Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Jumps in Equity Index Returns Before and During the Recent Financial Crisis: A Bayesian Analysis

Steven Kou, Cindy Yu, Haowen Zhong

To cite this article:
Steven Kou, Cindy Yu, Haowen Zhong (2017) Jumps in Equity Index Returns Before and During the Recent Financial Crisis: A Bayesian Analysis. Management Science 63(4):988-1010. https://doi.org/10.1287/mnsc.2015.2359

http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Vol. 63, No. 4, April 2017, pp. 988–1010
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Jumps in Equity Index Returns Before and During the Recent Financial Crisis: A Bayesian Analysis

Steven Kou,[a] **Cindy Yu,**[b] **Haowen Zhong**[c]

[a] Risk Management Institute and Department of Mathematics, NUS Risk Management Institute, National University of Singapore, Singapore 119613; [b] Department of Statistics, Iowa State University, Ames, Iowa 50011; [c] Department of Industrial Engineering and Operations Research, Columbia University, New York, New York 10027

Contact: matsteve@nus.edu.sg (SK); cindyyu@iastate.edu (CY); hz2193@columbia.edu (HZ)

Received: June 10, 2013
Revised: September 30, 2014; May 21, 2015
Accepted: September 16, 2015
Published Online in Articles in Advance: March 28, 2016; updated October 11, 2016

https://doi.org/10.1287/mnsc.2015.2359

Abstract. We attempt to answer two questions in this paper: (i) How did jumps in equity returns change after the 2008–2009 financial crisis—in particular, were there significant changes in jump rates or in jump sizes, or both? (ii) Can the performance of affine jump-diffusion models be improved if jump sizes are larger, i.e., jumps with tails heavier than those of the normal distribution? To answer the second question, we find that a simple affine jump-diffusion model with both stochastic volatility and double-exponential jumps fits both the S&P 500 and the NASDAQ-100 daily returns from 1980 to 2013 well; the model outperforms existing ones (e.g., models with variance-gamma jumps or jumps in volatility) during the crisis and is at least comparable before the crisis. For the first question, on the basis of the model and the data sets, we observe that during the crisis, negative jump rate increased significantly, although there was little change in the average negative jump size.

History: Accepted by Jerome Detemple, finance.
Funding: The research of S. Kou is in part supported by the Singapore Ministry of Education [Grant MOE2014-T2-1-006].

Keywords: affine jump-diffusion models • equity returns • financial crisis • Bayesian Markov chain Monte Carlo

1. Introduction

It is well known that jump risk affects equity returns significantly; see, e.g., Duffie et al. (2000) and Singleton (2006) and the references therein. We attempt to answer two questions about jumps in equity returns: (i) How did jumps in equity returns change after the 2008 financial crisis—in particular, were there significant changes in jump rates or in jump sizes, or both? (ii) Can the performance of affine jump-diffusion models be improved if jump sizes are larger, i.e., jumps with tails heavier than those of the normal distribution, such as the double-exponential distribution? For the first question, the increases of jump rates when the market is in distress, especially in the 1987 crash and the tech bubble burst in 2001–2002, are documented in Eraker (2004) and Johannes et al. (1999). However, whether there are significant changes in jump sizes during financial crises has not been addressed in the existing literature; also, it is not clear whether the previous conclusion about jump rates still holds for the financial crisis of 2008. In addition, the previous empirical studies in general do not distinguish positive and negative jump rates. The second question is interesting because different jump size specifications (either in affine jump-diffusion models or Lévy jump models) lead to different option pricing results for equity options. For instance, even with the same total volatility, affine jump-diffusion models with a finite number of double-exponential jumps tend to give larger out-of-the-money option prices (especially in short maturity) than those based on Gaussian jumps; see the discussion in Chen and Kou (2009). The resulting model from answering the second question also helps answer the first question.

To answer the two questions, we focus on the data sets of the S&P 500 and the NASDAQ-100 daily returns from January 1980 to October 2013, covering the recent financial crisis. We compare the empirical performance of an extensive set of nested models: models with normal and double-exponential jump sizes, models with Lévy-type jumps (the variance-gamma process), and models with or without jumps in stochastic volatility; see Section 2.2 for the list of models considered.

The existing literature only compares certain special cases of affine jump-diffusion models and Lévy jump models in shorter time periods and the results are mixed. For example, Eraker et al. (2003) find an affine jump-diffusion model with stochastic volatility and correlated Gaussian jumps in returns and volatility fits the S&P 500 data from 1980 to 1999 well, which suggests there are a finite number of medium jumps. However, Li et al. (2008) stress the importance of infinitely many small jumps by fitting a stochastic volatility model with jumps in returns modeled by the variance-gamma process, which is a special case of Lévy processes. Li et al. (2008) show that the variance-gamma jump model outperforms the affine

jump-diffusion model in Eraker et al. (2003) for the S&P 500 data from 1980 to 2000.

There could be an explanation for the disparities above. The conclusions in the two aforementioned papers are based on particular modeling assumptions about the jump-size distributions. In particular, the affine jump-diffusion models studied in both papers above have (conditionally) Gaussian distributed jumps in returns, and Li et al. (2008) also compare the affine jump-diffusion model of this sort to models with Lévy jumps. Neither of these two papers covers the period of the 2008 crisis. In this paper, we find that the performance of an affine jump-diffusion model with larger jumps, i.e., jumps with heavier tails such as the double-exponential distribution, is comparable to the variance-gamma process before the 2008 crisis and is better afterward.

More precisely, we find that a simple affine jump-diffusion model with both stochastic volatility and double-exponential jump sizes in returns fits both the S&P 500 and the NASDAQ-100 daily returns well. In fact, the model outperforms existing ones (in particular, models with the variance-gamma jumps, affine jump-diffusion models with normal jump sizes, and models with jumps in volatility) for the returns during the crisis, and it is comparable for the returns before the crisis (see Table 6 and Figures 3 and 4 for the S&P 500 results). The relevant results for the NASDAQ-100 are quite similar and hence omitted for simplicity unless otherwise stated. The omitted results are available upon request.

There are some intuitive explanations for our findings. (1) There is a drawback of using normal distribution with a negative mean to model jump sizes because such a distribution does not have a monotone decreasing density for negative jumps.[1] For example, if the jump mean is −3%, then with the normal distribution it is more likely to see a −2% negative jump than a −0.5% negative jump. This lack of monotonicity may lead to a poor fitting for small jumps, which is intuitively why Li et al. (2008) find that Lévy-type jumps fit the data better than the affine jump-diffusion models in Eraker et al. (2003) with normal jump sizes. However, with the double-exponential distribution for jump sizes, the density is monotonically decreasing for negative jumps, resulting in a potentially better fit for small jumps. (2) The heavy-tail feature of the double-exponential distribution also helps fit large jumps during the crisis period; for a more detailed discussion of both the small and the large jumps, see Section 2.1. (3) More complex structure of volatilities, such as jumps in volatility, may not necessarily result in better performance in models with double-exponential jumps. This is partly because, as volatility cannot rise forever, to compensate sudden jumps in volatility the mean-reverting speed in stochastic volatility tends to be higher for models with jumps in volatility. However, higher mean-reverting speed may hinder the model capability of generating enough volatility clustering effects to match those observed in the data during the 2008 financial crisis (see Figures 5 and 6 and the discussion in Section 2.1).

In summary, our answer to the second question is affirmative: the performance of affine jump-diffusion models can be significantly improved if the jump sizes are larger, i.e., jumps with heavier tails than those of the normal distribution. An affine jump-diffusion model with both stochastic volatility and the double-exponential jump-size distribution can fit equity return data well both before and after the crisis.

For the first question, based on the best-fitted model in the paper, we find, for both the S&P 500 and the NASDAQ-100 (see Table 7), that (a) during the crisis, negative jump rate has increased significantly whereas there is little change in the average negative jump sizes, and (b) jump rates can decrease even when there is no change in volatility. We should be very cautious about the empirical conclusions aforementioned regarding the changes in the jump rates and jump sizes, because there are only about five years' (2008–2013) worth of data after the financial crisis of 2008. However, (b) the conclusions above tentatively suggest the previous findings of Eraker (2004) and Johannes et al. (1999), that jump rate is positively correlated with volatility, do not necessarily hold after the crisis of 2008; see a detailed discussion in Section 5.

There is a vast amount of literature on jump risk of equity returns. Besides different objectives and answering different questions, the current paper is distinctive from the existing literature in terms of data, model specification, and econometric methodology. (1) For data selection, existing studies typically focus on the short episodes following the market turmoil from 1980s to early 2000s—say, the 1987 crash, the Long-Term Capital Management crisis, and the tech bubble burst. One exception is Bates (2012), who covers a period of 2007–2010; in that subsample, however, he does not explicitly report model fits or parameter estimates. Our paper uses the data of the prolonged period of 1980–2013, including the period after the outbreak of the recent financial crisis, for which we report model fits as well as the changes of parameters before and during the crisis. Moreover, existing work typically relies on options data or uses a combination of both equity returns and options. This paper uses only returns data. As pointed out by Bates, studies focusing only on returns data are "of interest in their own right" and are important in testing the compatibility of option prices with their underlying asset prices. (2) In terms of model specification, we estimate a series of new models with double-exponential jumps, which have never been studied before except in Bates,

who imposes other complex features in the model besides stochastic volatility and double-exponential jumps. (3) In terms of econometric methods, we use the Bayesian Markov chain Monte Carlo (MCMC) inference, which imputes latent variables of the models, such as jump times and jump sizes, and is beneficial to our analysis (for imputed jumps, see Figures 8 and 9). Bayesian MCMC inference is different from the maximum likelihood estimate or generalized method of moments approaches in Bakshi et al. (1997), Chernov et al. (2003), Pan (2002), and Ramezani and Zeng (2007), among others. Also, our study compliments the nonparametric test of Aït-Sahalia and Jacod (2011), who focus on high-frequency data.

The rest of the paper is organized as follows. In Section 2, we first provide intuition of why double-exponentially distributed jump sizes may fit the equity index returns better, and then we formally introduce the models that have been studied in the literature and propose a series of new models with double-exponential jumps. In Section 3, we outline the Bayesian MCMC inference procedure utilized in our study and discuss model diagnostics to evaluate model performance. Section 4 reports empirical results for the S&P 500 index returns. Changes of the jump sizes and jump rates are studied in Section 5. Section 6 concludes. The appendix provides all technical details.

2. The Models

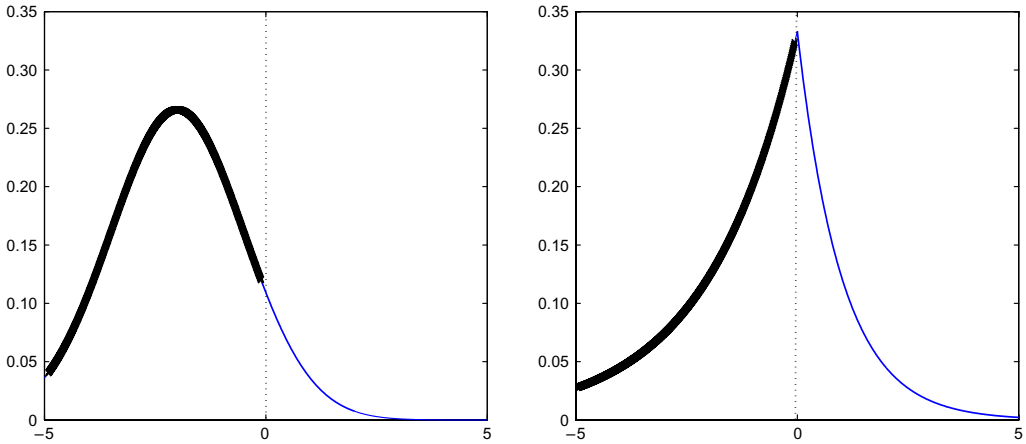
2.1. Intuition

Our empirical studies in Section 4 show that an affine jump-diffusion model with double-exponentially distributed jump sizes and stochastic volatility outperforms models with Gaussian distributed jumps in returns and models with jumps in volatility. Below are some intuitive explanations for this.

1. Empirically, it is found that jumps in asset returns tend to have a negative mean (see Eraker et al. 2003, Li et al. 2008). If a normal distribution is used to model the jump sizes, to incorporate the negative jump mean one has to introduce a location shift, leading to a nonmonotone density for negative jumps. For example, it can be seen from Figure 1 that if jump sizes are normally distributed with mean −2%, jumps down −2% are more likely to occur (i.e., with a higher density) than jumps down just −1%. This nonmonotonicity might be problematic in modeling small jumps. To the contrary, with the double-exponential distribution there is no need to introduce a location shift parameter to incorporate the negative mean; indeed, one can simply use different parameters for positive and negative jumps. The absence of the location shift parameter helps to preserve the monotonicity of the jump-size density; i.e., in the model, larger negative jumps are always less likely to occur than smaller negative jumps. Intuitively, the monotone structure of double-exponential jumps should help to fit small jumps. Li et al. (2008) found that Lévy-type jumps fit small jumps in the equity index returns better than the affine jump-diffusion model in Eraker et al. (2003) with normally distributed jump sizes. Consistent with this, we find that the model with double-exponential jumps has better performance than normally distributed jumps and has similar performance before the crisis as the model with Lévy-type jumps (see Section 4).

2. In terms of large jumps, the double-exponential distribution is also suitable because it has heavier tails than the normal distribution. A good intuition may be obtained by simply looking at the quantiles for both standardized Laplace (with a symmetric density of exponential-type tails $f(x) = (1/2)e^{-x}1(x>0) + (1/2)\cdot e^{x}1(x\le 0)$) and standardized Student's t distributions

Figure 1. (Color online) The Lack of Monotonicity for Normal Densities



Notes. The panels show that normal distribution with negative mean does not have monotone structure, especially for small negative values. The left panel is a normal density with mean −2 and standard deviation 1.5. The right panel is a double-exponential density $f(x) = (p/\eta^{+})e^{-x/\eta^{+}}1(x>0) + ((1-p)/\eta^{-})e^{x/\eta^{-}}1(x\le 0)$ with $p = 1/3$, $\eta^{+} = 1$, and $\eta^{-} = 2$. The negative parts of the densities are highlighted.

Table 1. The (Right) Quantiles for the Laplace and Normalized t Densities

Probability (%)	Laplace	$t7$	$t6$	$t5$	$t4$	$t3$
1	**2.77**	2.53	2.57	2.61	2.65	2.62
0.1	**4.39**	4.04	4.25	**4.57**	5.07	5.90
0.01	**6.02**	5.97	**6.55**	7.50	9.22	12.82
0.001	**7.65**	**8.54**	9.82	12.04	16.50	27.67

Note. Far at the right tail—say, at the probability of 0.01%—the quantile of the Laplace distribution is larger than the quantile of $t7$; this contrasts to the fact that the Student's t distribution should have asymptotically heavier tails than the exponential-type tails.

(with power-type tail) with the same mean and same variance. The right quantiles for the Laplace and standardized Student's t densities with degrees of freedom from 3 to 7 are given in Table 1 (taken from Heyde and Kou 2004), which show the heavy-tail feature of the exponential distribution.

