion of forecast error variance explained by innovations to the spot variance, V_t, and the spot market illiquidity, L_t, factors. The model parameters are obtained from their MLE estimates in Table 2.3, and their state variables are set equal to their long-run values. We do not plot the proportion of forecast error variance explained by Ψ_t because it is trivially small. The top two panels of Figure 2.7 plot the proportion of forecast error variance explained by V_t and L_t one day ahead, i.e., $\tau = 1$. These results provide insight on the source of information that most importantly impacts unexpected changes in the crash probability at the very short horizon. We also plot the proportion of forecast error variance at a longer horizon, i.e., $\tau = 250$, which approximately corresponds to one year ahead. These results are shown in the bottom two panels.

We find that errors in the short-term forecast of the crash probability are mainly explained by shocks to the market illiquidity factor, with the exception of the crisis period when shocks to the spot variance dominate. This finding is consistent with the results shown in Figure 2.5. In contrast, when we look at the sources of risk that explain errors in the long-term forecasted crash probability, we find that the market spot variance dominates. The bottom panels of Figure 2.7 show that the diffusive variance component in index returns contributes about 62% to unexpected changes in the long-run crash probability, while the market illiquidity component contributes around 38%. This finding suggests that changes in investors' perception about the long-run stock market crash risk is associated with the market uncertainty level. On the other hand, unexpected changes in crash probability at the near horizon are mostly explained by shocks to market illiquidity.

2.4.5 Option Fit

We also compare the three models based on their in-sample option fit. We define in-sample option pricing error as the vega-weighted root mean squared error (VWRMSE) in fitting observed Black-Scholes vega-weighted option prices obtained from the UKF procedure

$$\text{VWRMSE(ATM)} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(ATM_{t+1}^{O} - A\bar{T}M_{t+1}^{M})^2}, \tag{2.28}$$

where $A\bar{T}M_{t+1}^{M}$ denotes the ex ante forecast of vega-weighted ATM option price at time $t+1$. Option pricing error for OTM options are computed in a similar way.

Table 2.5 reports in-sample option pricing errors for the three models. We separate the sample into three sub-periods of three years each. Overall pricing errors are very similar in magnitude between SJ and SJV models, where the SJV model performs better during the crisis period while the SJ model has lower pricing errors during normal times. More importantly, the SJVI model has a superior in-sample option fit in most of the periods for both ATM and OTM options. Thus, the improvement in fitting OTM options using the SJVI model suggests that its jump intensity specification is better-suited for capturing the jump intensity dynamic embedded in the index options.

2.4.6 Risk Premiums

Using the risk-neutral parameter estimates in Table 2.3 and daily filtered states variables $\{\hat{V}_t, \hat{L}_t, \hat{\Psi}_t\}$ estimated previously, we infer the risk premium parameters. This is done by estimating the model on daily S&P 500 index returns from 2004–2012, and keeping the parameters that are not affected by the change of probability measures fixed. This approach to identify risk premiums was also employed in Andersen, Fusari, and Todorov (2015b).

We assume the conventional form of the pricing kernel that preserves the affine structure of the model under the physical measure. The prices of risk associated with the four Brownian motions are given by

$$dW_t^{1,\mathbb{P}} = dW_t^1 - \nu_1\sqrt{V_t}dt \tag{2.29}$$

$$dW_t^{2,\mathbb{P}} = dW_t^2 - \nu_V\sqrt{V_t}dt \tag{2.30}$$

$$dW_t^{3,\mathbb{P}} = dW_t^3 \tag{2.31}$$

$$dW_t^{4,\mathbb{P}} = dW_t^4 \tag{2.32}$$

The parameter ν_1 in equation (2.29) corresponds to the price of risk parameter for the first Brownian innovation in the return process. Recall that ν_V is the price of risk parameter for the volatility innovation that we estimated from options and realized spot variance as part of the UKF. Its estimate is reported in Table 2.3. Recall also that we do not impose any risk premium assumptions on the third and fourth Brownian motions corresponding to the liquidity and latent jump intensity innovations, respectively.

We follow Pan (2002) and assume the difference between jump distributions under the physical and risk-neutral measures derives from the jump-size risk premium, ν_θ, defined as the difference between jump-size means, $\theta^P - \theta$. The dynamic of log-stock price under the physical probability measure can be

written as

$$\begin{aligned} d\log(S_t) &= (r - \frac{1}{2}V_t - \xi^P\lambda_t + (\sqrt{1-\rho^2}\nu_1 + \rho\nu_v)V_t)dt + \\ &\quad \sqrt{V_t}(\sqrt{1-\rho^2}dW_t^{1,\mathbb{P}} + \rho dW_t^{2,\mathbb{P}}) + q_t dN_t^{\mathbb{P}} \end{aligned} \tag{2.33}$$

where $\xi^P = \exp(\theta^P + \frac{1}{2}\delta^2)$ is the jump compensator under the physical measure. Comparing the $\mathbb{P}$-measure return dynamic in equation (2.33) to the $\mathbb{Q}$-measure return dynamic in equation (2.8) shows that the equity risk premium, π_t, can be written as

$$\pi_t = (\xi^P - \xi)\lambda_t + (\sqrt{1-\rho^2}\nu_1 + \rho\nu_V)V_t \tag{2.34}$$

$$= (\xi^P - \xi)\lambda_t + \nu_S V_t, \tag{2.35}$$

where we define $\nu_S = \sqrt{1-\rho^2}\nu_1 + \rho\nu_v$ in equation (2.35).

Using the filtered state variables, $\{\hat{V}_t, \hat{L}_t, \hat{\Psi}_t\}$, we apply daily discretization to the return process and estimate the risk premium parameters ν_θ and ν_S using MLE while fixing all other parameters. The estimate for ν_1 is then inferred from ν_S. Appendix 2.6 shows the discretization of the continuous-time model, and presents the log-likelihood function for fitting the return process.

Table 2.6 reports estimation results of the risk premium parameters. We find that the jump risk premium parameter ν_θ is well identified in all models. The estimates for ν_θ are statistically significant at a confidence level of 1% or greater. On the other hand, estimates of the diffusive risk premium parameter ν_S are only marginally significant. These findings are consistent with Pan (2002) who finds that the jump risk premium is more easily identified from index option prices, while risk premiums associated with the diffusive and variance risks are more difficult to estimate. Table 2.6 also reports estimates for the price of risk coefficient ν_1 associated with the first Brownian motion. Their values are inferred from the corresponding estimates of ν_S in Table 2.6, and ν_V in Table 2.3. Because ν_1 is indirectly inferred, we do not report its t-statistic. This parameter can be usefully thought of as the price risk for exposure to the diffusive component in index return.

Using the estimates reported in Tables 2.6 and 2.3, we quantify the economic magnitude of each risk premium component in terms of annualized excess returns. Equation (2.35) shows that the equity risk premium can be decomposed into two main components. The first component represents the compensation for bearing stock market crash risk, $(\xi^P - \xi)\lambda_t$. The second component represents the compensation for bearing stock market's diffusive return and variance risks, $\nu_S V_t$. For brevity, we refer to $\nu_S V_t$ as diffusive risk in the equity risk premium.

We first look at the compensation for bearing stock market crash risk. For each model, we calculate the long-run jump risk premium level $(\xi^P - \xi)\bar{\lambda}_t$, where $\bar{\lambda}_t$ is the annualized time-series mean of the jump intensity dynamic reported in Table 2.4. We find that the compensation for bearing the market's crash risk for the SJ, SJV, and SJVI models are 3.0%, 4.8%, and 4.8% in annualized excess returns, respectively. The jump risk premium estimate implied by the SJ model is lower than the other two models. This finding reflects the relatively lower jump intensity levels that we find for the SJ model. The jump risk premium estimates implied by the SJV and SJVI models are mostly consistent with prior studies that estimate a time-varying jump risk model on the S&P 500 index over a similar sample period. For instance, Ornthanalai (2014) estimates the jump risk premium implied by the compound Poisson jump process over the 1996–2012 period and finds that its magnitude is 4.5% per year. Using index options and returns data from an earlier time period, i.e., 1989–1996, Pan (2002) finds that the implied jump risk premium is 3.5% per annum.

We next look at the compensation for bearing stock market diffusive risk. This is calculated as $\nu_S \bar{V}_t$, where $\bar{V}_t$ is the time-series mean of the annualized variance reported in Table 2.4. We find the compensation for bearing diffusive risk for the SJ, SJV, and SJVI models is 7.48%, 4.15%, and 2.39% in annualized excess returns, respectively. The relatively larger magnitude of diffusive risk premium found in the SJ model is expected. This is because the SJ model has the lowest jump risk premium level and hence it must rely on the diffusive risk premium component to match the level of equity risk premium found in the data.

The realized equity premium calculated using daily index returns data over the 2004–2012 period is 8.7% per year. The total equity premiums that we find for the SJ, SJV, and SJVI models are 10.48%, 8.92%, and 7.14%, in annualized terms, respectively. Our estimates of the total equity premium are therefore consistent with the value calculated using daily returns data. This finding suggests that the magnitudes of equity risk premium implied by our models are economically plausible.

2.5 Robustness

2.5.1 Circuit Breakers

Following the flash crash incident on May 6, 2010, the Securities and Exchange Commission (SEC) installed "circuit breakers" on 404 NYSE-listed S&P 500 stocks on June 16, 2010 to halt trading for five minutes if any stock experiences more than 10% movement, either up or down, in a five-minute period. This new trading rule potentially affects our aggregate illiquidity measure constructed from individual

firms' effective spreads, thereby altering the impact of market illiquidity on jump probability. We test whether this change in market-trading rules alters our findings on the influence of market illiquidity on time-varying crash risk.

We take June 16, 2010 as the date of exogenous shift in the market-trading structure. Specifically, we divide our sample into two periods, one starting on January 4, 2004 and ending on June 15, 2010, and the other starting on June 16, 2010 and ending on December 31, 2012. We use the same jump intensity specification as in the SJVI model for both subsamples, as below. The model parameters are estimated separately, yielding two sets of parameter estimates. We summarize the results below. For brevity, we report only coefficient loadings on the covariates in the jump intensity dynamic. The t-statistic for each parameter is reported in parentheses underneath its estimate.[17]

$$\text{Before the circuit breaker:} \qquad \lambda_t = \Psi_t + \underset{(1.71)^*}{19.69\, V_t} + \underset{(9.04)^{***}}{9.24\, L_t} \tag{2.36}$$

$$\text{After the circuit breaker:} \qquad \lambda_t = \Psi_t + \underset{(1.47)}{18.47\, V_t} + \underset{(1.98)^{**}}{7.98\, L_t}. \tag{2.37}$$

We find that the loading coefficient on the market illiquidity factor is smaller in latter period, being 7.98, relative to the estimate of 9.24 in the earlier period. This suggests that the introduction of circuit breakers has slightly reduced the impact of market illiquidity on jump intensity, perhaps, by eliminating sudden contiguous large movements in equity prices that were often identified as symptoms of liquidity shortage. Nevertheless, equations (2.36)–(2.37) show that the implementation of circuit breakers does not materially impact the importance of the market illiquidity channel. Both coefficients are statistically significant at 1% and 5% levels, respectively, with a lower t-statistic for the latter period because of its much smaller sample size. The loading coefficients on market spot variance V_t for the pre- and post-circuit breaker periods are 19.69 and 18.47, respectively. These magnitudes are similar to the estimate of 18.38 found using the full sample period. In both periods, the impact of the coefficient estimates on V_t are statistically weak. These results confirm the robustness of our estimates in Table 2.3.

2.5.2 Alternative Iliquidity Measure

We have so far defined market illiquidity using the aggregate *relative* effective spread of S&P 500 constituents. This measure captures the aggregate transaction cost of participating in the stock market and has been shown in Aït-Sahalia and Yu (2009), and Goyenko, Holden, and Trzcinka (2009) to be a

[17] ***, **, and * denote statistical significance at the 1, 5, and 10 confidence levels, respectively.

good proxy for market illiquidity. This section tests whether our results are robust to other illiquidity measures.

First, we use dollar effective spread as an alternative measure for the market-trading cost. It is calculated as the absolute dollar difference between the transaction price and the prevailing mid-price of each transaction instead of the relative percentage to the mid-price as before. More precisely, the dollar effective spread measure associated with each transaction k on day t for firm i is defined as

$$\$ILQ^i_{t,k} = 2|S^{i,P}_{t,k} - S^{i,M}_{t,k}|. \tag{2.38}$$

We aggregate $\$ILQ^i_{t,k}$ across S&P 500 index constituents to construct the daily measure of market illiquidity.

Our second alternative market illiquidity measure is from Amihud (2002). On each day t, we compute the Amihud illiquidity measure for each firm i in the S&P 500 index as a fraction of absolute return, $|r_i|$, over dollar trading volume, $DVol_{i,t}$, that day:

$$ALIQ^i_t = \sum_{i=1}^{N} \frac{|r_{i,t}|}{DVol_{i,t}}. \tag{2.39}$$

The daily Amihud market illiquidity measure for the stock market is then calculated as an equally-weighted average of individual firms' Amihud illiquidity measure.

The middle and bottom panels of Figure 2A.2 plot the time-series dynamic of the two alternative market illiquidity measures. For a quick comparison, we plot the relative effective spread in the top panel. All illiquidity measures are normalized to have the same sample mean. This normalization method does not impact our results because the absolute level does not matter for our specification. Figure 2A.2 shows that the dollar effective spread measure is similar to the relative effective spread measure, although with some small differences during the crisis period. The Amihud illiquidity measure is much noisier than the other two measures calculated from intraday bid-ask spreads.

We re-estimate the SJVI model using the market dollar effective spread and the Amihud illiquidity measure. We report the results in Table 2.7. Overall parameter estimates are fairly consistent compared with those estimated using relative effective spreads shown in Table 2.3. We find the coefficient estimate γ_V is similar in magnitude and is not statistically significant in either alternative illiquidity measures that we use. Meanwhile, the coefficient estimate for γ_L is similar for the Amihud illiquidity measure, and even larger for the dollar effective spread measure. In all cases, γ_L remains statistically significant. We find that in-sample option pricing errors are higher using the Amihud illiquidity measure, which is

expected due to the noisiness of the measure. On the other hand, the in-sample options fit shows a small marginal improvement using the dollar effective spread. To save space, the option pricing results are not reported here. Overall, our main conclusions remain qualitatively unchanged. We conclude that our main results are robust to different definitions of market illiquidity.

2.6 Conclusion

We study the role of market liquidity in explaining the time-varying market crash risk in the S&P 500 index. We estimate a continuous-time model with stochastic volatility and crash probability. We introduce market illiquidity as an observable variable to the model by allowing it to affect the dynamics of spot variance and jump risk intensity. We follow the recent empirical literature and measure the daily stock market illiquidity level using volume-weighted intraday bid-ask spreads of all securities constituting the S&P500 index. We estimate the model over 2004–2012 using daily S&P 500 index options, realized spot variance and market illiquidity measure, and find that 64% of time-varying crash risk is due to the stock market's exposure to market illiquidity. The influence of market illiquidity dominates other factors that we examined, including the market's spot variance. This is with an exception of the 2008 crisis, when the influence of spot variance dominates and the contribution of market illiquidity falls to about 30%.

Table 2.1: Regression Model on Realized Jump Variation (RJV)

	Realized Jump Variation Next Day: RJV_{t+1}					
	(1)	(2)	(3)	(4)	(5)	(6)
$MinRV_t$	0.071***			0.038		0.037
	(2.63)			(1.36)		(1.31)
ILQ_t		0.091***		0.061***	0.070**	0.045**
		(3.97)		(4.25)	(2.13)	(2.01)
VIX_t			0.046***		0.014	0.011
			(4.50)		(1.19)	(0.94)
Ret_t	-0.058	-0.057	-0.038	-0.054	-0.050	-0.049
	(-0.97)	(-0.94)	(-0.63)	(-0.92)	(-0.79)	(-0.79)
RJV_t	-0.196	-0.075	-0.006	-0.185	-0.068	-0.177
	(-1.19)	(-0.94)	(-0.09)	(-1.22)	(-0.85)	(-1.15)
Adjusted R^2	29.3%	30.5%	29.0%	32.1%	30.7%	32.2%

Notes: We report estimated coefficients and t-statistics from the predictive OLS regression on the non-parametrically estimated variance component in daily S&P 500 index returns that is due to jumps. The sample period is from January 2, 2004 to December 31, 2012. The dependent variable is the realized jump variation (RJV) calculated using high-frequency intraday trades on S&P 500 cash index. The independent variables include lagged realized variance estimator, *MinRV*, from Andersen, Dobrev, and Schaumburg (2012), which measures variations in daily S&P 500 index returns that are associated with non-jump risk; market illiquidity proxy, *ILQ*, measured by daily averaged effective spreads across firms in the S&P 500 constituents; option-implied volatility index, *VIX*, obtained from the CBOE; and log return of S&P 500 index. All variables are lagged by one day. *RJV*, *MinRV*, *ES*, and *VIX* are expressed in annualized terms by multiplying their daily measure by 252. We also include the autoregressive term for *RJV* in the regression. Year and day-of-the week fixed effects are included. Coefficient estimates on the fixed-effect terms are not reported here to save space. We report heteroskedasticity-consistent t-statistic in parenthesis "( )" below each parameter estimate. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively, based on the heterosedasticity-consistent t-statistic. Adjusted R-squared for each regression model is reported in the bottom row.

Table 2.2: Regression Model on Change in Risk-Neutral Skewness

	Change in Risk-Neutral Skewness: $\Delta RNSkew_t$			
	(1)	(2)	(3)	(4)
$\Delta MinRV_t$		-0.005		0.450*
		(-0.03)		(1.69)
ΔILQ_t			-0.622*	-1.144**
			(-1.79)	(-2.48)
$Return_t$	-0.549***	-0.551***	-0.826***	-0.857***
	(-3.28)	(-3.00)	(-3.40)	(-3.70)
$\Delta RNSkew_{t-1}$	0.455***	0.455***	0.449***	0.451***
	(18.93)	(18.86)	(18.59)	(18.74)
AICC	-2.181	-2.180	-2.182	-2.182
R^2	23.4%	23.4%	23.5%	23.7%

Notes: We report regression results on daily changes in risk-neutral skewness of S&P 500 index returns, $\Delta RSkew_t$. The sample period is from January 2, 2004 to December 31, 2012. Risk-neutral skewness, $RSkew_t$, on day t is calculated from end-of-the-day S&P 500 index option prices with maturity closest to 30 days. We use the nonparametric method of Bakshi and Madan (2000) to calculate the 30-day forward-looking risk-neutral skewness measure. The independent variables include change in realized variance estimator, $\Delta MinRV$, from Andersen, Dobrev, and Schaumburg (2012); change in market illiquidity proxy; ΔILQ, measured by daily averaged effective spreads across firms in the S&P 500 constituents; *Return*, log S&P 500 return. Each specification is estimated using maximum likelihood. We use an ARMA(2,1) structure in the regression model, which is determined by the LjungBox test to sufficiently remove cross-correlations in the residuals. We control for seasonality due to the day-of-the-week effect. We report coefficient estimates on the two autoregressive terms. For brevity, we do not report coefficient estimates on the moving-average error term and day-of-the-week fixed effects. Robust t-statistic is reported in parenthesis below each parameter estimate. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. The last row reports regression diagnostics based on the Akaike information criterion (AICC) and R^2 metric.

Table 2.3: Maximum Likelihood Estimates: 2004–2012

	(1) SJ Model $\lambda_t = \Psi_t$	(2) SJV Model $\lambda_t = \Psi_t + \gamma_V V_t$	(3) SJVI Model $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$
Parameter	Estimate	Estimate	Estimate
Panel A. Filtered state dynamics			
κ_V	3.422	3.565	3.549
	(7.75)	(8.39)	(4.73)
θ_V	0.031	0.031	0.031
	(8.86)	(11.00)	(5.66)
ξ_V	0.336	0.343	0.346
	(99.50)	(22.97)	(42.49)
ν_V	1.604	1.559	1.554
	(1.06)	(0.42)	(0.71)
κ_L	2.416	2.344	2.353
	(3.93)	(1.27)	(4.61)
θ_L	0.178	0.182	0.171
	(7.64)	(5.16)	(6.69)
ξ_L	0.149	0.151	0.158
	(37.50)	(6.10)	(36.80)
κ_Ψ	0.972	0.661	0.662
	(5.55)	(0.83)	(2.25)
θ_Ψ	1.619	0.102	0.101
	(5.00)	(3.34)	(1.66)
ξ_Ψ	0.402	0.204	0.204
	(2.47)	(7.22)	(1.79)
ρ	-0.343	-0.351	-0.353
	(19.03)	(2.90)	(5.92)
Panel B. Jump-size parameters			
θ	-0.059	-0.037	-0.037
	(34.98)	(10.41)	(26.28)
δ	0.047	0.033	0.031
	(55.16)	(13.48)	(29.84)
Panel C. Loadings on covariates			
γ	0.120	0.117	0.118
	(5.22)	(3.52)	(7.95)
γ_V		52.723	18.380
		(2.17)	(1.61)
γ_L			9.259
			(13.87)
Log-Likelihood:	7,307.32	7,641.24	7,707.29

Notes: We report MLEs of the three time-varying jump models: SJ, SJV, and SJVI. The sample period is from January 2, 2004 to December 31, 2012. Each model is estimated using daily OTM and ATM S&P 500 index options, averaged effective spreads of S&P 500 constituents, and spot variance estimated from one-minute high-frequency S&P 500 futures data. We maximize the log likelihood function in equation (2.26). The state variables are estimated using the UKF. We report t-statistic calculated using the outer product of the gradient in parenthesis below each parameter estimate.

Table 2.4: Descriptive Statistics of Filtered Jump Intensities and Spot Variances

	(1) SJ Model $\lambda_t = \Psi_t$	(2) SJV Model $\lambda_t = \Psi_t + \gamma_V V_t$	(3) SJVI Model $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$
Panel A. Jump intensity λ_t			
Mean	1.1922	2.2231	2.8772
Median	0.4367	1.3128	2.1420
Std. Dev.	5.9503	3.1282	2.6284
25 percentile	0.1818	0.7964	1.7096
75 percentile	0.8955	2.4995	3.1357
Panel B. Spot variance V_t			
Mean	0.0258	0.0266	0.0213
Median	0.0159	0.0145	0.0147
Std. Dev.	0.0334	0.0390	0.0234
25 percentile	0.0099	0.0090	0.0090
75 percentile	0.0280	0.0251	0.0236
Panel C. Spot illiquidity L_t			
Mean	0.1681	0.1681	0.1682
Median	0.1539	0.1538	0.1538
Std. Dev.	0.0558	0.0558	0.0560
25 percentile	0.1374	0.1375	0.1375
75 percentile	0.1741	0.1742	0.1742

Notes: We report the descriptive statistics of filtered jump intensities λ_t, spot variances V_t, and spot illiquidity L_t for three models: SJ, SJV, and SJVI. The variables are reported in annualized terms by multiplying their daily values by 252. We obtain the filtered state variables from the UKF step in the MLE estimation. Parameter estimates of the three models are reported in Table 2.3.

Table 2.5: Vega-weighted Root Mean Squared Error of Different Models

	(1) SJ Model $\lambda_t = \Psi_t$	(2) SJV Model $\lambda_t = \Psi_t + \gamma_V V_t$	(3) SJVI Model $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$
Panel A. VWRMSE by sub-period for OTM options			
2004–2006	3.21%	3.38%	3.21%
2007–2009	10.41%	9.99%	8.72%
2010–2012	6.57%	6.57%	6.00%
2004–2012	7.34%	7.17%	6.39%
Panel B. VWRMSE by sub-period for ATM options			
2004–2006	1.12%	1.68%	1.21%
2007–2009	7.69%	7.58%	5.82%
2010–2012	4.50%	4.70%	4.00%
2004–2012	5.18%	5.24%	4.14%

Notes: We report in-sample fit for the three models: SJ, SJV, and SJVI. The models are estimated using MLE. Equation (2.26) shows the log likelihood function. Panel A reports in-sample option pricing errors for OTM and panel B reports the pricing errors for ATM options. Option pricing errors are obtained from the measurement equations in the UKF step. The numbers reported are vega-weighted root mean squared error (VWRMSE). For ATM options, the VWRMSE is calculated as

$$\text{VWRMSE(ATM)} \quad = \quad \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(ATM_{t+1}^{O} - \bar{ATM}_{t+1}^{M}\right)^2}$$

where $\bar{ATM}_{t+1}^{M}$ denotes the ex ante forecast of vega-weighted ATM option price at time $t+1$, and ATM_{t+1}^{O} denotes the vega-weighted ATM option price observed in the data. We use option vega reported in the Ivey Optionmetrics database to scale option prices, which makes their levels more comparable across moneyness and time (Trolle and Schwartz, 2009). The VWRMSE for OTM options is calculated similarly.

Table 2.6: Risk Premium Parameters Estimated from Daily Returns: 2004–2012

	(1) SJ Model $\lambda_t = \Psi_t$	(2) SJV Model $\lambda_t = \Psi_t + \gamma_V V_t$	(3) SJVI Model $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$
Parameter	Estimate	Estimate	Estimate
$\nu_s = \sqrt{1-\rho^2}\nu_1 + \rho\nu_v$	2.900	1.562	1.121
	(1.89)	(1.39)	(1.54)
ν_θ	0.048	0.025	0.019
	(11.76)	(4.23)	(2.24)
ν_1	3.674	2.253	1.784
Log-Likelihood:	7,125.88	7,190.94	7,204.12

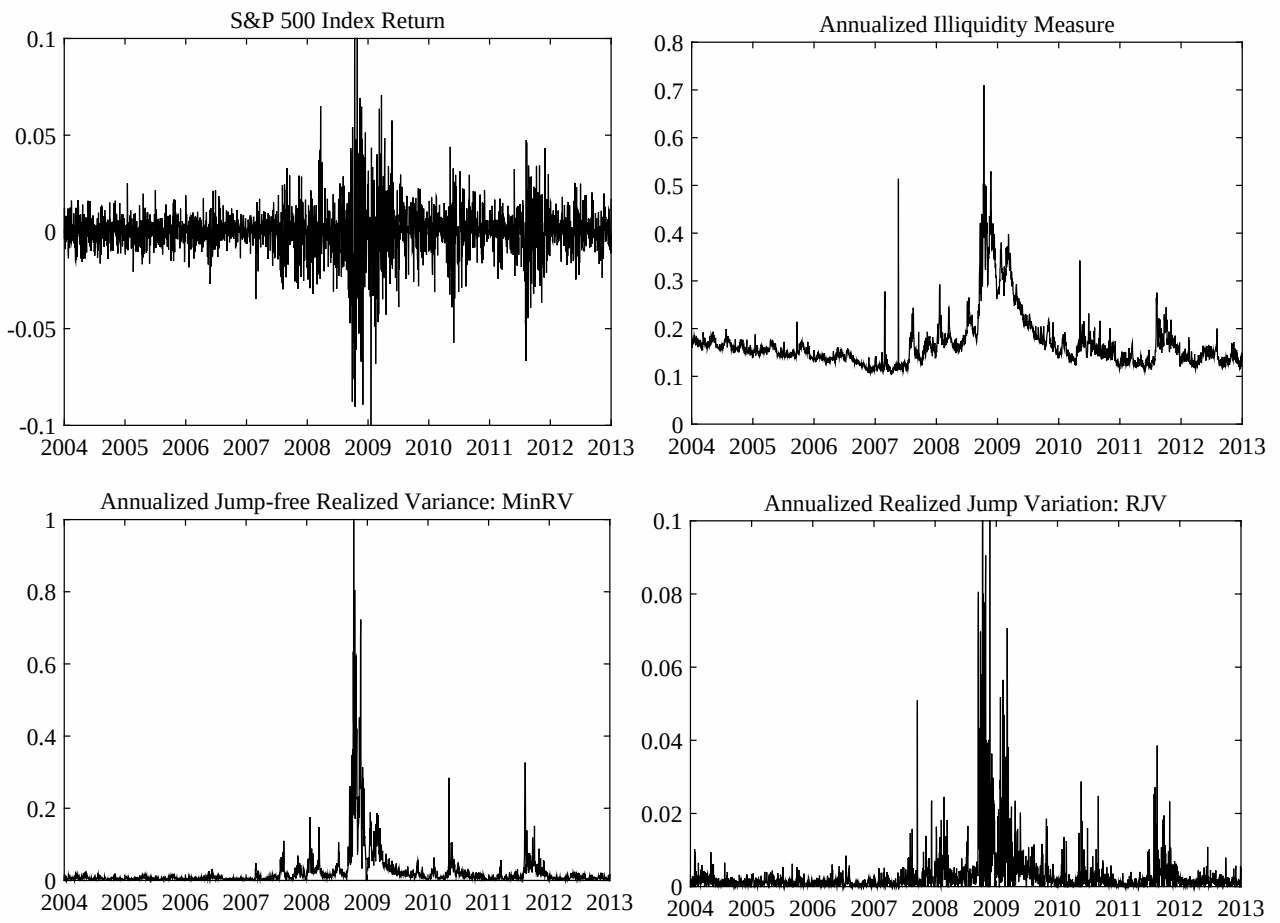
Notes: We report MLE estimates of the risk premium parameters for the three time-varying jump models: SJ, SJV, and SJVI. Each model is fitted to daily S&P 500 return daily returns data from January 2, 2004 to December 31, 2012. We obtain daily state values V_t, L_t, and Ψ_t, as well as $\mathbb{Q}$-measure parameters from the first-stage estimation results reported in Table 2.3. The parameter ν_θ is the difference between jump-size means under the physical and risk-neutral measures, i.e., $\theta^P - \theta$. The parameter ν_1 corresponds to the price of risk coefficient associated with the Brownian innovation in the return process; see equation (2.29). We report t-statistic calculated using the outer product of the gradient in parenthesis below parameter estimates for ν_s and ν_θ . To facilitate econometric identification, we estimate $\nu_s = \sqrt{1-\rho^2}\nu_1 + \rho\nu_V$ from daily returns MLE and later infer ν_1 from its estimate together with the value of ν_V reported in Table 2.3.

