Journal of Financial Economics 106 (2012) 447–472

Contents lists available at SciVerse ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options ☆

Peter Christoffersen [a,b], Kris Jacobs [c,d,*], Chayawat Ornthanalai [a]

[a] Rotman School of Management, University of Toronto, Canada
[b] Copenhagen Business School and CREATES, Denmark
[c] C.T. Bauer College of Business, 334 Melcher Hall, University of Houston, Houston, TX 77204-6021, United States
[d] Tilburg University, The Netherlands

ARTICLE INFO

Article history:
Received 13 June 2010
Received in revised form
23 September 2011
Accepted 27 November 2011
Available online 7 June 2012

JEL classification:
G13

Keywords:
Compound Poisson jumps
Analytical filtering
Fat tails
Risk premiums

ABSTRACT

We build a new class of discrete-time models that are relatively easy to estimate using returns and/or options. The distribution of returns is driven by two factors: dynamic volatility and dynamic jump intensity. Each factor has its own risk premium. The models significantly outperform standard models without jumps when estimated on S&P500 returns. We find very strong support for time-varying jump intensities. Compared to the risk premium on dynamic volatility, the risk premium on the dynamic jump intensity has a much larger impact on option prices. We confirm these findings using joint estimation on returns and large option samples.

1. Introduction

This paper provides a new modeling framework that allows for general specifications of dynamic volatility models with fat-tailed innovations that are potentially time-varying themselves. These models are much easier to implement than existing models. As suggested in Fleming and Kirby (2003), we directly use GARCH processes as filters for the unobservable state variables. We allow for dynamic fat tails in the standard GARCH model by adding a dynamic Compound Poisson jump component to the conventional normal innovation. Our framework allows for dynamic jumps in returns as well as in volatility. The resulting specifications of complex fat-tailed models can be estimated on return data using a standard maximum likelihood estimation procedure and an analytical filtering technique to identify the two return innovations. Estimation using extensive option data sets is also feasible. Our approach allows for separate risk premiums for the fat-tailed jump and normal innovations. We provide the risk-neutral processes for use in option valuation, and conduct an extensive empirical investigation of return and option fit.

Because implementation of our models is relatively straightforward, we are able to estimate models with very complex tail characteristics. We consider four different

☆ Christoffersen and Jacobs are grateful for financial support from FQRSC, IFM2 and SSHRC. Ornthanalai was supported by CIREQ and J.W. McConnell Fellowships. Part of this research was done while Ornthanalai was at the Georgia Institute of Technology. For helpful comments we would like to thank Pedro Santa-Clara (the referee), Giovanni Barone-Adesi, David Bates, Tom McCurdy, Anders Trolle, participants at the AFA, the EFA, the NFA, the EFMA, the FDIC Derivatives Conference, CESG, and seminar participants at the Federal Reserve Board, the Atlanta Fed, CU Boulder, HKUST, the Helsinki School of Economics, York, INSEAD, and Rotman. Any remaining inadequacies are ours alone.

* Corresponding author at: C.T. Bauer College of Business, 334 Melcher Hall, University of Houston, Houston, TX 77204-6021, United States. Tel.: +1 713 743 2826; fax: +1 713 743 4622.
E-mail address: kjacobs@bauer.uh.edu (K. Jacobs).

0304-405X/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jfineco.2012.05.017

but nested models. The simplest model, labeled DVCJ (dynamic volatility with constant jumps), has a constant jump intensity. The CVDJ (constant volatility with dynamic jumps) model has time-varying jump intensity, but the normal innovation to the return process is assumed to be homoskedastic. The DVSDJ (dynamic volatility with separate dynamic jumps) model is the most general model we investigate: both the jump intensity and the normal innovations are time-varying, and the dynamics are parameterized separately. The DVDJ (dynamic volatility with dynamic jumps) model is a special case of DVSDJ: both the jumps and the normal innovation are time-varying, but parameterized identically. We first estimate these models using 47 years of daily returns. After estimating the models, we risk-neutralize the parameters and compare their option valuation performance using 13 years of index option data. Subsequently, we estimate the models using returns and option data jointly.

Our empirical results emphasize the importance of time-varying jump intensities to generate dynamic fat tails, fit returns, and value options. The index return data reveal strong support for the DVSDJ and DVDJ models. When using the risk-neutralized estimates from returns to value options, the DVSDJ model yields a 45% improvement in the option-implied volatility root mean squared errors over the standard GARCH model. The DVDJ model also performs well. It yields a 34% improvement in the option-implied volatility root mean squared errors over the GARCH model. When estimating the models using return data and option data jointly, the DVSDJ model again performs best, and the DVDJ model also performs well.

Our findings therefore indicate the benefit of specifications with time-varying jump intensities and dynamic fat tails, with the more richly parameterized model performing relatively better, in-sample as well as out-of-sample. Our results also indicate that fat-tailed models provide superior option pricing performance, particularly during medium and high volatility periods. The contribution of the jump component to total equity volatility varies with the specification of the jump innovation. The more flexible the jump innovation, the more important it is, and the larger is the contribution of the jump component to total equity volatility.