In light of the above, Heyde and Kou (2004) show that it is difficult to distinguish exponential-type tails from power-type tails from empirical data unless one has extremely large sample size, perhaps in the order of tens of thousands or even hundreds of thousands. The heavy-tail feature of the double-exponential distribution helps fit the large jumps during the crisis, leading to a better performance than that of the model with Lévy-type jumps in Li et al. (2008).

3. In terms of the volatility clustering effect, more complicated models, such as those with jumps in volatility, may not necessarily result in better performance than models with double-exponential jumps and no jumps in volatility. This seems counterintuitive since volatility with jumps has the capability to increase rapidly. But there is also another consequence of adding jumps to volatility: to compensate the sudden jumps in volatility, the estimated value of the mean-reverting speed appears to be higher compared with the same parameter in models without jumps in volatility, as volatility cannot rise forever; see Table 5. The increase of mean-reverting speed hinders the ability of models with jumps in volatility to generate sufficient volatility clustering: we show in Section 4 that models with jumps in volatility have smaller sample autocorrelation functions (ACFs) of absolute and square returns. On the contrary, the ACF plots of absolute and squared returns in Figures 5 and 6 suggest that an affine-jump diffusion models with stochastic volatility, double-exponential jumps in returns, and no jumps in volatility can generate the level of volatility clustering closer to that observed in the data during the 2008 crisis.

However, we emphasize that the above discussion is far from a final claim that there are no jumps in volatility. There are three potential caveats. (1) In this paper, option data are not utilized, which is arguably more informative of the volatility structure. (2) Only some popular setups of jumps in volatility, e.g., the jumps in volatility having an exponential distribution and concurring with jumps in returns (see Eraker et al. 2003), are considered in our empirical study. Other settings may lead to different conclusions. (3) Our conclusion is made assuming a daily frequency in the data. It is possible that in the high-frequency domain the conclusion would become utterly different, but it is beyond the scope of the current study. Our point here is that adding jumps in volatility may not always help improve the performance of a model.

2.2. Model Formulation

A general model for asset returns y_t in continuous time is

$$\begin{cases} dy_t = \mu\, dt + \sqrt{v_{t-}}\, dW_t^y + dJ_{t+1}^y, \\ dv_t = \kappa(\theta - v_t)\, dt + \sigma_v \sqrt{v_{t-}}\, dW_t^v + dJ_{t+1}^v, \end{cases}$$

where the correlation between the Brownian motions (W_t^y, W_t^v) is ρ. In the subsequent study, we work on the discrete-time version of the above:

$$\begin{cases} y_{t+1} = y_t + \mu\Delta + \sqrt{v_t\Delta}\epsilon_{t+1}^y + J_{t+1}^y, \\ v_{t+1} = v_t + \kappa(\theta - v_t)\Delta + \sigma_v\sqrt{v_t\Delta}\epsilon_{t+1}^v + J_{t+1}^v, \end{cases} \tag{1}$$

where Δ is the time length, v_t is the stochastic volatility, and ϵ_{t+1}^y and ϵ_{t+1}^v are diffusion noises with normal distribution such that $\mathrm{corr}(\epsilon_{t+1}^y$ and $\epsilon_{t+1}^v) = \rho$. The jump part (J_{t+1}^y, J_{t+1}^v) is independent of the diffusion part $(\epsilon_{t+1}^y, \epsilon_{t+1}^v)$. Various specifications of J_{t+1}^y and J_{t+1}^v lead to different models, representing a finite or infinite number of small/medium/exponential (large) jumps.

Model 1: Finite Number of Moderate Jumps (Stochastic Volatility with Correlated Merton's Jumps)**.** In the stochastic volatility with correlated Merton's jumps (SV-MJ-JV) model, $J_{t+1}^y = \xi_{t+1}^y N_{t+1}$, $J_{t+1}^v = \xi_{t+1}^v N_{t+1}$, where[2] $P(N_{t+1} = 1) = 1 - P(N_{t+1} = 0) = \lambda\Delta$, $\xi_{t+1}^v \sim \exp(\mu_v)$, $\xi_{t+1}^y \mid \xi_{t+1}^v \sim \mathrm{N}(\mu_y + \rho_J \xi_{t+1}^v, \sigma_y^2)$. This model is studied in Eraker et al. (2003).[3] Jump-size distribution is medium in the sense that it is conditional normal, a probability distribution without heavy tails.

Model 2: Infinite Number of Small Jumps (Stochastic Volatility with Variance-Gamma Jumps)**.** In this model, stochastic volatility with variance-gamma jumps (SV-VG),

jumps in return follow a discrete variance gamma process:[4] $J^y_{t+1} = \gamma G_{t+1} + \sigma\sqrt{G_{t+1}}\epsilon^J_{t+1}$, where $\epsilon^J_{t+1} \sim \mathrm{N}(0,1)$ and $G_{t+1} \sim \Gamma(\Delta/\nu, \nu)$. Note that ϵ^J_{t+1} and G_{t+1} are independent of each other and independent of everything else; there are no jumps in volatility, i.e., $J^v_{t+1} = 0$. The continuous variance gamma process is an infinite-activity Lévy process and has an infinite number of small jumps in any given finite-time horizon. Li et al. (2008) find that SV-VG fits the S&P 500 returns from 1980 to 2000 better than SV-MJ-JV.

The two models above feature small or medium jump sizes in returns. In this paper, we also consider a series of models with large jump sizes (i.e., heavier tails) and monotone structure in returns. In particular, the jump sizes have a double-exponential distribution. We are interested to see whether we can fit equity returns better both before and during the financial crisis with this new model. To see whether a model benefits from more complex volatilities structure, we also incorporate jumps in volatilities similar as in the SV-MJ-JV model.

Model 3: Finite Number of Large and Monotonic Jumps in Returns (Stochastic Volatility with Double-Exponential Jumps). In the stochastic volatility with double-exponential jumps (SV-DEJ) model, jumps sizes J^y_{t+1} are double-exponentially distributed; i.e., $\xi^{y+}_{t+1} \sim \exp(\eta^+)$, $\xi^{y-}_{t+1} \sim \exp(\eta^-)$,

$$J^y_{t+1} = 1_{(N_{t+1}=1)}\xi^{y+}_{t+1} + 1_{(N_{t+1}=-1)}(-\xi^{y-}_{t+1}),$$
$$N_{t+1} = \begin{cases} 1 & \text{w.p. } \lambda^+\Delta, \\ 0 & \text{w.p. } 1-\lambda^+\Delta-\lambda^-\Delta, \\ -1 & \text{w.p. } \lambda^-\Delta. \end{cases}$$

There are no jumps in volatility; i.e., $J^v_{t+1} = 0$.

Model 4: Finite Number of Exponential Jumps in Returns and Jumps in Volatility (Stochastic Volatility with Double-Exponential Jumps and Jumps in Volatility). In this model, stochastic volatility with double-exponential jumps and jumps in volatility (SV-DEJ-JV), jumps sizes J^y_{t+1} are double-exponentially distributed and there are jumps in volatility; i.e.,

$$J^y_{t+1} = 1_{(N_{t+1}=1)}(\xi^{y+}_{t+1} + \rho_J\xi^v_{t+1}) + 1_{(N_{t+1}=-1)}(-\xi^{y-}_{t+1} + \rho_J\xi^v_{t+1}),$$
$$J^v_{t+1} = (1_{(N_{t+1}=1)} + 1_{(N_{t+1}=-1)})\xi^v_{t+1},$$
$$\xi^{y+}_{t+1} \sim \exp(\eta^+), \quad \xi^{y-}_{t+1} \sim \exp(\eta^-), \quad \xi^v_{t+1} \sim \exp(\mu_v),$$
$$N_{t+1} = \begin{cases} 1 & \text{w.p. } \lambda^+\Delta, \\ 0 & \text{w.p. } 1-\lambda^+\Delta-\lambda^-\Delta, \\ -1 & \text{w.p. } \lambda^-\Delta. \end{cases}$$

An important feature of Models 3 and 4, different from existing ones, is that the jump sizes have the double-exponential distribution. A simpler model without stochastic volatility is proposed by Kou (2002) with an emphasis on option pricing for path-dependent options. With stochastic volatility, pricing vanilla options is still possible with double-exponential jumps. For instance, under the conditions specified in Appendix B, the vanilla call option price for the SV-DEJ model with initial log price Y_0, initial volatility v_0, maturity T, and strike K is then given by (assuming constant interest rate r)

$$C(Y_0, v_0, T, K) = \frac{e^{-rT}}{\pi} \times \mathrm{Re}\left(\int_0^\infty e^{-ix\log K}\frac{\phi(x-i, Y_0, v_0, T)}{-x^2+ix}\,dx\right).$$

The nontrivial part $\phi(u, y, v, t)$ in the integration has an explicit expression and is also given in Appendix B. The infinite integral is an inverse Laplace transform and can be efficiently evaluated numerically; see Abate and Whitt (1995).

Note that SV-DEJ, SV-DEJ-JV, and SV-MJ-JV are special cases of affine jump-diffusion models, but SV-VG is not. Affine jump-diffusion models provide analytical tractability, and model parameters have natural economic interpretations. For example, in Model 3, λ^+ and λ^- are the positive and negative jump rates, respectively, and η^+ and η^- are the means of positive and negative jumps, respectively. To carry the comparison between double-exponential jumps (DEJ) and variance-gamma jumps (VG) farther, we also consider the following two models.

Model 5: Both Finite Number of Exponential Jumps and Infinite Number of Small Jumps in Returns and Jumps in Volatility (Stochastic Volatility with Double-Exponential Jumps, Variance-Gamma Jumps, and Jumps in Volatility). The stochastic volatility with double-exponential jumps, variance-gamma jumps, and jumps in volatility (SV-DEJ-VG-JV) model has both a finite number of large and monotonic jumps and an infinite number of small jumps in returns as well as jumps in volatility:

$$J^y_{t+1} = J^y_{DE,t+1} + J^y_{VG,t+1},$$
$$J^y_{DE,t+1} = 1_{(N_{t+1}=1)}(\xi^{y+}_{t+1} + \rho_J\xi^v_{t+1}) + 1_{(N_{t+1}=-1)}(-\xi^{y-}_{t+1} + \rho_J\xi^v_{t+1}),$$
$$J^v_{t+1} = (1_{(N_{t+1}=1)} + 1_{(N_{t+1}=-1)})\xi^v_{t+1},$$
$$\xi^{y+}_{t+1} \sim \exp(\eta^+), \quad \xi^{y-}_{t+1} \sim \exp(\eta^-); \quad \xi^v_{t+1} \sim \exp(\mu_v),$$
$$N_{t+1} = \begin{cases} 1 & \text{w.p. } \lambda^+\Delta, \\ 0 & \text{w.p. } 1-\lambda^+\Delta-\lambda^-\Delta, \\ -1 & \text{w.p. } \lambda^-\Delta, \end{cases}$$
$$J^y_{VG,t+1} = \gamma G_{t+1} + \sigma\sqrt{G_{t+1}}\epsilon^J_{t+1},$$
$$\epsilon^J_{t+1} \sim \mathrm{N}(0,1), G_{t+1} \sim \Gamma(\Delta/\nu, \nu).$$

Jumps in volatilities follow an exponential distribution, which is similar to the SV-MJ-JV model. The positive jump sizes of $J^y_{DE,t+1} = \xi^{y+}_{t+1} + \rho_J\xi^v_{t+1}$ are correlated with

the jump sizes in volatilities ξ^v_{t+1} via the parameter ρ_J and similar structure is in place for negative jumps. The arrival times of the large jumps in returns and volatilities are governed by the same Poisson process, which are the most popular setup for jumps in volatility models; see Eraker et al. (2003).

Model 6: Both Finite Number of Exponential Jumps and Infinite Number of Small Jumps in Returns, But No Jump in Volatility. This model is the same as the previous one, except there is no jump in volatility. Hence it is be denoted as SV-DEJ-VG.

It is clear that both the SV-DEJ and the SV-VG models are special cases of the SV-DEJ-VG model. The SV-DEJ is also a special case of the SV-DEJ-JV model. The biggest model, the SV-DEJ-VG-JV model, nests all other models with double-exponential (DE) and variance-gamma (VG) jumps.

3. Econometric Methodology

3.1. Bayesian MCMC Inference

To perform statistical inference for the models above, we use the Bayesian MCMC method; see a survey in Johannes and Polson (2010). In MCMC, statistical inference is done via samples from the posterior distribution, which are obtained by constructing a Markov chain $X(m)$, where m denotes the number of iterations, such that $X(m) \mid X(0) \to X(\infty) \sim F$, as m goes to infinity, where F is the posterior distribution and independent of $X(0)$. In practice, we choose a large number m and use $X(m), X(m+1), \ldots, X(m+k)$ as approximated samples from F. To ensure that the results are not sensitive to initial value $X(0)$, we run multiple (e.g., 100) Markov chains with different starting points $X(0)$.

Bayesian MCMC imputes latent variables of the model, such as jump times and jump sizes, which is beneficial to our analysis. For the same reason, Bayesian MCMC is applied in Eraker et al. (2003) and Li et al. (2008) to provide inference for the SV-MJ-JV and SV-VG models. Table 2 lists the conditional posteriors for the model parameters of the SV-DEJ and SV-VG models, of which priors are mostly conjugate. The priors for both the SV-MJ-JV and SV-VG models are the same as in Eraker et al. (2003) and Li et al. (2008). For the detail calculation of the posteriors, refer to Appendix A.

To check the validity of the MCMC method, we conduct a simulation study using the data of 5,000 observations (i.e., about 20 years' worth of daily data); details can be found in the caption of Table 3. The estimators for all the models using the priors in Table 2 are reported in Table 3, which suggests that the MCMC inference performs well.

3.2. Model Diagnostics

To check the goodness of fit of the models, we use the following diagnostic methods:

Table 2. Prior Distribution, Prior Means, and Standard Deviations for the SV-DEJ and SV-VG Models

Parameter	Prior distribution	Prior mean	Prior standard deviation
η^+	(conjugate) Inverse gamma	1.0	1.0
η^-	(conjugate) Inverse gamma	1.0	1.0
(λ^+, λ^-)	(conjugate) Dirichlet	(0.0455, 0.0455)	(0.0311, 0.0311)
μ	(conjugate) Normal	0	1
θ	(conjugate) Truncated normal	0.7979	0.6028
κ	(conjugate) Truncated normal	0.7979	0.6028
(σ_v, ρ)	$\phi_v \mid w_v$ Normal, w_v Inverse gamma, where $\sigma_v = \sqrt{\phi_v^2 + w_v}$, $\rho = \phi_v / \sqrt{\phi_v^2 + w_v}$	0, 100	$\sqrt{w_v/2}$, ∞
γ	(conjugate) Normal	0	1
σ	(conjugate) Inverse gamma	0.1333	0.1086
ν	Inverse gamma	1.1111	0.3704

Notes. The parameters η^+, η^-, λ^+, and λ^- are unique to the SV-DEJ model. For the details of the conditional posterior distributions (such as the values for hyperparameters) as well as rigorous proofs, see Appendix A. The parameters μ, θ, κ, σ_v, and ρ are the parameters of the SV-DEJ model that are shared with the SV-MJ-JV and the SV-VG models. Proofs of these conditional posteriors are provided in Li et al. (2008) and Eraker et al. (2003); hence, in the appendix, the details of these conditional posteriors are listed without proofs. The parameters γ, σ, and ν are parameters special to the SV-VG model; the details of these conditional posteriors can also be found in Li et al. (2008).