Table 2.7: Maximum Likelihood Estimates: Alternative Illiquidity Measures

	(1) SJVI: Amihud-ILQ	(2) SJVI: $ES-ILQ
Parameter	Estimate	Estimate
Panel A. Filtered state dynamics		
κ_V	3.553	3.553
	(1.31)	(4.44)
θ_V	0.032	0.031
	(28.83)	(227.78)
ξ_V	0.345	0.346
	(7.55)	(13.01)
ν_v	1.552	1.554
	(0.55)	(0.21)
κ_L	2.357	2.355
	(1.18)	(0.77)
θ_L	0.170	0.171
	(1.75)	(20.61)
ξ_L	0.158	0.144
	(0.85)	(5.62)
κ_Ψ	0.661	0.661
	(0.35)	(0.32)
θ_Ψ	0.101	0.101
	(1.35)	(5.23)
ξ_Ψ	0.203	0.206
	(1.43)	(25.36)
ρ	-0.344	-0.352
	(1.38)	(2.01)
Panel B. Jump-size parameters		
θ	-0.034	-0.037
	(7.41)	(30.05)
δ	0.032	0.027
	(8.25)	(16.34)
Panel C. Loadings on covariates		
γ	0.119	0.118
	(1.17)	(1.04)
γ_V	18.180	19.372
	(0.38)	(0.47)
γ_L	9.686	13.584
	(11.81)	(3.21)

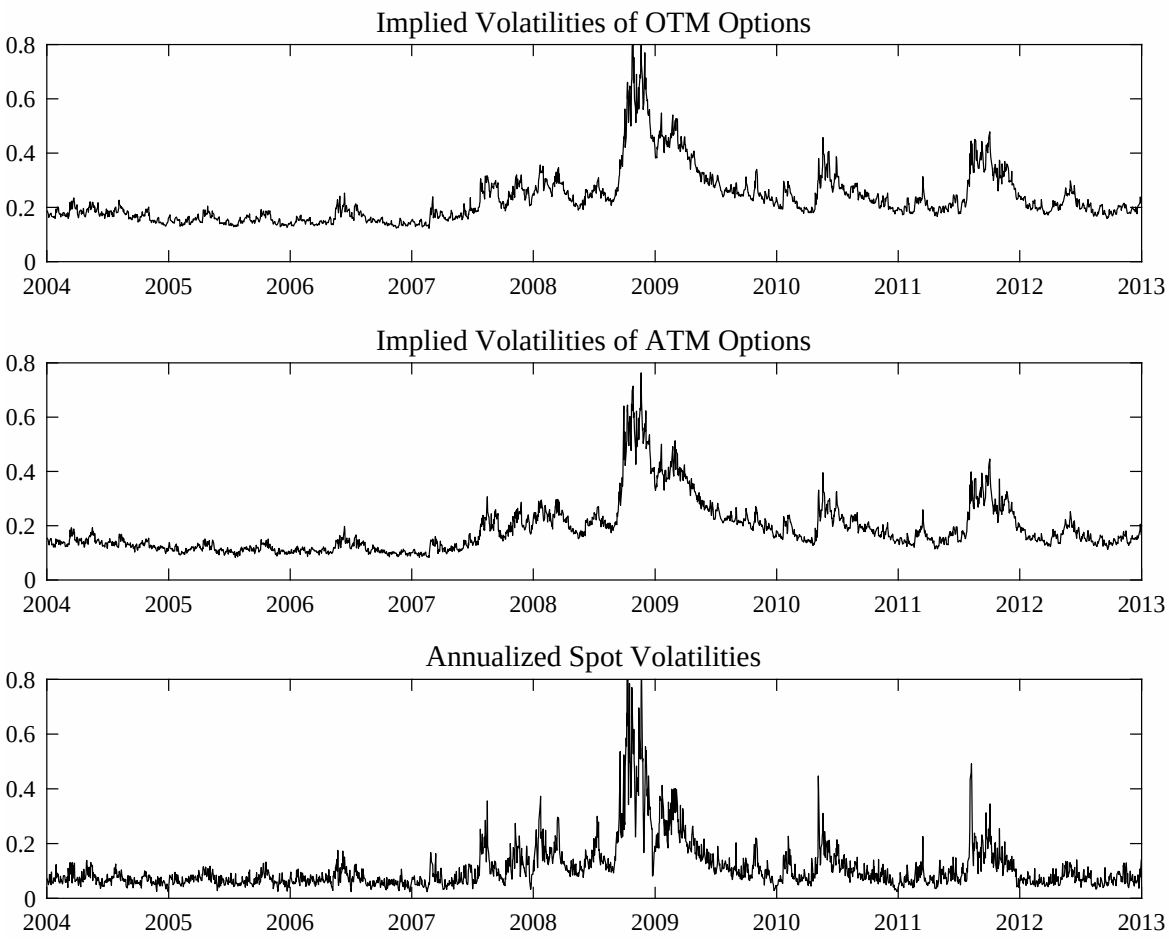
Notes: We report MLE parameter estimates for the SJVI model estimated using two alternative illiquidity measures. The sample period is from January 2, 2004 to December 31, 2012. Each model is estimated using daily OTM and ATM options, daily spot variance calculated from high-frequency index returns, and daily illiquidity measure. In the first column, the daily illiquidity measure is calculated using Amihud (2002). The second column reports results using dollar effective spread as a measure of illiquidity. The daily illiquidity measure is calculated at the stock level, and then aggregated across firms constituting the S&P 500 index to yield the daily market illiquidity measure. See Section 2.5.2 for more details. All models are estimated by maximizing log likelihood from UKF. T-statistic calculated using the outer product of the gradients is reported in parentheses.

Figure 2.1: Daily Time Series of the Stock Market Variables



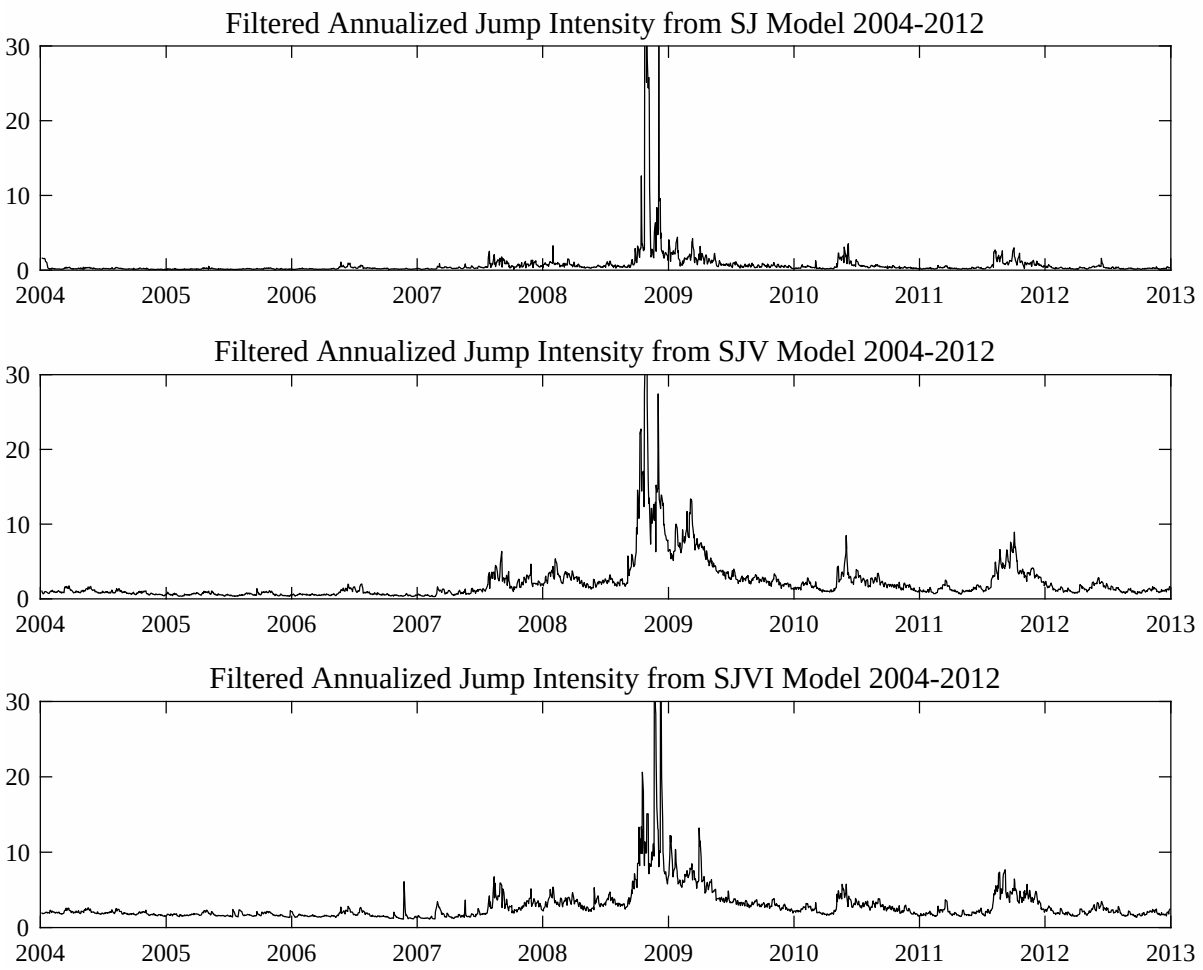
Notes: This figure plots four daily time series of selected variables for the S&P 500 index. The sample period is from January 2, 2004 to December 31, 2012. The top-left panel plots the daily returns on the S&P 500 index. In the top-right panel, we plot the annualized illiquidity measure calculated as the equally weighted average effective spread of intraday trades across firms constituting the S&P 500 index. The bottom-left panel plots the annualized jump-robust variance, *MinRV*, estimated using the one-minute grid returns of S&P 500 cash index. It is calculated following the approach of Andersen, Dobrev, and Schaumburg (2012). In the bottom-right panel, we plot the annualized realized jump variation, *RJV*, of daily index returns.

Figure 2.2: Implied Volatilities of OTM and ATM Options and Spot Volatility



Notes: In the top panel, we plot the daily implied volatilities of OTM put options written on the S&P 500 index from January 2, 2004 to December 31, 2012. In the second panel, we plot the implied volatilities of ATM call options. Both options are chosen to have the time to maturity to be closest to 30 calendar days. OTM options are chosen to have forward price-to-strike ratio to be closest to 0.95 while ATM options have the same ratio being closest to 1. The last panel plots the time series of spot volatility measure constructed using one-minute grid of intraday returns at 4:30 p.m. each day, following Andersen, Fusari, and Todorov (2015b).

Figure 2.3: Filtered Jump Intensity: λ_t

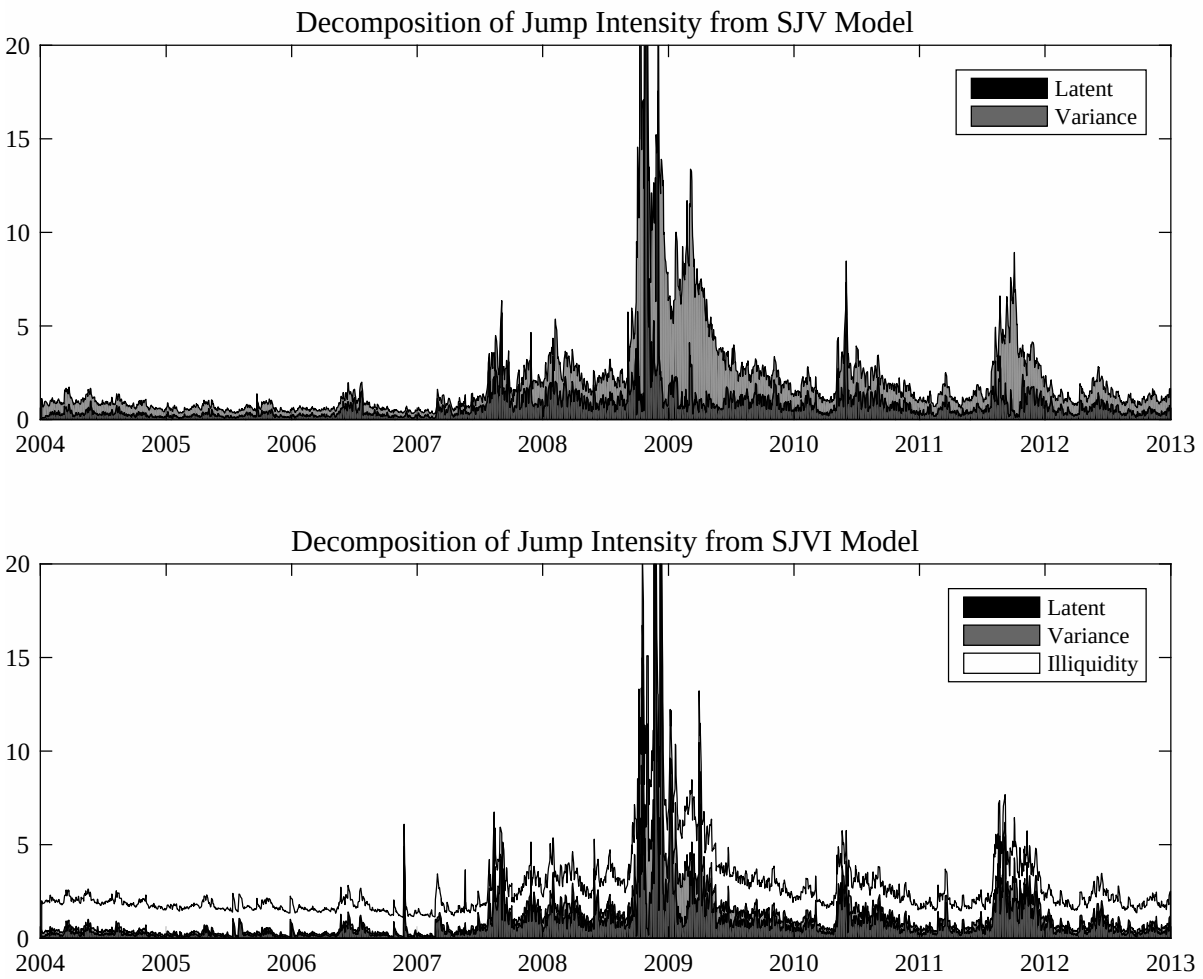


Notes: We plot daily annualized jump intensities λ_t filtered for the three models that we study from January 2, 2004 to December 31, 2012. The jump intensity specifications in the three models can be summarized as follows:

$$\begin{aligned} \text{SJ} \quad : \quad \lambda_t &= \Psi_t \\ \text{SJV} \ : \quad \lambda_t &= \Psi_t + \gamma_V V_t \\ \text{SJVI} : \quad \lambda_t &= \Psi_t + \gamma_V V_t + \gamma_L L_t. \end{aligned}$$

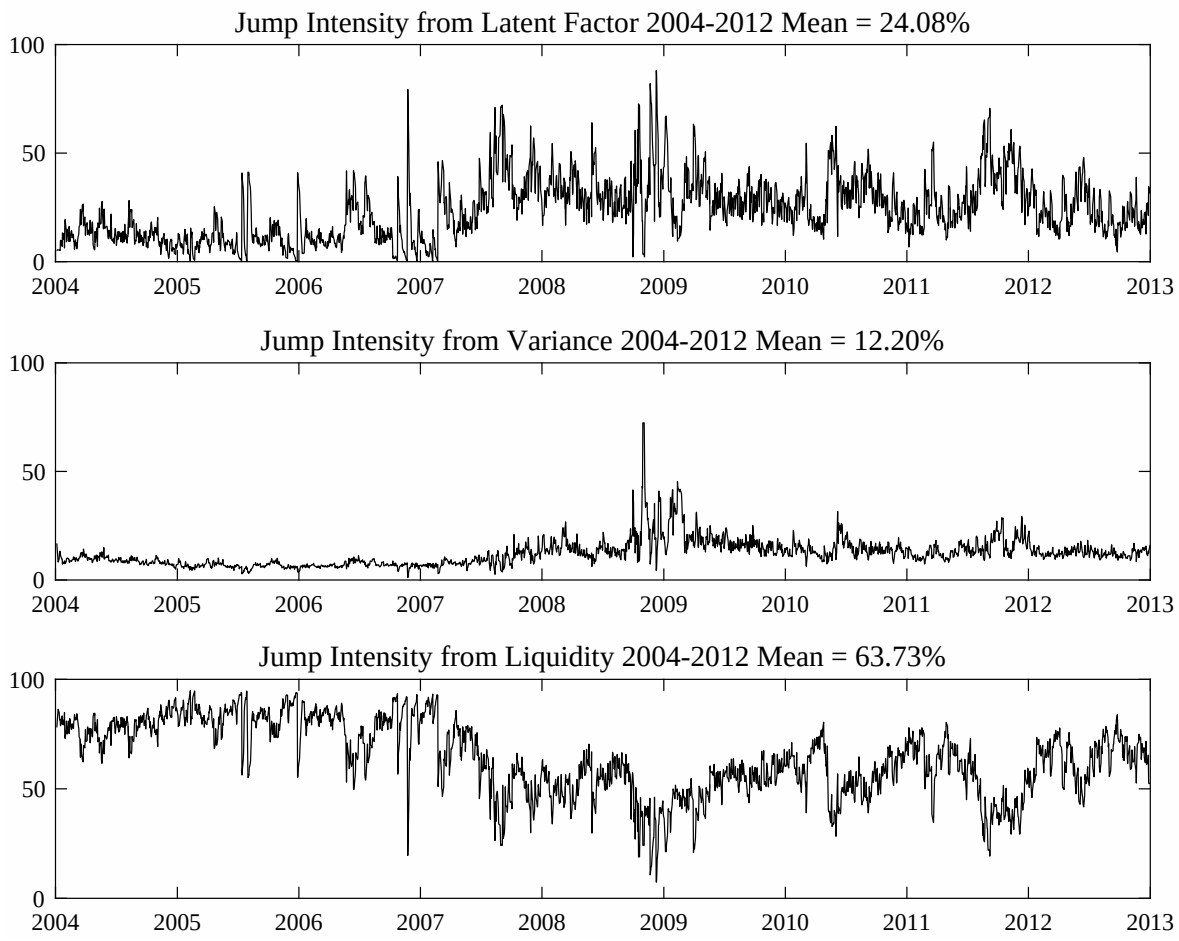
The top panel corresponds to the SJ model that has jump intensities solely driven by a latent jump intensity term; the middle panel corresponds to the SJV model that has jump intensity being driven by latent stochastic jump intensity and variance; and the bottom panel corresponds to the SJVI model that has jump intensity being driven by latent stochastic jump intensity, variance, and illiquidity.

Figure 2.4: Decomposition of Jump Intensity: SJV vs. SJVI



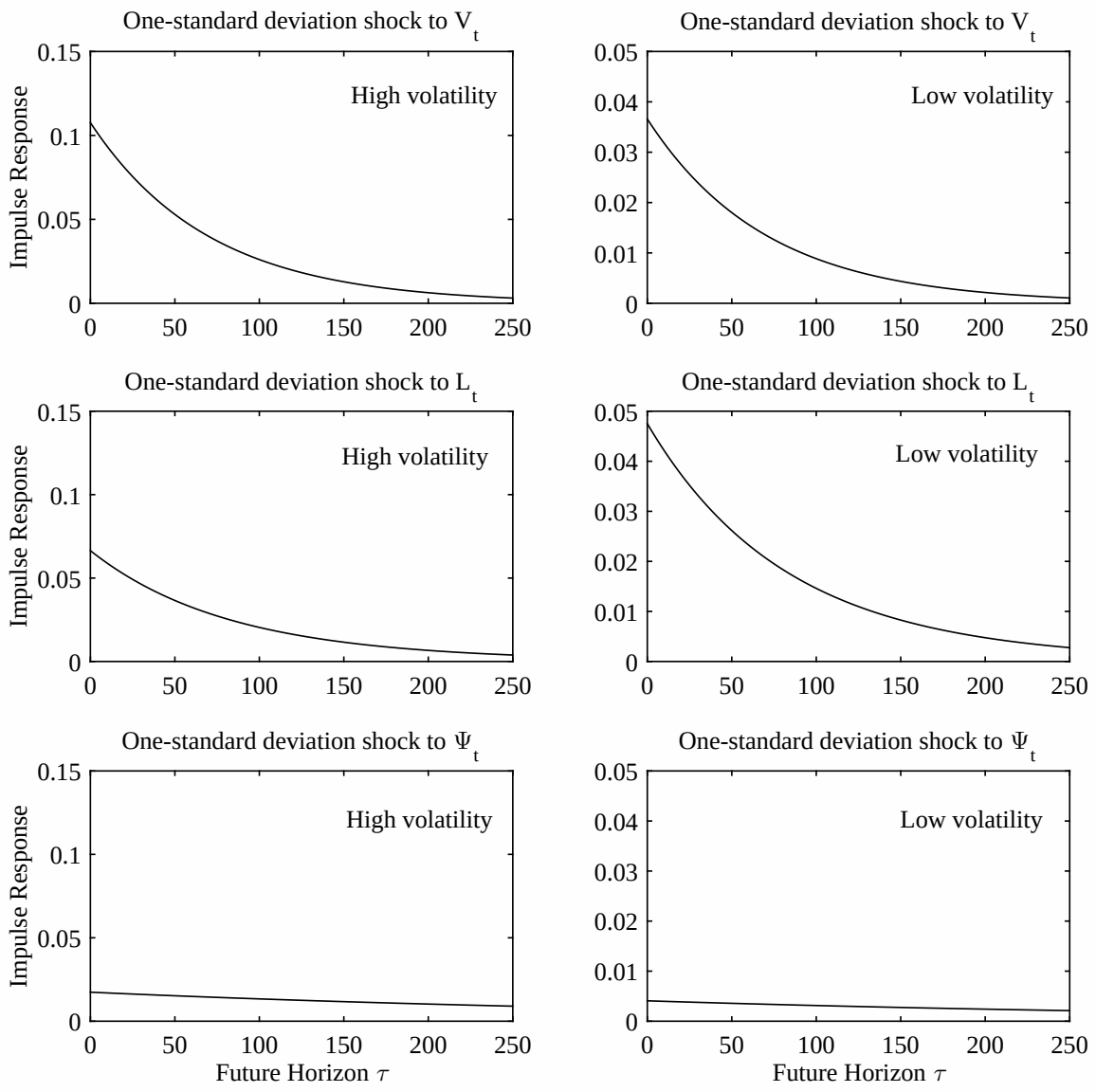
Notes: We plot the decomposition of daily annualized jump intensities λ_t filtered from the SJV model (top panel) and the SJVI model (bottom panel) from January 2, 2004 to December 31, 2012. The top panel decomposes daily jump intensity dynamics of the SJV model into the portion coming from the latent stochastic jump intensity term, Ψ_t, and the portion that is due to the daily spot variance, $\gamma_V V_t$. In the bottom panel, we decompose daily jump intensity dynamics of the SJVI model into the portion coming from the latent stochastic jump-intensity-specific term Ψ_t, the portion that is due to the daily spot variance, $\gamma_V V_t$, and the portion that is due to daily spot market illiquidity, $\gamma_L L_t$.

Figure 2.5: Relative Contribution to Jump Intensity: SJVI Model



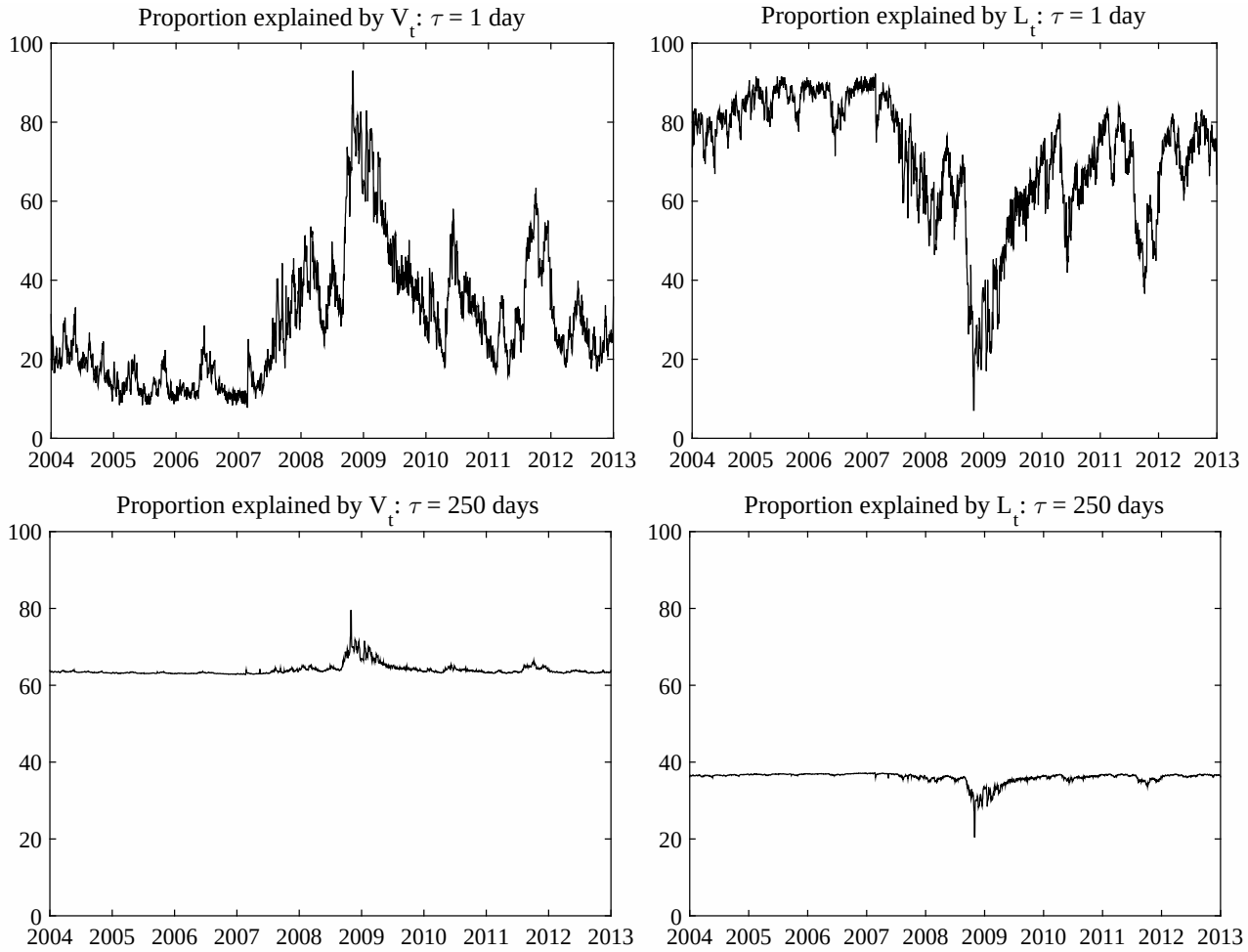
Notes: We plot the breakdown of daily annualized jump intensity, $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$, filtered from the SJVI model from January 2, 2004 to December 31, 2012. The top panel plots the percentage contribution coming from the latent stochastic jump intensity term, Ψ_t/λ_t. The middle panel plots the contribution coming from the variance term, $\gamma_V V_t/\lambda_t$. The bottom panel plots the contribution from the illiquidity term, $\gamma_l L_t$.

Figure 2.6: Impulse Response Function of λ_τ: High vs. Low Volatility Days



Notes: This figure plots the IRF describing the impact of one-standard-deviation shock to V_t, L_t, and Ψ_t on the jump intensity $\lambda_{t+\tau}$ in the τ days ahead. The x-axis displays the horizon τ in number of days and the y-axis corresponds to the response of $\lambda_{t+\tau}$. We plot the IRF from two days where the spot volatility levels are relatively high (left-column panels) and low (right-column panels). In each column, the top, middle, and bottom panels plot the IRF examining the impact of shocks to the spot volatility, spot illiquidity, and the latent state variable, respectively. All model parameters and filtered state variables are taken from the SJVI model estimates in Table 2.3. The date with high spot volatility (26.9%) corresponds to March 11, 2009. The date with log spot volatility (9.2%) corresponds to January 8, 2004.

Figure 2.7: Proportion of Forecast Error Variance ($\hat{\epsilon}_{\lambda,t+\tau}$) Explained by V_t and L_t



Notes: This figure plots the time series of FEVD for the jump intensity $\lambda_{t+\tau}$ at the short- and long-run horizons. The top-row panels plot FEVD for the short-run horizon, i.e., $\tau = 1$ day, while the bottom-row panels plot FEVD at the long-run horizon, i.e., $\tau = 250$ days. The y-axis displays the proportion of the forecast error variance explained by the factors. The forecast error for the jump intensity τ days ahead conditional on time t is defined as $\hat{\epsilon}_{\lambda,t+\tau} = \lambda_\tau - E_t[\lambda_\tau]$. We decompose the variance of the forecast error $\text{Var}_t\left[\hat{\epsilon}_{\lambda,t+\tau}\right]$ into components associated with shocks to the illiquidity factor L_t (right-column panels), and the spot variance V_t (left-column panels). The contribution of the latent state variable Ψ_t in the variance $\hat{\epsilon}_{\lambda,t+\tau}$ is very small and for brevity, is not reported here. All model parameters and filtered state variables are taken from the SJVI model estimates in Table 2.3.