1. The Kolmogorov–Smirnov (KS) test for normality of normalized residuals:

$$\epsilon^y_{t+1} = \frac{y_{t+1} - y_t - \mu\Delta - J^y_{t+1}}{\sqrt{v_t \Delta}}.$$

This is the same method used in Li et al. (2008). If a model in question has good fit to the data, then normalized residuals ϵ^y_{t+1} from that model should follow a standard normal distribution. In our simulation study, for each of the models, one KS test is performed on the residuals $\tilde{\epsilon}^y_{t+1} = (y_{t+1} - y_t - \mu^{(m)}\Delta - (J^y_{t+1})^{(m)}) \cdot (v_t^{(m)}\Delta)^{-1/2}$, $t = 1, \ldots, T$, where $\mu^{(m)}$, $(J^y_{t+1})^{(m)}$, and $v_t^{(m)}$ are obtained from the last iteration of each of the $m = 1, \ldots, 100$ simulation runs. A lower percentage of rejection by the KS test and a higher average p-value thus suggest a better goodness of fit of the model. The percentage of rejection and average p-values of the KS test for our simulation are also reported in Table 3.

2. We examine Q-Q plots of actual data versus model-simulated returns, where Q-Q plots close to a straight line suggest a good model fit. More specifically, we compare the quantiles of $Y_{0:T}$ simulated using the true parameter values versus quantiles of model-simulated returns using the point estimators in Table 3. The Q-Q plots of the simulation study are reported in Figure 2.

3. The time-series implication of heavy-tailed distributions as a result of stochastic volatility versus jumps

Table 3. Simulation Study of Bayesian MCMC

Panel A: SV-MJ-JV

	μ	κ	θ	σ_v	ρ	μ_y	σ_y	λ	μ_v	ρ_J
True	0.05	0.015	0.81	0.1	−0.4	−3.0	3.5	0.015	1.0	−0.4
Posterior mean	0.0503	0.0162	0.9677	0.1071	−0.3647	−3.0968	3.0923	0.0146	1.0527	−0.4274
Posterior S.D.	0.0065	0.0027	0.0528	0.0032	0.0126	0.9038	0.2940	0.0038	0.1556	0.5398

Percentage of rejection of the KS test: 9%, average p-value: 0.4587

Panel B: SV-VG

	μ	κ	θ	σ_v	ρ	γ	σ	ν
True	0.05	0.015	0.81	0.1	−0.4	−0.01	0.4	3.0
Posterior mean	0.0551	0.0169	0.7949	0.1012	−0.3920	−0.0103	0.4075	2.8996
Posterior S.D.	0.0181	0.0025	0.0346	0.0068	0.0631	0.0098	0.0231	0.1546

Percentage of rejection of the KS test: 19%, average p-value: 0.2239

Panel C: SV-DEJ

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-
True	0.05	0.015	0.81	0.1	−0.4	4.0	4.6	0.004	0.016
Posterior mean	0.0512	0.0158	0.8211	0.0964	−0.3819	4.1254	4.4898	0.0042	0.0158
Posterior S.D.	0.0084	0.0025	0.0127	0.0057	0.0214	0.1602	0.1683	0.0056	0.0006

Percentage of rejection of the KS test: 11%, average p-value: 0.3755

Panel D: SV-DEJ-JV

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	μ_v	ρ_J
True	0.05	0.015	0.81	0.1	−0.4	4.0	4.6	0.03	0.06	0.5	−0.4
Posterior mean	0.0525	0.0177	0.9160	0.0905	−0.3879	3.5923	4.4763	0.0219	0.0496	0.6874	−0.3478
Posterior S.D.	0.0290	0.0153	0.0972	0.0045	0.1442	0.2082	0.4976	0.0044	0.0085	0.3294	0.4698

Percentage of rejection of the KS test: 7%, average p-value: 0.4623

Panel E: SV-DEJ-VG

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-
True	0.05	0.015	0.81	0.1	−0.4	4.0	4.6	0.03	0.06
Posterior mean	0.0534	0.0157	0.8910	0.0909	−0.3782	3.3122	5.1223	0.0339	0.0554
Posterior S.D.	0.0891	0.0212	0.0882	0.0056	0.1262	0.6111	0.4976	0.0064	0.0077

	γ	σ	ν
True	−0.01	0.4	3.0
Posterior mean	−0.0110	0.4014	2.9208
Posterior S.D.	0.0034	0.0181	0.1323

Percentage of rejection of the KS test: 8%, average p-value: 0.2611

Panel F: SV-DEJ-VG-JV

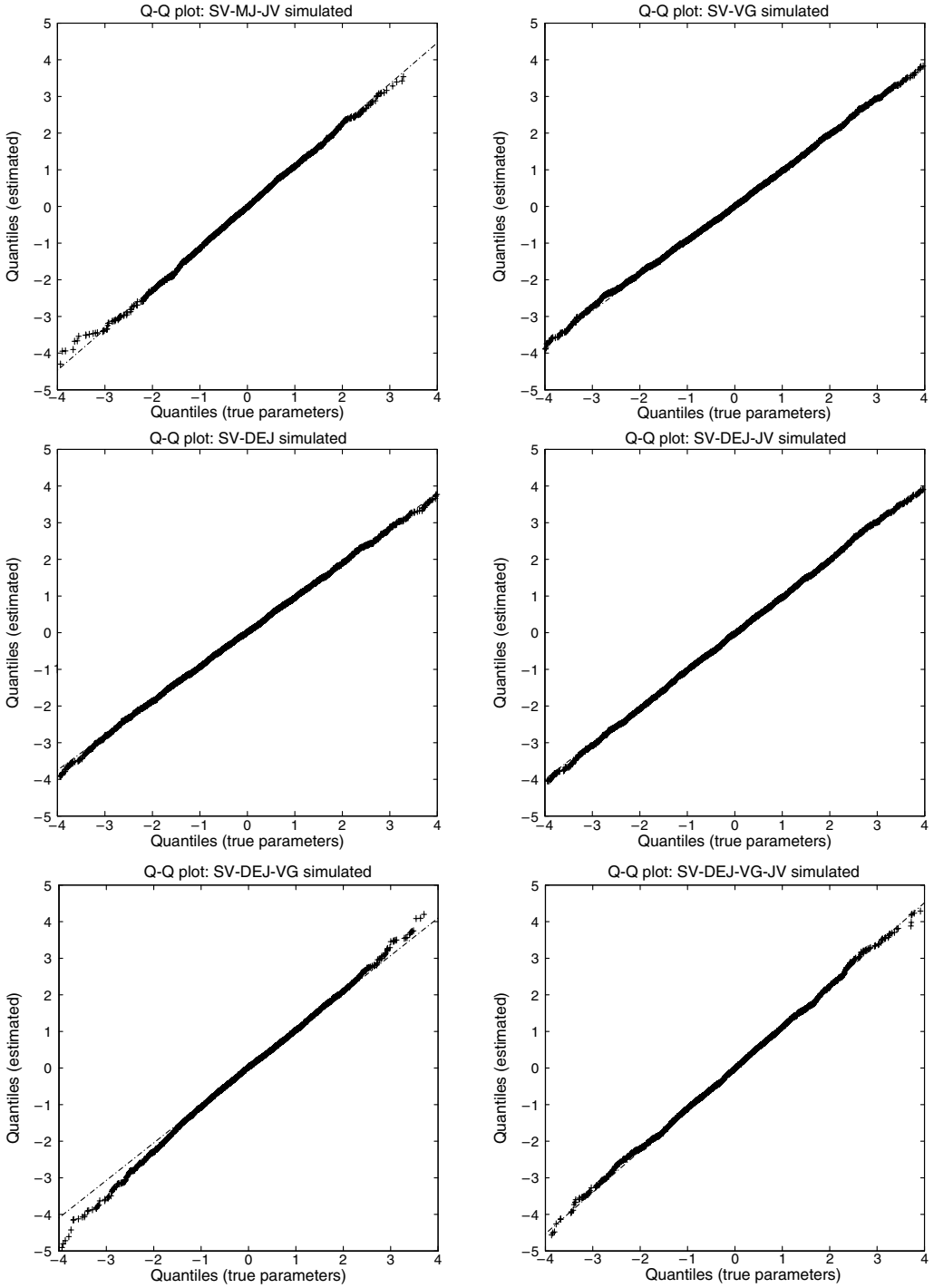
	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-
True	0.05	0.015	0.81	0.1	−0.4	4.0	4.6	0.03	0.06
Posterior mean	0.0611	0.0144	0.7798	0.0915	−0.3903	4.1110	4.9977	0.0311	0.0578
Posterior S.D.	0.0987	0.0198	0.0783	0.0088	0.0925	0.9817	0.5616	0.0084	0.0072

	γ	σ	ν	μ_v	ρ_J
True	−0.01	0.4	3.0	1.0	−0.4
Posterior mean	−0.0090	0.4125	3.1138	0.8897	−0.4
Posterior S.D.	0.0064	0.0192	0.1891	0.4392	0.5587

Percentage of rejection of the KS test: 10%, average p-value: 0.2427

Notes. In the simulation study, we use one simulated data set of 5,000 observations (i.e., about 20 years' worth of daily data). On the same data set with the same priors, we run 100 Markov chains (with different starting points) each with 50,000 iterations. For each Markov chain, the first 49,900 were discarded as burn-in. The resulting $100 \times 100 = 10{,}000$ samples of parameters in total from the 100 chains are considered 10,000 random draws from the posterior distribution, from which we obtain the estimates of the posterior mean and standard deviation. We also obtain 100 sets of model residuals from the last iterations of the 100 Markov chains, on which we perform the KS test and report its average p-values (see Section 3.2).

Figure 2. Q-Q Plots of Simulated Returns



is different, especially in terms of the joint distributions of returns. Therefore, a diagnostic for the joint distribution of consecutive returns is needed; we thank a referee for suggesting this. We investigate which model is capable of generating the pattern of ACFs observed in the data. More specifically, it is a well-known stylized fact that sample ACFs of stock returns are insignificant, whereas the ACFs of absolute stock returns and squared stock returns are significantly positive. A good model of equity returns should seek to generate the level of ACFs of absolute and squared returns consummate to what is observed in the data.

To implement this diagnostic, we need to compute and compare the 95% confidence interval (CI) of model-implied ACFs of absolute (and squared) returns with the sample ACF of absolute (and squared) returns of the S&P 500. Let the length of the S&P 500 time series in question be T. Then given any model

with its estimated parameters, we simulate N independent time series of absolute returns with length T: $|r_1^{(n)}|, |r_2^{(n)}|, \ldots, |r_T^{(n)}|$, $n = 1, \ldots, N$. The sample ACF of lag h for the nth time series is then

$$a_h^{(n)} := \frac{1}{(T-h)\sigma_n^2} \sum_{t=1}^{T-h} (|r_{t+h}^{(n)}| - \mu_n)(|r_t^{(n)}| - \mu_n),$$

where $\mu_n = (1/T)\sum_{t=1}^{T} |r_t^{(n)}|$ and $\sigma_n^2 = (1/T)\sum_{t=1}^{T}(|r_t^{(n)}| - \mu_n)^2$. Given N sufficiently large, the lower and upper bounds of the theoretical 95% CI model-implied ACFs at lag h can be approximated by the 2.5% and 97.5% quantiles of $\{a_h^{(1)}, a_h^{(2)}, \ldots, a_h^{(N)}\}$, respectively. The above is repeated for the squared returns of the S&P 500.

4. Empirical Results

We study the daily returns of the S&P 500 from January 1980 to October 2013. To better relate to Li et al. (2008) and Eraker et al. (2003), for the S&P 500 we consider three separate periods: (i) January 1980–December 2000 (which coincides exactly with the data set considered in Li et al. (2008) and roughly with Eraker et al. (2003), who consider 1980–1999), (ii) January 2001–July 2007 (before the financial crisis), and (iii) August 2007–October 2013 (after the outbreak of the financial crisis).[5] Although there appears to be no consensus in previous literature concerning when the crisis started, we choose August 2007 in this paper, since both the CBOE Volatility Index and the spread between Eurodollar forward rate (the London Interbank Offered Rate) and the U.S. Treasury bill peaked in July and August of 2007. On the other hand, June 2009, deemed as the end of the recession by the National Bureau of Economic Research, is chosen to be the end of the crisis. To check the robustness of our findings, we also use a second data set of daily returns of the NASDAQ-100 from January 2001 to October 2013. To match the study of the S&P 500, we consider two separate periods of the NASDAQ-100: January 2001–July 2007 and August 2007–October 2013. Table 4 provides summary statistics for the daily (log) returns.[6]

It is found that the empirical results for the NASDAQ-100 are quite similar to those from the S&P 500. Hence in what follows, the results for the NASDAQ-100 are omitted unless otherwise mentioned. The point estimates of the model parameters—in particular, the positive and negative jump rates and the average positive and negative jump sizes before and during the crisis—are reported in Table 5. For the period 1980–2000, the results resemble those in Li et al. (2008) closely. In particular, we find a significant leverage effect in all three models (ρ between −0.4 and −0.9), consistent with previous studies (e.g., Duffee 1995).

The results of the KS test are summarized in Table 6. The Q-Q plots for comparing the sample paths with actual returns are given in Figures 3 and 4. Figures 5 and 6 show the sample ACFs of the S&P 500 against the corresponding confidence intervals constructed using each of the six models we study in the period of August 2007–October 2013. In the precrisis period, the resulting histograms and Q-Q plots using periods January 1980–December 2000, January 2001–July 2007, and January 1980–July 2007 seem similar. Hence only the results for January 2001–July 2007 are reported.

To carry the comparison between VG jumps and DE jumps even further, we investigate in the SV-DEJ-VG model of whether the VG jumps may "drive out" the DE jumps in returns. To that end, we examine the jump rates of the DE jumps when VG jumps are present—that is, the posterior distributions of λ^+ and λ^- in the SV-DEJ-VG model. Suppose that, indeed, DE jumps are not needed when VG jumps are present in the returns, Bayesian asymptotic theory implies that the posterior distribution of λ^+ and λ^- in the SV-DEJ-VG model should concentrate in a neighborhood of 0, which should then be the mode of the posterior distribution. But as can be seen in Figure 7, this does not seem to be the case. Furthermore, the posterior means and standard deviations of the jump-related parameters in the SV-DEJ-VG model (see Table 5 for the S&P 500) show that in the three periods considered, at least one of λ^+ and λ^- values is significantly different from 0. The results for the NASDAQ-100 are similar and omitted. In summary, the conclusion seems to be that the VG jumps cannot replace DE jumps in fitting the data.

Some cautionary remarks are in order for comparing the nested models. In the above, SV-DEJ is a submodel to both SV-DEJ-VG and SV-DEJ-JV. It is then natural

Table 4. Summary Statistics of the Data Sets

Data period	Mean	Volatility	Skewness	Kurtosis	Minimum	Maximum
			Panel A: S&P 500			
Jan 1980–Dec 2000	0.0476	1.0435	−2.3584	55.6080	−22.8997	8.7089
Jan 2001–Jul 2007	0.0076	1.0531	0.1159	5.9645	−5.0468	5.5744
Aug 2007–Oct 2013	0.0120	1.5392	−0.2778	10.8497	−9.4695	10.9572
			Panel B: NASDAQ-100			
Jan 2001–Jul 2007	−0.0206	1.9063	0.0976	6.2757	−8.6139	10.2727
Aug 2007–Oct 2013	0.0346	1.6014	−0.1282	9.7984	−11.1149	11.8493

Table 5. S&P 500: Model Parameters Estimates

Panel A: Jan 1980–Dec 2000

SV-MJ-JV model

	μ	κ	θ	σ_v	ρ	μ_y	σ_y	λ	μ_v	ρ_J
Posterior mean	0.0503	0.0277	0.6697	0.1032	−0.4830	−1.4946	2.1938	0.0067	1.0192	−0.8103
Posterior S.D.	0.0053	0.0030	0.0302	0.0029	0.0158	0.7107	0.2711	0.0034	0.1215	0.7930

SV-VG model

	μ	κ	θ	σ_v	ρ	γ	σ	ν
Posterior mean	0.0755	0.0162	0.8953	0.1051	−0.5687	−0.0468	0.3307	5.9248
Posterior S.D.	0.0150	0.0024	0.0437	0.0095	0.0635	0.0096	0.0203	0.1826

SV-DEJ model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-
Posterior mean	0.0479	0.0139	0.9144	0.1076	−0.4527	0.2690	2.5991	0.0257	0.0090
Posterior S.D.	0.0071	0.0010	0.0123	0.0032	0.0210	0.1597	0.1719	0.0060	0.0007

SV-DEJ-JV model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	μ_v	ρ_J
Posterior mean	0.0501	0.0594	0.8669	0.2048	−0.3824	0.2844	3.5648	0.0080	0.0040	0.9669	−0.5222
Posterior S.D.	0.0105	0.0062	0.0403	0.0077	0.0276	0.0646	0.3552	0.0030	0.0012	0.1871	0.5894

SV-DEJ-VG model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	γ	σ	ν
Posterior mean	0.0261	0.0163	0.8201	0.1151	−0.2738	0.2697	3.0399	0.0173	0.0066	0.0320	0.2310	5.5629
Posterior S.D.	0.0144	0.0027	0.0863	0.0026	0.0164	0.1099	0.0078	0.0022	0.0128	0.0322	0.1079	0.0497

SV-DEJ-VG-JV model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	γ	σ	ν	μ_v	ρ_J
Posterior mean	−0.0018	0.0554	0.8199	0.1109	−0.4472	0.2930	1.9471	0.0068	0.0039	0.0569	0.2322	5.4878	0.3033	0.1302
Posterior S.D.	0.0139	0.0025	0.0956	0.0018	0.0162	0.6867	0.0086	0.0024	0.0128	0.0163	0.0974	0.0707	0.2237	0.5589

Panel B: Jan 2001-Jul 2007

SV-MJ-JV model

	μ	κ	θ	σ_v	ρ	μ_y	σ_y	λ	μ_v	ρ_J
Posterior mean	0.0042	0.0193	1.0444	0.1404	−0.7553	−1.3753	2.0348	0.0037	1.0000	0.0918
Posterior S.D.	0.0023	0.0033	0.0462	0.0073	0.0269	2.9737	0.0817	0.0012	0.9528	0.7758

SV-VG model

	μ	κ	θ	σ_v	ρ	γ	σ	ν
Posterior mean	0.0331	0.0141	1.4864	0.1780	−0.8617	−0.0389	0.1875	4.8425
Posterior S.D.	0.0118	0.0038	0.0623	0.0138	0.0294	0.0085	0.0127	0.0866

SV-DEJ model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-
Posterior mean	0.0269	0.0176	0.9734	0.1336	−0.7208	0.1029	0.1338	0.0576	0.0776
Posterior S.D.	0.0021	0.0010	0.0261	0.0045	0.0134	0.0496	0.0170	0.0285	0.0226

Table 5. (Continued)

SV-DEJ-JV model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	μ_v	ρ_J
Posterior mean	0.0141	0.0526	1.0902	0.2583	−0.5123	0.1584	0.4799	0.0026	0.0050	1.1636	0.0528
Posterior S.D.	0.0206	0.0113	0.1272	0.0364	0.0761	0.0361	0.0515	0.0009	0.0022	0.6297	0.7692

SV-DEJ-VG model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	γ	σ	ν
Posterior mean	0.0423	0.0117	1.1943	0.1379	−0.7036	0.1150	0.5502	0.0276	0.0437	−0.0080	0.1603	5.7951
Posterior S.D.	0.0248	0.0026	0.2594	0.0035	0.0200	0.0443	0.2261	0.0114	0.0204	0.0148	0.0186	0.1837

SV-DEJ-VG-JV model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	γ	σ	ν	μ_v	ρ_J
Posterior mean	0.0276	0.0149	1.2646	0.1712	−0.8317	0.1363	0.5852	0.0177	0.0529	0.0017	0.1526	5.7133	0.4013	−0.0784
Posterior S.D.	0.0196	0.0026	0.2411	0.0048	0.0145	0.0246	0.0885	0.0202	0.0206	0.0107	0.0193	0.1674	0.3382	0.5360

Panel C: Aug 2007–Oct 2013

SV-MJ-JV model

	μ	κ	θ	σ_v	ρ	μ_y	σ_y	λ	μ_v	ρ_J
Posterior mean	0.0448	0.0215	1.3964	0.2114	−0.7785	−1.7506	1.8322	0.0084	1.4726	0.2924
Posterior S.D.	0.0203	0.0028	0.2200	0.0127	0.0247	0.4622	0.4445	0.0034	0.2694	0.7287

SV-VG model

	μ	κ	θ	σ_v	ρ	γ	σ	ν
Posterior mean	0.0425	0.0231	2.0891	0.3014	−0.7733	−0.0095	0.1869	5.2794
Posterior S.D.	0.0249	0.0033	0.2879	0.0059	0.0122	0.0208	0.0292	0.1191

SV-DEJ model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-
Posterior mean	0.0442	0.0182	1.8904	0.2612	−0.7760	0.1500	0.1725	0.0394	0.0646
Posterior S.D.	0.0205	0.0038	0.2713	0.0166	0.0353	0.0343	0.0567	0.0235	0.0271

SV-DEJ-JV model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	μ_v	ρ_J
Posterior mean	0.0614	0.0398	1.8091	0.3313	−0.5605	0.1408	0.5374	0.0037	0.0064	0.9436	0.1844
Posterior S.D.	0.0226	0.0060	0.2331	0.0135	0.0547	0.0407	0.1982	0.0016	0.0029	0.2247	0.7266

SV-DEJ-VG model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	γ	σ	ν
Posterior mean	0.0410	0.0213	2.1958	0.2960	−0.7320	0.6153	0.1309	0.0114	0.0615	0.0095	0.1659	6.1008
Posterior S.D.	0.0255	0.0034	0.3614	0.0055	0.0142	0.2023	0.0518	0.0059	0.0284	0.0136	0.0217	0.1818

SV-DEJ-VG-JV model

	μ	κ	θ	σ_v	ρ	η^+	η^-	λ^+	λ^-	γ	σ	ν	μ_v	ρ_J
Posterior mean	0.0457	0.0258	2.3101	0.3328	−0.7690	0.1830	0.2328	0.0141	0.0167	−0.0067	0.1609	5.9743	0.2721	−0.2931
Posterior S.D.	0.0234	0.0039	0.3259	0.0073	0.0120	0.0658	0.1374	0.0175	0.0213	0.0118	0.0231	0.1889	0.1014	0.4612

Notes. We use 100 Markov chains with different starting points. For each of the models, we estimate the posterior mean and posterior standard deviation using the last 100 random draws from each of the Markov chains (each Markov chain has 50,000 iterations in total, though we discard the first 49,900 as burn-in).

Table 6. S&P 500: Kolmogorov–Smirnov Tests

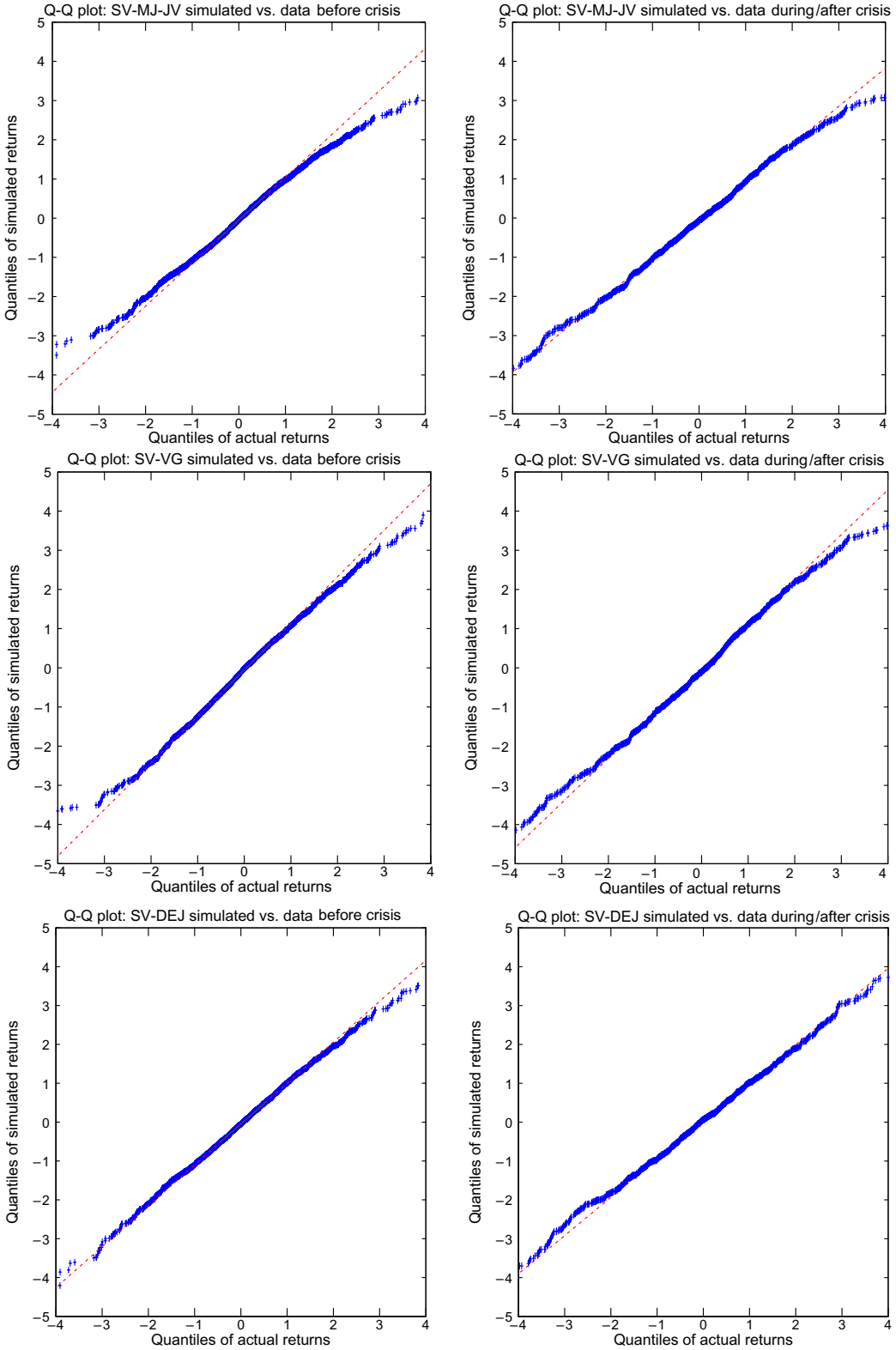
Model	Percentage of rejection by KS	Average p-value of KS
	Panel A: Jan 1980–Dec 2000	
SV-MJ-JV	81	0.0034
SV-VG	9	0.4212
SV-DEJ	15	0.4011
SV-DEJ-JV	96	0.0002
SV-DEJ-VG	10	0.4196
SV-DEJ-VG-JV	11	0.3883
	Panel B: Jan 2001–Jul 2007	
SV-MJ-JV	78	0.0035
SV-VG	10	0.4134
SV-DEJ	19	0.3715
SV-DEJ-JV	91	0.0005
SV-DEJ-VG	9	0.4196
SV-DEJ-VG-JV	13	0.3267
	Panel C: Aug 2007–Oct 2013	
SV-MJ-JV	62	0.0071
SV-VG	59	0.0082
SV-DEJ	32	0.1425
SV-DEJ-JV	69	0.0041
SV-DEJ-VG	78	0.0031
SV-DEJ-VG-JV	80	0.0023

Notes. The Kolmogorov–Smirnov test seems to favor the SV-DEJ model during and after the crisis. The SV-DEJ also performs reasonably well before the crisis. More specifically, the table reports the percentage of rejection of the KS tests out of the 100 Markov chains and the average p-value of the KS tests for each of the models for the S&P 500 returns in each of the three periods considered. It can be seen that the SV-DEJ model has with largest average p-value during and after the crisis.

to conceive that the parent models should perform at least as well as the submodels when model complexity is not penalized. However, it seems from Table 6 that the SV-DEJ-VG model outperforms the SV-DEJ model only in the sample of the S&P 500 before the crisis, i.e., only in periods where the SV-VG model outperforms the SV-DEJ model. Why does the SV-DEJ perform better than its parent models when the parent models have more parameters? This is partly because of the certain constraints imposed on the more complex parent models. (This is similar to the negative adjusted R^2 that one may observe in a linear regression model with constraints on parameters.) For example, the SV-DEJ-JV model may not outperform the SV-DEJ model, since by construction, jumps in volatility occur at the same time as jumps in returns, a constraint that impedes the double-exponential distribution in picking up small jumps. The failure to capture the small jumps, in the same spirit of Li et al. (2008), seems to explain the poor performance of the SV-DEJ-JV and the SV-DEJ-VG-JV models. Similarly, even if the SV-DEJ-VG model reduces to the SV-DEJ model when $\gamma = \sigma = \nu = 0$, the prior distributions for γ and σ, as taken from Li et al. (2008), though not extremely informative, are somehow strong enough to ensure that VG jumps exist in the SV-DEJ-VG model; see the prior specifications in Table 2. When small jumps are needed for a good model fit, the VG jumps component enhances model performance, consistent with the main message in Li et al. (2008). To the contrary, during the crisis period, it appears that large jumps dominate small jumps. Thus the VG component becomes less beneficial to the SV-DEJ-VG model. Of course, under these circumstances if one chooses weaker prior distributions of γ and σ, it is expected that the VG component disappears in posterior and the model SV-DEJ-VG essentially becomes the submodel SV-DEJ.