Appendix

Appendix 2A: Risk-Neutral Skewness Measure

We use the model-free methodology implemented by Bakshi and Madan (2000) and Kozhan, Neuberger, and Schneider (2014), among others, to compute the risk-neutral moments. A key insight of this approach is that one can replicate any desired payoff by designing a portfolio of OTM European call and put options over a continuum of strike prices.

For an overview of this approach, let S denote the underlying asset value and let $\mathcal{G}[S]$ denote the payoff at maturity τ for a generic contingent claim written on S. By discounting the contingent claim with the risk-free rate r, its price can be evaluated under the risk-neutral expectation as $\mathbb{E}_t^{\mathbb{Q}}\{e^{-r\tau}\mathcal{G}[S]\}$. Bakshi and Madan (2000) show that for any twice-continuously differentiable payoff function $\mathcal{G}[S]$ with bounded expectation, the price of this contingent claim contract can be spanned according to the formula

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}\{e^{-r\tau}\mathcal{G}[S]\} &= e^{-r\tau}(\mathcal{G}[\overline{S}]-\overline{S}\mathcal{G}_S[\overline{S}])+\mathcal{G}_S[\overline{S}]S_t+\int_{\overline{S}}^{\infty}\mathcal{G}_{SS}[K]C(t,\tau;K)dK \\ &\quad+\int_0^{\overline{S}}\mathcal{G}_{SS}[K]P(t,\tau;K)dK, \end{aligned} \tag{2.40}$$

where $\mathcal{G}_s[\overline{S}]$ and $\mathcal{G}_{ss}[K]$ represent the first and second derivatives of the payoff function $\mathcal{G}$ evaluated at some asset value $\overline{S}$ and at the strike price K, respectively. The above equation shows that the contingent claim price can be replicated using a portfolio consisting of a risk-free bond, an underlying asset, and OTM calls and puts. The integrals in equation (2.40) can be evaluated numerically. We use a cubic spline method to calculate the integrals across moneyness.

To construct higher risk-neutral moments, we focus on the payoff function $\mathcal{G}$ with power contracts. That is,

$$\mathcal{G}[S] = \begin{cases} r_{t,\tau}^2 \text{ the volatility contract} \\ r_{t,\tau}^3 \text{ the cubic contract,} \end{cases} \tag{2.41}$$

where $r_{t,\tau}$ denotes the log-return of asset price S from time t to $t+\tau$. The risk-neutral volatility and skewness are then computed as

$$Vol_{t,\tau}^{\mathbb{Q}} = \{\mathbb{E}_t^{\mathbb{Q}}[(r_{t,\tau}-\mathbb{E}_t^{\mathbb{Q}}[r_{t,\tau}])^2]\}^{1/2}, \tag{2.42}$$

$$Skew_{t,\tau}^{\mathbb{Q}} = \frac{\mathbb{E}_t^{\mathbb{Q}}[(r_{t,\tau}-\mathbb{E}_t^{\mathbb{Q}}[r_{t,\tau}])^3]}{\{\mathbb{E}_t^{\mathbb{Q}}[(r_{t,\tau}-\mathbb{E}_t^{\mathbb{Q}}[r_{t,\tau}])^2]\}^{3/2}}. \tag{2.43}$$

We obtain data on S&P 500 index options between 2004 and 2012 from OptionMetrics. We use the average of the bid and ask quotes for each option contract and filter out options with bids of \$0 as well as those whose average quotes are less than \$3/8. We also filter out quotes that do not satisfy standard no-arbitrage conditions. Finally, we eliminate in-the-money options because they are less liquid than OTM and ATM options. We only estimate the moments for days that have at least two OTM call prices and two OTM put prices available. Finally, for any given maturity of interest, i.e., 30-day, we implement a linear interpolation to calculate the corresponding risk-neutral moments.

Appendix 2B: Realized Skewness Measure

We construct the daily realized skewness measure, *RSkew*, following the method in Amaya, Christoffersen, Jacobs, and Vasques (2015), which has been shown to significantly predict stock returns. This

realized skewness is calculated using one-minute log returns of the S&P500 cash index as follows:

$$RSkew_t^N = \frac{\sqrt{N}\sum_{i=1}^{N} r_{i,t}^3}{(RV_t^N)^{3/2}}, \tag{2.44}$$

where N is the number of time intervals in a trading day. As N goes to infinity, the above two measures converge to the cubic variations of jump component in the daily return, i.e., the diffusive component is excluded in their measurement.

Appendix 2C: Realized Skewness Regression

We also examine the impact of market illiquidity on daily realized skewness measure, *RSkew*. Unlike risk-neutral skewness that represents a forward-looking measure of the stock market crash risk, *RSkew* is calculated using historically observed high-frequency intraday index returns. Therefore, a more negative daily realized skewness level would indicate an increasing probability that a crash in the stock market has occurred during that trading day.

We estimate a time-series regression for the change in realized skewness, $\Delta RSkew_{t+1}$, similar to the general model shown in equation (2.7). However, we use a predictive regression model for the change in realized skewness by lagging all independent variables by one day. This is because *RSkew* is calculated from intraday trades observed during the day, which is the same data period used for calculating *MinRV* and *ILQ*.[18] This concern, however, does not apply to the risk-neutral skewness regression because *RNSkew* is calculated using end-of-day option prices and is derived from a different data source.

Table 2A.1 reports four sets of regression results on changes in daily realized skewness. We use the ARMA model with $p = 1$ in the autoregressive term and $q = 2$ in the error term. These lags are determined by the LjungBox test. The results shown in Table 2A.1 strongly support the findings in Table 2.2, which are obtained using daily changes in risk-neutral skewness. That is, an increase in market illiquidity is negatively related to the realized skewness. Column (2) shows the negative coefficient on $\Delta MinRV$ is negligible in magnitude as well as in statistical significance. However, when both $\Delta MinRV$ and ΔILQ are added to the regression model, the coefficient estimate on $\Delta MinRV$ becomes positive and statistically significant. These findings are highly consistent with the results obtained in Table 2.2. Therefore, we find the effect of market illiquidity on crash probability is robust to whether we measure the stock market crask risk using the forwarding-looking risk-neutral skewness or the historical realized skewness.

Appendix 2D: High Frequency Spot Variance Measure

Following Andersen, Fusari, and Todorov (2015b) and Mancini (2009), we construct the consistent estimator of spot variance at the end of each trading day using the one-minute grid of S&P 500 futures returns as follows:

$$\hat{V}_t^{(n,m_n)} = \frac{n}{m_n}\sum_{i=n-m_n+1}^{n} (r_{i,t})^2 \mathbb{I}(|r_{i,t}| \leq \alpha n^{-\omega}). \tag{2.45}$$

We use one-minute-grid returns over 6.5 hours in a trading day, thus resulting in $n = 390$ observations. The value of m_n is set to be 75% of n for each day. Other tuning parameters are set as follows: $\alpha = 4\sqrt{BPV_t}$ and $\omega = 0.49$ where BPV denotes the bi-power variation of day t computed using full one-minute grid of returns.

[18] Our conclusion is unaffected when we use a contemporaneous regression instead of a predictive regression.

Appendix 2E: Coefficients in the Affine Characteristic Function

The model that we study is casted in affine framework, the conditional characteristic function is exponential affine in the state variables following Duffie, Pan, and Singleton (2000). Its function form is given by

$$E_t[\exp(i\phi \log(S_T))] = \exp\left(\alpha(\tau) + \beta_0(\tau)\log(S_t) + \beta_1(\tau)V_t + \beta_2(\tau)L_t + \beta_3(\tau)\Psi_t\right) \tag{2.46}$$

We use the notation $\tau = T - t$ for simplicity. The coefficients satisfy the following system of Ricatti ordinary differential equation (ODE) with the boundary conditions $\beta_0(0) = i\phi$ and $\alpha(0) = \beta_1(0) = \beta_2(0) = \beta_3(0) = 0$

$$\begin{aligned}
\frac{d\beta_0}{d\tau} &= 0 \\
\frac{d\alpha}{d\tau} &= ir\phi + (\kappa_V\theta_V + \gamma\kappa_L\theta_L)\beta_1 + \kappa_L\theta_L\beta_2 + \kappa_\Psi\theta_\Psi\beta_3 \\
\frac{d\beta_1}{d\tau} &= \frac{1}{2}\xi_V^2\beta_1^2 + (\xi_V\rho i\phi - \kappa_V)\beta_1 + (\frac{1}{2}(i\phi)^2 - (\frac{1}{2} + \gamma_v\xi)i\phi + \gamma_v\theta_u) \\
\frac{d\beta_2}{d\tau} &= \frac{1}{2}\xi_L^2\beta_2^2 + (\gamma\xi_L^2\beta_1 - \kappa_L)\beta_2 + (\frac{1}{2}\gamma^2\xi_L^2\beta_1^2 - \gamma\kappa_L\beta_1 - \gamma_l\xi i\phi + \gamma_l\theta_u) \\
\frac{d\beta_3}{d\tau} &= \frac{1}{2}\xi_\Psi^2\beta_3^2 - \kappa_\Psi\beta_3 + \theta_u - \xi i\phi
\end{aligned}$$

where $\theta_u = (e^{\theta i\phi + \frac{1}{2}\delta^2(i\phi)^2} - 1)$. Equations for β_0, β_1, and β_3 can be solved analytically in terms of elementary functions while α and β_2 need to be solved numerically. We employ fourth-order Runge-Kutta method with the step size of $\Delta t = 1/252$.

Appendix 2F: Discretization of Daily Returns and Estimation

We apply daily discretization to the physical return process in (2.33). This yields

$$r_{t+1} \simeq (r + (\nu_S - \frac{1}{2})\hat{V}_t - \xi^P\hat{\lambda}_t)\Delta t + \sqrt{\hat{V}_t}\sqrt{\Delta t}\epsilon_t + \sum_{i=1}^{N_t} y_{i,t}, \tag{2.47}$$

where $\nu_S = \sqrt{1-\rho^2}\nu_1 + \rho\nu_v$, and ϵ_t is the standard normal innovation. The jump component is represented a compound Poisson process $\sum_{i=1}^{N_t} y_{i,t}$, where N_t is the number of jump arrival with intensity λ_t on day t, and $y_{i,t}$ is i.i.d. normal with mean θ^P and variance δ^2. Conditional on the number of jumps $N_t = j$, we can write the likelihood as conditionally normal, thus, the daily return likelihood can be analytically computed.

Appendix 2G: Impulse Response Function

In this section we construct the impulse response function of the discretized SJVI model under the physical measure. We follow the same Euler-discretization scheme applied to the UKF procedure; see the main text. The discretized system under the $\mathbb{P}$ measure is written as

$$\begin{aligned}
V_{t+1} &= V_t + \kappa_V(\theta_V - V_t)\Delta t + \gamma\kappa_L(\theta_L - L_t)\Delta t + \xi_V\sqrt{\Delta t V_t}\epsilon_{t+1}^1 + \gamma\xi_L\sqrt{\Delta t L_t}\epsilon_{t+1}^2 \\
L_{t+1} &= L_t + \kappa_L(\theta_L - L_t)\Delta t + \xi_L\sqrt{\Delta t L_t}\epsilon_{t+1}^2 \\
\Psi_{t+1} &= \Psi_t + \kappa_\Psi(\theta_\Psi - \Psi_t)\Delta t + \xi_\Psi\sqrt{\Delta t \Psi_t}\epsilon_{t+1}^3,
\end{aligned}$$

where error terms ϵ^i_{t+1}, for $i = 1$ to 3, are i.i.d. standard normal with the step size $\Delta t = 1/252$.

We next expand the above system and rewrite them in terms of past innovation terms only. The expansion for L_{t+1} and Ψ_{t+1} is straightforward and is given by

$$L_{t+1} = \theta_L + \sum_{j=0}^{\infty} \rho_L^j \eta^L_{t+1-j} \tag{2.48}$$

$$\Psi_{t+1} = \theta_\Psi + \sum_{j=0}^{\infty} \rho_\Psi^j \eta^\Psi_{t+1-j}, \tag{2.49}$$

where the new coefficients are $\rho_L = 1 - \kappa_L \Delta t$, $\rho_\Psi = 1 - \kappa_\Psi \Delta t$, $\eta^L_{t+1} = \xi_L \sqrt{\Delta t} \sqrt{L_t} \epsilon^2_{t+1}$, and $\eta^\Psi_{t+1} = \xi_\Psi \sqrt{\Delta t} \sqrt{\Psi_t} \epsilon^3_{t+1}$. The expansion for V_{t+1} is a bit more involved because there are two independent shocks. After some algebraic work, we obtain

$$V_{t+1} = \theta_V + \gamma \eta^L_{t+1} + \sum_{j=0}^{\infty} \rho_V^j \eta^V_{t+1-j} + \gamma \sum_{j=1}^{\infty} \left[\frac{\rho_L^{j+1} - \rho_V^{j+1} - \left(\rho_L^j - \rho_V^j \right)}{\rho_L - \rho_V} \right] \eta^L_{t+1-j}, \tag{2.50}$$

where $\rho_V = 1 - \kappa_V \Delta t$ and $\eta^V_{t+1} = \xi_V \sqrt{\Delta t} \sqrt{V_t} \epsilon^1_{t+1}$.

Plugging the expansions shown in equations (2.48)–(2.50) into the jump intensity dynamic, $\lambda_{t+1} = \Psi_{t+1} + \gamma_V V_{t+1} + \gamma_L L_{t+1}$, we can express λ_{t+1} only in terms of shocks to the system as

$$\begin{aligned} \lambda_{t+1} &= \theta_\Psi + \gamma_L \theta_L + \gamma_V \theta_V + \sum_{j=0}^{\infty} \rho_\Psi^j \eta^\Psi_{t+1-j} + \gamma_V \sum_{j=0}^{\infty} \rho_V^j \eta^V_{t+1-j} + \left(\gamma_L + \gamma_V \gamma \right) \eta^L_{t+1} \\ &+ \sum_{j=1}^{\infty} \left(\gamma_L \rho_L^j + \gamma_V \gamma \left[\frac{\rho_L^{j+1} - \rho_V^{j+1} - \left(\rho_L^j - \rho_V^j \right)}{\rho_L - \rho_V} \right] \right) \eta^L_{t+1-j}, \end{aligned} \tag{2.52}$$

where η^V_{t+1}, η^L_{t+1}, and η^Ψ_{t+1} represent shocks specific to the variance, illiquidity, and latent factors, respectively. Thus, the impulse response of a specific shock for τ periods ahead can be calculated by simply setting $j = \tau$ in the coefficient associated with that specific shock in equation (2.52).

Appendix 2H: Forecast Error Variance Decomposition

This section presents the variance decomposition of forecast error in the conditional jump intensity. The error from forecasting the jump intensity $\lambda_{t+\tau}$ with τ-period horizon conditional on day t is defined as

$$\hat{\epsilon}_{\lambda, t+\tau} = \lambda_{t+\tau} - E_t[\lambda_{t+\tau}]. \tag{2.53}$$

In the SJVI model, changes in jump intensity are driven by shocks to the spot illiquidity L_t, the latent factor Ψ_t, and the spot variance V_t. Under a mild assumption of zero autocorrelation among the three shocks, we can approximate the variance in the forecast error, $\hat{\epsilon}_{\lambda, t+\tau}$, associated with each shock as:

$$\begin{aligned} \mathrm{Var}_t \left[\hat{\epsilon}_{\lambda, t+\tau} \right] \approx & \left\{ \begin{array}{c} \left(\gamma_V + \gamma_L \gamma \right)^2 \mathrm{Var}_t \left[\eta^L_{t+\tau} \right] \\ + \sum_{j=1}^{\tau-1} \left(\gamma_L \rho_L^j + \gamma_V \gamma 1_{[\tau>1]} \left[\frac{\rho_L^{j+1} - \rho_V^{j+1} - \left(\rho_L^j - \rho_V^j \right)}{\rho_L - \rho_V} \right] \right)^2 \mathrm{Var}_t \left[\eta^L_{t+\tau-j} \right] \end{array} \right\} \\ & + \left\{ \sum_{j=0}^{\tau-1} \left(\rho_\Psi^j \right)^2 \mathrm{Var}_t \left[\eta^\Psi_{t+\tau-j} \right] \right\} + \left\{ \sum_{j=0}^{\tau-1} \left(\gamma_V \rho_V^j \right)^2 \mathrm{Var}_t \left[\eta^V_{t+\tau-j} \right] \right\}. \end{aligned} \tag{2.54}$$

Expressions in the first, second, and third brackets in equation (2.54) represent the approximate forecast error variance that is associated with shocks to the illiquidity L_t, the latent factor Ψ_t, and the variance

V_t, respectively. The notations that we use in equation (2.54) are shown in Appendix 2.6.

We obtain the closed-form expression for each contribution factor in equation (2.54) by computing the conditional variance of each shock explicitly. For example, the proportion of the error variance explained by the variation in illiquidity is given by

$$\frac{\left\{\begin{array}{c}\left(\gamma_V+\gamma_L\gamma\right)^2\left[\xi_L^2\Delta t\left(\theta_L+\rho_L^{\tau-1}\left(L_t-\theta_L\right)\right)\right]\\ +\sum_{j=1}^{\tau-1}\left(\gamma_L\rho_L^j+\gamma_V\gamma 1_{[\tau>1]}\left[\frac{\rho_L^{j+1}-\rho_V^{j+1}-\left(\rho_L^j-\rho_V^j\right)}{\rho_L-\rho_V}\right]\right)^2\left[\xi_L^2\Delta t\left(\theta_L+\rho_L^{\tau-j-1}\left(L_t-\theta_L\right)\right)\right]\\ +\xi_V^2\Delta t\sum_{j=0}^{\tau-1}\left(\gamma_V\rho_V^j\right)^2\left(\gamma\left(\rho_L-1\right)\frac{\rho_L^{\tau-j-1}-\rho_V^{\tau-j-1}}{\rho_L-\rho_V}\left(L_t-\theta_L\right)\right)\end{array}\right\}}{\mathrm{Var}_t\left[\hat{\epsilon}_{\lambda,t+\tau}\right]}. \quad (2.56)$$

The proportion explained by the variation in diffusive variance, V_t, can be written as

$$\frac{\left\{\xi_V^2\Delta t\sum_{j=0}^{\tau-1}\left(\gamma_V\rho_V^j\right)^2\left(\theta_V+\rho_V^{\tau-j-1}\left(V_t-\theta_V\right)\right)\right\}}{\mathrm{Var}_t\left[\hat{\epsilon}_{\lambda,t+\tau}\right]}. \quad (2.58)$$

Lastly, the proportion explained by the variation in latent factor, Ψ_t, is given by

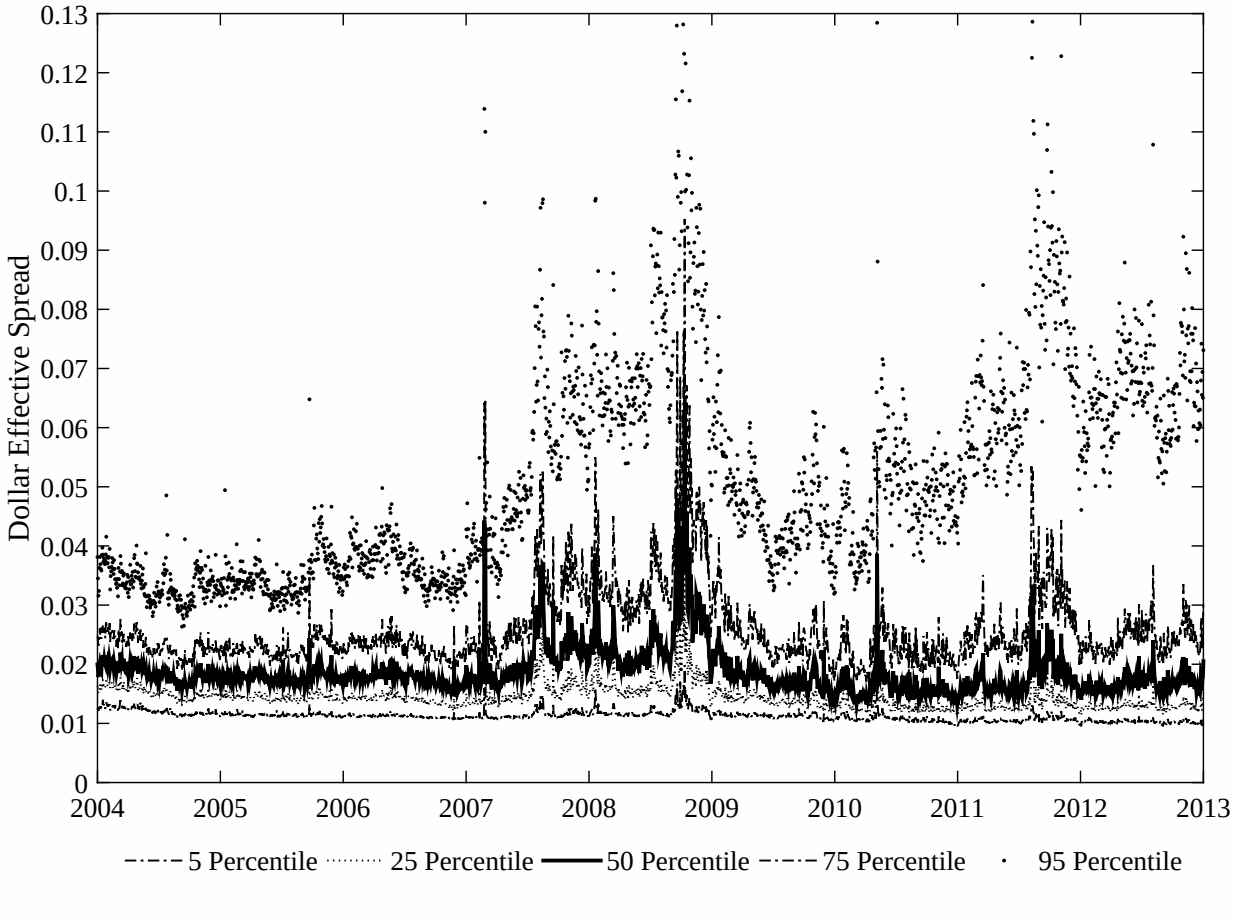
$$\frac{\left\{\sum_{j=0}^{\tau-1}\left(\rho_\Psi^j\right)^2\left[\xi_\Psi^2\Delta t\left(\theta_\Psi+\rho_\Psi^{\tau-j-1}\left(\Psi_t-\theta_\Psi\right)\right)\right]\right\}}{\mathrm{Var}_t\left[\hat{\epsilon}_{\lambda,t+\tau}\right]}. \quad (2.60)$$

Table 2A.1: Regression Model on Changes in Realized Skewness

	Change in Realized Skewness: $\Delta RSkew_{t+1}$			
	(1)	(2)	(3)	(4)
$\Delta MinRV_t$		-0.066		0.625*
		(-0.33)		(1.94)
ΔILQ_t			-0.580***	-1.261***
			(-2.63)	(-2.75)
$Return_t$	0.010	0.006	-0.102	-0.189**
	(0.12)	(0.08)	(-1.37)	(-2.03)
$\Delta RSkew_t$	0.766***	0.767***	0.775***	0.773***
	(13.62)	(13.91)	(14.16)	(14.10)
AICC	1.630	1.642	1.638	1.636
R^2	54.2%	54.2%	54.3%	54.4%

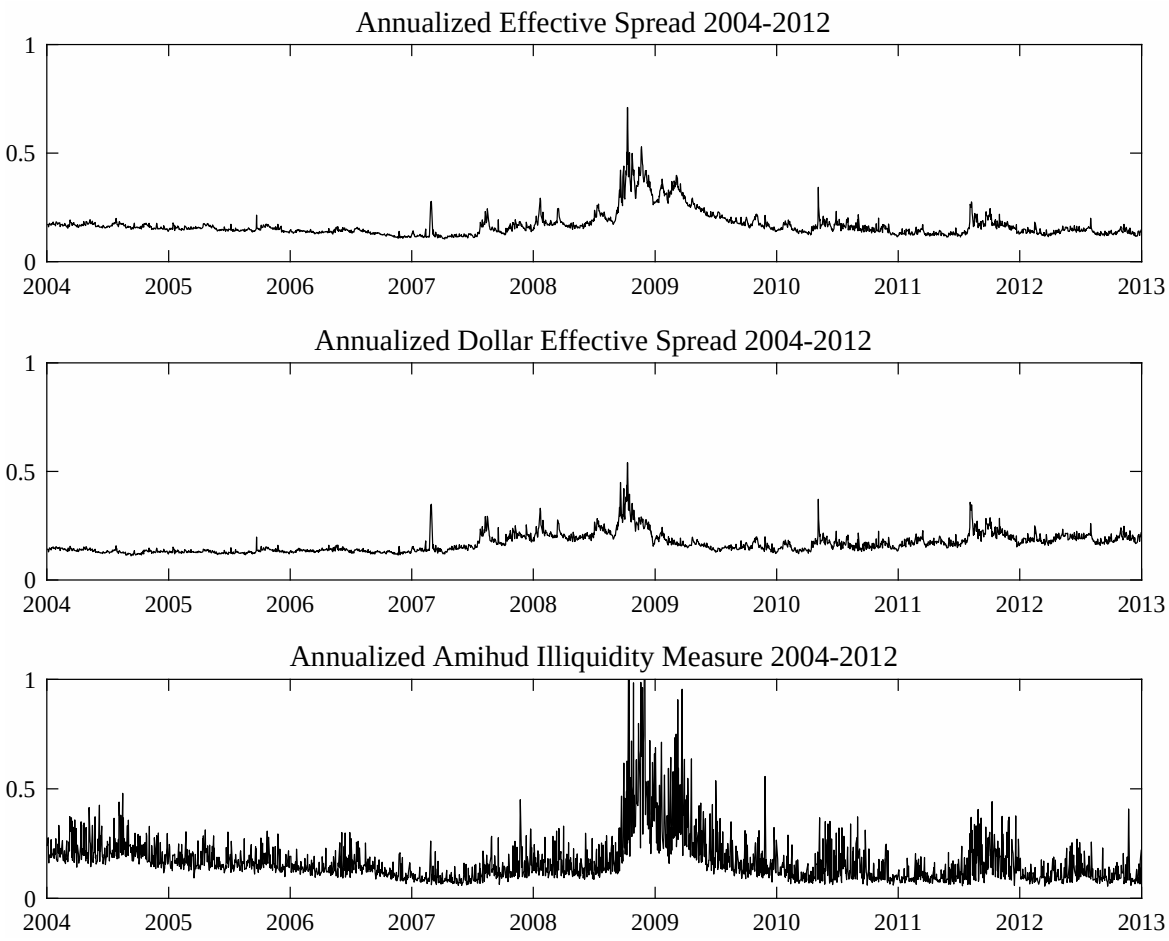
Notes: We report regression results on daily changes in realized skewness of S&P 500 index returns, $\Delta RSkew_{t+1}$. The sample period is from January 2, 2004 to December 31, 2012. The daily realized skewness measure, $RSkew$, on each trading day is constructed from high-frequency data following the method in Amaya, Christoffersen, Jacobs, and Vasquez (2015). The independent variables include lagged change in realized variance estimator, $\Delta MinRV$, from Andersen, Dobrev, and Schaumburg (2012); change in market illiquidity proxy; ΔILQ, measured by daily averaged effective spreads across firms in the S&P 500 constituents; *Return*, log S&P 500 return. We lag all independent variables by one day because the daily realized skewness measure is calculated from intraday trades observed over each day, which overlap with the data period used for constructing independent variables. Each specification is estimated using maximum likelihood. We use an ARMA(1,2) structure in the regression model, which is determined by the LjungBox test to sufficiently remove cross-correlations in the residuals. We control for seasonality due to the day-of-the-week effect. We report coefficient estimates on the autoregressive term. We do not report coefficient estimates on the moving-average error term and day-of-the-week fixed effects for brevity. Robust t-statistic is reported in parenthesis below each parameter estimate. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. The last row reports regression diagnostics based on the Akaike information criterion (AICC) and R^2 metric.

Figure 2A.1: Percentiles of Dollar Effective Spread: S&P 500 Constituents



Notes: We plot the 5th, 25th, 50th, 75th, and 95th percentiles of the daily effective spreads (in dollars) from the constituents of the S&P 500 index. For the majority of firms in the S&P 500 index, trades are executed with an effective spread above one cent, which is the minimum tick size in the NYSE. This finding holds throughout our sample period going from January 2, 2004 to December 31, 2012.

Figure 2A.2: Alternative Illiquidity Measures



Notes: We plot three daily market illiquidity measures from January 2, 2004 to December 31, 2012 that we use to verify the robustness of our results. The top panel plots the annualized relative effective spread measure defined in equation (2.1), which is the main illiquidity measure that we use in the paper. The bottom two panels plot the annualized illiquidity measure that we use to verify the robustness of our results. In the middle panel, we plot the daily market illiquidity measure calculated from dollar effective spreads shown in equation (2.38). In the bottom panel, we plot the Amihud (2002) illiquidity measure. It is calculated as the equally weighted average Amihud illiquidity measure of all securities constituting the S&P 500 index on each day; see equation (2.39). We normalize the dollar effective spread and Amihud illiquidity measures to have the same in-sample mean as the illquidity measure that we calculated using relative effective spreads.