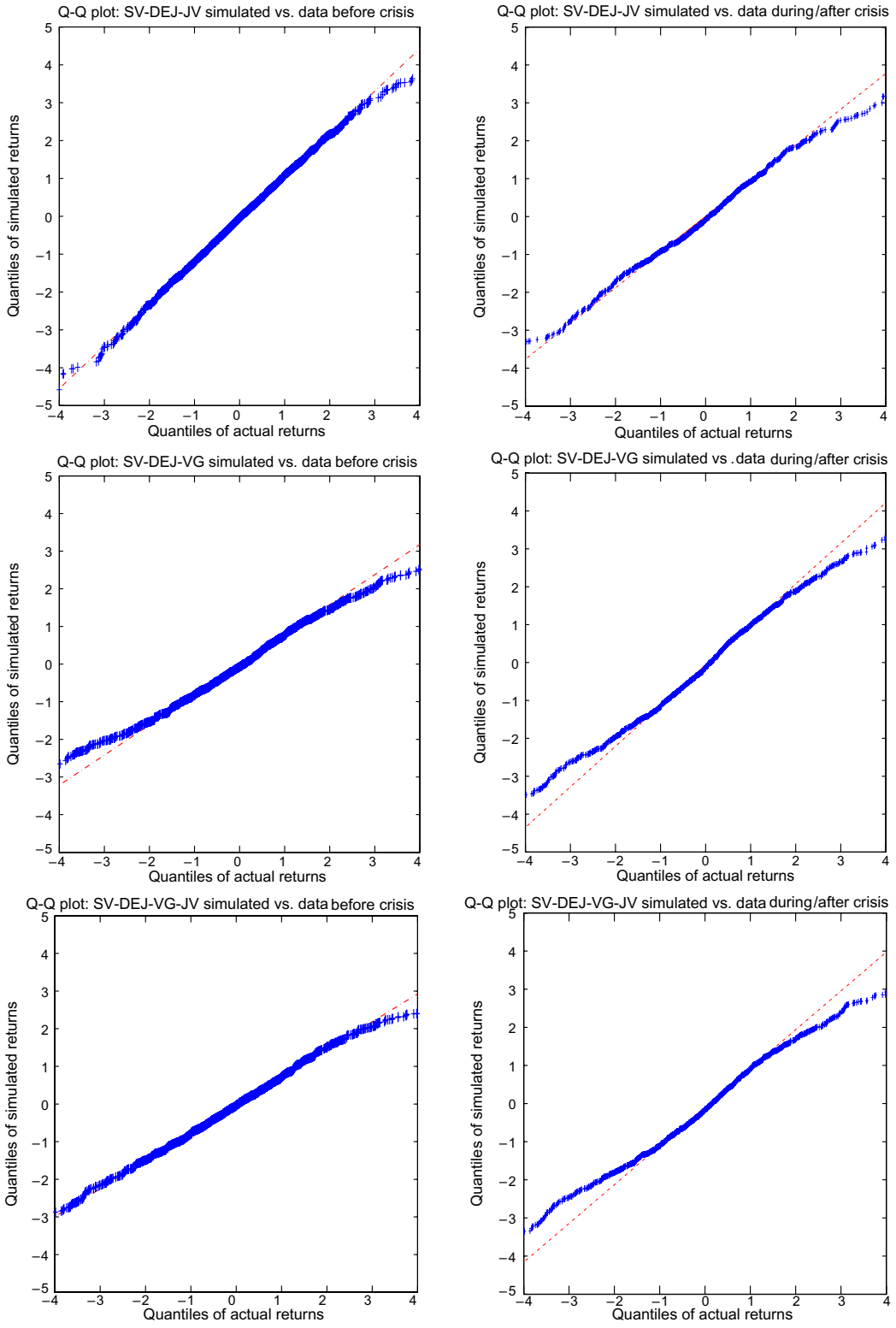
We summarize the empirical findings as follows. (i) Judging from the results of the KS test, the SV-DEJ model outperforms all other models during the financial crisis. (ii) The Q-Q plots seem to confirm that the SV-DEJ model provides the best tail fit. In particular, better fit in the tails by the SV-DEJ model becomes apparent for the postcrisis data, suggesting that the better performance of the SV-DEJ model comes at least partly from capturing large jumps. (iii) During the normal time period and before the crisis, from 1980 to 2007, the SV-DEJ model and the SV-VG model seem to perform equally well, by and large, with SV-VG slightly better, which is consistent with the results in Bates (2012) for the Center for Research in Security Prices (CRSP) value-weighted returns up to 2006. Hence it appears that the monotonic structure of double-exponential jumps can also provide reasonably good fit to small jumps in the data, as prescribed by Lévy models. (iv) It seems that the SV-DEJ model is

Figure 3. (Color online) Tail Fitness of Models to the S&P 500 (Part I): Q-Q Plots Before and During/After the Crisis—The SV-MJ-JV, SV-VG, and SV-DEJ Models



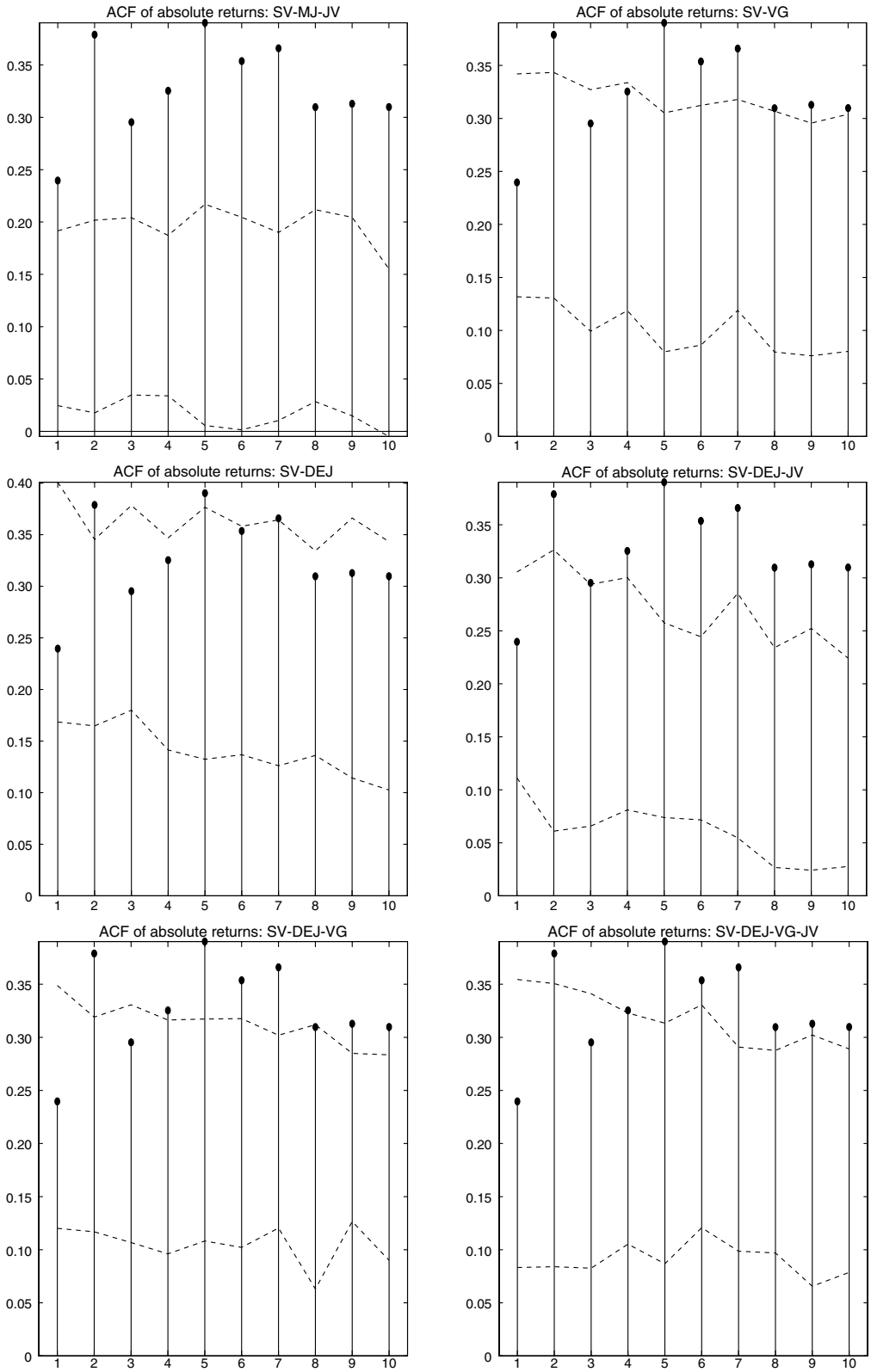
Notes. Presented are Q-Q plots between a sample path simulated from the corresponding model using point estimates in Table 5 and actual data (returns of the S&P 500 before the crisis (January 2001–July 2007) and during/after the crisis (August 2007–October 2013); results for January 1980–July 2007 are found to be similar to those for January 2001–July 2007). In all of the graphs, quantiles are compared in the range of $-4 \le Z \le 4$, where Z represents the actual returns. It appears that the SV-DEJ model provides a good fit in both tails, whereas other models may not.

Figure 4. (Color online) Tail Fitness of Models to the S&P 500 (Part II): Q-Q Plots Before and During/After the Crisis—The SV-DEJ-JV, SV-DEJ-VG, and SV-DEJ-VG-JV Models



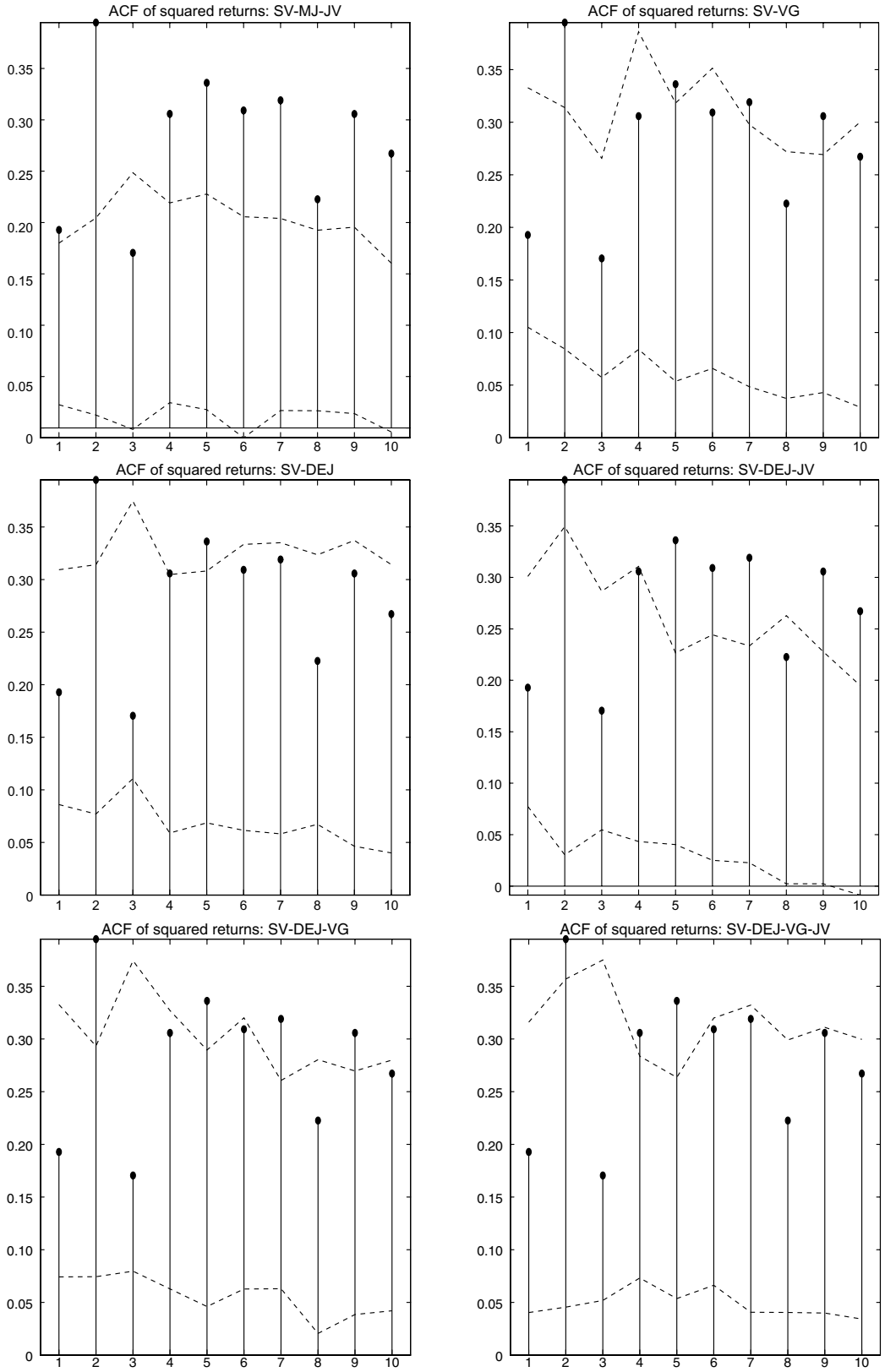
Notes. Presented are Q-Q plots between a sample path simulated from the corresponding model using point estimates in Table 5 and actual data (returns of the S&P 500 before the crisis (January 2001–July 2007) and during/after the crisis (August 2007–October 2013); results for January 1980–July 2007 are found to be similar to those for January 2001–July 2007). In all of the graphs, quantiles are compared in the range of $-4 \le Z \le 4$, where Z represents the actual returns. It appears that the SV-DEJ model provides a good fit in both tails, whereas other models may not.

Figure 5. S&P 500: Autocorrelation Functions of Absolute Values of Simulated Returns vs. Actual Returns During and After the Crisis



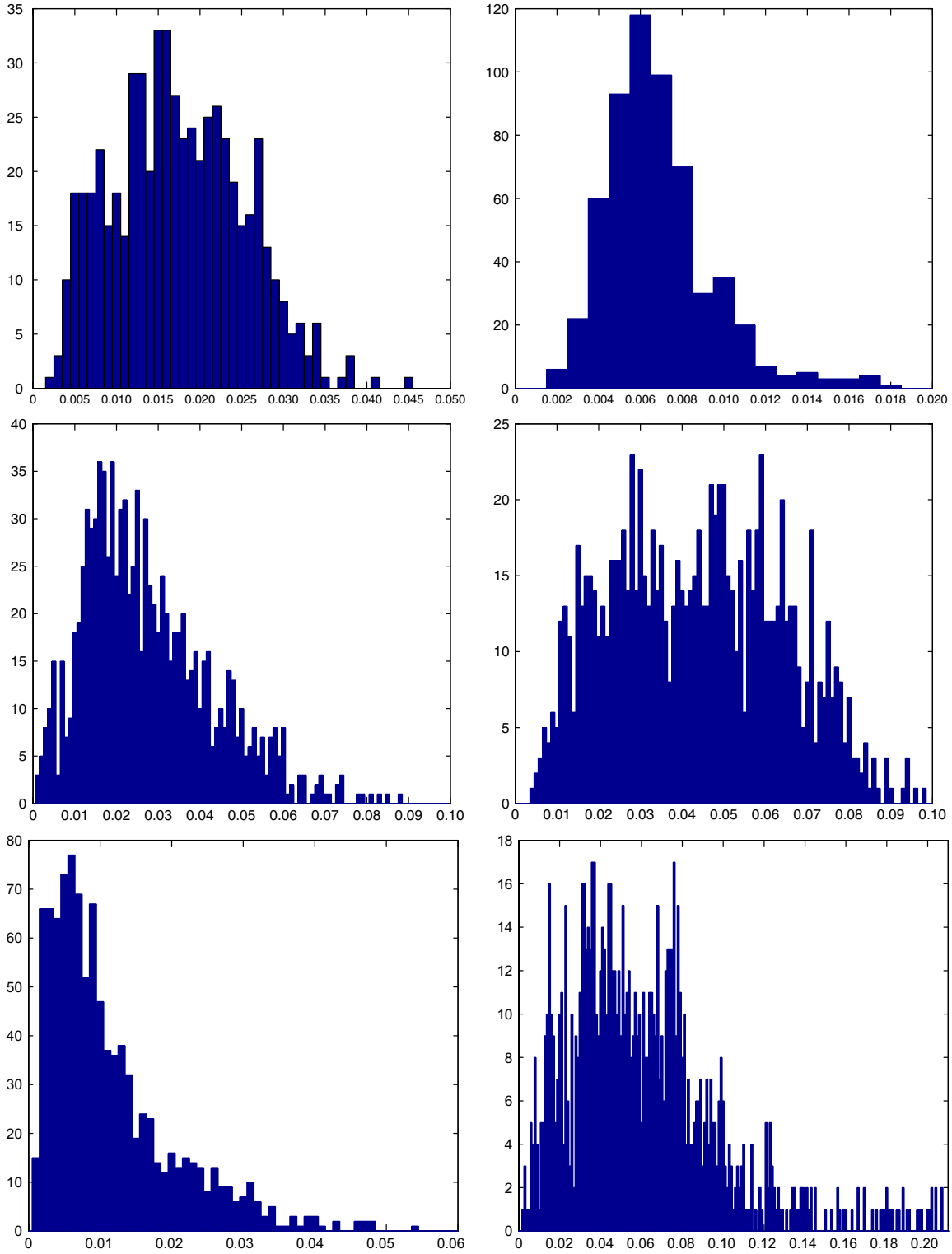
Notes. Shown are the sample ACFs of *absolute actual returns* (in dots) and 95% confident interval bands of sample ACFs of *absolute* model-simulated returns. The data set used is the S&P 500 daily log returns from August 2007 to October 2013. The SV-DEJ model seems to be the best, as the 95% CI from the model covers more sample ACFs.

Figure 6. S&P 500: Autocorrelation Functions of Squares of Simulated Returns vs. Actual Returns During and After the Crisis



Notes. Shown are the sample ACFs of *squared actual returns* (in dots) and 95% confident interval bands of sample ACFs of *squared* model-simulated returns The data set used is the S&P 500 daily log returns from August 2007 to October 2013. The SV-DEJ model seems to be the best, as the 95% CI from the model covers more sample ACFs.

Figure 7. (Color online) Posterior Distributions of λ^+ and λ^- in the SV-DEJ-VG Model for the S&P 500



Notes. The left graphs are the histograms of 1,000 posterior draws of λ^+ of the SV-DEJ-VG model, and the right the λ^-. The data periods are January 1980–December 2000 (top), January 2001–July 2007 (middle), and August 2007–October 2013 (bottom).

Table 7. Change of Model Parameters

Parameter	Jan 2001–Jul 2007 (Period A)	Aug 2007–Jun 2009 (Period B)	Jul 2009–Oct 2013 (Period C)	95% CI of the percentage change from period A to B	95% CI of the percentage change from period A to C
			Panel A: S&P 500		
Positive jump rate λ^+	0.0576 (daily) 14.5152 (annually)	0.0573 (daily) 14.4396 (annually)	0.0310 (daily) 7.8120 (annually)	[5.00, 14.32]	[−51.88, −32.97]
Negative jump rate λ^-	0.0776 (daily) 19.5552 (annually)	0.0758 (daily) 19.1016 (annually)	0.0590 (daily) 14.8680 (annually)	[5.87, 15.65]	[−34.31, −10.07]
Mean positive jump size η^+	0.1029	0.1587	0.1407	[12.36, 52.87]	[−8.35, 67.91]
Mean negative jump size η^-	0.1338	0.1296	0.1845	[−32.49, 47.52]	[13.27, 25.94]
Positive jump impact $\eta^+ * \lambda^+$	0.0067	0.0084	0.0050	[15.58, 26.43]	[−52.99, −8.30]
Negative jump impact $\eta^- * \lambda^-$	0.0082	0.0089	0.0094	[1.57, 20.06]	[−9.77, 33.83]
Annual mean volatility $\sqrt{\theta \times 252}$	15.3888	21.3099	16.2926	[11.60, 52.50]	[−21.43, 10.85]
			Panel B: NASDAQ-100		
Positive jump rate λ^+	0.0346 (daily) 8.7192 (annually)	0.0430 (daily) 10.8360 (annually)	0.0317 (daily) 7.9884 (annually)	[10.82, 104.82]	[41.18, 47.40]
Negative jump rate λ^-	0.1058 (daily) 26.6616 (annually)	0.1045 (daily) 17.3376 (annually)	0.0588 (daily) 14.8176 (annually)	[2.76, 34.18]	[−50.94, −36.76]
Average positive jump size η^+	0.2859	0.1857	0.2214	[−39.18, −5.10]	[−36.66, 0.00]
Average negative jump size η^-	0.6421	0.6687	0.7591	[−37.16, 38.14]	[5.40, 49.98]
Positive jump impact $\eta^+ * \lambda^+$	0.0078	0.0083	0.0062	[−42.17, 34.86]	[−27.79, −13.36]
Negative jump impact $\eta^- * \lambda^-$	0.0652	0.0619	0.0420	[−51.74, 5.72]	[−49.89, −32.75]
Annual mean volatility $\sqrt{\theta \times 252}$	19.2969	24.9843	17.7192	[−17.14, 72.30]	[−18.82, 29.11]

Notes. It seems that (1) during the crisis, negative jump rate λ^- has increased significantly, whereas there is little change in the average negative jump sizes η^-; and (2) jump rate does not necessarily increase with volatility, somewhat different from previous findings in Eraker (2004) and Kumar et al. (1999). We report in the table the changes of model parameters estimated for SV-DEJ in January 2001–July 2007, August 2007–June 2009, and July 2009–October 2013 for the S&P 500 (panel A) and the NASDAQ-100 (panel B) returns. The left and right endpoints of 95% credible intervals for the percentage of change are approximated by their simulation values. That is, we report the 2.5 and 97.5 quantiles of $(\theta_Y^{(m)} - \theta_X^{(m)})/|\theta_X^{(m)}|$, where $\theta_X^{(m)}$ and $\theta_Y^{(m)}$ denote the value of the mth draw of the corresponding posterior distribution in any two periods X and Y, respectively. Here, $m = 1, \ldots, M$ and $M = 10{,}000$ in the current empirical study.

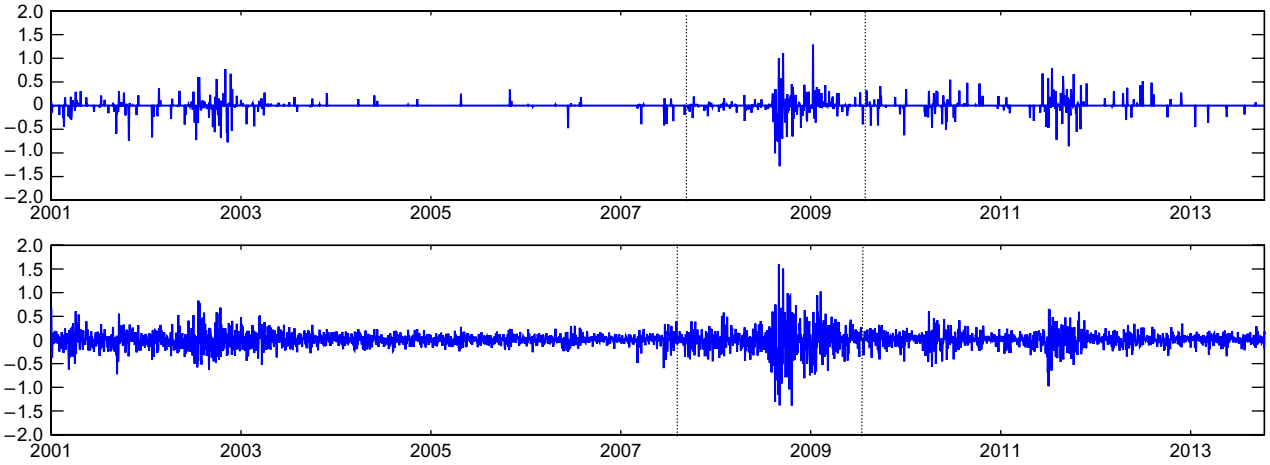
also the best-performing model in generating the ACFs of absolute returns (see Figure 5) and squared returns (see Figure 6) after the outbreak of the recent crisis, since the confidence intervals (the bands) contain most of the sample ACFs from the data (the dots). A more complex volatility structure by introducing jumps in volatility does not necessarily lead to better capturing of the volatility clustering effects, probably because of the increase of mean-reverting speed.

5. Change of Jump Patterns Before, During, and After the Recent Crisis

Based on the SV-DEJ model, the best-fitted model, we report the changes in model parameters in different time periods. We distinguish three periods: before, during, and after the crisis, using the breakpoints chosen as in Section 4.

The changes of the model parameters are reported via their 95% credible intervals. The left and right endpoints of the 95% credible intervals for the percentage of changes from a period X to period Y are approximated by their simulation values. More precisely, denote as $\theta_Y^{(m)}$ the mth draw of the parameter of interest in the posterior simulation in period Y and $\theta_X^{(m)}$ the mth draw in period X, for $m = 1, \ldots, M$ and $M = 10{,}000$ in the current empirical study; we simply report the 2.5% and 97.5% quantiles in the pool of random draws of $(\theta_Y^{(m)} - \theta_X^{(m)})/|\theta_X^{(m)}|$. To better compare the movement in returns resulting from jumps during different periods, we also look at the jump impact defined as $\lambda \times \eta$, where λ and η are the rate and average size for either positive or negative jumps, respectively. The 95%

Figure 8. (Color online) S&P 500: Returns vs. Jumps



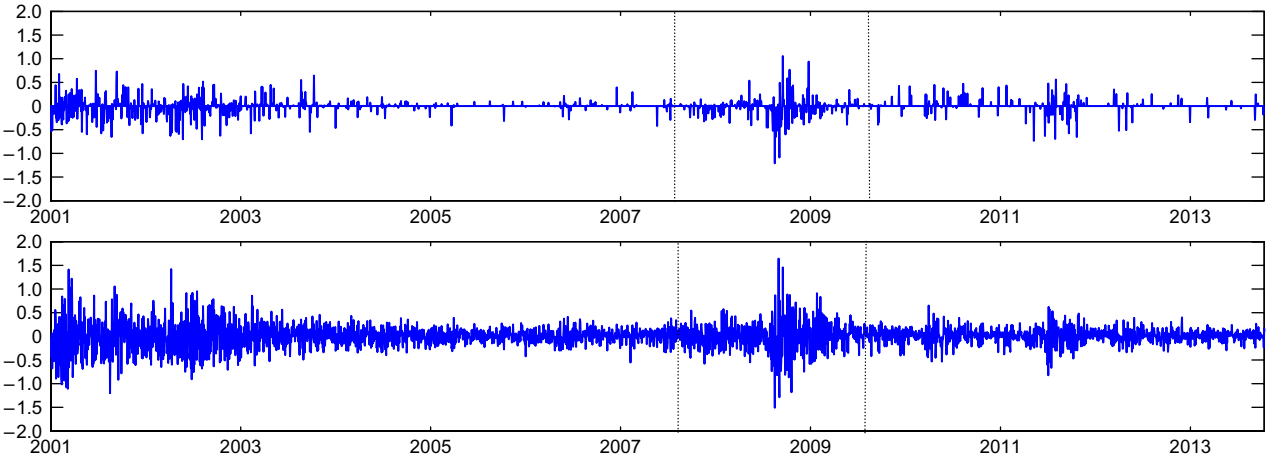
Notes. Jumps (i.e., $1_{(N_{t+1}=1)}(\xi^+_{t+1}) + 1_{(N_{t+1}=-1)}(-\xi^-_{t+1})$, in the top panel) imputed by the SV-DEJ model for the S&P 500 (in the bottom panel). The dotted vertical lines indicate the outbreak of the financial crisis (August 2007) and its end (June 2009).

credible intervals for changes in both positive and negative jump impact for different periods, together with change in jump rates and average sizes, are reported in Table 7. For a visual impression of the imputed jumps, see Figure 8 for the S&P 500 and Figure 9 for the NASDAQ-100.

Several comments are in order. (1) It appears that during the crisis, negative jump rate λ^- has increased significantly, whereas there is little change in the average negative jump sizes η^-. (2) Jump rates can decrease even when there is no change in volatility. Indeed, although volatility in the precrisis period is on average comparable with that in the postcrisis period (see the credible interval of change for average square root of annualized volatility $\sqrt{252\times\theta}$ from period A to period C), negative jump rate λ^- has significantly decreased after the crisis for both the S&P 500 and the NASDAQ-100, and positive jump rate λ^+ has also decreased after the crisis for the S&P 500. In other words, λ^+ and λ^- might decrease when θ remains unchanged. This is somewhat different from previous findings in Eraker (2004) and Johannes et al. (1999), who study models with jump rate $\lambda_t = \lambda_0 + \lambda_1 V_t$ (λ_0, λ_1 are constants) and find $\lambda_1 > 0$. Note that in their models, λ_t is the time-varying jump rate of both positive and negative jumps and V_t is the volatility.

We note that negative jump sizes on average become larger after June 2009 compared with the level before July 2007. Hence the precrisis pattern does not seem to reassert itself after the crisis has settled. Larger negative jumps are observed in both the S&P 500 and the NASDAQ-100 returns in the after-crisis period with implications for market participants. In particular, out-of-the-money equity vanilla options become more expensive, whereas at-the-money options become cheaper, since larger jumps are less likely to end up in the neighborhood of the strike price.

Figure 9. (Color online) NASDAQ-100: Returns vs. Jumps



Notes. Jumps (i.e., $1_{(N_{t+1}=1)}(\xi^+_{t+1}) + 1_{(N_{t+1}=-1)}(-\xi^-_{t+1})$, in the top panel) imputed by the SV-DEJ model for the NASDAQ-100 (in the bottom panel). The dotted vertical lines indicate the outbreak of the financial crisis (August 2007) and its end (June 2009).