Chapter 3

Option Valuation with Observable Volatility and Jump Dynamics

3.1 Introduction

State-of-the-art derivative valuation models assume that price changes in the underlying asset are driven by a diffusive component as well as a jump component.[1] The volatility of the diffusive component is typically assumed to be stochastic and the jump intensity is sometimes assumed to be constant. The econometric literature has developed powerful model-free methods for detecting statistically significant jumps and for separating the daily total diffusive volatility from the total quadratic variation via the use of high-frequency observations.[2] Our contribution is to combine these insights and develop a new derivative valuation model that directly uses the observable realized diffusive volatility and realized jump variation to model dynamics in the diffusive volatility and in the jump intensity. We cast our model within the broad class of affine discrete time models which implies that volatility and jump intensity filtering is straightforward and that derivative valuation can be done without relying on simulation-based methods. We develop a stochastic discount factor for the model that enables us to compute European option values using Fourier inversion of the conditional characteristic function.

The development of rigorous statistical foundations for the use of intraday returns to construct daily realized volatility measures is arguably one of the most successful branches of financial econometrics. For early references, see Andersen and Bollerslev (1998), Barndorff-Nielsen and Shephard (2002), Andersen, Bollerslev, Diebold, and Labys (2003), and Zhang, Mykland, and Aït-Sahalia (2005). For an early application of realized volatility in finance, see for example Bakshi, Cao, and Chen (1997).

The finance literature has recently developed models that use daily total quadratic variation from in-

[1] See for example Bates (2000), Bates (2012), Eraker (2004), Huang and Wu (2004), and Santa-Clara and Yan (2010).

[2] See Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen and Shephard (2006), Huang and Tauchen (2005), and the recent survey in Aït-Sahalia and Jacod (2012).

traday data to specify and estimate daily models of option valuation which outperform models estimated only on daily returns. See for example Stentoft (2015), Corsi, Fusari, and Vecchia (2013), Christoffersen, Feunou, Jacobs, and Meddahi (2014), and Amaya, Bégin, and Gauthier (2016). However, we are the first to develop an option valuation model with separate dynamics for observable realized diffusive volatility and realized jump variation.

The econometric literature has shown that decomposing total quadratic variation into its diffusive and jump variation parts leads to improved forecasts of future volatility. See for example Andersen, Bollerslev, and Diebold (2007), and Busch, Christensen, and Nielsen (2011).[3] Our goal is to assess if the improvements found in the volatility forecasting literature carry over to option valuation. We find that they do.

Our paper is part of a larger research agenda applying realized volatility measures in financial decision making. The development of tools for computing highly informative daily realized volatility from noisy intraday data is arguable one of the great success stories of financial econometrics. The application of realized volatility in option valuation is particularly intriguing. For many assets, including equities, the underlying contract is traded very actively at a high frequency whereas the option is typically traded much less often (see, for example, Christoffersen, Goyenko, Jacobs, and Karoui (2015)). This setting is ideal for the application of realized volatility which presumes that the frequency of interest (say daily) is lower than the frequency of observation (say each minute). Realized volatility and jumps, which are generally unobserved at the one-minute frequency, can be reliably estimated at the daily frequency. This insight enables the implementation of a new class of option valuation models which we develop in this paper.

When estimating the new model on returns, realized diffusive volatility, and realized jump variation we find that it outperforms standard benchmark models in the literature including the Heston and Nandi (2000) affine GARCH model which is a special case of our model. The general model also outperforms a special case that models only the total quadratic variation dynamic, as well a special case that assumes the entire quadratic variation is attributable to the jump component.

When estimating the new model on S&P500 index options as well as returns and realized variation measures and evaluating the option fit then the model again performs well. The option implied volatility root mean squared error of the new model is 17% below that of the affine GARCH model. The improvement in option fit arises in virtually all the moneyness, maturity and market volatility categories that we consider.

[3]Dynamic models for daily returns and volatility using high-frequency information have been developed in Forsberg and Bollerslev (2002), Engle and Gallo (2006), Bollerslev, Kretschmer, Pigorsch, and Tauchen (2009), Shephard and Sheppard (2010), and Hansen, Huang, and Shek (2012).

One key advantage of our approach is that we avoid the filtering issues that arise in related discrete time jump models, for example, Maheu and McCurdy (2004), Christoffersen, Jacobs, and Ornthanalai (2012), and Ornthanalai (2014) who either rely on particle filtering or ignore the impact of estimated state variables when constructing the likelihood. More generally, we argue that using high-frequency information to discern between daily jumps and diffusive volatility is likely to lead to a much more accurate identification of the two components than relying only on daily returns, or only on daily returns and options.

The remainder of the paper proceeds as follows: In Section 2 we briefly review the theory for separating diffusive volatility from jumps and we show the two time series for the S&P500 index which is the underlying asset in our empirical study. In Section 3 we develop the physical return process. Section 4 estimates the physical process on returns, realized bipower and jump variation measures. In Section 5, we derive an option valuation formula for the model. In Section 6 we estimate the model on options and analyzes its fit. Finally, Section 6 concludes. The proofs of our propositions are relegated to the appendices.

3.2 Daily Returns and Realized Variation Measures

In this section we first briefly review the key theoretical results that allows us to separate daily diffusive volatility and jump variation using intraday data. We then construct empirical measures of realized diffusive volatility and realized jumps and plot the daily realized variation series along with daily returns.

3.2.1 Separating Volatility and Jumps: Theory

Barndorff-Nielsen and Shephard (2004) assume the stock price follows a jump-diffusion process of the form

$$d\log(S_t) = \sqrt{V}_t dW_t + J_t dq_t \tag{3.1}$$

where q_t is a Poisson process with intensity $\lambda_{J\,t}$, and J_t is the normally distributed log jump size with mean μ_J and standard deviation σ_J. Under this very general assumption about the instantaneous return process, Barndorff-Nielsen and Shephard (2004) show the following limit result as the sampling frequency goes to infinity

$$\begin{aligned} RV_t &\rightarrow \int_{t-1}^{t} V_s ds + \int_{t-1}^{t} J_s^2 dq_s \\ RBV_t &\rightarrow \int_{t-1}^{t} V_s ds, \end{aligned} \tag{3.2}$$

where RV_t denotes realized variance measuring total quadratic variation, and RBV_t denotes bipower variation measuring diffusive volatility. These quantities will be defined in detail below. We can now define realized jump variation (RJV_t) using

$$RJV_t \equiv RV_t - RBV_t \rightarrow \int_{t-1}^{t} J_s^2 dq_s,$$

which provides the decomposition of total quadratic variation that we need.

The next step in our analysis is to construct empirical measures of RV_t, RBV_t, and RJV_t.

3.2.2 Separating Volatility and Jumps: Empirics

Our empirical investigation begins by obtaining intraday S&P500 cash index data from TickData.com. Using the last observation each minute we construct a grid of one-minute equity index values each day from which we compute five series of overlapping five minute log-returns. Each day we can compute five realized variance measures from the sum of squared five-minute returns. The five overlapping realized variance series are then averaged to obtain a single market microstructure robust measure of total quadratic variation as follows

$$RV_{t+1}^{'} = \frac{1}{5}\sum_{i=0}^{4} RV_{t+1}^{5,i} = \frac{1}{5}\sum_{i=0}^{4}\sum_{j=1}^{m/5} R_{t+(i+5j)/m}^2$$

where $R_{t+(i+5j)/m}$ denotes the j^{th} period 5-min intraday return, and m denotes the number of 1-minute returns available on day $t+1$. Following Hansen and Lunde (2005) the RV_{t+1} computed above is finally rescaled so that the average value of RV_{t+1} is equal to the sample variance of daily log-returns.

$$RV_{t+1} = \frac{\sum_{t=1}^{T} R_t^2}{\sum_{t=1}^{T} RV_t^{'}} RV_{t+1}^{'}$$

where $R_t = \log(S_t) - \log(S_{t-1})$ is the daily log return computed from closing prices.

Diffusive volatility is computed using realized bipower variation defined from

$$RBV_{t+1}^{'} = \frac{1}{5}\sum_{i=0}^{4} RBV_{t+1}^{5,i} = \frac{1}{5}\sum_{i=0}^{4}\frac{\pi}{2}\sum_{j=1}^{m/5-1}|R_{t+(i+5j)/m}||R_{t+(i+5(j+1))/m}|$$

Then, in order to ensure the empirical version of the theoretical relationship in equation (3.2) holds, namely,

$$RV_{t+1} = RBV_{t+1} + RJV_{t+1}$$

and also in order to ensure that $RJV_{t+1} \geq 0$, we use the following definitions,

$$\begin{aligned} RBV_{t+1} &= \min(RV_{t+1}, RBV_{t+1}^{'}) \\ RJV_{t+1} &= RV_{t+1} - RBV_{t+1} \end{aligned}$$

Figure 1 plots the four R_t (top left), RV_t (top right), RBV_t (bottom left), and RJV_t (bottom right) series from January 2, 1990 through December 31, 2013. Note from Figure 1 that the RV_t, RBV_t, and RJV_t series share broadly similar patterns including the fact that their largest values occur during the 2008 financial crises. This commonality suggests that when RV_t is high then both RBV_t and RJV_t are high and vice versa. Note also that RBV_t is an order of magnitude larger than RJV_t.

Figure 2 plots the sample autocorrelation functions for the four series. Note that, as expected, the autocorrelation of returns (top-left) are close to zero across lag orders. Also as expected, the autocorrelations of realized variance (top-right) and bipower variation (bottom-left) are both very high and statistically significant throughout the 60 trading-day period considered. More interestingly, the realized jump variation measure in the bottom-right panel shows strong evidence of persistence as well. To be sure, the autocorelations for realized jump variation are lower at short lags than for realized variance and bipower variation, but they are very persistent. It is thus clear that the realized jump measure requires a dynamic specification of its own and likely one that is different from the dynamic specification of bipower variation. Building a dynamic return model with such features is our next task.

3.3 A New Dynamic Model for Asset Returns

The goal of this section is to build a model for end-of-day t option valuation that incorporates the information in the R_t, RBV_t, and RJV_t series computed at the end of the day. We want to build a model in which state variables are explicitly filtered using our observables and in which option valuation

can be done without Monte Carlo simulation.

3.3.1 The Asset Return Process

Consider first the following generic specification of daily log returns

$$R_{t+1} = r + \left(\lambda_z - \tfrac{1}{2}\right) h_{z,t} + \left(\lambda_y - \xi\right) h_{y,t} + z_{t+1} + y_{t+1} \tag{3.3}$$

where r denotes the risk-free rate, and the first innovation, z_{t+1}, denotes a heteroskedastic Gaussian innovation

$$z_{t+1} = \sqrt{h_{z,t}}\varepsilon_{1,t+1}, \text{ with } \varepsilon_{1,t+1} \overset{iid}{\sim} N\left(0,1\right). \tag{3.4}$$

The second innovation, y_{t+1}, denotes a compound jump process

$$y_{t+1} = \sum_{j=0}^{n_{t+1}} x_{t+1}^{j}, \text{ with } x_{t+1}^{j} \overset{iid}{\sim} N(\theta, \delta^2), \tag{3.5}$$

where the number of Gaussian jumps per day is Poisson distributed

$$n_{t+1} \sim Ps\left(h_{y,t}\right). \tag{3.6}$$

Note that this general framework allows for dynamic volatility via $h_{z,t}$ and dynamic jump intensity via $h_{y,t}$. These dynamics still need to be specified and crucially for us they need to be linked with the daily realized bipower and jump variation measures.

Finally, note that in our timing convention, $h_{z,t}$ denotes the expected "diffusive" variance for day $t+1$ constructed at the end of day t. Similarly, $h_{y,t}$ denotes the expected number of jumps on day $t+1$ constructed at the end of day t.

The parameters we estimate on daily data in our discrete time model are reported in daily units below. When estimating continuous time models of the type in (3.1), the literature often reports annualized parameters. In that case we have the following mapping,

$$\theta \approx \frac{1}{252}\mu_J, \quad \delta \approx \frac{1}{\sqrt{252}}\sigma_J, \quad h_{y,t} \approx \frac{1}{252}\int_{t-1}^{t} \lambda_{J,s} ds.$$

3.3.2 Incorporating Realized Bipower and Jump Variation

Each day the realized bipower variation provides new information about diffusive volatility, $h_{z,t}$. However, RBV_{t+1} is measured with error and we therefore specify the following measurement equation

$$RBV_{t+1} = h_{z,t} + \sigma \left[\left(\varepsilon_{2,t+1} - \gamma \sqrt{h_{z,t}} \right)^2 - \left(1 + \gamma^2 h_{z,t} \right) \right], \tag{3.7}$$

where we have introduced a measurement error variable

$$\varepsilon_{2,t+1} \overset{iid}{\sim} N(0,1),$$

which has a correlation ρ with the diffusive return shock, $\varepsilon_{1,t+1}$, defined in equation (3.4). The innovation term inside the brackets in equation (3.7) is constructed to have zero mean ensuring that

$$E_t \left[RBV_{t+1} \right] = h_{z,t}.$$

Note also that equation (3.7) allows for a nonlinear impact of $\varepsilon_{2,t+1}$ on RBV_{t+1} via γ.

Our daily realized jump variation measure constructed from intraday data is naturally linked with the sum of squared daily jump variation in the model as follows:

$$RJV_{t+1} = \sum_{j=0}^{n_{t+1}} \left(x_{t+1}^j \right)^2.$$

This relationship implies that

$$E_t \left[RJV_{t+1} \right] = \left(\theta^2 + \delta^2 \right) h_{y,t},$$

where we have used the second moment of the Poisson distribution. Notice that while we do allow for measurement error in RBV we assume for simplicity that RJV is free of error. We plan to relax this assumption in future work.

3.3.3 Volatility and Jump Dynamics

We are now ready to specify the dynamics of the expected volatility and jump intensity. In the empirical sections below, we will focus on a special case of our modeling framework in which we simply pose that

$$h_{z,t+1} = \omega_z + b_z h_{z,t} + a_z RBV_{t+1}, \text{ and} \tag{3.8}$$

$$h_{y,t+1} = \omega_y + b_y h_{y,t} + a_y RJV_{t+1}. \tag{3.9}$$

Note that in this specification, $h_{z,t+1}$ and $h_{y,t+1}$ are both univariate $AR(1)$ processes, which we can write as

$$\begin{aligned} h_{z,t+1} &= \omega_z - a_z\sigma + \left(b_z + a_z - a_z\sigma\gamma^2\right) h_{z,t} + a_z\sigma\left(\varepsilon_{2,t+1} - \gamma\sqrt{h_{z,t}}\right)^2 \\ h_{y,t+1} &= \omega_y + b_y h_{y,t} + a_y \sum_{j=0}^{n_{t+1}} \left(x_{t+1}^j\right)^2. \end{aligned}$$

The dynamics in (3.8-3.9) imply that RBV_{t+1} and RJV_{t+1} are both univariate $ARMA(1,1)$ processes. We will refer to this as the BPJVM model.

3.3.4 The General Case

Our dynamic modelling framework is more general than the BPJVM model. Define the bivariate processes

$$\begin{aligned} h_t &\equiv (h_{z,t},\ h_{y,t})', \text{ and} \\ RVM_{t+1} &\equiv (RBV_{t+1},\ RJV_{t+1})'. \end{aligned}$$

The general dynamic vector process is then of the form,

$$h_{t+1} = \omega + bh_t + aRVM_{t+1},$$

where the parameter vector and matrices are

$$\omega = (\omega_z, \omega_y)',\ b = \begin{pmatrix} b_z & b_{z,y} \\ b_{y,z} & b_y \end{pmatrix},\ a = \begin{pmatrix} a_z & a_{z,y} \\ a_{y,z} & a_y \end{pmatrix}.$$

Note that by construction h_{t+1} is a vector autoregressive process of order one, $VAR(1)$, and RVM_{t+1} is a vector autoregressive moving average model, $VARMA(1,1)$. In particular, note that the expected

value of the vector h_{t+1} is

$$\begin{aligned} E_t\left[h_{t+1}\right] &= \omega + bh_t + a\begin{pmatrix} h_{z,t} \\ \left(\theta^2+\delta^2\right)h_{y,t} \end{pmatrix} \\ &\equiv \omega + \begin{bmatrix} b_z + a_z & b_{z,y} + \left(\theta^2+\delta^2\right)a_{z,y} \\ b_{y,z} + a_{y,z} & b_y + \left(\theta^2+\delta^2\right)a_y \end{bmatrix} h_t. \end{aligned}$$

Below we will focus on the BPJVM version of the model in which $a_{z,y} = a_{y,z} = b_{z,y} = b_{y,z} = 0$.

3.3.5 Expected Returns and Risk Premiums

It is clear from equation (3.3) that the one-day-ahead conditionally expected log returns in the model is simply

$$E_t\left[R_{t+1}\right] = r + \left(\lambda_z - \tfrac{1}{2}\right)h_{z,t} + \left(\lambda_y - \xi + \theta\right)h_{y,t}.$$

The jump compensator parameter, ξ, in our model is itself a particular function of other parameters

$$\xi = e^{\theta+\frac{1}{2}\delta^2} - 1. \tag{3.10}$$

This functional form ensures that conditionally expected total return is

$$E_t\left[\exp\left(R_{t+1}\right)\right] = \exp\left(r + \lambda_z h_{z,t} + \lambda_y h_{y,t}\right), \tag{3.11}$$

which in turn ensures that λ_z and λ_y can be viewed as compensation for diffusive volatility and jump exposure, respectively. Substituting equation (3.3) into (3.11), taking expectations, and solving for ξ yields equation (3.10). The ξ parameter will therefore not be estimated below but instead simply set to is value implied by equation (3.10).

3.3.6 Conditional Second Moments

From the model above, it is relatively straightforward to derive the following one-day-ahead conditional second moments

$$\begin{aligned} Var_t\left[R_{t+1}\right] &= h_{z,t} + \left(\theta^2 + \delta^2\right) h_{y,t} \\ Var_t\left[RBV_{t+1}\right] &= 2\sigma^2\left(1 + 2\gamma^2 h_{z,t}\right) \\ Var_t\left[RJV_{t+1}\right] &= \left(\theta^4 + 3\delta^4 + 6\theta^2\delta^2\right) h_{y,t} \\ Cov_t\left(R_{t+1}, RBV_{t+1}\right) &= -2\rho\gamma\sigma h_{z,t} \\ Cov_t\left(R_{t+1}, RJV_{t+1}\right) &= \theta\left(\theta^2 + 3\delta^2\right) h_{y,t} \\ Cov_t\left(RBV_{t+1}, RJV_{t+1}\right) &= 0 \end{aligned} \tag{3.12}$$

Note that the model allows for two types of "leverage" effects: One via the return covariance with bipower variation and another via the return covariance with jumps.

3.4 Physical Parameter Estimates

Above we have laid out the general framework for incorporating bipower variation and realized jump variation when modeling return dynamics. In this section we develop a likelihood-based estimation method that enables us to estimate the physical parameters using daily observations on returns, as well as the realized variation measures from Figure 1. We also develop two special cases of the general model and we briefly describe the Heston and Nandi (2000) benchmark GARCH model as well.

3.4.1 Deriving the Likelihood Function

When deriving the conditional quasi-likelihood function note first that the contribution to the total conditional likelihood by day $t+1$ can be obtained by summing over the number of jumps occurring on that day. We can write

$$\begin{aligned} f_t\left(R_{t+1}, RBV_{t+1}, RJV_{t+1}\right) &= \sum_{j=0}^{\infty} f_t\left(R_{t+1}, RBV_{t+1}, RJV_{t+1}, n_{t+1} = j\right) \\ &= \sum_{j=0}^{\infty} f_t\left(R_{t+1}, RBV_{t+1}, RJV_{t+1} | n_{t+1} = j\right) P_t\left[n_{t+1} = j\right] \end{aligned}$$

with the number of jumps drawn from the Poisson distribution,

$$P_t\left[n_{t+1} = j\right] = \frac{e^{-h_{y,t}} h_{y,t}^j}{!j}.$$

Separating out the days with exactly zero jumps, we get

$$f_t\left(R_{t+1}, RBV_{t+1}, RJV_{t+1} \middle| n_{t+1}=j\right) = \begin{cases} f_t\left(R_{t+1}, RBV_{t+1}\right), & \text{if } j=0 \\ f_t\left(j\right), & \text{if } j>0 \end{cases}$$

In order to save on notation, define the variable vectors

$$X_{t+1} \equiv \left(R_{t+1}, RBV_{t+1}, RJV_{t+1}\right)' \quad X_{t+1}^{(1,2)} \equiv \left(R_{t+1}, RBV_{t+1}\right)'$$

and the corresponding conditional first and second moments

$$\begin{array}{ll} \mu_t\left(n_{t+1}\right) \equiv E_t\left[X_{t+1} \middle| n_{t+1}\right] & \mu_t^{(1,2)} \equiv E_t\left[X_{t+1}^{(1,2)} \middle| n_{t+1}\right] \\ \Omega_t\left(n_{t+1}\right) \equiv Var_t\left[X_{t+1} \middle| n_{t+1}\right] & \Omega_t^{(1,2)}\left(n_{t+1}\right) \equiv Var_t\left[X_{t+1}^{(1,2)} \middle| n_{t+1}\right] \end{array} \tag{3.13}$$

Then we can write the marginal likelihood for returns and bipower variation when $n_{t+1}=0$ as

$$\begin{aligned} f_t\left(R_{t+1}, RBV_{t+1}\right) &= \left(2\pi\right)^{-1}\left|\Omega_t^{(1,2)}\left(0\right)\right|^{-1/2} \\ &\quad \cdot \exp\left(-\frac{1}{2}\left(X_{t+1}^{(1,2)}-\mu_t^{(1,2)}\left(0\right)\right)'\Omega_t^{(1,2)}\left(0\right)^{-1}\left(X_{t+1}^{(1,2)}-\mu_t^{(1,2)}\left(0\right)\right)\right) \end{aligned}$$

and when $n_{t+1}>0$ we have

$$f_t\left(j\right) = \left(2\pi\right)^{-3/2}\left|\Omega_t\left(j\right)\right|^{-1/2}\exp\left(-\frac{1}{2}\left(X_{t+1}-\mu_t\left(j\right)\right)'\Omega_t\left(j\right)^{-1}\left(X_{t+1}-\mu_t\left(j\right)\right)\right).$$

The log-likelihood is now defined by

$$\ln L^P = \sum_{t=1}^{T-1}\ln(f_t(R_{t+1}, RBV_{t+1}, RJV_{t+1})). \tag{3.14}$$

We maximize the likelihood function in (3.14) using the fminunc routine in Matlab with the following settings:

optimset('display','iter','MaxIter',1500,'TolFun',$1e-5$,'MaxFunEvals',$1e+06$,'TolX',$1e-20$).

3.4.2 Conditional Moments

The likelihood function above requires that we derive the first two moments conditional on time and on the number of jumps, n_{t+1}. For the conditional first moments we have

$$\begin{aligned} E_t\left[R_{t+1}|\, n_{t+1}\right] &= r + \left(\lambda_z - \frac{1}{2}\right) h_{z,t} + \left(\lambda_y - \xi\right) h_{y,t} + \theta n_{t+1} \\ E_t\left[RBV_{t+1}|\, n_{t+1}\right] &= h_{z,t} \\ E_t\left[RJV_{t+1}|\, n_{t+1}\right] &= \left(\theta^2 + \delta^2\right) n_{t+1} \end{aligned}$$

For the conditional second moments we have

$$\begin{aligned} Var_t\left[R_{t+1}|\, n_{t+1}\right] &= h_{z,t} + \delta^2 n_{t+1} \\ Var_t\left[RBV_{t+1}|\, n_{t+1}\right] &= 2\sigma^2\left(1 + 2\gamma^2 h_{z,t}\right) \\ Var_t\left[RJV_{t+1}|\, n_{t+1}\right] &= 2\delta^2\left(\delta^2 + 2\theta^2\right) n_{t+1} \\ Cov_t\left[R_{t+1}, RBV_{t+1}|\, n_{t+1}\right] &= -2\rho\gamma\sigma h_{z,t} \\ Cov_t\left[R_{t+1}, RJV_{t+1}|\, n_{t+1}\right] &= 2\theta\delta^2 n_{t+1} \\ Cov_t\left[RBV_{t+1}, RJV_{t+1}|\, n_{t+1}\right] &= 0 \end{aligned}$$

From these moments we can easily construct the μ_t vectors and Ω_t matrices in equation (3.13) needed for the likelihood function in equation (3.14).

Before turning to estimation of the new model we define three special cases of interest which we also estimate below.

3.4.3 The Heston-Nandi GARCH Model as a Special Case

First, by setting $h_{y,t} = 0$, and $\rho = 1$, we obtain one of the standard GARCH(1,1) models in the literature.

Specifically, note that $\rho = 1$ implies that $\varepsilon_{1,t+1} = \varepsilon_{2,t+1}$ and the realized variance therefore becomes irrelevant. We then get

$$\begin{aligned} h_{z,t+1} &= \omega_z - a_z\sigma + \left(b_z + a_z - a_z\sigma\gamma^2\right) h_{z,t} + a_z\sigma\left(\varepsilon_{2,t+1} - \gamma\sqrt{h_{z,t}}\right)^2 \\ &\equiv \omega + \beta h_{z,t} + \alpha\left(\varepsilon_{1,t+1} - \gamma\sqrt{h_{z,t}}\right)^2, \end{aligned}$$

which is exactly the Heston and Nandi (2000) affine GARCH(1,1) model.

3.4.4 The RVM Model as a Special Case

Second, we can shut down the separate jump variation by setting $h_{y,t} = 0$ in the new model. We then get

$$\begin{aligned} R_{t+1} &\equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r + \left(\lambda_z - \frac{1}{2}\right) h_{z,t} + z_{t+1}, \text{ with } z_{t+1} = \sqrt{h_{z,t}}\varepsilon_{1,t+1} \\ RV_{t+1} &= RBV_{t+1} + RJV_{t+1} = h_{z,t} + \sigma\left[\left(\varepsilon_{2,t+1} - \gamma\sqrt{h_{z,t}}\right)^2 - \left(1 + \gamma^2 h_{z,t}\right)\right]. \end{aligned}$$

This is exactly the autoregressive RV model in Christoffersen, Feunou, Jacobs, and Meddahi (2014). We will refer to this as the RVM model below.

3.4.5 The RJM Model as a Special Case

Third, we can shut down the bipower variation channel by setting $h_{z,t} = 0$. We then get

$$\begin{aligned} R_{t+1} &= r - \theta - \tfrac{\delta^2}{2} + (\lambda_y - \xi)\, h_{y,t} + y_{t+1} \\ y_{t+1} &= \sum_{j=1}^{n_{t+1}} x_{t+1}^j, \text{ where } x_{t+1}^j \overset{iid}{\sim} N(\theta, \delta^2) \\ P[n_{t+1} &= j | I_t] = \frac{e^{-h_{y,t}} h_{y,t}^{j-1}}{!(j-1)} \end{aligned}$$

and furthermore we set

$$\begin{aligned} RV_{t+1} &= \sum_{j=1}^{n_{t+1}} \left(x_{t+1}^j\right)^2 - \theta^2 \\ h_{y,t+1} &= \omega_y + b_y h_{y,t} + a_y RV_{t+1}. \end{aligned}$$

Note that in this case the entire quadratic variation is assumed to be driven by jumps so that each day has at least one jump. We will refer to this as the RJM model below.

3.4.6 Parameter Estimates and Model Properties

Table 1 contains the maximum likelihood estimation results for the physical return processes developed above. One year prior to our estimation sample we set the conditional variance equal to the unconditional variance and then burn-in the model on the pre-sample year to get an appropriate conditional variance on the first day of the sample. Note that the ω parameters do not have standard errors as they are computed by variance targeting thus exactly matching the observed sample variance of returns. The

parameter estimates are generally significant except for λs which are always difficult to pin down in relatively short return-based samples.

Note that volatility persistence is very high in the RVM and BPJVM models and considerably lower in the GARCH and RJM model. Unconditional volatility and volatility persistence is defined in the GARCH model as

$$E[h_t] = \frac{\omega + \alpha}{1 - (\beta + \alpha\gamma^2)} \equiv \frac{\omega + \alpha}{1 - Persist},$$

in the RVM model as

$$E[h_{z,t}] = \frac{\omega_z}{1 - (b_z + a_z)} \equiv \frac{\omega_z}{1 - Persist},$$

in the JVM model as

$$E[h_{y,t}] = \frac{\omega_y}{1 - (b_y + (\theta^2 + \delta^2)a_y)} \equiv \frac{\omega_y}{1 - Persist},$$

and in the BPJVM model as

$$E[h_t] = \frac{\omega_z}{1 - (b_z + a_z)} + \frac{(\theta^2 + \delta^2)\omega_y}{1 - (b_y + (\theta^2 + \delta^2)a_y)}.$$

Persistence for the two variance components in the BPJVM model are thus equal to the RVM and JVM cases.

When comparing model fit, we are faced with the challenge that the GARCH model is only fit to returns, the RVM and RJM models are fit to returns and RV, and the general BPJVM model is fit to returns, RBV and RJV. Table 1 shows that the likelihood value for the general model is $129,226$ but this is not readily comparable to the other models which are fit to different quantities. We therefore re-estimate the BPJVM model maximizing only the joint likelihood of returns and RV.[4] The second row of log likelihoods contains the results. From this perspective, the BPJVM model by far performs the best with a likelihood of $69,656$ compared with $68,783$ for the JVM model and $68,212$ for the RVM model.

When maximizing only the return likelihood the BPJVM model again performs the best with a likelihood of $19,522$. The improvement over the RVM and JVM model is not dramatic here but returns are unlikely to be informative about all the parameters of the model and so this set of results is only provided to enable comparison with GARCH. Note that the RVM, JVM and BPJVM models all perform very well compared with the benchmark affine GARCH model.

In Figure 3 we plot the daily conditional volatility computed as the square root of h_{t+1} for each model. Note that the volatility spikes are much more dramatically in the RVM, JVM and BPJVM

[4]See Appendix 3A for the details.

models than in the GARCH model. It is interesting and perhaps surprising that the JVM model is able to produce a spot volatility time path which is quite similar to that from the RVM and BPJVM models. This is partly because the RJM model is fit to returns and RV and not returns and RJV.

In Figure 4 we plot the daily conditional volatility of variance computed as the square root of

$$Var_t(h_{t+1}) = 2a_z^2\sigma^2(1+2\gamma^2 h_{z,t}) + a_y^2(\theta^2+\delta^2)^2((2\delta^2(\theta^2+2\delta^2)+(\theta^2+\delta^2)^2)h_{y,t} \tag{3.15}$$

for the BPJVM model. The variance of variance expressions for the other models are similar. Note from Figure 4 that the conditional volatility of variance is relatively low and almost constant in the GARCH model whereas in the other models it tends to be large when volatility is high thus matching the empirical evidence. Note that the volatility of volatility is slightly lower in the RJM than in the RVM and BPJVM models.

In Figure 5 we plot the conditional correlation of returns and variance, which are computed for the BPJVM model using

$$Corr_t(R_{t+1}, h_{t+1}) = \frac{-2\rho\sigma\gamma a_z h_{z,t} + a_y\theta(\theta^2+\delta^2)(\theta^2+3\delta^2)h_{y,t}}{\sqrt{Var_t\left[R_{t+1}\right]Var_t(h_{t+1})}} \tag{3.16}$$

where the terms in the denominator can be obtained from equations (3.12) and (3.15). The conditional correlation expressions for the other models are similar. Figure 5 shows that the differences across models are quite large from this perspective. The GARCH model implies a correlation of almost negative one. The other models imply correlations around -0.2. The RJM and the BPJVM models imply some time series variation in the correlation whereas the RJM model does not.

The negative correlation between return and volatility is a strong empirical regularity. It is often coined the "leverage" effect but there is actually not much evidence that it is driven by financial leverage (see, for example, Hasanhodzic and Lo (2011).) and its economic source is as of yet largely unknown. In our models the negative daily correlation generates negative skewness across horizons which in turn creates higher prices for out-of-the-money puts (and in-the-money calls) which the Black-Scholes model is well-known to under-price. In the Heston-Nandi GARCH model, a large estimate of γ generates the very large negative correlation evident in Figure 5, but it is also needed to generate variance of variance dynamics in that model. We suspect that the large value of γ in the Heston-Nandi GARCH model is partly driven by the need for a sufficiently high variance of variance (and thus kurtosis) and that it in turn causes the perhaps unrealistically large negative "leverage" correlation seen in Figure 5.

Figure 6 presents evidence on the different models' ability to forecast one-day ahead realized variance.

The ex-post realized variance is on the vertical axis and the model-predicted variance is on the horizontal axis. The corresponding regression fit is 49% for the GARCH model, 85% for the JVM model and 87% for the RVM and BPJVM models.[5] The slope coefficient on the volatility forecast, which ideally should be 1, is 2.5 in the GARCH model, 1.3 in the JVM model and 1.1 in the RVM and BPJVM models. The RVM and BPJVM models are thus able to predict realized variance quite well.

The properties we have investigated above are likely to have important implications for the models' ability to fit large panels of options. This is the task to which we now turn.

3.5 Option Valuation

In this section we show how the physical model developed above can be used for option valuation. We first derive the moment generating function of the physical process and show that it is affine. We then define a pricing kernel which implies that the risk-neutral moment generating function is of the same form as its physical counterpart. This in turn implies that we can compute option prices using Fourier inversion. Empirical results from estimating the model jointly on returns, realized measures and options follow.

3.5.1 The Physical Moment Generating Function

Using the vector notation $h_t \equiv (h_{z,t},\ h_{y,t})'$ defined above, and further defining the coefficients $v \equiv (v_z,\ v_y)'$, Appendix 3B shows that we can write the physical moment generating function (MGF) as

$$\begin{aligned} E_t\left[\exp\left(uR_{t+1} + v'h_{t+1}\right)\right] &= \exp\left(\begin{array}{c} u\left(r + \left(\lambda_z - \frac{1}{2}\right)h_{z,t} + \left(\lambda_y - \xi\right)h_{y,t}\right) + v'\left(\omega + bh_t\right) \\ +v_1\left(h_{z,t} - \sigma\left(1 + \gamma^2 h_{z,t}\right)\right) - \frac{1}{2}\ln\left(1 - 2\sigma v_1\right) \\ +\left(v_1\sigma\gamma^2 + \frac{1}{2}\left(1-\rho^2\right)u^2 + \frac{\left(u\rho - 2\sigma v_1\gamma\right)^2}{2\left(1-2\sigma v_1\right)}\right)h_{z,t} + \left(e^{v_3} - 1\right)h_{y,t} \end{array}\right) \\ &\equiv \exp\left(A\left(u,v\right)'h_t + B\left(u,v\right)\right) \end{aligned} \tag{3.17}$$

where we have further defined

$$v'a = (v_z,\ v_y)\begin{pmatrix} a_z & a_{z,y} \\ a_{y,z} & a_y \end{pmatrix} = \left(v_z a_z + v_y a_{y,z},\ v_z a_{z,y} + v_y a_y\right) \equiv (v_1,\ v_2),$$

[5] The detailed regression results are not reported in the tables but are available from the authors upon request.

and

$$v_3 = -\frac{1}{2}\ln\left(1-2v_2\delta^2\right) + u\theta + v_2\theta^2 + \frac{\left(u+2\theta v_2\right)^2\delta^2}{2\left(1-2v_2\delta^2\right)}.$$

Note that the physical MGF is of an exponentially affine form which will greatly facilitate option valuation below.

3.5.2 Risk Neutralization

We follow Christoffersen, Elkamhi, Feunou, and Jacobs (2010) and assume an exponential pricing kernel of the form

$$\begin{aligned}\zeta_{t+1} &= \frac{M_{t+1}}{E_t\left[M_{t+1}\right]} \equiv \frac{\exp\left(\nu_{1,t}\varepsilon_{1,t+1}+\nu_{2,t}\varepsilon_{2,t+1}+\nu_{3,t}\sum_{j=0}^{n_{t+1}}x_{t+1}^j\right)}{E_t\left[\exp\left(\nu_{1,t}\varepsilon_{1,t+1}+\nu_{2,t}\varepsilon_{2,t+1}+\nu_{3,t}\sum_{j=0}^{n_{t+1}}x_{t+1}^j\right)\right]} \\ &= \exp\left(\begin{array}{c}\nu_{1,t}\varepsilon_{1,t+1}+\nu_{2,t}\varepsilon_{2,t+1}+\nu_{3,t}\sum_{j=0}^{n_{t+1}}x_{t+1}^j \\ -\frac{1}{2}\nu_{1,t}^2-\frac{1}{2}\nu_{2,t}^2-\rho\nu_{1,t}\nu_{2,t}-\left(e^{\theta\nu_{3,t}+\frac{1}{2}\delta^2\nu_{3,t}^2}-1\right)h_{y,t}\end{array}\right)\end{aligned} \tag{3.18}$$

In order to ensure that the model is affine under Q, it is necessary and sufficient to impose the following conditions

$$\begin{aligned}\nu_{2,t} &= \left(\gamma-\gamma^*\right)\sqrt{h_{z,t}}-\rho\nu_{1,t} \\ \nu_{3,t} &= \nu_3.\end{aligned}$$

Appendix 3C shows that the risk-neutral probability measure for the BPJVM model is then

$$\begin{aligned}R_{t+1} &\equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r-\frac{1}{2}h_{z,t}-\xi^*h_{y,t}^*+\sqrt{h_{z,t}}\varepsilon_{1,t+1}^*+y_{t+1} \\ y_{t+1} &= \sum_{j=0}^{n_{t+1}}x_{t+1}^j;\quad x_{t+1}^j \overset{iid}{\sim}{}^Q\ N(\theta^*,\delta^2);\quad n_{t+1}|I_t \sim^Q Poisson\left(h_{y,t}^*\right) \\ RBV_{t+1} &= h_{z,t}+\sigma\left(\left(\gamma^*\right)^2-\gamma^2\right)h_{z,t}+\sigma\left[\left(\varepsilon_{2,t+1}^*-\gamma^*\sqrt{h_{z,t}}\right)^2-\left(1+\left(\gamma^*\right)^2h_{z,t}\right)\right] \\ RJV_{t+1} &= \sum_{j=0}^{n_{t+1}}\left(x_{t+1}^j\right)^2\end{aligned}$$

where $\varepsilon_{1,t+1}^*$ and $\varepsilon_{2,t+1}^*$ are bivariate Gaussian under Q, and where

$$\begin{aligned}h_{y,t}^* &= e^{\theta\nu_3+\frac{1}{2}\delta^2\nu_3^2}h_{y,t} \\ \theta^* &= \theta+\delta^2\nu_3,\quad \xi^* = e^{\theta^*+\frac{1}{2}\delta^2}-1\end{aligned}$$

Hence we have the risk premiums

$$\begin{aligned} E_t^Q\left[RBV_{t+1}\right] - E_t\left[RBV_{t+1}\right] &= \sigma\left(\left(\gamma^*\right)^2 - \gamma^2\right) h_{z,t} \\ E_t^Q\left[RJV_{t+1}\right] - E_t\left[RJV_{t+1}\right] &= \left(\left(\theta^*\right)^2 + \delta^2\right) h_{y,t}^* - \left(\theta^2 + \delta^2\right) h_{y,t}. \end{aligned}$$

where γ^* and ν_3 are additional parameters to be estimated. Below we will use the notation $\chi = \gamma - \gamma^*$ and report estimates of χ instead of γ^*.

By the nature of the model, risk-neutralization of the JVM model is slightly different from the other models. Appendix 3D provides the details.

3.5.3 Computing Option Values

Above we have shown that the risk-neutral distribution is of the same form as physical distribution. The risk-neutral MGF will therefore be of the form shown in Appendix 3B but with risk-neutral parameters used instead of their physical counterparts. We can therefore write the one-period risk-neutral conditional MGF as

$$\begin{aligned} \Psi_{t,t+1}^Q &\equiv E_t^Q\left[\exp\left(uR_{t+1} + v'h_{t+1}\right)\right] \\ &= \exp\left(\begin{array}{c} u\left(r - \frac{1}{2}h_{z,t} - \xi^* h_{y,t}^*\right) + v'\left(\omega + bh_t\right) \\ +v_1\left(h_{z,t} + \sigma\left(\left(\gamma^*\right)^2 - \gamma^2\right)h_{z,t} - \sigma\left(1 + \left(\gamma^*\right)^2 h_{z,t}\right)\right) - \frac{1}{2}\ln\left(1 - 2\sigma v_1\right) \\ + \left(v_1\sigma\left(\gamma^*\right)^2 + \frac{1}{2}\left(1-\rho^2\right)u^2 + \frac{\left(u\rho - 2\sigma v_1^*\gamma^*\right)^2}{2\left(1-2\sigma v_1\right)}\right)h_{z,t} + \left(e^{v_3} - 1\right)e^{\theta\nu_3 + \frac{1}{2}\delta^2\nu_3^2}h_{y,t} \end{array}\right) \\ &\equiv \exp\left(A^*\left(u,v\right)'h_t + B^*\left(u,v\right)\right) \end{aligned} \tag{3.19}$$

Call option values can now be computed via standard Fourier inversion techniques

$$\begin{aligned} Call &= S_t P_1(t,M) - \exp(-rM)XP_2(t,M), \text{ where} \\ P_1(t,M) &= \frac{1}{2} + \int_0^{+\infty} \mathrm{Re}\left(\frac{\Psi_{t,t+M}^Q(1+iu)\exp(-rM - iu\ln(\frac{X}{S_t}))}{\pi iu}\right)du \\ P_2(t,M) &= \frac{1}{2} + \int_0^{+\infty} \mathrm{Re}\left(\frac{\Psi_{t,t+M}^Q(iu)\exp(-iu\ln(\frac{X}{S_t}))}{\pi iu}\right)du \end{aligned} \tag{3.20}$$

where $\Psi_{t,t+M}^Q$ denotes the risk-neutral M-period MGF (see Appendix 3B) corresponding to the one-day MGF in equation (3.19). Put option values can be computed from put-call parity. Matlab code for computing option values is provided in Appendix 3E.

Armed with the quasi-closed form option-pricing formula in equation (3.20) we are now ready to

embark on a large-scale empirical investigation of the four models.

3.5.4 Fitting Options and Returns

From OptionMetrics we obtain Wednesday closing mid-quotes on SPX options data starting on January 2, 1996 and ending on August 28, 2013 which was the last date available at the time of writing.

We apply some commonly-used option data filters to the raw data. We restrict attention to out-of-the-money options with maturity between 15 and 180 calendar days. We omit contracts that do not satisfy well-known no-arbitrage conditions. We use only the six strikes with highest trading volume for each maturity quoted on Wednesdays. Finally, we convert puts to calls using put-call parity so as to ease the computation and interpretation below.

Table 2 provides descriptive statistics of the resulting data set consisting of 21,283 options. The top panel shows the contracts sorted by moneyness defined using the Black-Scholes delta. The persistent "smile" pattern in implied volatility is readily apparent from the top panel. The middle panel sorts the contracts by maturity and shows that there is not a persistent maturity pattern in implied volatilities: The term-structure of implied volatility is roughly flat on average. The bottom panel sorts by the VIX level. Table 2 shows that roughly half the contracts have a Delta above 0.6, a time-to-maturity between 30 and 90 days and are recorded on days when the VIX is between 15 and 25.

Joint estimation is performed by following Trolle and Schwartz (2009) who assume that the vega-weighted option errors, e_j, are i.i.d. Gaussian. We can then define the option likelihood, $\ln L^O$, and the joint likelihood, $\ln L$, as follows

$$\begin{aligned} VWRMSE &= \sqrt{\frac{1}{N}\sum_{j=1}^{N} e_j^2} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}((C_j^{Mkt} - C_j^{Mod})/BSV_j^{Mkt})^2} \\ \ln L^O &= -\frac{1}{2}\sum_{j=1}^{N}[\ln(VWRMSE^2) + e_j^2/VWRMSE^2] \\ \ln L &= \ln L^P + \ln L^O, \end{aligned} \tag{3.21}$$

where $\ln L^P$ denotes the log of likelihood function of the physical process defined in equation (3.14). We now estimate all physical parameters and risk premia by maximizing the joint likelihood function, $\ln L$.

Table 3 contains the parameter estimates and log likelihoods for our four models. We again calibrate the ω parameters by targeting the unconditional model variance to the sample variance of returns. As in Table 1, the physical parameters tend to be estimated precisely whereas some of the risk premium parameters continue to be difficult to pin down. A sequential estimation procedure in which only risk

premia are estimated from options may lead to more precise estimates. We leave this for future work.

The log-likelihoods reported in Table 3 are from joint estimation on returns and options for the GARCH model; from returns, RV and options for the RVM and JVM models; and from returns, RBV, RJV and options for the BPJVM model. They are therefore not directly comparable.

The option errors at the bottom of the table, however, are comparable. They show that in terms of implied volatility root mean squared error ($IVRMSE$) the RVM and JVM models offer a 12% and 18% improvement over the standard GARCH model, respectively. The BPJVM model offers a 21% improvement which is quite impressive. The $VWRMSE$ metric is broadly consistent with the $IVRMSE$ metric again showing a 21% improvement of BPJVM over GARCH.

3.5.5 Exploring the Results

In Table 4 we decompose the overall $IVRMSE$ fit in Table 3 by moneyness, maturity and VIX level following the layout of Table 2. The top panel of Table 4 shows that the BPJVM model performs the best in all but one moneyness category, namely deep out-of-the-money calls where RVM is best. The BPJVM model performs particularly well for deep in-the-money calls (corresponding to deep out-of-the-money puts) which have proven notoriously difficult to fit in the literature. The middle panel of Table 4 shows that the BPJVM model performs the best in all maturity categories including one virtual tie with the JVM model, namely for maturities between 30 and 60 days. The bottom panel shows that the BPJVM model is best in five of six VIX categories and virtually tied in the sixth when VIX is between 15 and 20%.

All together, Table 4 shows that the overall improvement in option fit by the BPJVM model evident in Table 3 is not due to any particular subset of the data set. The superior fit is obtained virtually everywhere.

Figure 7 reports the weekly time series of $IVRMSE$ for at-the-money options only. The figure is thus designed to reveal the models' ability to match the pattern of market volatility through time. Figure 7 shows that the RVM, JVM and BPJVM models are all much better than the GARCH model at capturing the dramatic dynamics in volatility unfolding during the 2008 financial crisis. It is indeed quite remarkable that the recent financial crisis does not appear as an outlier for the RVM, JVM and BPJVM models in Figure 7.

Figure 8 plots the model-implied risk neutral higher moments over time for the six-month horizon. Note that the BPJVM model is able to generate higher skewness (middle panel) and excess kurtosis (lower panel) values than are the three other models. This feature of the model is likely a key driver in

its success in fitting observed option prices as evident from Tables 3 and 4.

The top panel of Figure 8 shows that the RVM, JVM and BPJVM models generate much higher six-month risk-neutral volatility values than GARCH during the financial crisis in 2008. This is likely driving the at-the-money $IVRMSE$ performance of these models evident from Figure 7.

3.5.6 Option Error Specification

The option-based objective function in (3.21) implicitly assumes that option errors are independent and identically distributed (iid) across contracts. In reality, option errors have complicated dependence structures that are likely a function of time, moneynes, time-to-maturity, and potentially also the level of the market and its volatility. Figure 7 above which plots the weekly $IVRMSE$ indeed suggests that option errors are persistent across time. To study this issue further we plot in Figure 9 the autocorrelations of the weekly $VWRMSE$ series which is ultimately part of the input into equation (3.21). As in Figure 7 we consider only ATM options in Figure 9. The autocorrelation plot using all options looks similar. Figure 9 confirms that the option error magnitudes are indeed persistent over time.

The persistence in option errors implies that our maximum likelihood estimation procedure is not fully efficient. However, as long as the option errors are ergodic stationary (Hayashi (2000), p. 465), we will still obtain consistent parameter estimates. Our estimates should thus be viewed as quasi maximum likelihood as opposed to fully maximum likelihood.

While we have not pursued them there, it is important to acknowledge that the literature has offered various approaches to account for the non iid property of option errors. For example, Bates (2000) allows for heteroskedastic and autocorrelated errors in a Kalman filter approach. More recently, Kaeck and Alexander (2012) assume a multiplicative autoregressive error structure when estimating stochastic volatility models and Kanniainen, Lin, and Yang (2014) assume additive autoregressive errors when estimating GARCH models.

Our focus is primarily on comparing different models and we conjecture that our comparisons are little affected by the quasi maximum likelihood estimation strategy that we apply uniformly across models. Nevertheless, specifying a more accurate error structure for the model we have suggested will likely lead to more efficient parameter estimates and therefore remains an important topic for future research.

3.6 Summary and Conclusions

Under very general conditions, the total quadratic variation of a stochastic volatility process can be decomposed into diffusive variation and squared jump variation. We have used this result to develop a new class of option valuation models in which the underlying asset price exhibits volatility and jump intensity dynamics. The first key feature of our model is that the volatility and jump intensity dynamics in the model are directly driven by model-free empirical measures of diffusive volatility and jump variation. Second, because the empirical measures are observed in discrete intervals, our option valuation model is cast in discrete time, allowing for straightforward estimation of the model. Third, our model belongs to the affine class enabling us to derive the conditional characteristic function so that option values can be computed rapidly without relying on simulation methods. When estimated on S&P500 index options, realized measures, and returns the new model performs well compared with standard benchmarks.

Our analysis points to some interesting avenues for future research. First, a sequential estimation of physical parameters and then risk premia would be interesting. Second, implementing a more efficient estimation methodology would be of value as discussed in Section 3.5.6. Third, several alternatives exist to the nonparametric measures of jumps explored in this paper. For example, Li (2013) employs hedging errors implied by delta-hedged positions in European-style options to identify jumps. Applying these alternative jump measures in our modeling framework could be useful. Fourth, we have allowed for measurement errors on some but not all variables. Extending our model in this direction would likely lead to even better empirical performance. Fifth, so far we have only used model-free physical measures of jumps and diffusive volatility. However, Du and Kapadia (2012) have recently proposed model-free risk-neutral counterparts to the realized bipower variation and realized jump variation measures we employ. Using the risk-neutral measures in our modeling framework may well lead to an even better fit of our model to observed option prices. We leave these avenues for future exploration.

Appendix

Appendix 3A: A Special Case of the Likelihood Function

In this section, we compute a special case of the likelihood function used to fit BPJVM model to the observed returns and RV only. Denote

$$f_t\left(R_{t+1}, RV_{t+1}\right) = f_t\left(R_{t+1}, RBV_{t+1} + RJV_{t+1}\right)$$

Using the methodology from the general case, we have

$$\begin{aligned} f_t\left(R_{t+1}, RBV_{t+1} + RJV_{t+1}\right) &= \sum_{j=0}^{\infty} f_t\left(R_{t+1}, RBV_{t+1} + RJV_{t+1}, n_{t+1} = j\right) \\ &= \sum_{j=0}^{\infty} f_t\left(R_{t+1}, RBV_{t+1} + RJV_{t+1} \middle| n_{t+1} = j\right) P_t\left[n_{t+1} = j\right] \end{aligned}$$

with

$$P_t\left[n_{t+1} = j\right] = \frac{e^{-h_{y,t}} h_{y,t}^{j}}{!j}$$

$$f_t\left(R_{t+1}, RBV_{t+1} + RJV_{t+1} \middle| n_{t+1} = j\right) = \begin{cases} f_t\left(R_{t+1}, RBV_{t+1}\right) & \text{if} \quad j = 0 \\ \bar{f}_t\left(j\right) & \text{if} \quad j > 0 \end{cases}$$

where

$$\bar{f}_t\left(j\right) = \left(2\pi\right)^{-1} \left|\Omega_t^{(r,rv)}\left(0\right)\right|^{-1/2} \exp\left(-\frac{1}{2}\left(X_{t+1}^{(r,rv)} - \mu_t^{(r,rv)}\left(j\right)\right)' \Omega_t^{(r,rv)}\left(j\right)^{-1} \left(X_{t+1}^{(r,rv)} - \mu_t^{(r,rv)}\left(j\right)\right)\right)$$

$$\mu_t^{(r,rv)}\left(j\right) = \begin{pmatrix} r + \left(\lambda_z - \frac{1}{2}\right) h_{z,t} + \left(\lambda_y - \xi\right) h_{y,t} + \theta j \\ h_{z,t} + \left(\theta^2 + \delta^2\right) j \end{pmatrix}$$

$$\Omega_t^{(r,rv)}\left(j\right) = \begin{bmatrix} h_{z,t} + \delta^2 j & -2\rho\gamma\sigma h_{z,t} + 2\theta\delta^2 j \\ -2\rho\gamma\sigma h_{z,t} + 2\theta\delta^2 j & 2\sigma^2\left(1 + 2\gamma^2 h_{z,t}\right) + 2\delta^2\left(\delta^2 + 2\theta^2\right) j \end{bmatrix}$$

Appendix 3B: Physical MGF for the BPJVM Model

In this section, we derive the closed-form MGF for the BPJVM model under the physical measure. Using the vector notation $h_t \equiv \left(h_{z,t},\ h_{y,t}\right)'$ and further defining the coefficients $v \equiv \left(v_z,\ v_y\right)'$, we can write the physical moment generating function as

$$E_t\left[\exp\left(uR_{t+1} + v'h_{t+1}\right)\right] = E_t\left[\exp\begin{pmatrix} u\left(r + \left(\lambda_z - \frac{1}{2}\right) h_{z,t} + \left(\lambda_y - \xi\right) h_{y,t} + z_{t+1} + y_{t+1}\right) \\ +v'\left(\omega + bh_t + aRVM_{t+1}\right) \end{pmatrix}\right] =$$

$$\exp\left(u\left(r + \left(\lambda_z - \frac{1}{2}\right) h_{z,t} + \left(\lambda_y - \xi\right) h_{y,t}\right) + v'\left(\omega + bh_t\right)\right) E_t\left[\exp\left(u\left(z_{t+1} + y_{t+1}\right) + v'aRVM_{t+1}\right)\right]$$

We further define

$$v'a = \left(v_z,\ v_y\right) \begin{pmatrix} a_z & a_{z,y} \\ a_{y,z} & a_y \end{pmatrix} = \left(v_z a_z + v_y a_{y,z},\ v_z a_{z,y} + v_y a_y\right) \equiv \left(v_1,\ v_2\right)$$

Then, we can write

$$\begin{aligned}
& E_t\left[\exp\left(u\left(z_{t+1}+y_{t+1}\right)+v'aRVM_{t+1}\right)\right] \\
= \; & E_t\left[\exp\left(u\left(z_{t+1}+y_{t+1}\right)+v_1RBV_{t+1}+v_2RJV_{t+1}\right)\right] \\
= \; & \exp\left(v_1\left(h_{z,t}-\sigma\left(1+\gamma^2h_{z,t}\right)\right)\right)E_t\left[\exp\left(u\sqrt{h_{z,t}}\varepsilon_{1,t+1}+v_1\sigma\left(\varepsilon_{2,t+1}-\gamma\sqrt{h_{z,t}}\right)^2\right)\right]\times \\
& E_t\left[\exp\left(\sum_{j=0}^{n_{t+1}}ux_{t+1}^j+v_2\left(x_{t+1}^j\right)^2\right)\right]
\end{aligned}$$

Where the expectations can be computed explicitly as

$$\begin{aligned}
& E_t\left[\exp\left(u\sqrt{h_{z,t}}\varepsilon_{1,t+1}+v_1\sigma\left(\varepsilon_{2,t+1}-\gamma\sqrt{h_{z,t}}\right)^2\right)\right] \\
= \; & \exp\left(-\frac{1}{2}\ln\left(1-2\sigma v_1\right)+\left(v_1\sigma\gamma^2+\frac{1}{2}\left(1-\rho^2\right)u^2+\frac{\left(u\rho-2\sigma v_1\gamma\right)^2}{2\left(1-2\sigma v_1\right)}\right)h_{z,t}\right)
\end{aligned}$$

and

$$E_t\left[\exp\left(\sum_{j=0}^{n_{t+1}}ux_{t+1}^j+v_2\left(x_{t+1}^j\right)^2\right)\right]=E_t\left[E_t\left[\exp\left(\sum_{j=0}^{n_{t+1}}ux_{t+1}^j+v_2\left(x_{t+1}^j\right)^2\right)\middle|n_{t+1}\right]\right]$$

$$E_t\left[\exp\left(\sum_{j=0}^{n_{t+1}}ux_{t+1}^j+v_2\left(x_{t+1}^j\right)^2\right)\middle|n_{t+1}\right]=\exp\left(v_3n_{t+1}\right)$$

where

$$v_3=-\frac{1}{2}\ln\left(1-2v_2\delta^2\right)+u\theta+v_2\theta^2+\frac{\left(u+2\theta v_2\right)^2\delta^2}{2\left(1-2v_2\delta^2\right)}$$

hence

$$E_t\left[\exp\left(\sum_{j=0}^{n_{t+1}}ux_{t+1}^j+v_2\left(x_{t+1}^j\right)^2\right)\right]=E_t\left[\exp\left(v_3n_{t+1}\right)\right]=\exp\left(h_{y,t}\left(e^{v_3}-1\right)\right)$$

Therefore, we have the following expression

$$E_t\left[\exp\left(u\left(z_{t+1}+y_{t+1}\right)+v'aRVM_{t+1}\right)\right]=\exp\left(\begin{array}{c}v_1\left(h_{z,t}-\sigma\left(1+\gamma^2h_{z,t}\right)\right)-\frac{1}{2}\ln\left(1-2\sigma v_1\right)\\+\left(v_1\sigma\gamma^2+\frac{1}{2}\left(1-\rho^2\right)u^2+\frac{\left(u\rho-2\sigma v_1\gamma\right)^2}{2\left(1-2\sigma v_1\right)}\right)h_{z,t}\\+\left(e^{v_3}-1\right)h_{y,t}\end{array}\right)$$

Substituting the above back to the original MGF, we get

$$\begin{aligned}
E_t\left[\exp\left(uR_{t+1}+v'h_{t+1}\right)\right] &= \exp\left(\begin{array}{c}u\left(r+\left(\lambda_z-\frac{1}{2}\right)h_{z,t}+\left(\lambda_y-\xi\right)h_{y,t}\right)+v'\left(\omega+bh_t\right)\\+v_1\left(h_{z,t}-\sigma\left(1+\gamma^2h_{z,t}\right)\right)-\frac{1}{2}\ln\left(1-2\sigma v_1\right)\\+\left(v_1\sigma\gamma^2+\frac{1}{2}\left(1-\rho^2\right)u^2+\frac{\left(u\rho-2\sigma v_1\gamma\right)^2}{2\left(1-2\sigma v_1\right)}\right)h_{z,t}+\left(e^{v_3}-1\right)h_{y,t}\end{array}\right) \\
&\equiv \exp\left(A\left(u,v\right)'h_t+B\left(u,v\right)\right)
\end{aligned}$$

which shows that the physical one-step-ahead moment generating function is exponentially affine.