We should be very cautious about the empirical conclusion aforementioned about the changes in the jump rates and jump sizes, because the length of the data series is only about five years (2008–2013) since the financial crisis of 2008. However, our findings suggest that the previous pattern of positive correlation between jump rates and volatility does not necessarily hold after the crisis of 2008.

6. Conclusion

We attempt to answer two questions about jumps in equity returns: (i) How did jumps in equity returns change after the 2008 financial crisis—in particular, were there significant changes in jump rates or in jump sizes, or both? (ii) Can the performance of affine jump-diffusion models be improved if the jump sizes are larger, i.e., jumps with tails heavier than those of the normal distribution, such as the double-exponential distribution? To answer these questions, we first find that a simple affine jump-diffusion model with both stochastic volatility and double-exponential jump sizes in returns fits both the S&P 500 and the NASDAQ-100 daily return data well, both before and during the crisis. In fact, the model outperforms existing ones (in particular, models with Lévy-type small jumps, affine jump-diffusion models with normal jump sizes, or models with jumps in volatility) during the crisis, and it is at least as good before the crisis. On the basis of the model and the data, we answer the aforementioned two questions as follows: (i) We find that (a) negative jump rate during the crisis increased significantly, whereas there is little change in the average negative jump sizes; and (b) jump rate can decrease even when volatility stays unchanged, somewhat different from previous findings. (ii) Our empirical study favors double-exponential jumps in equity returns; in short, an affine jump-diffusion model with a proper jump-size distribution and stochastic volatility can fit equity return data well both before and after the crisis.

Acknowledgments

The authors thank two anonymous referees and the associate editor for helpful comments.

Appendix A. MCMC Subroutines and Proofs for the SV-DEJ Model

The MCMC procedures for the SV-MJ-JV model and the SV-VG model have been provided in Eraker et al. (2003) and Li et al. (2008). Moreover, the conditional posteriors of parameters of the SV-DEJ-JV, the SV-DEJ-VG, and the SV-DEJ-VG-JV models that have not been provided in the literature are those associated with the double-exponential jumps in the returns. Thus in what follows, we focus on the MCMC procedures for the SV-DEJ model. More details of the conditional posteriors of the SV-DEJ-VG-JV model can be found in the online supplement (available at http://www.rmi.nus.edu.sg/aboutus/profile/stevenkou/stevenkou.html). We mostly choose conjugate priors to ease computation. The priors are relatively not informative and are used in previous study; see the values of hyperparameters below and also the prior mean and standard deviation in Table 2. We also check the robustness of the results reported in Section 4 by specifying different combinations of hyperparameters. Results do not seem to differ significantly. Bayesian MCMC inferences are done on the parameter space Θ (say, under the SV-DEJ model, $\Theta = \{\mu, \kappa, \theta, \rho, \sigma_v, \eta^+, \eta^-, \lambda^+, \lambda^-\}$), latent variables $N_{1:T}$ (the jump times), $V_{0:T}$ (the volatilities), $\xi^+_{1:T}$ (positive jump sizes), and $\xi^-_{1:T}$ (negative jump sizes), given the returns data $Y_{0:T}$. Below are details of all conditional posteriors.

First, we introduce the conditional posteriors for parameters and latent variables of the SV-DEJ model that are shared with the SV-MJ-JV model. One can easily discover the similarity by comparing the results here and those in Section A.2 of Li et al. (2008). Thus the proofs are omitted for the following conditional posteriors.

1. *Posterior for μ:* The posterior of μ follows a normal distribution, $\mu \sim \mathrm{N}(S^{(\mu)}/W^{(\mu)}, 1/W^{(\mu)})$, where $W^{(\mu)} = (\Delta/(1-\rho^2)) \cdot (\sum_{t=0}^{T-1}(1/v_t)) + 1/M^{(\mu)^2}$, $S^{(\mu)} = (1/(1-\rho^2)) \sum_{t=0}^{T-1}(1/v_t)(C^{(\mu)}_{t+1} - \rho(D^{(\mu)}_{t+1}/\sigma_v)) + m^{(\mu)}/M^{(\mu)^2}$, $C^{(\mu)}_{t+1} = y_{t+1} - y_t - J^y_{t+1}$, and $D^{(\mu)}_{t+1} = v_{t+1} + (\kappa\Delta - 1)v_t - \kappa\theta\Delta$. Here, the prior for μ is $\mathrm{N}(m^{(\mu)}, M^{(\mu)^2})$, where $m^{(\mu)} = 0$, $M^{(\mu)} = 1$ in our study.

2. *Posterior for θ:* The posterior of θ follows a truncated normal distribution, $\theta \sim \mathrm{N}(S^{(\theta)}/W^{(\theta)}, 1/W^{(\theta)})1_{(\theta>0)}$, where $W^{(\theta)} = \kappa^2\Delta/(\sigma_v^2(1-\rho^2)) \sum_{t=0}^{T-1}(1/v_t) + 1/(M^{(\theta)^2})$, $S^{(\theta)} = (\kappa/((1-\rho^2)\sigma_v)) \sum_{t=0}^{T-1}((D^{(\theta)}_{t+1}/\sigma_v - \rho C^{(\theta)}_{t+1})/v_t) + m^{(\theta)}/M^{(\theta)^2}$, $C^{(\theta)}_{t+1} = y_{t+1} - y_t - \mu\Delta - J^y_{t+1}$, and $D^{(\theta)}_{t+1} = v_{t+1} + (\kappa\Delta - 1)v_t$. Here, the prior for θ is $\mathrm{N}(m^{(\theta)}, M^{(\theta)^2})1_{(\theta>0)}$, where $m^{(\theta)} = 0$, $M^{(\theta)} = 1$ in our study.

3. *Posterior for κ:* The posterior of κ follows a truncated normal distribution, $\kappa \sim \mathrm{N}(S^{(\kappa)}/W^{(\kappa)}, 1/W^{(\kappa)})1_{(\kappa>0)}$, where $W^{(\kappa)} = (\Delta/(\sigma_v^2(1-\rho^2)))(\sum_{t=0}^{T-1}((\theta - v_t)^2/v_t)) + 1/M^{(\kappa)^2}$, $S^{(\kappa)} = (1/((1-\rho^2)\sigma_v)) \sum_{t=0}^{T-1}(((\theta - v_t)(D^{(\kappa)}_{t+1}/\sigma_v - \rho C^{(\kappa)}_{t+1}))/v_t) + m^{(\kappa)}/M^{(\kappa)^2}$, $C^{(\kappa)}_{t+1} = y_{t+1} - y_t - \mu\Delta - J^y_{t+1}$, and $D^{(\kappa)}_{t+1} = v_{t+1} - v_t$. Here, the prior for κ is $\mathrm{N}(m^{(\kappa)}, M^{(\kappa)^2})1_{(\kappa>0)}$, where $m^{(\kappa)} = 0$, $M^{(\kappa)} = 1$ in our study.

4. *Posterior for σ_v and ρ:* Following Jacquier et al. (1994), we transform (ρ, σ_v) one-to-one into (ϕ_v, w_v), where $\phi_v = \sigma_v\rho$ and $w_v = \sigma_v^2(1-\rho^2)$. The priors are a normal–inverse-gamma distribution: $\phi_v \mid w_v \sim \mathrm{N}(0, \frac{1}{2}w_v)$ and $w_v \sim \mathrm{IG}(m^{(\mathrm{RS})}, M^{(\mathrm{RS})})$, where $m^{(\mathrm{RS})} = 2$, $M^{(\mathrm{RS})} = 200$ in our study. Given this reparameterization, the joint posteriors of (ϕ_v, w_v) are the conjugate of the priors

$$w_v \sim \mathrm{IG}\left(\frac{T}{2} + m^{(\mathrm{RS})}, \frac{1}{(1/2)\sum_{t=0}^{T-1}(D^{(\mathrm{RS})}_{t+1})^2 + 1/M^{(\mathrm{RS})} - S^{(\mathrm{RS})^2}/(2W^{(\mathrm{RS})})}\right)$$

and

$$\phi_v \mid w_v \sim \mathrm{N}\left(\frac{S^{(\mathrm{RS})}}{W^{(\mathrm{RS})}}, \frac{w_v}{W^{(\mathrm{RS})}}\right),$$

where $W^{(\mathrm{RS})} = \sum_{t=0}^{T-1}(C^{(\mathrm{RS})}_{t+1})^2 + 2$, $S^{(\mathrm{RS})} = \sum_{t=0}^{T-1} C^{(\mathrm{RS})}_{t+1} D^{(\mathrm{RS})}_{t+1}$, $C^{(\mathrm{RS})}_{t+1} = (y_{t+1} - y_t - \mu\Delta - J^y_{t+1})/\sqrt{v_t\Delta}$, and $D^{(\mathrm{RS})}_{t+1} = (v_{t+1} - v_t - \kappa(\theta - v_t)\Delta)/\sqrt{v_t\Delta}$.

5. *Posterior for the latent variable* v_{t+1}: For $0 < t+1 < T$, the posterior of v_{t+1} is

$$p(v_{t+1}\,|\,\cdot) \propto \exp\left[-\frac{-2\rho\epsilon^y_{t+1}\epsilon^v_{t+1}+(\epsilon^v_{t+1})^2}{2(1-\rho^2)}\right]\times\frac{1}{v_{t+1}} \times\exp\left[-\frac{(\epsilon^y_{t+2})^2-2\rho\epsilon^y_{t+2}\epsilon^v_{t+2}+(\epsilon^v_{t+2})^2}{2(1-\rho^2)}\right],$$

where $\epsilon^y_{t+1}=(y_{t+1}-y_t-\mu\Delta-J^y_{t+1})/\sqrt{v_t\Delta}$, $\epsilon^v_{t+1}=(v_{t+1}-v_t-\kappa(\theta-v_t)\Delta)/(\sigma_v\sqrt{v_t\Delta})$. For $t+1=T$, the above posterior only has the first exponential part because v_T depends only on v_{T-1}. Similarly, the posterior of $p(v_0\,|\,\cdot)$ depends on $1/v_0$ and the second exponential part. Following Li et al. (2008), we make use of the adaptive rejection metropolitan sampling method of Gilks et al. (1995) to draw from this conditional posterior.

Next, we turn to parameters and variables that are special to SV-DEJ. The conditional posteriors are new, and they are derived as follows.

1. *Posterior for* η^+, η^-: The posterior of η^+ follows an inverse-gamma distribution,

$$\eta^+ \sim \mathrm{IG}\left(T+m^{(\mathrm{EP})}, \frac{1}{\sum_{t=0}^{T-1}\xi^+_{t+1}+1/M^{(\mathrm{EP})}}\right),$$

where the prior of η^+ is $\mathrm{IG}(m^{(\mathrm{EP})}, M^{(\mathrm{EP})})$ and $m^{(\mathrm{EP})}=3$, $M^{(\mathrm{EP})}=0.5$ in our study. Similarly, the posterior of η^- also follows an inverse-gamma distribution,

$$\eta^- \sim \mathrm{IG}\left(T+m^{(\mathrm{EM})}, \frac{1}{\sum_{t=0}^{T-1}\xi^-_{t+1}+1/M^{(\mathrm{EM})}}\right),$$

where the prior of η^- is $\mathrm{IG}(m^{(\mathrm{EM})}, M^{(\mathrm{EM})})$ and $m^{(\mathrm{EM})}=3$, $M^{(\mathrm{EM})}=0.5$ in our study.

Proof. Since the two cases for η^+ and η^- are almost the same, we focus on the case of η^+ only. To ease notation, in this proof, we denote $m=m^{(\mathrm{EP})}$ and $M=M^{(\mathrm{EP})}$. By Bayes' rule,

$$\begin{aligned}
&p(\eta^+\,|\,\Theta\backslash\{\eta^+\}, N_{1:T}, Y_{0:T}, V_{0:T}, \xi^+_{1:T}, \xi^-_{1:T})\\
&\quad= p(\eta^+\,|\,\xi^+_{1:T}) \propto p(\xi^+_{1:T}\,|\,\eta^+)p(\eta^+)\\
&\quad\propto \Pi_{t=0}^{T-1}[(\eta^+)^{-1}\exp(-\xi^+_{t+1}/\eta^+)]\cdot(\eta^+)^{-m-1}\exp\left(-\frac{1/M}{\eta^+}\right)\\
&\quad\propto (\eta^+)^{-(m+T)-1}\exp\left(-\frac{\sum_{t=0}^{T-1}\xi^+_{t+1}+1/M}{\eta^+}\right),
\end{aligned}$$

where the equality results from independency. This is readily recognized as a standard exponential–gamma conjugate pair model.

2. *Posterior for the latent variables* ξ^+_{t+1}, ξ^-_{t+1}: For $1 \le t+1 \le T$, the posterior of ξ^+_{t+1} follows a truncated normal distribution when $N_{t+1}=1$:

$$\xi^+_{t+1} \sim \mathrm{N}\left(\frac{S^{(\mathrm{XP})}}{W^{(\mathrm{XP})}}, \frac{1}{W^{(\mathrm{XP})}}\right)1_{(\xi^+_{t+1}>0)},$$

where $S^{(\mathrm{XP})}=(C^{(\mathrm{XP})}_{t+1}-\rho D^{(\mathrm{XP})}_{t+1}/\sigma_v)/(v_t\Delta(1-\rho^2))-1/\eta^+$, $W^{(\mathrm{XP})}=1/(v_t\Delta(1-\rho^2))$, $C^{(\mathrm{XP})}_{t+1}=y_{t+1}-y_t-\mu\Delta$, and $D^{(\mathrm{XP})}_{t+1}=v_{t+1}-v_t-\kappa(\theta-v_t)\Delta$. When $N_{t+1}\neq 1$, the posterior is the same as its prior, $\xi^+_{t+1}\sim\exp(\eta^+)$. Similarly, for $1 \le t+1 \le T$, the posterior of ξ^-_{t+1} follows a truncated normal distribution when $N_{t+1}=-1$:

$$\xi^-_{t+1} \sim \mathrm{N}\left(\frac{S^{(\mathrm{XM})}}{W^{(\mathrm{XM})}}, \frac{1}{W^{(\mathrm{XM})}}\right)1_{(\xi^-_{t+1}>0)},$$

where $S^{(\mathrm{XM})}=-(C^{(\mathrm{XP})}_{t+1}-\rho D^{(\mathrm{XP})}_{t+1}/\sigma_v)/(v_t\Delta(1-\rho^2))-1/\eta^-$, and $W^{(\mathrm{XM})}=1/(v_t\Delta(1-\rho^2))$. And when $N_{t+1}\neq -1$, the posterior is the same as its prior, $\xi^-_{t+1}\sim\exp(\eta^-)$.