We conjecture that the multi-step moment generating function is also of the affine form. First, define

$$\begin{aligned}\Psi_{t,t+M}(u) &= E_t[\exp(u\sum_{j=1}^{M} R_{t+j})] \\ &= \exp(C(u,M)'h_t + D(u,M))\end{aligned}$$

From this we can compute

$$\begin{aligned}\Psi_{t,t+M+1}(u) &= E_t[\exp(u\sum_{j=1}^{M} R_{t+j})] = E_t[E_{t+1}[\exp(u\sum_{j=1}^{M} R_{t+j})]] \\ &= E_t[\exp(uR_{t+1})E_{t+1}[\exp(u\sum_{j=2}^{M} R_{t+j})]] \\ &= E_t[\exp(uR_{t+1} + C(u,M)'h_{t+1} + D(u,M))] \\ &= \exp(A(u,C(u,M))'h_t + B(u,C(u,M)) + D(u,M))\end{aligned}$$

which yields the following recursive relationship

$$\begin{aligned}C(u,M+1) &= A(u,C(u,M)) \\ D(u,M+1) &= B(u,C(u,M)) + D(u,M)\end{aligned}$$

using the following initial conditions

$$\begin{aligned}C(u,1) &= A(u,\mathbf{0}) \\ D(u,1) &= B(u,\mathbf{0})\end{aligned}$$

where A and C are 2-by-1 vector-valued functions.

Appendix 3C: Risk Neutralization of the BPJVM Model

In this appendix, we derive the risk-neutralization of the BPJVM model. We assume an exponential pricing kernel of the following form

$$\begin{aligned}\zeta_{t+1} &= \frac{M_{t+1}}{E_t[M_{t+1}]} \equiv \frac{\exp\left(\nu_{1,t}\varepsilon_{1,t+1} + \nu_{2,t}\varepsilon_{2,t+1} + \nu_{3,t}\sum_{j=0}^{n_{t+1}} x_{t+1}^j\right)}{E_t\left[\exp\left(\nu_{1,t}\varepsilon_{1,t+1} + \nu_{2,t}\varepsilon_{2,t+1} + \nu_{3,t}\sum_{j=0}^{n_{t+1}} x_{t+1}^j\right)\right]} \\ &= \exp\begin{pmatrix} \nu_{1,t}\varepsilon_{1,t+1} + \nu_{2,t}\varepsilon_{2,t+1} + \nu_{3,t}\sum_{j=0}^{n_{t+1}} x_{t+1}^j \\ -\frac{1}{2}\nu_{1,t}^2 - \frac{1}{2}\nu_{2,t}^2 - \rho\nu_{1,t}\nu_{2,t} - \left(e^{\theta\nu_{3,t}+\frac{1}{2}\delta^2\nu_{3,t}^2} - 1\right)h_{y,t}\end{pmatrix}\end{aligned}$$

We need to impose the no-arbitrage condition

$$E_t^Q[\exp(R_{t+1})] \equiv E_t[\zeta_{t+1}\exp(R_{t+1})] = \exp(r)$$

where

$$\begin{aligned} E_t\left[\zeta_{t+1}\exp\left(R_{t+1}\right)\right] &= E_t\left[\zeta_{t+1}\exp\left(r+\left(\lambda_z-\frac{1}{2}\right)h_{z,t}+\left(\lambda_y-\xi\right)h_{y,t}+z_{t+1}+y_{t+1}\right)\right] \\ &= \exp\left(\begin{array}{c} r+\left(\lambda_z-\frac{1}{2}\right)h_{z,t}+\left(\lambda_y-\xi\right)h_{y,t} \\ -\frac{1}{2}\nu_{1,t}^2-\frac{1}{2}\nu_{2,t}^2-\rho\nu_{1,t}\nu_{2,t}-\left(e^{\theta\nu_{3,t}+\frac{1}{2}\delta^2\nu_{3,t}^2}-1\right)h_{y,t} \\ +\frac{1}{2}\left(\nu_{1,t}+\sqrt{h_{z,t}}\right)^2+\frac{1}{2}\nu_{2,t}^2+\rho\left(\nu_{1,t}+\sqrt{h_{z,t}}\right)\nu_{2,t} \\ +\left(e^{\theta(\nu_{3,t}+1)+\frac{1}{2}\delta^2(\nu_{3,t}+1)^2}-1\right)h_{y,t} \end{array}\right) \\ &= \exp\left(\begin{array}{c} r+\lambda_z h_{z,t}+\left(\lambda_y-\xi\right)h_{y,t} \\ e^{\theta\nu_{3,t}+\frac{1}{2}\delta^2\nu_{3,t}^2}\left(e^{\theta+\frac{1}{2}\delta^2+\delta^2\nu_{3,t}}-1\right)h_{y,t} \\ +\nu_{1,t}\sqrt{h_{z,t}}+\rho\sqrt{h_{z,t}}\nu_{2,t} \end{array}\right). \end{aligned}$$

Setting this expression equal to the risk-free rate, and taking logs, yields the condition

$$\lambda_z h_{z,t}+\left(\lambda_y-\xi\right)h_{y,t}+\nu_{1,t}\sqrt{h_{z,t}}+\rho\sqrt{h_{z,t}}\nu_{2,t}+e^{\theta\nu_{3,t}+\frac{1}{2}\delta^2\nu_{3,t}^2}\left(e^{\theta+\frac{1}{2}\delta^2+\delta^2\nu_{3,t}}-1\right)h_{y,t}=0$$

In order to determine the form of the risk-neutral distribution of the shocks we consider the moment generating function

$$E_t^Q\left[\exp\left(u_1\varepsilon_{1,t+1}+u_2\varepsilon_{2,t+1}+u_3 y_{t+1}\right)\right]=\exp\left(\begin{array}{c} u_1\left(\nu_{1,t}+\rho\nu_{2,t}\right)+u_2\left(\nu_{2,t}+\rho\nu_{1,t}\right)+\frac{u_1^2}{2}+\frac{u_2^2}{2}+\rho u_1 u_2 \\ +\left(e^{\left(\theta+\delta^2\nu_{3,t}\right)u_3+\frac{1}{2}\delta^2u_3^2}-1\right)e^{\theta\nu_{3,t}+\frac{1}{2}\delta^2\nu_{3,t}^2}h_{y,t} \end{array}\right)$$

In order to obtain an affine model under the Q measure, we set $\nu_{3,t}$ to a constant, i.e. $\nu_{3,t}=\nu_3$.Under the Q measure we have

$$\begin{aligned} \varepsilon_{1,t+1}^* &= \varepsilon_{1,t+1}-\left(\nu_{1,t}+\rho\nu_{2,t}\right);\ \ \varepsilon_{1,t+1}^* \overset{iid^Q}{\sim} N\left(0,1\right) \\ \varepsilon_{2,t+1}^* &= \varepsilon_{2,t+1}-\left(\nu_{2,t}+\rho\nu_{1,t}\right);\ \ \varepsilon_{2,t+1}^* \overset{iid^Q}{\sim} N\left(0,1\right) \\ y_{t+1} &= \sum_{j=0}^{n_{t+1}} x_{t+1}^j;\ \ x_{t+1}^j \overset{iid^Q}{\sim} N(\theta+\delta^2\nu_3,\delta^2);\ \ n_{t+1}|I_t \sim^Q Poisson\left(e^{\theta\nu_3+\frac{1}{2}\delta^2\nu_3^2}h_{y,t}\right) \end{aligned}$$

We thus see that under the Q measure, $\varepsilon_{1,t+1}^*$ and $\varepsilon_{2,t+1}^*$ follow a bivariate standard normal distribution with correlation ρ.

The realized bipower variation equation can be written as follows

$$\begin{aligned} RBV_{t+1} &= h_{z,t}+\sigma\left[\left(\varepsilon_{2,t+1}-\gamma\sqrt{h_{z,t}}\right)^2-\left(1+\gamma^2h_{z,t}\right)\right] \\ &= h_{z,t}+\sigma\left[\left(\varepsilon_{2,t+1}^*+\nu_{2,t}+\rho\nu_{1,t}-\gamma\sqrt{h_{z,t}}\right)^2-\left(1+\gamma^2h_{z,t}\right)\right] \end{aligned}$$

In order to ensure that the model is affine under Q, it is necessary and sufficient to fix

$$\nu_{2,t}+\rho\nu_{1,t}-\gamma\sqrt{h_{z,t}}=-\gamma^*\sqrt{h_{z,t}},$$

which yields the condition

$$\nu_{2,t}=\left(\gamma-\gamma^*\right)\sqrt{h_{z,t}}-\rho\nu_{1,t}.$$

Using the no-arbitrage condition above implies that

$$\nu_{1,t}\sqrt{h_{z,t}}+\rho\sqrt{h_{z,t}}\left(\left(\gamma-\gamma^*\right)\sqrt{h_{z,t}}-\rho\nu_{1,t}\right)=-\lambda_z h_{z,t}-\left(\lambda_y-\xi\right)h_{y,t}-e^{\theta\nu_3+\frac{1}{2}\delta^2\nu_3^2}\left(e^{\theta+\frac{1}{2}\delta^2+\delta^2\nu_3}-1\right)h_{y,t}$$

thus we have

$$\nu_{1,t} = \frac{\left(\rho\left(\gamma^* - \gamma\right) - \lambda_z\right) h_{z,t} - \left(\lambda_y - \xi + e^{\theta\nu_3 + \frac{1}{2}\delta^2\nu_3^2}\left(e^{\theta + \frac{1}{2}\delta^2 + \delta^2\nu_3} - 1\right)\right) h_{y,t}}{\left(1 - \rho^2\right)\sqrt{h_{z,t}}}$$

$$\left(\nu_{1,t} + \rho\nu_{2,t}\right)\sqrt{h_{z,t}} = -\lambda_z h_{z,t} - \left(\left(\lambda_y - \xi\right) + e^{\theta\nu_3 + \frac{1}{2}\delta^2\nu_3^2}\left(e^{\theta + \frac{1}{2}\delta^2 + \delta^2\nu_3} - 1\right)\right) h_{y,t}$$

Now we can re-write the returns equation under the risk-neutral measure as follows

$$\begin{aligned} R_{t+1} &\equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r + \left(\lambda_z - \frac{1}{2}\right) h_{z,t} + \left(\lambda_y - \xi\right) h_{y,t} + z_{t+1} + y_{t+1} \\ &= r - \frac{1}{2} h_{z,t} - e^{\theta\nu_3 + \frac{1}{2}\delta^2\nu_3^2}\left(e^{\theta + \frac{1}{2}\delta^2 + \delta^2\nu_3} - 1\right) h_{y,t} + \sqrt{h_{z,t}}\varepsilon^*_{1,t+1} \\ &= r - \frac{1}{2} h_{z,t} - \xi^* h^*_{y,t} + \sqrt{h_{z,t}}\varepsilon^*_{1,t+1} + y_{t+1} \end{aligned}$$

Hence under the risk-neutral measure, we have

$$\begin{aligned} R_{t+1} &\equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r - \frac{1}{2} h_{z,t} - \xi^* h^*_{y,t} + \sqrt{h_{z,t}}\varepsilon^*_{1,t+1} + y_{t+1} \\ y_{t+1} &= \sum_{j=0}^{n_{t+1}} x^j_{t+1}; \quad x^j_{t+1} \sim^Q iidN(\theta^*, \delta^2); \quad n_{t+1} | I_t \sim^Q Poisson\left(h^*_{y,t}\right) \\ RBV_{t+1} &= h_{z,t} + \sigma\left(\left(\gamma^*\right)^2 - \gamma^2\right) h_{z,t} + \sigma\left[\left(\varepsilon^*_{2,t+1} - \gamma^*\sqrt{h_{z,t}}\right)^2 - \left(1 + \left(\gamma^*\right)^2 h_{z,t}\right)\right] \\ RJV_{t+1} &= \sum_{j=0}^{n_{t+1}} \left(x^j_{t+1}\right)^2 \end{aligned}$$

with the following parameter mappings

$$\begin{aligned} h^*_{y,t} &= e^{\theta\nu_3 + \frac{1}{2}\delta^2\nu_3^2} h_{y,t} \\ \theta^* &= \theta + \delta^2\nu_3, \quad \xi^* = e^{\theta^* + \frac{1}{2}\delta^2} - 1 \end{aligned}$$

Appendix 3D: Risk Neutralization of the RJM Model

In this appendix, we derive the risk-neutralization of the RJM model. We use the following particular form of the pricing kernel to ensure the affine structure is preserved under the risk-neutral measure.

$$\zeta_{t+1} = \frac{M_{t+1}}{E_t\left[M_{t+1}\right]} \equiv \frac{\exp\left(\nu_1 \sum_{j=1}^{n_{t+1}} x^j_{t+1} + \nu_2 \sum_{j=1}^{n_{t+1}} \left(x^j_{t+1}\right)^2 + \nu_3 n_{t+1}\right)}{E_t\left[\exp\left(\nu_1 \sum_{j=1}^{n_{t+1}} x^j_{t+1} + \nu_2 \sum_{j=1}^{n_{t+1}} \left(x^j_{t+1}\right)^2 + \nu_3 n_{t+1}\right)\right]}$$

which can be written as

$$\begin{aligned}
& E_t\left[\exp\left(\nu_1\sum_{j=1}^{n_{t+1}}x_{t+1}^j+\nu_2\sum_{j=1}^{n_{t+1}}\left(x_{t+1}^j\right)^2+\nu_3 n_{t+1}\right)\right] \\
= \; & E_t\left[E_t\left[\exp\left(\nu_1\sum_{j=1}^{n_{t+1}}x_{t+1}^j+\nu_2\sum_{j=1}^{n_{t+1}}\left(x_{t+1}^j\right)^2+\nu_3 n_{t+1}\right)\middle| n_{t+1}\right]\right] \\
= \; & E_t\left[\exp\left(\nu_3 n_{t+1}\right)E_t\left[\exp\left(\sum_{j=1}^{n_{t+1}}\left(\nu_1 x_{t+1}^j+\nu_2\left(x_{t+1}^j\right)^2\right)\right)\middle| n_{t+1}\right]\right] \\
= \; & E_t\left[\exp\left(\nu_3 n_{t+1}\right)\prod_{j=1}^{n_{t+1}}E_t\left[\exp\left(\nu_1 x_{t+1}^j+\nu_2\left(x_{t+1}^j\right)^2\right)\middle| n_{t+1}\right]\right] \\
= \; & E_t\left[\exp\left(\nu_3 n_{t+1}\right)\left(E_t\left[\exp\left(\nu_1 x_{t+1}^j+\nu_2\left(x_{t+1}^j\right)^2\right)\right]\right)^{n_{t+1}}\right] \\
= \; & E_t\left[\exp\left(\nu_4 n_{t+1}\right)\right]
\end{aligned}$$

with the notation

$$\nu_4=-\frac{1}{2}\ln\left(1-2\nu_2\delta^2\right)+\nu_1\theta+\nu_2\theta^2+\frac{\left(\nu_1+2\theta\nu_2\right)^2\delta^2}{2\left(1-2\nu_2\delta^2\right)}+\nu_3$$

hence we have

$$\zeta_{t+1}=\exp\left(\nu_1\sum_{j=1}^{n_{t+1}}x_{t+1}^j+\nu_2\sum_{j=1}^{n_{t+1}}\left(x_{t+1}^j\right)^2-\nu_4-\left(e^{\nu_4}-1\right)h_{y,t}\right)$$

We need to impose the no-arbitrage condition

$$E_t^Q\left[\exp\left(R_{t+1}\right)\right]\equiv E_t\left[\zeta_{t+1}\exp\left(R_{t+1}\right)\right]=\exp\left(r\right)$$

which gives us the following parameter restriction

$$\begin{aligned}
E_t\left[\zeta_{t+1}\exp\left(R_{t+1}\right)\right] &= E_t\left[\zeta_{t+1}\exp\left(\bar{r}+\left(\lambda_y-\xi\right)h_{y,t}+\sum_{j=1}^{n_{t+1}}x_{t+1}^j\right)\right] \\
&= E_t\left[\exp\left(\begin{array}{c}\bar{r}+\left(\lambda_y-\xi\right)h_{y,t}+\left(1+\nu_1\right)\sum_{j=1}^{n_{t+1}}x_{t+1}^j+\\ \nu_2\sum_{j=1}^{n_{t+1}}\left(x_{t+1}^j\right)^2-\nu_4-\left(e^{\nu_4}-1\right)h_{y,t}\end{array}\right)\right] \\
&= \exp\left(\bar{r}+\nu_5-\nu_4+\left(\lambda_y-\xi+e^{\nu_5}-e^{\nu_4}\right)h_{y,t}\right)
\end{aligned}$$

with

$$\nu_5=-\frac{1}{2}\ln\left(1-2\nu_2\delta^2\right)+\left(1+\nu_1\right)\theta+\nu_2\theta^2+\frac{\left(1+\nu_1+2\theta\nu_2\right)^2\delta^2}{2\left(1-2\nu_2\delta^2\right)}+\nu_3$$

Hence, the following relationships need to hold

$$\begin{aligned}
\nu_5-\nu_4 &= \theta+\frac{\delta^2}{2} \\
e^{\nu_5}-e^{\nu_4} &= \xi-\lambda_y
\end{aligned}$$

thus

$$\begin{aligned} e^{\nu_5} - e^{\nu_4} &= e^{\nu_4}\left(e^{\nu_5-\nu_4}-1\right) \\ &= e^{\nu_4}\left(e^{\theta+\frac{\delta^2}{2}}-1\right) \\ &= e^{\nu_4}\xi \end{aligned}$$

$$e^{\nu_4} = \frac{\xi - \lambda_y}{\xi}$$

and

$$\nu_4 = \ln\left(1 - \frac{\lambda_y}{\xi}\right)$$

$$\begin{aligned} \nu_5 - \nu_4 &= \theta + \frac{\left[\left(1+\nu_1+2\theta\nu_2\right)^2 - \left(\nu_1+2\theta\nu_2\right)^2\right]\delta^2}{2\left(1-2\nu_2\delta^2\right)} \\ &= \theta + \frac{\left(1+2\nu_1+4\theta\nu_2\right)\delta^2}{2\left(1-2\nu_2\delta^2\right)} \end{aligned}$$

$$\nu_5 - \nu_4 = \theta + \frac{\delta^2}{2}$$

which implies that

$$\theta + \frac{\delta^2}{2} = \theta + \frac{\left(1+2\nu_1+4\theta\nu_2\right)\delta^2}{2\left(1-2\nu_2\delta^2\right)}$$

hence

$$1 + 2\nu_1 + 4\theta\nu_2 = 1 - 2\nu_2\delta^2$$

which can be written as

$$\nu_1 = -\left(\delta^2 + 2\theta\right)\nu_2$$

and

$$\begin{aligned} \nu_3 &= \nu_4 - \left(-\frac{1}{2}\ln\left(1-2\nu_2\delta^2\right) + \nu_1\theta + \nu_2\theta^2 + \frac{\left(\nu_1+2\theta\nu_2\right)^2\delta^2}{2\left(1-2\nu_2\delta^2\right)}\right) \\ &= \ln\left(1-\frac{\lambda_y}{\xi}\right) + \frac{1}{2}\ln\left(1-2\nu_2\delta^2\right) + \theta\left(\delta^2+\theta\right)\nu_2 - \frac{\delta^6\nu_2^2}{2\left(1-2\nu_2\delta^2\right)} \end{aligned}$$

To determine the risk-neutral distribution of the shocks, we consider

$$\begin{aligned} E_t^Q\left[\exp\left(un_{t+1}\right)\right] &= E_t\left[\exp\left(un_{t+1} + \nu_1\sum_{j=1}^{n_{t+1}} x_{t+1}^j + \nu_2\sum_{j=1}^{n_{t+1}}\left(x_{t+1}^j\right)^2 + \nu_3 n_{t+1} - \nu_4 - \left(e^{\nu_4}-1\right)h_{y,t}\right)\right] \\ &= \exp\left(-\nu_4 - \left(e^{\nu_4}-1\right)h_{y,t}\right)E_t\left[\exp\left(vn_{t+1}\right)\right] \\ &= \exp\left(-\nu_4 - \left(e^{\nu_4}-1\right)h_{y,t} + v + \left(e^v-1\right)h_{y,t}\right) \\ &= \exp\left(v - \nu_4 + \left(e^v - e^{\nu_4}\right)h_{y,t}\right) = \exp\left(v - \nu_4 + e^{\nu_4}\left(e^{v-\nu_4}-1\right)h_{y,t}\right) \end{aligned}$$

with

$$v = -\frac{1}{2}\ln\left(1-2\nu_2\delta^2\right) + \nu_1\theta + \nu_2\theta^2 + \frac{\left(\nu_1+2\theta\nu_2\right)^2\delta^2}{2\left(1-2\nu_2\delta^2\right)} + u + \nu_3$$

$$v - \nu_4 = u$$

hence

$$E_t^Q\left[\exp\left(un_{t+1}\right)\right] = \exp\left(u + \left(e^u - 1\right)\left(1 - \frac{\lambda_y}{\xi}\right)h_{y,t}\right)$$

$$n_{t+1} = n^*_{t+1} + 1$$

where

$$n^*_{t+1} | I_t \tilde{}^Q Poisson(h^*_{y,t})$$

$$h^*_{y,t} = \left(1 - \frac{\lambda_y}{\xi}\right) h_{y,t}$$

Next, we compute the conditional moment generating function of individual jumps under the risk-neutral measure

$$\begin{aligned} E^Q_t\left[\exp\left(ux^{j_0}_{t+1}\right)\right] &= E_t\left[\exp\left(\begin{array}{c} ux^{j_0}_{t+1} + \nu_1 \sum_{j=1}^{n_{t+1}} x^j_{t+1} + \nu_2 \sum_{j=1}^{n_{t+1}} \left(x^j_{t+1}\right)^2 + \\ \nu_3\left(n_{t+1} - 1\right) + \nu_3 - \nu_4 - \left(e^{\nu_4} - 1\right) h_{y,t} \end{array}\right)\right] \\ &= \exp\left(\nu_3 - \nu_4 - \left(e^{\nu_4} - 1\right) h_{y,t}\right) E_t\left[\exp\left(v_4\left(n_{t+1} - 1\right)\right)\right] \times \\ &\quad \exp\left(-\frac{1}{2}\ln\left(1 - 2\nu_2\delta^2\right) + \left(u + \nu_1\right)\theta + \nu_2\theta^2 + \frac{\left(u + \nu_1 + 2\theta\nu_2\right)^2 \delta^2}{2\left(1 - 2\nu_2\delta^2\right)}\right) \end{aligned}$$

$$\begin{aligned} E^Q_t\left[\exp\left(ux^{j_0}_{t+1}\right)\right] &= \exp\left(\begin{array}{c} -\frac{1}{2}\ln\left(1 - 2\nu_2\delta^2\right) + \left(u + \nu_1\right)\theta + \nu_2\theta^2 + \frac{\left(u+\nu_1+2\theta\nu_2\right)^2\delta^2}{2\left(1-2\nu_2\delta^2\right)} \\ + \frac{1}{2}\ln\left(1 - 2\nu_2\delta^2\right) - \nu_1\theta - \nu_2\theta^2 - \frac{\left(\nu_1+2\theta\nu_2\right)^2\delta^2}{2\left(1-2\nu_2\delta^2\right)} \end{array}\right) \\ &\equiv \exp\left(u\theta^* + \frac{\left(\delta^*\right)^2}{2}u^2\right) \end{aligned}$$

with the following parameter mappings

$$\begin{aligned} \theta^* &= \theta + \frac{\left(\nu_1 + 2\theta\nu_2\right)\delta^2}{\left(1 - 2\nu_2\delta^2\right)} = \theta - \frac{\nu_2\delta^4}{\left(1 - 2\nu_2\delta^2\right)} \\ \left(\delta^*\right)^2 &= \frac{\delta^2}{1 - 2\nu_2\delta^2} \end{aligned}$$

Thus we can re-write the returns equation under the risk-neutral measure as

$$\begin{aligned} R_{t+1} &= \bar{r} + \left(\lambda_y - \xi\right) h_{y,t} + y_{t+1} \\ \bar{r} &= r - \theta - \frac{\delta^2}{2} \\ \xi &= e^{\theta + \frac{1}{2}\delta^2} - 1 \\ y_{t+1} &= \sum_{j=1}^{n_{t+1}} x^j_{t+1}, \quad x^j_{t+1} \tilde{}^Q iidN(\theta^*, \left(\delta^*\right)^2) \\ Q[n_{t+1} &= k | I_t] = \frac{e^{-h^*_{y,t}}\left(h^*_{y,t}\right)^{k-1}}{!(k-1)} \\ h^*_{y,t} &= \left(1 - \frac{\lambda_y}{\xi}\right) h_{y,t} \\ RV_{t+1} &= \sum_{j=1}^{n_{t+1}} \left(x^j_{t+1}\right)^2 - \theta^2 \end{aligned}$$

$$\begin{aligned} h_{y,t+1} &= \omega_y + b_y h_{y,t} + a_y RV_{t+1} \\ \theta^* &= \theta + \frac{\left(\nu_1 + 2\theta\nu_2\right)\delta^2}{\left(1 - 2\nu_2\delta^2\right)}, \quad \left(\delta^*\right)^2 = \frac{\delta^2}{1 - 2\nu_2\delta^2} \end{aligned}$$

where ν_2 is a parameter to be estimated.

Appendix 3E: Matlab Code for Option Pricing

```
% Computes the call option price using quadl numerical integration
% Code for BPJVM Model
% S_t = d vector of spot stock prices
% K = d vector of Strike Prices
% tau = d vector of time to maturity
% h_t = d by 2 matrix of variance processes
function [c] = cPrice_QL_BPJVM_PQ(S_t,K,h_t,rF,tau,param)
sigma = param(1); gamma = param(2); theta = param(3); delta = param(4);
rho = param(5);
b_z = param(6); b_y = param(7); a_z = param(8); a_y = param(9);
h_z0 = param(10); h_y0 = param(11); chi = param(12); nu_3 = param(13);
h_zt = h_t(:,1);
h_yt = h_t(:,2);
c = (0.5*(S_t'-K'.*exp(-rF'.*tau'))+(1/pi)*exp(-rF'.*tau').*...
quadl_v(@integ,0.00001,250,1e-06,[],sigma,gamma,theta,delta,rho,b_z,b_y,a_z,a_y,...
h_z0,h_y0,chi,nu_3,h_zt,h_yt,tau,S_t,K))';
function [f] = integ(u,sigma,gamma,theta,delta,rho,b_z,b_y,a_z,a_y,h_z0,h_y0,chi,nu_3,...
h_zt,h_yt,tau,S_t,K)
N = numel(u);
d = numel(K);
param = [sigma,gamma,theta,delta,rho,b_z,b_y,a_z,a_y,h_z0,h_y0,chi,nu_3];
x = log(K./S_t)';
h_t = [h_zt h_yt];
[Psi1 Psi2] = Psi_GARJV(u,h_t,tau,param);
f1 = real((Psi1.*exp(-1i*u*x))./(1i*repmat(u,1,d)));
f2 = real((Psi2.*exp(-1i*u*x))./(1i*repmat(u,1,d)));
f = repmat(S_t',N,1).*f1 - repmat(K',N,1).*f2;
% u is a N-column vector
% tau is a d-column vector of maturities
% Output is N by d matrix
function [Psi1 Psi2] = Psi_GARJV(u,h_t,tau,param)
u1 = 1+1i*u;
u2 = 1i*u;
h_zt = h_t(:,1);
h_yt = h_t(:,2);
N = numel(u);
T = max(tau);
C1Mat1 = zeros(N,T);
C2Mat1 = zeros(N,T);
DMat1 = zeros(N,T);
C1Mat2 = zeros(N,T);
C2Mat2 = zeros(N,T);
DMat2 = zeros(N,T);
% C and D for maturity 1
[C1Mat1(:,1) C2Mat1(:,1) DMat1(:,1)] = A12B(u1,zeros(N,1),zeros(N,1),param);
[C1Mat2(:,1) C2Mat2(:,1) DMat2(:,1)] = A12B(u2,zeros(N,1),zeros(N,1),param);
% Recursion up to M
for t = 2:T
[C1Mat1(:,t) C2Mat1(:,t) DMat1(:,t)] = A12B(u1,C1Mat1(:,t-1),C2Mat1(:,t-1),param);
[C1Mat2(:,t) C2Mat2(:,t) DMat2(:,t)] = A12B(u2,C1Mat2(:,t-1),C2Mat2(:,t-1),param);
DMat1(:,t) = DMat1(:,t) + DMat1(:,t-1);
DMat2(:,t) = DMat2(:,t) + DMat2(:,t-1);
end
Psi1 = exp(C1Mat1(:,tau).*repmat(h_zt',N,1) + C2Mat1(:,tau).*repmat(h_yt',N,1) + DMat1(:,tau));
Psi2 = exp(C1Mat2(:,tau).*repmat(h_zt',N,1) + C2Mat2(:,tau).*repmat(h_yt',N,1) + DMat2(:,tau));
% u,w_R,w_RV are N-column vectors
function [A1 A2 B] = A12B(u,v_z,v_y,param)
% Set rF = 0.05/365 for the ease of computation
rF = 0.05/365;
sigma = param(1); gamma = param(2); theta = param(3); delta = param(4);
rho = param(5);
b_z = param(6); b_y = param(7); a_z = param(8); a_y = param(9);
```

```
h_z0 = param(10); h_y0 = param(11); chi = param(12); nu_3 = param(13);
gamma_s = gamma - chi;
theta_s = theta + delta^2*nu_3;
xi_s = exp(theta_s+delta^2/2) - 1;
omega_z = h_z0*(1-(b_z+a_z));
omega_y = h_y0*(1-(b_y+a_y*(theta^2+delta^2)));
v1 = v_z*a_z;
v2 = v_y*a_y;
v3 = -0.5*log(1-2*v2*delta^2) + u*theta_s + v2*theta_s^2 +...
 ((u*delta+2*theta_s*v2*delta).^2)./(2*(1-2*v2*delta^2));
A1 = -0.5*u + v_z*b_z + v1 + sigma*(gamma_s^2-gamma^2)*v1 + 0.5*(1-rho^2)*u.^2 + ((u*rho-2*sigma*v1*gamma_s).^2)./(2*(1-
2*sigma*v1));
A2 = -xi_s*exp(theta*nu_3+0.5*delta^2*nu_3^2)*u + v_y*b_y +...
 (exp(v3)-1)*exp(theta*nu_3+0.5*delta^2*nu_3^2);
B = u*rF + v_z*omega_z + v_y*omega_y - v1*sigma - 0.5*log(1-2*sigma*v1);
```

Table 3.1: **Maximum Likelihood Estimation on Daily S&P500 Returns and Realized Measures. 1990-2013**

	GARCH		**RVM**		**JVM**		**BPJVM**	
Parameters	Estimate	Std Error	Estimate	Std Error	Estimate	Std Error	Estimate	Std Error
λ_z	4.30E-01	(6.87E-01)	4.40E-01	(1.12E+00)			4.19E-01	(1.86E+00)
λ_y					2.06E-06	(2.31E-05)	9.13E-05	(4.67E-05)
α	4.87E-06	(1.57E-07)						
β	8.50E-01	(1.13E-02)						
γ	1.53E+02	(7.66E+00)	7.40E+03	(1.81E+01)			1.45E+04	(6.23E+01)
ω_z	4.73E-14		2.35E-08				7.06E-08	
ω_y					3.38E-02		8.27E-02	
σ			5.28E-07	(2.12E-07)			2.51E-07	(1.71E-08)
θ					-7.98E-04	(3.08E-05)	1.42E-05	(1.98E-05)
δ					4.40E-03	(4.13E-06)	1.62E-03	(3.20E-06)
ρ			2.14E-01	(7.38E-02)			2.67E-01	(9.52E-02)
b_z			5.05E-01	(3.64E-02)			4.87E-01	(4.21E-02)
b_y					5.46E-01	(3.60E-02)	9.16E-01	(2.14E-02)
a_z			4.95E-01	(3.54E-02)			5.12E-01	(4.36E-02)
a_y					1.94E+04	(1.21E+02)	2.41E+04	(6.31E+02)
$E[h_{z,t}]$	1.16E-04		1.35E-04				1.24E-04	
$\mathrm{E}[h_{y,t}]$					5.70E+00		4.04E+00	
Model Properties								
Average Volatility	18.34		18.34		18.34		18.34	
Volatility Persistence								
From Returns	0.9635							
From RV			0.9998		0.9340			
From RBV							0.9998	
From RJV							0.9795	
Log Likelihoods								
Returns, RBV, and RJV							129,226	
Maximized on Returns and RV			68,212		68,783		69,656	
Maximized on Returns	19,312		19,515		19,515		19,522	

Notes: Using daily returns and daily realized variation measures we estimate our four models using maximum likelihood criteria. For comparison the last row reports likelihood values when all models are estimated on returns only. The second-to-last row reports likelihood values when the RVM, JVM, and BPJVM models are estimated on returns and realized variance. The third-to-last row reports the likelihood value when the BPJVM model is estimated on returns, bipower variation and jump variation. The parameter values reported correspond to the second-last row for RVM and JVM and to the third-last row for the BPJVM model. The sample is from January 2, 1990 through December 31, 2013. Standard errors are reported in parentheses. Variance targeting is used to fix the ω parameters.

Table 3.2: **S&P500 Index Option Data by Moneyness, Maturity and VIX Level. 1996-2013**

By Moneyness	Delta<0.3	0.3<Delta<0.4	0.4<Delta<0.5	0.5<Delta<0.6	0.6<Delta<0.7	Delta>0.7	All
Number of Contracts	3,788	1,391	1,781	2,846	2,746	8,731	21,283
Average Price	7.85	20.94	32.28	45.30	65.93	132.41	74.35
Average Implied Volatility	16.72	18.40	19.31	20.40	21.71	25.09	21.62
Average Bid-Ask Spread	1.046	1.674	1.955	2.018	1.834	1.228	1.470

By Maturity	DTM<30	30<DTM<60	60<DTM<90	90<DTM<120	120<DTM<150	DTM>150	All
Number of Contracts	2,725	6,480	5,053	2,869	1,974	2,182	21,283
Average Price	41.26	61.01	76.44	92.30	97.88	105.59	74.35
Average Implied Volatility	20.21	21.28	21.73	22.94	22.08	21.95	21.62
Average Bid-Ask Spread	0.820	1.231	1.579	1.872	1.800	1.910	1.470

By VIX Level	VIX<15	15<VIX<20	20<VIX<25	25<VIX<30	30<VIX<35	VIX>35	All
Number of Contracts	3,962	6,133	5,996	2,456	1,240	1,496	21,283
Average Price	57.95	66.90	80.75	85.77	85.33	94.86	74.35
Average Implied Volatility	13.61	18.04	22.45	26.24	30.22	39.42	21.62
Average Bid-Ask Spread	1.055	1.301	1.446	1.704	1.811	2.683	1.470

Notes: We use 21,283 S&P500 index option contracts from OptionMetrics. The contracts have been subjected to standard filters as described in the text. The top panel reports the contracts sorted by moneyness defined using the Black-Scholes delta. The second panel reports the contracts sorted by days to maturity (DTM). The third panel reports the contract sorted by the VIX level on the day corresponding to the option quote.

Table 3.3: **Maximum Likelihood Estimation on Daily S&P500 Returns, Realized Measures, and Options. 1996-2013**

	GARCH		RVM		JVM		BPJVM	
Parameters	Estimate	Std Error	Estimate	Std Error	Estimate	Std Error	Estimate	Std Error
λ_z	1.40E+01	(1.03E+01)	3.01E+00	(9.21E-01)			1.81E+01	(1.67E+00)
λ_y					1.45E-10	(4.20E-05)	4.98E-05	(4.03E-05)
α	9.01E-07	(1.86E-08)						
β	9.88E-01	(6.09E-04)						
γ	6.22E+01	(5.51E+00)	1.95E+02	(5.05E+00)			8.75E+02	(3.29E+00)
ω_z	1.64E-08		7.07E-07				1.68E-07	
ω_y					1.64E-06		1.20E-03	
σ			1.90E-05	(4.66E-07)			4.55E-06	(2.50E-09)
θ					-2.00E-03	(3.39E-05)	-2.03E-03	(1.58E-05)
δ					5.61E-03	(7.49E-06)	1.60E-03	(3.27E-06)
ρ			5.14E-01	(3.80E-03)			4.45E-01	(5.21E-03)
b_z			7.49E-01	(1.06E-03)			5.69E-01	(1.82E-03)
b_y					9.61E-01	(3.02E-04)	9.01E-01	(2.06E-03)
a_z			2.00E-01	(1.13E-03)			3.96E-01	(2.39E-03)
a_y					7.07E+02	(8.59E+00)	3.13E+04	(7.12E+02)
$E[h_{z,t}]$	1.03E-04	(9.44E-07)	4.76E-05	(8.07E-07)			7.67E-05	(4.07E-06)
$\mathrm{E}[h_{y,t}]$					1.63E+00	(4.91E-02)	1.20E+00	(9.84E-01)
χ			-3.45E-02	(1.30E-05)			-1.77E-02	(5.90E-02)
ν_2					1.33E-05	(2.54E+02)		
ν_3							2.20E-04	(5.37E+00)
Model Properties								
Average Physical Volatility	18.34		18.34		18.34		18.34	
Average Model IV	20.77		21.06		21.09		21.16	
Volatility Persistence								
From Returns	0.9911							
From RV			0.9878		0.9864			
From RBV							0.9890	
From RJV							0.9958	
Log Likelihoods								
Returns, RBV, RJV, and Options							163,571	
Returns	19,019		19,498		18,955		19,406	
Returns and Options	52,770		57,236		55,550		57,782	
Option Errors								
IVRMSE	5.74		4.72		5.05		4.50	
Ratio to GARCH	1.000		0.822		0.879		0.784	
VWRMSE	4.96		4.14		4.37		3.94	
Ratio to GARCH	1.000		0.835		0.881		0.796	

Notes: Using daily returns, daily realized variation measures and options we estimate our four models using a joint maximum likelihood criterion. The table reports the joint likelihood value as well as its decomposition into the various components. Option errors are reported using implied volatility root mean squared errors (IVRMSE) and vega-weighted root mean squared errors (VWRMSE) as defined in the text. The sample is from January 2, 1996 through August 28, 2013. Standard errors are reported in parentheses. Physical variance targeting is used to fix the ω parameters.

Table 3.4: **Implied Volatility Root Mean Squared Error (IVRMSE) by Moneyness, Maturity, and VIX Level. 1996-2013**

Panel A: IVRMSE by Moneyness

Model	Delta<0.3	0.3<Delta<0.4	0.4<Delta<0.5	0.5<Delta<0.6	0.6<Delta<0.7	Delta>0.7
GARCH	5.338	4.001	3.823	3.896	4.228	8.112
RVM	5.014	2.971	2.717	2.869	3.137	5.118
JVM	5.059	3.568	3.214	3.070	3.389	7.123
BPJVM	5.365	2.859	2.712	2.859	3.027	4.970

Panel B: IVRMSE by Maturity

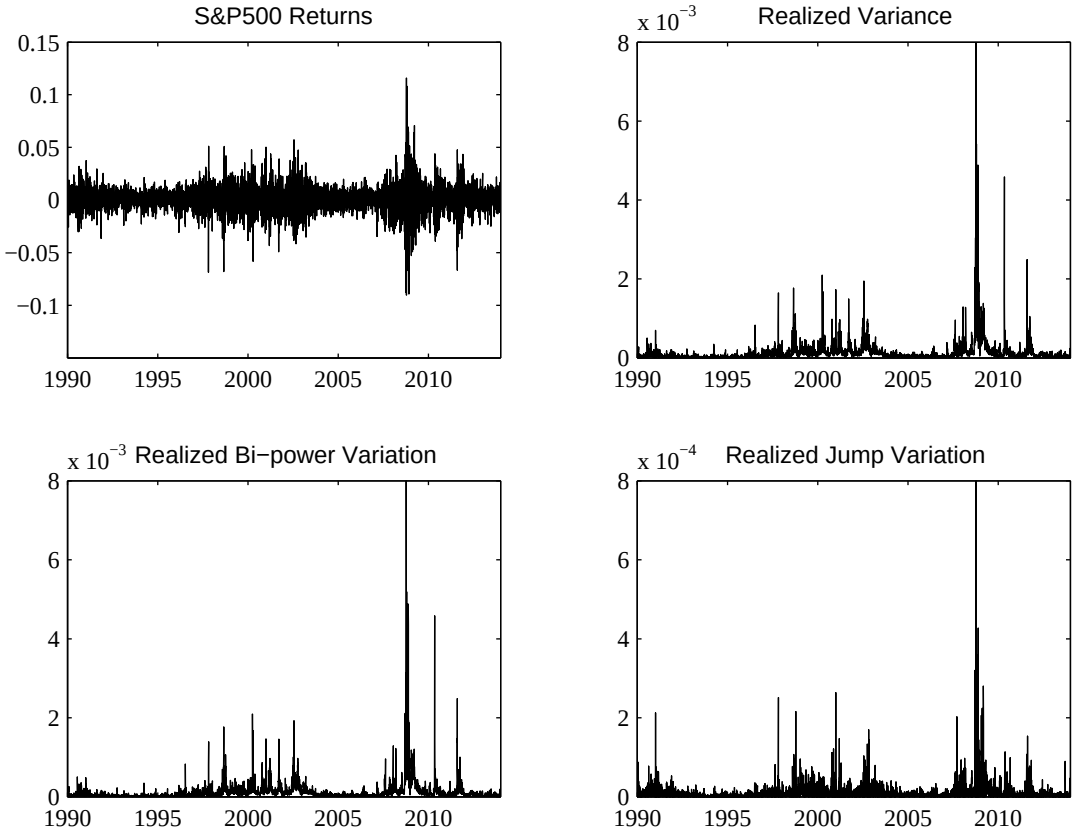
Model	DTM<30	30<DTM<60	60<DTM<90	90<DTM<120	120<DTM<150	DTM>150
GARCH	5.259	5.700	5.640	5.852	6.497	5.834
RVM	3.963	4.247	4.182	4.402	5.027	4.428
JVM	4.437	4.731	4.956	5.078	6.115	5.755
BPJVM	3.904	4.074	3.887	4.069	4.762	4.271

Panel C: IVRMSE by VIX Level

Model	VIX<15	15<VIX<20	20<VIX<25	25<VIX<30	30<VIX<35	VIX>35
GARCH	4.310	3.419	5.499	6.785	6.961	11.639
RVM	3.137	2.997	4.437	5.379	5.938	6.812
JVM	3.956	3.220	5.466	6.356	6.357	7.463
BPJVM	2.960	2.664	4.165	5.073	5.765	6.878

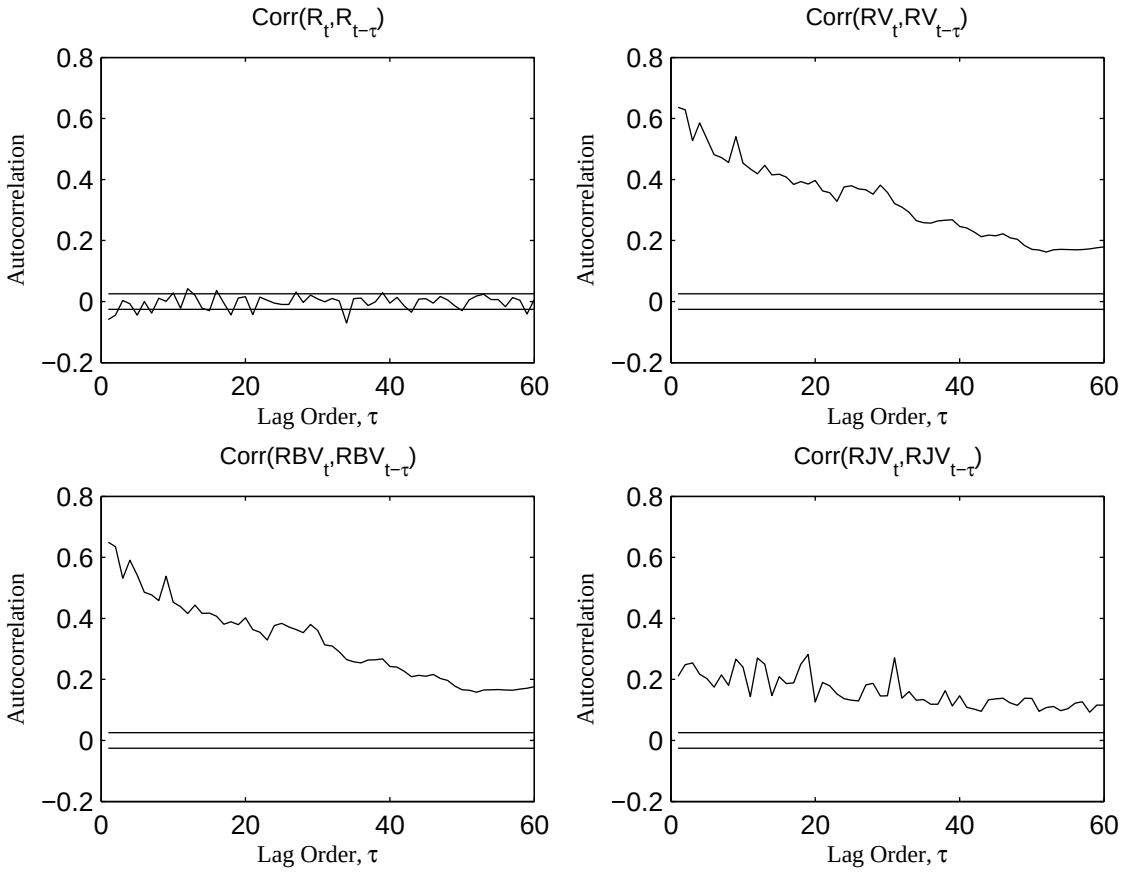
Notes: We use the parameter values in Table 3 to fit our four models to the 21,283 S&P500 index option contracts from OptionMetrics. The top panel reports IVRMSE for contracts sorted by moneyness defined using the Black-Scholes delta. The second panel reports IVRMSE for contracts sorted by days to maturity (DTM). The third panel reports the IVRMSE for contract sorted by the VIX level on the day corresponding to the option quote.

Figure 3.1: Daily Returns and Realized Variation Measures



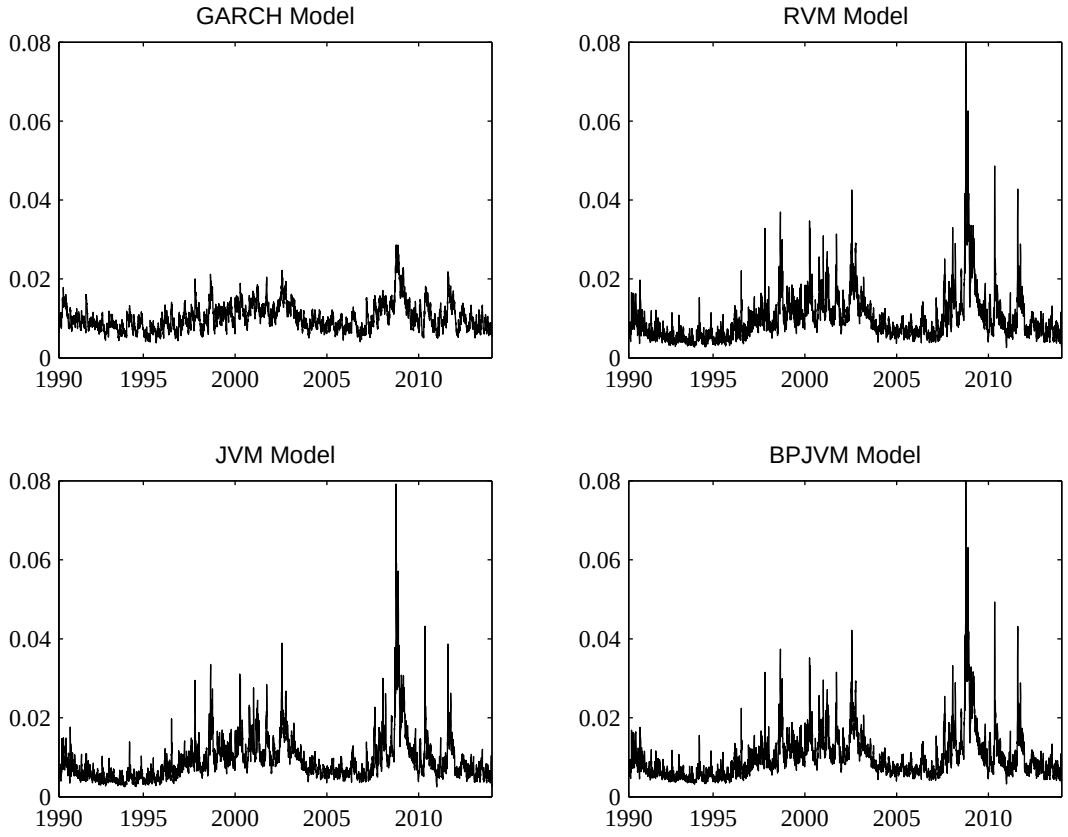
Notes: The top-left panel shows the daily log returns on the S&P500 index. The top-right panel shows the daily realized variance computed from averages of sum of squared overlapping 5-minute returns. The bottom left panel shows the realized bipower variation computed using the method in Barndorff-Nielsen and Shephard (2004). The bottom right panel shows the realized jump variation constructed as the residual between realized variance and realized bipower variation. The sample goes from January 2, 1990 through December 31, 2013.

Figure 3.2: Autocorrelations of Daily Returns and Realized Variation Measures



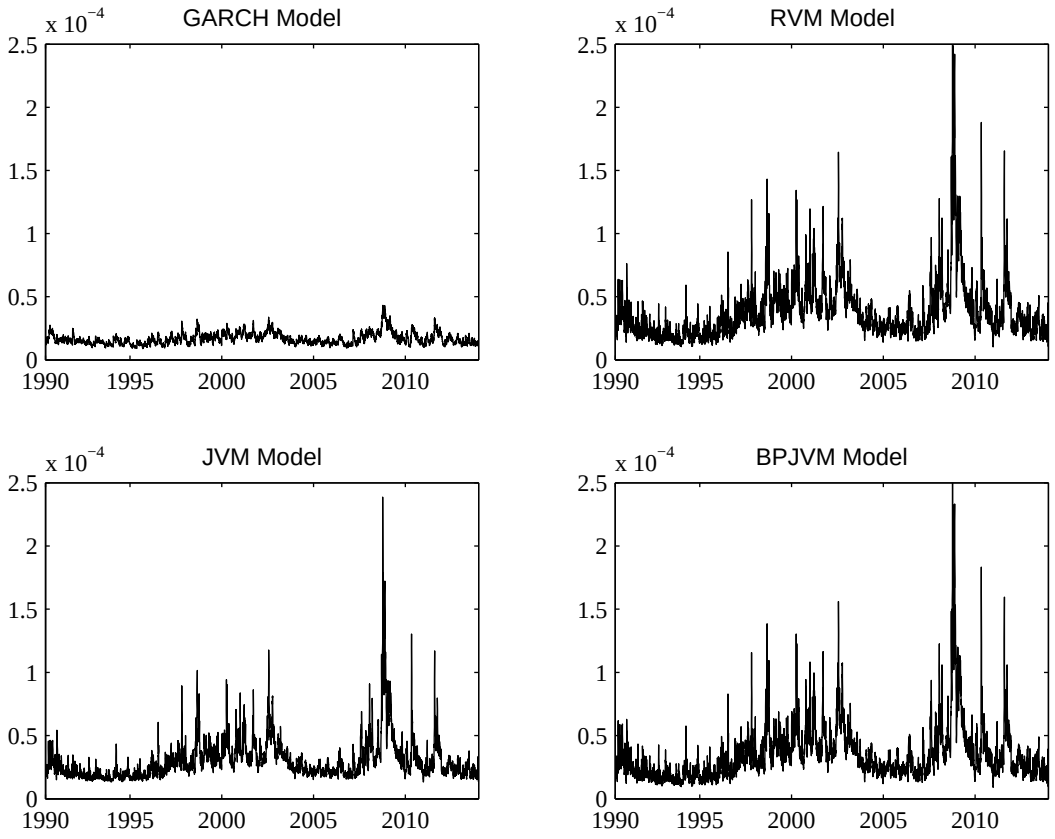
Notes: We report the sample autocorrelation functions for lag 1 through 60 trading days for returns (top-left panel), realized volatility (top-right panel), realized bipower variation (bottom-left panel), and realized jump variation (bottom-right panel). The sample goes from January 2, 1990 through December 31, 2013.

Figure 3.3: Daily Conditional Volatility



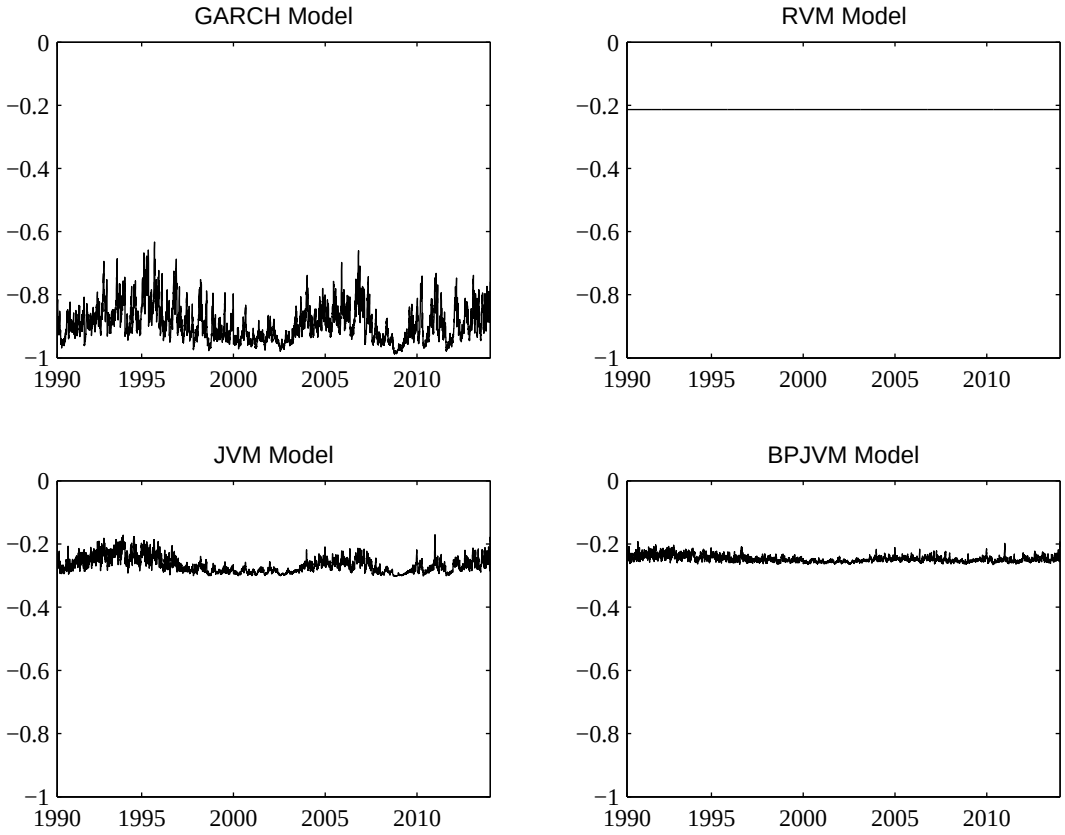
Notes: We plot the daily model-based conditional volatility for the four models we consider: The benchmark Heston-Nandi GARCH model (top-left), the RVM model based on realized volatility (top-right), the JVM model based on realized jump variation only (bottom-left), and the full BPJVM model that separately uses bipower variation and realized jump-variation (bottom-right). We use the parameter estimates from Table 1. The sample goes from January 2, 1990 through December 31, 2013.

Figure 3.4: Conditional Volatility of Variance



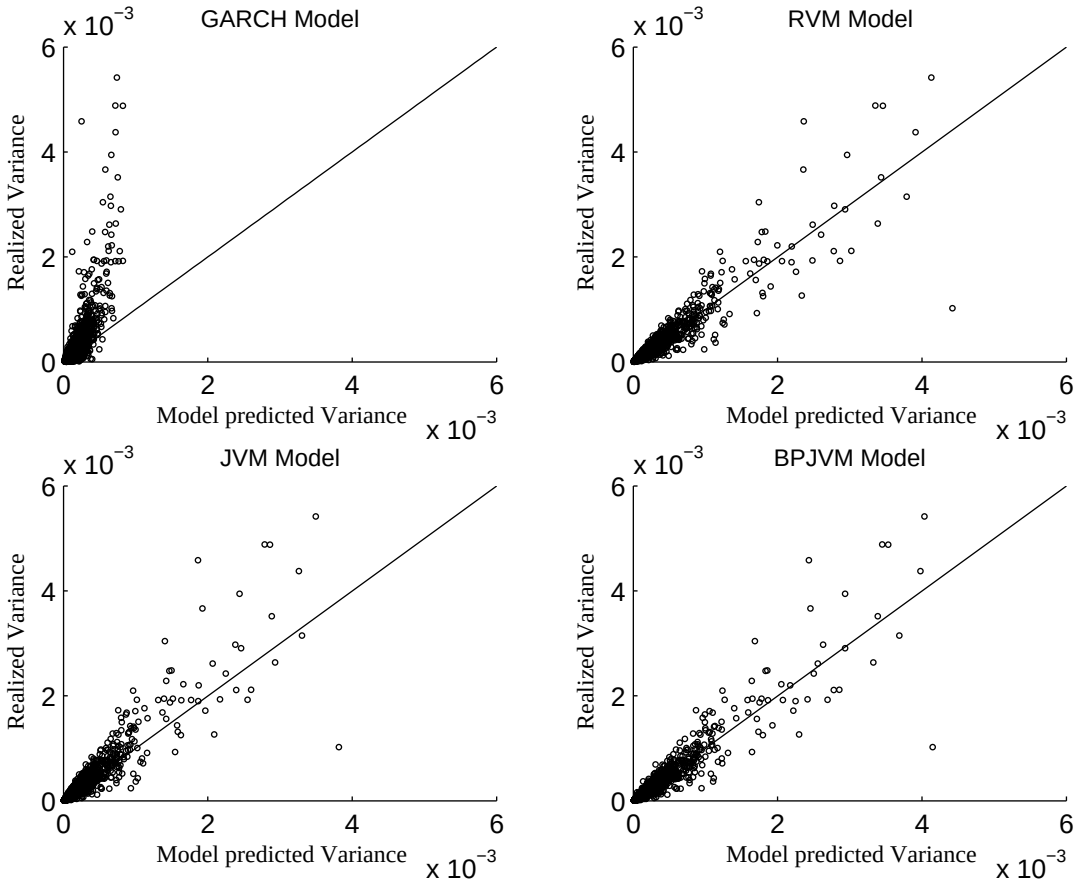
Notes: We plot the daily model-based conditional volatility of variance for the four models we consider: The benchmark Heston-Nandi GARCH model (top-left), the RVM model based on realized volatility (top-right), the JVM model based on realized jump variation only (bottom-left), and the full BPJVM model that separately uses realized bipower variation and realized jump-variation (bottom-right). We use the parameter estimates from Table 1. The sample goes from January 2, 1990 through December 31, 2013.