Proof. Since the cases for $N_{t+1}=1$ and $N_{t+1}=-1$ are similar, we focus on the case where $N_{t+1}=1$. To ease notation, in this proof, we let $m=m^{(\mathrm{XP})}$, $M=M^{(\mathrm{XP})}$, $C_{t+1}=C^{(\mathrm{XP})}_{t+1}$, $D_{t+1}=D^{(\mathrm{XP})}_{t+1}$, $S=S^{(\mathrm{XP})}$, and $W=W^{(\mathrm{XP})}$. It is easy to see that the posterior is $\xi^+_{t+1}\sim\exp(\eta^+)$ when $N_{t+1}\neq 1$ since the data provide no information. By Bayes' rule, when $N_{t+1}=1$,

$$\begin{aligned}
p(\xi^+_{t+1}\,|\,\cdot) &= p(\xi^+_{t+1}\,|\,y_{t+1}, y_t, v_{t+1}, v_t, N_{t+1}=1, \Theta)\\
&\propto p(y_{t+1}\,|\,y_t, \Theta, N_{t+1}=1, v_t, v_{t+1}, \xi^+_{t+1})p(\xi^+_{t+1}\,|\,\eta^+)\\
&\propto \exp\left[-\frac{1}{2v_t\Delta(1-\rho^2)}(C_{t+1}-\xi^+_{t+1}-\rho D_{t+1}/\sigma_v)^2\right]\\
&\quad\cdot\exp\left(-\frac{\xi^+_{t+1}}{\eta^+}\right)1_{(\xi^+_{t+1}>0)}\\
&\propto \exp\left(-\frac{1}{2}W(\xi^+_{t+1})^2+S\xi^+_{t+1}\right)1_{(\xi^+_{t+1}>0)}.
\end{aligned}$$

The required conclusion readily follows by completing the squares and comparing the parameters.

3. *Posterior for* $(\lambda^+, \lambda_0, \lambda^-)$: The posterior of $(\lambda^+, \lambda_0, \lambda^-)$ follows a Dirichlet distribution,

$$p((\lambda^+, \lambda_0, \lambda^-)\,|\,\cdot) \sim \mathrm{D}\left(\alpha_1+\sum_{t=0}^{T-1}1_{(N_{t+1}=1)}, \alpha_0+\sum_{t=0}^{T-1}1_{(N_{t+1}=0)}, \alpha_{-1}+\sum_{t=0}^{T-1}1_{(N_{t+1}=-1)}\right),$$

where the prior of $(\lambda^+, \lambda_0, \lambda^-)$ is $\mathrm{D}(\alpha_1, \alpha_0, \alpha_{-1})$, and $\alpha_1=\alpha_{-1}=2$ and $\alpha_0=40$ in our study.

Proof. By Bayes' rule,

$$\begin{aligned}
p((\lambda^+, \lambda_0, \lambda^-)\,|\,\cdot) &= p((\lambda^+, \lambda_0, \lambda^-)\,|\,N_{1:T})\\
&\propto p(N_{1:T}\,|\,(\lambda^+, \lambda_0, \lambda^-))p(\lambda^+, \lambda_0, \lambda^-)\\
&\propto (\lambda^+)^{\sum_{t=0}^{T-1}1(N_{t+1}=1)}\lambda_0^{\sum_{t=0}^{T-1}1(N_{t+1}=0)}\\
&\quad\cdot(\lambda^-)^{\sum_{t=0}^{T-1}1(N_{t+1}=-1)}(\lambda^+)^{\alpha_1-1}\lambda_0^{\alpha_0-1}(\lambda^-)^{\alpha_{-1}-1}\\
&= (\lambda^+)^{\sum_{t=0}^{T-1}1(N_{t+1}=1)+\alpha_1-1}\lambda_0^{\sum_{t=0}^{T-1}1(N_{t+1}=0)+\alpha_0-1}\\
&\quad\cdot(\lambda^-)^{\sum_{t=0}^{T-1}1(N_{t+1}=-1)+\alpha_{-1}-1}\\
&\sim \mathrm{D}\left(\alpha_1+\sum_{t=0}^{T-1}1(N_{t+1}=1), \alpha_0+\sum_{t=0}^{T-1}1(N_{t+1}=0),\right.\\
&\qquad\left.\alpha_{-1}+\sum_{t=0}^{T-1}1(N_{t+1}=-1)\right).
\end{aligned}$$

4. *Posterior for the latent variable* N_{t+1}: For $1 \le t+1 \le T$, the posterior of N_{t+1} follows a trinomial distribution,

$$p(N_{t+1}=i\,|\,\cdot)=\begin{cases}\dfrac{\lambda^+\exp(U_1)}{S^{(N)}} & \text{when } i=1,\\[2ex] \dfrac{\lambda_0\exp(U_0)}{S^{(N)}} & \text{when } i=0,\\[2ex] \dfrac{\lambda^-\exp(U_{-1})}{S^{(N)}} & \text{when } i=-1,\end{cases}$$

where

$$U_1 = -\frac{1}{2v_t\Delta(1-\rho^2)}\left[(C_{t+1}^{(N)} - \xi_{t+1}^+)^2 - 2\rho(C_{t+1}^{(N)} - \xi_{t+1}^+)D_{t+1}^{(N)}/\sigma_v\right],$$
$$U_0 = -\frac{1}{2v_t\Delta(1-\rho^2)}\left[(C_{t+1}^{(N)})^2 - 2\rho C_{t+1}^{(N)}D_{t+1}^{(N)}/\sigma_v\right],$$
$$U_{-1} = -\frac{1}{2v_t\Delta(1-\rho^2)}\left[(C_{t+1}^{(N)} + \xi_{t+1}^-)^2 - 2\rho(C_{t+1}^{(N)} + \xi_{t+1}^-)D_{t+1}^{(N)}/\sigma_v\right],$$
$$S^{(N)} = \lambda^+\exp(U_1) + \lambda_0\exp(U_0) + \lambda^-\exp(U_{-1}),$$
$$C_{t+1}^{(N)} = y_{t+1} - y_t - \mu\Delta,$$
$$D_{t+1}^{(N)} = v_{t+1} - v_t - \kappa(\theta - v_t)\Delta.$$

Proof. To ease notation, in this proof, we define the following: $p_1 := \lambda^+$, $p_0 := \lambda_0$, and $p_{-1} := \lambda^-$. By Bayes' rule, we have for $i = -1, 0, 1$

$$p(N_{t+1} = i \mid \cdot) = p(N_{t+1} = i \mid \Theta, y_t, y_{t+1}, v_t, v_{t+1}, \xi_{t+1}^+, \xi_{t+1}^-)$$
$$= \frac{p(y_{t+1} \mid y_t, \Theta, N_{t+1} = i, v_t, v_{t+1}, \xi_{t+1}^+, \xi_{t+1}^-)p(N_{t+1} = i \mid \Theta)}{\sum_j p(y_{t+1} \mid y_t, \Theta, N_{t+1} = j, v_t, v_{t+1}, \xi_{t+1}^+, \xi_{t+1}^-)p(N_{t+1} = j \mid \Theta)}$$
$$= \frac{p(y_{t+1} \mid y_t, \Theta, N_{t+1} = i, v_t, v_{t+1}, \xi_{t+1}^+, \xi_{t+1}^-)p_i}{\sum_j p(y_{t+1} \mid y_t, \Theta, N_{t+1} = j, v_t, v_{t+1}, \xi_{t+1}^+, \xi_{t+1}^-)p_j}.$$

Since all three cases are similar, only the case of $i = 1$ is considered here. When $i = 1$,

$$p(y_{t+1} \mid y_t, \Theta, N_{t+1} = i, v_t, v_{t+1}, \xi_{t+1}^+, \xi_{t+1}^-)$$
$$= p(y_{t+1} \mid y_t, \Theta, N_{t+1} = 1, v_t, v_{t+1}, \xi_{t+1}^+)$$
$$= \frac{1}{\sqrt{2\pi v_t\Delta(1-\rho^2)}}$$
$$\cdot\exp\left[-\frac{1}{2v_t\Delta(1-\rho^2)}(C_{t+1} - \xi_{t+1}^+ - \rho D_{t+1}/\sigma_v)^2\right].$$

The required conclusion readily follows when all common factors are canceled out.

Appendix B. Option Pricing Formula for the SV-DEJ Model

We provide the option pricing formula of the SV-DEJ model as follows. Following the assumptions in Yu et al. (2011), we let $\gamma_t^{(1)} = \eta^s\sqrt{v_t}$ and $\gamma_t^{(2)} = -(1/\sqrt{1-\rho^2})(\rho\eta^s + \eta^v/\sigma_v)\sqrt{v_t}$, where η^s and η^v are the market prices of risk. Then there exists a risk-neutral measure $\mathbb{Q}$ under which $W_t^{(1)}(\mathbb{Q})$ and $W_t^{(2)}(\mathbb{Q})$ are standard Brownian motions: $dW_t^{(i)}(\mathbb{Q}) := dW_t^{(1)} + \gamma_t^{(i)}\,dt$, $i = 1, 2$ (Yu et al. 2011, Equation (2.8)). Under this condition, Yu et al. (2011) provide the pricing formula of vanilla call for affine-jump diffusion models given the jump compensator $\psi_J(u) = \lambda(1 - E^{\mathbb{Q}}[e^{-iX}])$, where λ is the jump rate and X is the jump-size distribution. In the special case of double-exponential distributed jumps, the vanilla call option price with initial log price Y_0, initial volatility v_0, maturity T, and strike K is then given by (assuming constant interest rate r)

$$C(Y_0, v_0, T, K) = E^{\mathbb{Q}}\left[e^{-rT}(e^{y_T} - K)^+\right]$$
$$= \frac{e^{-rT}}{\pi} \times \mathrm{Re}\left(\int_0^\infty e^{-ix\log K}\frac{\phi(x - i, Y_0, v_0, T)}{-x^2 + ix}dx\right).$$

The nontrivial part $\phi(u, y, v, t)$ in the integration is given by

$$\phi(u, y, v, t) = e^{iuY_0 + iu(r + \psi_J(-i))t - t\psi_J(u) - b(t)v_0 - c(t)},$$

where

$$b(t) = \frac{(iu + u^2)(1 - e^{-\delta t})}{(\delta + \kappa_M) + (\delta - \kappa_M)e^{-\delta t}},$$
$$c(t) = \frac{\kappa\theta}{\sigma_v^2}\left[2\log\left(\frac{2\delta - (\delta - \kappa_M)(1 - e^{-\delta t})}{2\delta}\right) + (\delta - \kappa_M)t\right],$$
$$\kappa_M = \kappa - \eta^v - iu\sigma_v\rho, \quad \delta = \sqrt{\kappa_M^2 + (iu + u^2)\sigma_v^2},$$
$$\psi_J(u) = (\lambda^+ + \lambda^-) + \frac{\lambda^+}{iu\eta^+ - 1} - \frac{\lambda^-}{iu\eta^- + 1}.$$

Endnotes

[1] We only consider normal distribution with negative mean since it is a stylized fact that the Gaussian distributed jumps in asset returns have negative mean; see Eraker et al. (2003) and Li et al. (2008).

[2] The notation $\sim \exp(m)$ means that the distribution is exponential with mean m.

[3] Eraker et al. (2003) also consider an SV-IMJ model where jumps in returns and volatility are governed by two independent Poisson processes. They find that both SV-MJ-JV and SV-IMJ lead to similar fit to the data of the S&P 500 returns from 1980 to 1999. However, since SV-IMJ is much harder to estimate, Eraker et al. (2003) prefer SV-MJ-JV.

[4] The notation $\sim \Gamma(\alpha, \beta)$ means that the distribution is gamma with parameter α, β and density $(1/(\beta^\alpha\Gamma(\alpha)))x^{\alpha-1}e^{-x/\beta}$.

[5] Following one referee's advice, we also perform Bayesian MCMC analysis on the data set of 1980–2007; i.e., we do not split into two subperiods before the crisis. It is found that relevant results are similar to what is to be reported below.

[6] Following the convention in Li et al. (2008), in all data sets we multiply the daily returns by 100. Therefore, values of model parameters are in percent. For example, $\eta^- = 3.56$ means that the average downside exponential jump size is 3.56% in log asset returns, thanks to (1).

References

Abate J, Whitt W (1995) Numerical inversion of Laplace transforms of probability distributions. *ORSA J. Comput.* 7:36–43.

Aït-Sahalia Y, Jacod J (2011) Testing whether jumps have finite or infinite activity. *Ann. Statist.* 39:1689–1719.

Bakshi G, Cao C, Chen Z (1997) Empirical performance of alternative option pricing models. *J. Finance* 52:2003–2049.

Bates D (2012) U.S. stock market crash risk, 1926–2010. *J. Financial Econom.* 105:229–259.

Chen N, Kou S (2009) Credit spreads, optimal capital structure, and implied volatility with endogenous default and jump risk. *Math. Finance* 19:343–378.

Chernov M, Gallant R, Ghysels E, Tauchen G (2003) Alternative models for stock price dynamics. *J. Econometrics* 116:225–257.

Duffee G (1995) Stock returns and volatility a firm-level analysis. *J. Financial Econom.* 37:399–420.

Duffie D, Pan J, Singleton K (2000) Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68: 1343–1376.

Eraker B (2004) Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *J. Finance* 59: 1367–1404.

Eraker B, Johannes M, Polson N (2003) The impact of jumps in volatility and returns. *J. Finance* 58:1269–1300.

Gilks WR, Best NG, Tan KKC (1995) Adaptive rejection Metropolis sampling within Gibbs sampling. *Appl. Statist.* 44:455–472.

Heyde C, Kou S (2004) On the controversy over tailweight of distributions. *Oper. Res. Lett.* 32:399–408.

Jacquier E, Polson N, Rossi P (1994) Bayesian analysis of stochastic volatility models. *J. Bus. Econom. Statist.* 12:371–389.

Johannes M, Polson N (2010) MCMC methods for continuous-time financial econometrics. Aït-Sahalia Y, Hansen LP, eds. *Handbook of Financial Econometrics*, Vol. 2 (North-Holland, Oxford, UK), 1–72.

Johannes M, Kumar R, Polson N (1999) State dependent jump models: How do US equity indices jump? Working paper, University of Chicago, Chicago.

Kou S (2002) A jump diffusion model for option pricing. *Management Sci.* 48:1086–1101.

Li H, Wells M, Yu C (2008) A Bayesian analysis of return dynamics with Levy jumps. *Rev. Financial Stud.* 21:2345–2378.

Pan J (2002) The jump-risk premia implicit in options: Evidence from an integrated time-series study. *J. Financial Econom.* 63: 3–50.

Ramezani CA, Zeng Y (2007) Maximum likelihood estimation of the double exponential jump-diffusion process. *Ann. Finance* 3: 487–507.

Singleton K (2006) *Empirical Dynamic Asset Pricing* (Princeton University Press, Princeton, NJ).

Yu C, Li H, Wells M (2011) Estimation of Levy jump models under the risk neutral and physical measure using stock and option prices. *Math. Finance* 21:383–422.

CORRECTION

In this article, "Jumps in Equity Index Returns Before and During the Recent Financial Crisis: A Bayesian Analysis" by Steven Kou, Cindy Yu, and Haowen Zhong (first published in *Articles in Advance*, March 28, 2016, *Management Science*, DOI:10.1287/mnsc.2015.2359), Endnote 4, Table 2, and Appendix A have been corrected.