Figure 3.5: Daily Correlation of Return and Variance



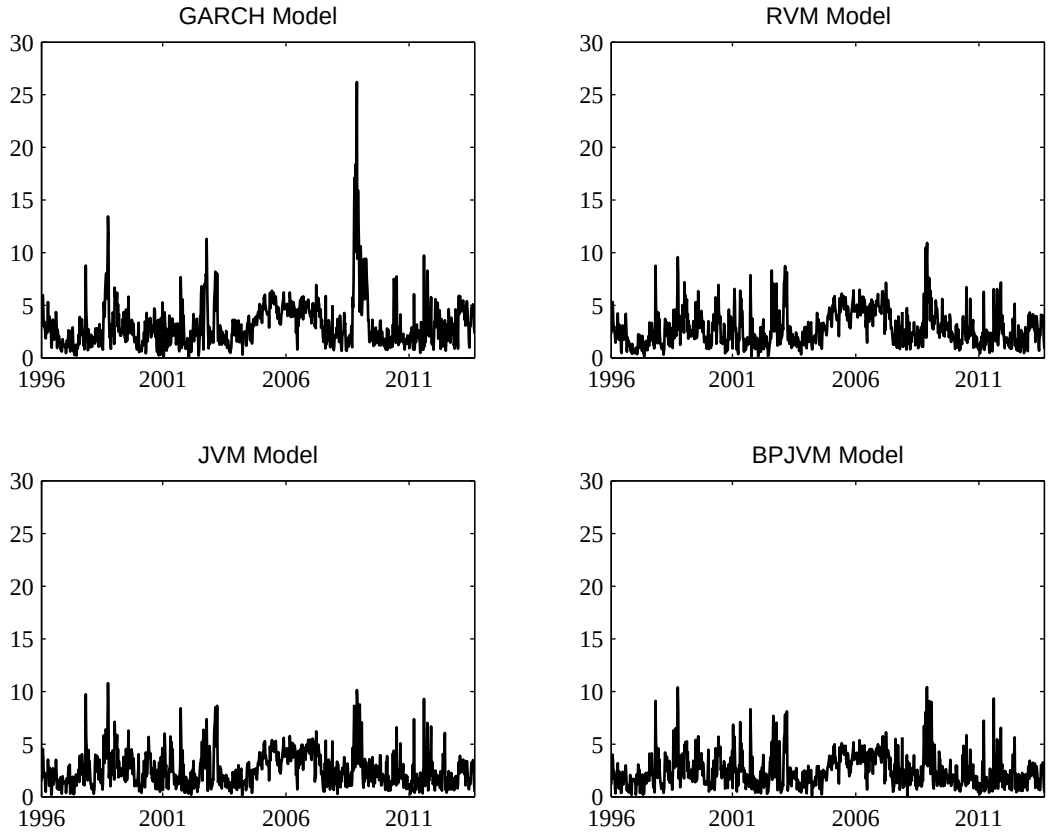
Notes: We plot the daily model-based conditional correlation of return and variance for the four models we consider: The benchmark Heston-Nandi GARCH model (top-left), the RVM model based on realized volatility (top-right), the JVM model based on realized jump variation only (bottom-left), and the full BPJVM model that separately uses realized bipower variation and realized jump-variation (bottom-right). We use the parameter estimates from Table 1. The sample goes from January 2, 1990 through December 31, 2013.

Figure 3.6: Realized Volatility and Predicted Volatility from Models



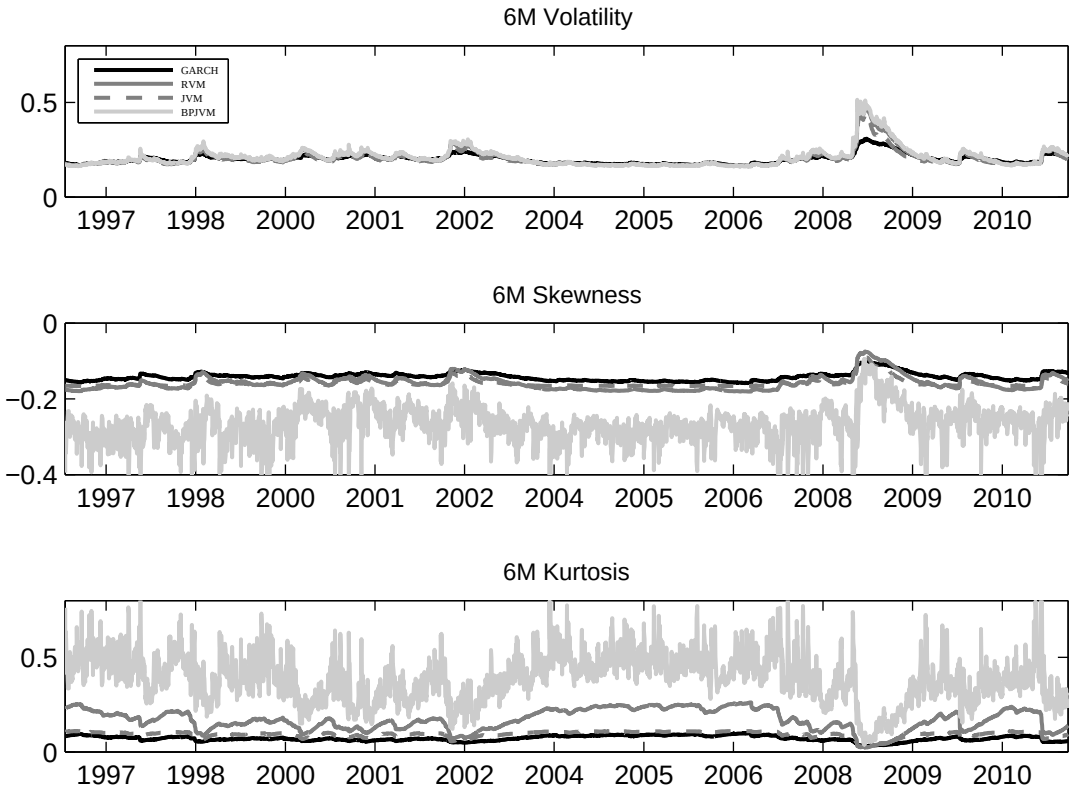
Notes: We scatter plot the ex-post realized variance (vertical axis) against the model-predicted total variance (horizontal axis) for each of our models: The benchmark Heston-Nandi GARCH model (top-left), the RVM model based on realized volatility (top-right), the JVM model based on realized jump variation only (bottom-left), and the full BPJVM model that separately uses realized bipower variation and realized jump-variation (bottom-right). We use the parameter estimates from Table 1. The sample goes from January 2, 1990 through December 31, 2013.

Figure 3.7: Weekly Implied Root Mean Squared Error from At-the-Money Options



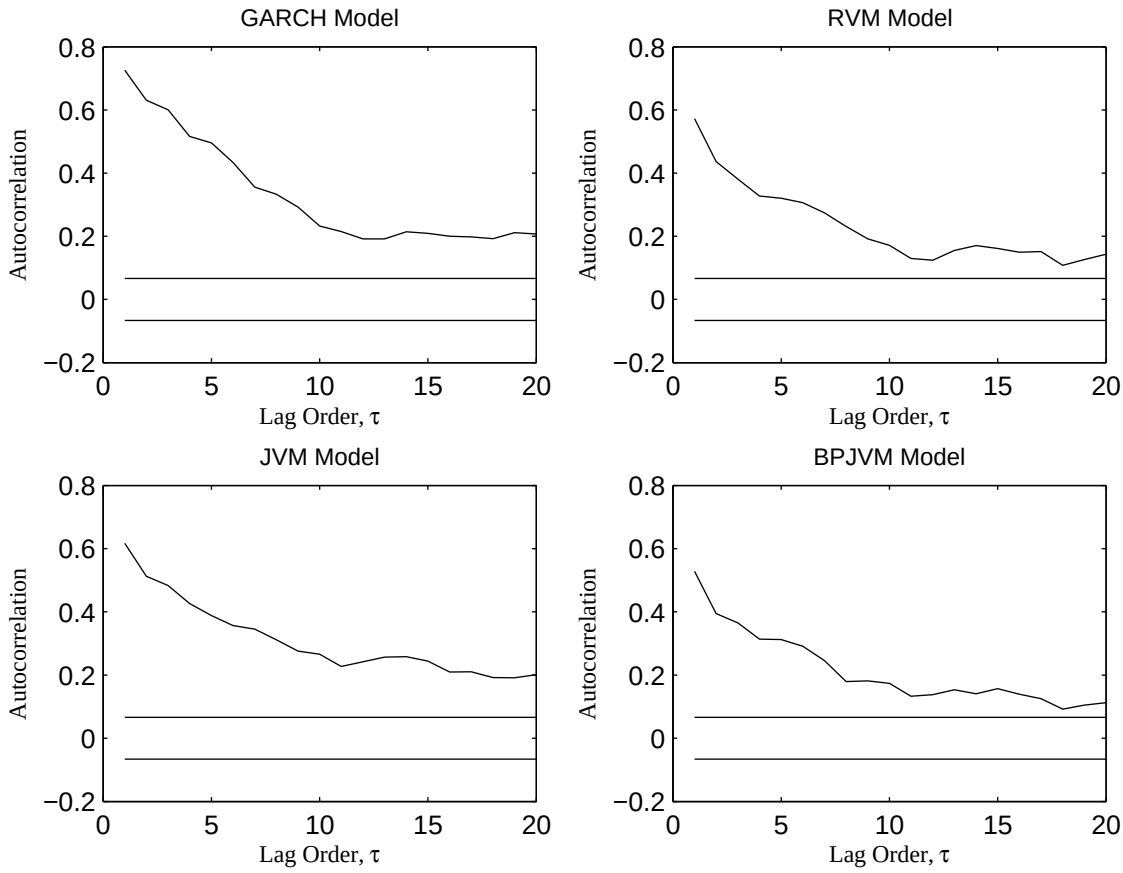
Notes: We plot the weekly implied volatility root mean squared error for at the money options for each of our models: The benchmark Heston-Nandi GARCH model (top-left), the RVM model based on realized volatility (top-right), the JVM model based on realized jump variation only (bottom-left), and the full BPJVM model that separately uses realized bipower variation and realized jump-variation (bottom-right). We use the parameter estimates from Table 2. The option sample goes from January 2, 1996 through August 28, 2013.

Figure 3.8: Model-Based, Risk-Neutral Higher Moments. Six-Month Horizon



Notes: We plot the six-month risk-neutral volatility, skewness and kurtosis implied by each of our models: The benchmark Heston-Nandi GARCH model (top-left), the RVM model based on realized volatility (top-right), the JVM model based on realized jump variation only (bottom-left), and the full BPJVM model that separately uses realized bipower variation and realized jump-variation. We use the parameter estimates from Table 2. The option sample goes from January 2, 1996 through August 28, 2013.

Figure 3.9: Autocorrelations of Weekly Vega-Weighted Root Mean Squared Errors of ATM Options



Notes: We plot autocorrelations of the weekly vega-weighted root mean squared error for at the money options from each of our models: The benchmark Heston-Nandi GARCH model (top-left), the RVM model based on realized volatility (top-right), the JVM model based on realized jump variation only (bottom-left), and the full BPJVM model that separately uses realized bipower variation and realized jump-variation (bottom-right). We use the parameter estimates from Table 2. The option sample goes from January 2, 1996 through August 28, 2013.

Bibliography

Aït-Sahalia, Y., and J. Jacod (2012): "Analyzing the Spectrum of Asset Returns: Jump and Volatility Components in High Frequency Data," *Journal of Economic Literature*, 50(4), 1007–1050.

Aït-Sahalia, Y., M. Karaman, and L. Mancini (2015): "The Term Structure of Variance Swaps and Risk Premia," *Working Paper*.

Aït-Sahalia, Y., and J. Yu (2009): "High Frequency Market Microstructure Noise Estimates and Liquidity Measures," *Annals of Applied Statistics*, 3(1), 422–457.

Allen, F., and D. Gale (1994): "Limited Market Participation and Volatility of Asset Prices," *The American Economic Review*, 84(4), 933–955.

Altman, E. I. (1968): "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *The Journal of Finance*, 23(4), 589–609.

Amaya, D., J.-F. Bégin, and G. Gauthier (2016): "Extracting Latent States with High Frequency Option Prices," *Working Paper*.

Amaya, D., P. Christoffersen, K. Jacobs, and A. Vasquez (2015): "Does Realized Skewness Predict the Cross-section of Equity Returns?," *Journal of Financial Economics*, 118(1), 135–167.

Amihud, Y. (2002): "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets*, 5(1), 31–56.

Andersen, T. G. (1996): "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility," *The Journal of Finance*, 51(1), 169–204.

Andersen, T. G., L. Benzoni, and J. Lund (2002): "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57, 1239–1284.

Andersen, T. G., and T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, (39), 885–905.

Andersen, T. G., T. Bollerslev, and F. Diebold (2007): "Roughing it Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89(4), 701–720.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71(2), 579–625.

Andersen, T. G., D. Dobrev, and E. Schaumburg (2012): "Jump-Robust Volatility Estimation Using Nearest Neighbor Truncation," *Journal of Econometrics*, 169(1), 75–93.

Andersen, T. G., N. Fusari, and V. Todorov (2015a): "Parametric Inference and Dynamic State Recovery From Option Panels," *Econometrica*, 83(3), 1081–1145.

——— (2015b): "The Risk Premia Embedded in Index Options," *Journal of Financial Economics*, 117(3), 558–584.

——— (2016): "The Pricing of Tail Risk and the Equity Premium: Evidence from International Option Markets," Working paper.

Bajgrowicz, P., and O. Scaillet (2011): "Jumps in high-frequency data: spurious detections, dynamics, and news," *Working Paper*.

Bakshi, G., C. Cao, and Z. Chen (1997): "Empirical Performance of Alternative Option Pricing Models," *The Journal of Finance*, 52(5), 2003–2049.

Bakshi, G., P. Carr, and L. Wu (2008): "Stochastic Risk Premiums, Stochastic Skewness in Currency Options, and Stochastic Discount Factors in International Economies," *Journal of Financial Economics*, 87(1), 132–156.

Bakshi, G., and D. Madan (2000): "Spanning and Derivative Security Valuation," *Journal of Financial Economics*, 55(2), 205–238.

Barndorff-Nielsen, O. E., and N. Shephard (2002): "Econometric analysis of realized volatility and its use in estimating stochastic volatility models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2), 253–280.

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

——— (2006): "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4(1), 1.

Bates, D. (2000): "Post-87 crash fears in the S&P500 futures option market," *Journal of Econometrics*, 94, 181 – 238.

——— (2006): "Maximum Likelihood Estimation of Latent Affine Processes," *Review of Financial Studies*, 26(9), 909–965.

——— (2012): "U.S. Stock Market Crash Risk, 1926-2010," *Journal of Financial Economics*, 105(2), 229–259.

Bégin, J.-F., C. Dorion, and G. Gauthier (2016): "Idiosyncratic Jump Risk Matters: Evidence from Equity Returns and Options," *Working Paper*.

Berry, T. D., and K. M. Howe (1994): "Public Information Arrival," *The Journal of Finance*, 49(4), 1331–1346.

Bollerslev, T., U. Kretschmer, C. Pigorsch, and G. Tauchen (2009): "A discrete-time model for daily S&P500 returns and realized variations: Jumps and leverage effects," *Journal of Econometrics*, 150(2), 151 – 166, Recent Development in Financial Econometrics.

Broadie, M., M. Chernov, and M. Johannes (2007): "Model Specification and Risk Premia: Evidence from Futures Options," *Journal of Finance*, 63, 1453–1490.

Brogaard, J., D. Li, and Y. Xia (2016): "Stock Liquidity and Default Risk," *Journal of Financial Economics, forthcoming*.

Busch, T., B. J. Christensen, and M. Nielsen (2011): "The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets," *Journal of Econometrics*, 160(1), 48 – 57, Realized Volatility.

Campbell, J. Y., S. J. Grossman, and J. Wang (1993): "Trading Volume and Serial Correlation in Stock Returns," *The Quarterly Journal of Economics*, 108(4), 905–939.

Carr, P., and D. B. Madan (1999): "Option Valuation Using the Fast Fourier Transform," *Journal of Computational Finance*, 2(4), 61–73.

Carr, P., and L. Wu (2009): "Variance Risk Premiums," *Review of Financial Studies*, 22(3), 1311–1341.

Cespa, G., and T. Foucault (2014): "Illiquidity Contagion and Liquidity Crashes," *Review of Financial Studies*, 27(6), 1615–1660.

Chae, J. (2005): "Trading Volume, Information Asymmetry, and Timing Information," *The Journal of Finance*, 60(1), 413–442.

Chang, X., Y. Chen, and L. Zolotoy (2016): "Stock Liquidity and Stock Price Crash Risk," *Journal of Financial and Quantitative Analysis, forthcoming*.

Chen, J., H. Hong, and J. C. Stein (2001): "Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices," *Journal of Financial Economics*, 61(3), 345–381.

Chen, X., and E. Ghysels (2011): "News-Good or Bad-and Its Impact on Volatility Predictions over Multiple Horizons," *Review of Financial Studies*, 24(1), 46–81.

Christoffersen, P., C. Dorion, K. Jacobs, and L. Karoui (2014): "Nonlinear Kalman Filtering in Affine Term Structure Models," *Management Science*, 60(9), 2248–2268.

Christoffersen, P., R. Elkamhi, B. Feunou, and K. Jacobs (2010): "Option Valuation with Conditional Heteroskedasticity and Nonnormality," *The Review of Financial Studies*, 23(5), 2139.

Christoffersen, P., B. Feunou, K. Jacobs, and N. Meddahi (2014): "The Economic Value of Realized Volatility: Using High-Frequency Returns for Option Valuation," *Journal of Financial and Quantitative Analysis*, 49(3), 663–697.

Christoffersen, P., M. Fournier, and K. Jacobs (2015): "The Factor Structure in Equity Options," *Rotman School of Management Working Paper*.

Christoffersen, P., R. Goyenko, K. Jacobs, and M. Karoui (2015): "Illiquidity Premia in the Equity Options Market," *Working Paper, University of Toronto*.

Christoffersen, P., S. Heston, and K. Jacobs (2013): "Capturing Option Anomalies with a Variance-Dependent Pricing Kernel," *Review of Financial Studies*, 26(8), 1963–2006.

Christoffersen, P., K. Jacobs, and C. Ornthanalai (2012): "Dynamic Jump Intensities and Risk Premiums: Evidence from S&P 500 Returns and Options," *Journal of Financial Economics*, 106(3), 447–472.

Chung, K. H., and C. Chuwonganant (2014): "Uncertainty, Market Structure, and Liquidity," *Journal of Financial Economics*, 113(3), 476–499.

Clark, P. K. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41(1), 135–155.

Corsi, F., N. Fusari, and D. L. Vecchia (2013): "Realizing smiles: Options pricing with realized volatility," *Journal of Financial Economics*, 107(2), 284 – 304.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990): "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98(4), 703–738.

Dionne, G., G. Gauthier, K. Hammami, M. Maurice, and J.-G. Simonato (2011): "A reduced form model of default spreads with Markov-switching macroeconomic factors," *Journal of Banking and Finance*, 35(8), 1984 – 2000.

Drissen, J., P. J. Maenhout, and G. Vilkov (2009): "The Price of Correlation Risk: Evidence from Equity Options," *The Journal of Finance*, 64(3), 1377–1406.

Du, J., and N. Kapadia (2012): "Tail and Volatility Indices from Option Prices," *Working Paper, University of Massachusetts, Amherst*.

Dubinsky, A., and M. Johannes (2006): "Earnings Announcements and Equity Options," *Working Paper*.

Duffie, D., J. Pan, and K. Singleton (2000): "Transform Analysis and Asset Pricing for Affine Jump-diffusions," *Econometrica*, 68(6), 1343–1376.

Duffie, D., L. Saita, and K. Wang (2007): "Multi-period corporate default prediction with stochastic covariates," *Journal of Financial Economics*, 83(3), 635 – 665.

Easley, D., M. López de Prado, and M. O'Hara (2010): "The microstructure of the Flash Crash: Flow toxicity, liquidity crashes and the probability of informed trading," *The Journal of Portfolio Management*, 37(2), 118–128.

Easley, D., and M. O'hara (2004): "Information and the Cost of Capital," *The Journal of Finance*, 59(4), 1553–1583.

Ederington, L. H., and J. H. Lee (1993): "How Markets Process Information: News Releases and Volatility," *The Journal of Finance*, 48(4), 1161–1191.

Egloff, D., M. Leippold, and L. Wu (2010): "The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments," *Journal of Financial and Quantitative Analysis*, 45, 1279–1310.

Engle, R. F., and G. M. Gallo (2006): "A multiple indicators model for volatility using intra-daily data," *Journal of Econometrics*, 131, 3 – 27.

Engle, R. F., M. K. Hansen, and A. Lunde (2012): "And Now, The Rest of the News: Volatility and Firm Specific News Arrival," *Working Paper*.

Epps, T. W., and M. L. Epps (1976): "The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis," *Econometrica*, 44(2), 305–321.

Eraker, B. (2004): "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," *Journal of Finance*, 59(3), 1367–1404.

Eraker, B., M. Johannes, and N. Polson (2003): "The Impact of Jumps in Volatility and Returns," *Journal of Finance*, 58, 1269–1300.

Filipović, D., E. Gourier, and L. Mancini (2016): "Quadratic Variance Swap Models," *Journal of Financial Economics*, 119(1), 44–68.

Fleming, M. J. (2003): "Measuring Treasury Market Liquidity," *Economic Policy Review*, (9), 83–108.

Forsberg, L., and T. Bollerslev (2002): "Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): the GARCH-NIG model," *Journal of Applied Econometrics*, 17(5), 535–548.

Gennotte, G., and H. Leland (1990): "Market Liquidity, Hedging, and Crashes," *The American Economic Review*, 80(5), 999–1021.

Gilder, D., M. B. Shackleton, and S. J. Taylor (2014): "Cojumps in stock prices: Empirical evidence," *Journal of Banking & Finance*, 40, 443 – 459.

Gourier, E. (2016): "Pricing of Idiosyncratic Equity and Variance Risk," *Working Paper*.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka (2009): "Do Liquidity Measures Measure Liquidity?," *Journal of Financial Economics*, 92(2), 153–181.

Grossman, S. J., and M. H. Miller (1988): "Liquidity and Market Structure," *The Journal of Finance*, 43(3), 617–633.

Han, B., Y. Tang, and L. Yang (2016): "Public information and uninformed trading: Implications for market liquidity and price efficiency," *Journal of Economic Theory*, 163, 604 – 643.

Hansen, P. R., Z. Huang, and H. H. Shek (2012): "Realized GARCH: a joint model for returns and realized measures of volatility," *Journal of Applied Econometrics*, 27(6), 877–906.

Hansen, P. R., and A. Lunde (2005): "A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data," *Journal of Financial Econometrics*, 3(4), 525.

Hasanhodzic, J., and A. Lo (2011): "Black's Leverage Effect Is Not Due To Leverage," *Working Paper, MIT*.

Hayashi, F. (2000): "Econometrics," *Princeton University Press*.

Heston, S. (1993): "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6(2), 327–343.

Heston, S. L., and S. Nandi (2000): "A Closed-Form GARCH Option Valuation Model," *The Review of Financial Studies*, 13(3), 585.

Hiemstra, C., and J. D. Jones (1994): "Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation," *The Journal of Finance*, 49(5), 1639–1664.

Huang, J., and J. Wang (2009): "Liquidity and Market Crashes," *Review of Financial Studies*, 22(7), 2607–2643.

Huang, J.-z., and L. Wu (2004): "Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes," *The Journal of Finance*, 59(3), 1405–1439.

Huang, X., and G. Tauchen (2005): "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3(4), 456–499.

Jeon, Y., T. H. McCurdy, and X. Zhao (2016): "News as Sources of Jumps in Stock Returns: Evidence from 26 Million News Articles," *Working Paper*.

Jiang, G. J., and T. Yao (2013): "Stock Price Jumps and Cross-Sectional Return Predictability," *Journal of Financial and Quantitative Analysis*, 48, 1519–1544.

Jones, C. M. (2002): "A Century of Stock Market Liquidity and Trading Costs," *Working Paper*.

Kaeck, A., and C. Alexander (2012): "Volatility dynamics for the S&P 500: Further evidence from non-affine, multi-factor jump diffusions," *Journal of Banking and Finance*, 36(11), 3110 – 3121, International Corporate Finance Governance Conference.

Kanniainen, J., B. Lin, and H. Yang (2014): "Estimating and using {GARCH} models with {VIX} data for option valuation," *Journal of Banking and Finance*, 43, 200 – 211.

Kozhan, R., A. Neuberger, and P. Schneider (2014): "The Skew Risk Premium in the Equity Index Market," *Review of Financial Studies*, 26(9), 2174–2203.

Lamoureux, C. G., and W. D. Lastrapes (1990): "Heteroskedasticity in Stock Return Data: Volume Versus GARCH Effects," *Journal of Finance*, 45(1), 221–229.

——— (1994): "Endogenous Trading Volume and Momentum in Stock-Return Volatility," *Journal of Business & Economic Statistics*, 12(2), 253–260.

Lee, C. M. C., B. Mucklow, and M. J. Ready (1993): "Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis," *Review of Financial Studies*, 6(2), 345–374.

Lee, S. S. (2012): "Jumps and Information Flow in Financial Markets," *Review of Financial Studies*, 25(2), 440–479.

Lee, S. S., and P. A. Mykland (2008): "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics," *Review of Financial Studies*, 21(6), 2535–2563.

Li, J. (2013): "Robust Estimation and Inference for Jumps in Noisy High Frequency Data: A Local-to-Continuity Theory for the Pre-Averaging Method," *Econometrica*, 81(4), 1673–1693.

Lo, A. W., H. Mamaysky, and J. Wang (2004): "Asset Prices and Trading Volume under Fixed Transactions Costs," *Journal of Political Economy*, 112(5), 1054–1090.

Loughran, T., and B. McDonald (2011): "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *The Journal of Finance*, 66(1), 35–65.

Maheu, J., and T. McCurdy (2004): "News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns," *Journal of Finance*, 59(2), 755–793.

Maheu, J. M., T. H. McCurdy, and X. Zhao (2013): "Do Jumps Contribute to the Dynamics of the Equity Premium?," *Journal of Financial Economics*, 110(2), 457–477.

Mancini, C. (2009): "Non-Parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps," *Scandinavian Journal of Statistics*, 36(2), 270–296.

Mitchell, M. L., and J. H. Mulherin (1994): "The Impact of Public Information on the Stock Market," *The Journal of Finance*, 49(3), 923–950.

Ornthanalai, C. (2014): "Lévy Jump Risk: Evidence from Options and Returns," *Journal of Financial Economics*, 112(1), 69–90.

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

Richardson, M., and T. Smith (1994): "A Direct Test of the Mixture of Distributions Hypothesis: Measuring the Daily Flow of Information," *Journal of Financial and Quantitative Analysis*, 29, 101–116.

Rogers, J. L., D. J. Skinner, and A. V. Buskirk (2009): "Earnings guidance and market uncertainty," *Journal of Accounting and Economics*, 48(1), 90 – 109.

Santa-Clara, P., and S. Yan (2010): "Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options," *Review of Economics and Statistics*, 92(2), 435–451.

Schwert, G. W. (1989): "Why Does Stock Market Volatility Change Over Time?," *The Journal of Finance*, 44(5), 1115–1153.

Shephard, N., and K. Sheppard (2010): "Realising the future: forecasting with high-frequency-based volatility (HEAVY) models," *Journal of Applied Econometrics*, 25(2), 197–231.

Shumway, T. (2001): "Forecasting Bankruptcy More Accurately: A Simple Hazard Model," *The Journal of Business*, 74(1), 101–124.

Stentoft, L. (2015): "What We Can Learn from Pricing 139,879 Individual Stock Options," *Journal of Derivatives*, 22(4), 54–78.

Tauchen, G., and M. Pitts (1983): "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51(2), 485–505.

Tetlock, P. C. (2007): "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *The Journal of Finance*, 62(3), 1139–1168.

Trolle, A. B., and E. S. Schwartz (2009): "Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives," *Review of Financial Studies*, 22(11), 4423–4461.

Van der Merwe, R., and E. A. Wan (2001): "The Square-Root Unscented Kalman Filter for State and Parameter-Estimation," in *Proceedings. 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3461–3464.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48(4), 817–838.

Xing, Y., X. Zhang, and R. Zhao (2010): "What Does the Individual Option Volatility Smirk Tell Us About Future Equity Returns?," *Journal of Financial and Quantitative Analysis*, 45(3), 641–662.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005): "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100(472), 1394–1411.

Zhao, X. (Forthcoming): "Does Information Intensity Matter for Stock Returns? Evidence from Form 8-K Filings," *Management Science*.