We also investigate the importance of the risk premium on the jump innovation. We conclude that to produce significant improvements in option valuation, models must allow for a risk premium on the jump innovation. We investigate if risk premiums can generate plausible shapes and levels of the implied volatility term structure, and we find that the implied volatility term structure is highly responsive to the jump risk premium. On the other hand, unrealistically large risk premiums for the normal innovation are required to generate levels and slopes of implied volatility found in the data.

Our approach is closely related to the discrete-time jump approach in Maheu and McCurdy (2004), who find strong evidence of time-varying jumps in various individual equity and index returns. We find similar evidence using S&P500 equity index returns, and importantly, we provide theory and empirics on option valuation using our fat-tailed models. Hansen (1994) develops a class of dynamic fat-tailed density models that has subsequently been generalized by Jondeau and Rockinger (2003). However, neither study considered option valuation. Duan, Ritchken, and Sun (2006) provide a risk-neutralization of a discrete-time model with jumps, but they do not allow for time-varying jump intensities and higher-moment dynamics. Option valuation using GARCH models with normal innovations was initially developed in Duan (1995), and explored further in Ritchken and Trevor (1999) and Heston and Nandi (2000).

Our results are also closely related to the existing literature on continuous-time stochastic volatility jump-diffusions (SVJ). The DVCJ model has features similar to the stochastic volatility with correlated jumps in returns and volatility (SVCJ) model underlying most existing empirical estimates. Our DVDJ and DVSDJ models and estimation results are related to the most complex dynamics investigated in the continuous-time literature, as in Santa-Clara and Yan (2010) and Eraker (2004), who specify time-varying intensities, and estimate their models using returns and options jointly.[1]

The remainder of the paper proceeds as follows. Section 2 presents our modeling approach and discusses the four nested specifications. Section 3 provides empirical results from estimating the models on daily returns. Section 4 develops the theoretical framework for risk-neutralization and option valuation. Section 5 provides the empirical results on option valuation using parameters estimated on returns, and Section 6 estimates the models using a joint likelihood composed of returns and options data. Section 7 concludes.

2. Daily returns with jump dynamics

In this section we present the proposed asset return process, including the structure of the jump innovation and the dynamics for volatility and jump intensity.

2.1. The return process

The process contains two components. The first we model using a normal innovation and the second is a jump component. Here, we discuss some aspects of the general structure of the jump and normal components. The variance and jump intensity dynamics are discussed in more detail in Section 2.5.

The return process is given by

$$R_{t+1} \equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r + \left(\lambda_z - \frac{1}{2}\right) h_{z,t+1} + (\lambda_y - \xi) h_{y,t+1} + z_{t+1} + y_{t+1}, \quad (1)$$

[1] For other empirical estimates of jump processes in the SVJ literature using returns and/or options, see, for example, Bates (2000, 2006), Andersen, Benzoni, and Lund (2002), Pan (2002), Huang and Wu (2004), Eraker, Johannes, and Polson (2003), Broadie, Chernov, and Johannes (2007), Li, Wells, and Yu (2007, 2011), and Chernov, Gallant, Ghysels, and Tauchen (2003). Bollerslev and Todorov (2011) demonstrate the need for a time-varying risk premium with a model-free investigation.

where S_{t+1} denotes the underlying asset price at the close of day $t+1$, and r the risk free rate. Shocks to returns are generated by the normal component z_{t+1} and the jump component y_{t+1}, which are assumed to be independent. The normal component z_{t+1} is assumed to be distributed as $N(0,h_{z,t+1})$, where $h_{z,t+1}$ is the conditional variance.

We model the jump component using a compound Poisson process, which has been the standard process used in the continuous-time jump literature since Merton (1976). We let y_{t+1} be conditionally distributed as compound Poisson $J(h_{y,t+1},\theta,\delta^2)$, where $h_{y,t+1}$ denotes the jump intensity, θ the mean jump size, and δ^2 the jump size variance. The convexity adjustment terms $\frac{1}{2}h_{z,t+1}$ and $\xi h_{y,t+1}\equiv(e^{\theta+\delta^2/2}-1)h_{y,t+1}$ in (1) act as compensators to the normal and jump components, respectively. They ensure that when taking conditional expectations of the gross rate of return, we get

$$E_t\left[\frac{S_{t+1}}{S_t}\right]=e^{r+\lambda_z h_{z,t+1}+\lambda_y h_{y,t+1}}, \tag{2}$$

which shows that $\lambda_z h_{z,t+1}+\lambda_y h_{y,t+1}$ is the conditional equity premium, with λ_z and λ_y denoting the market prices of risk for the normal and jump components, respectively. Our setup allows for a time-varying equity premium, and its dynamic will depend on the specification of $h_{z,t+1}$ and $h_{y,t+1}$.

The specification in (1) is similar in spirit to the return process of Heston and Nandi (2000), with the addition of a jump component. This fat-tailed jump innovation allows for more flexible modeling of the higher moments of the return distribution.

Other studies have formulated generalizations of the basic GARCH setup that are related to our approach. Barone-Adesi, Engle, and Mancini (2008) propose a nonparametric approach to modeling conditionally nonnormal innovations. Maheu and McCurdy (2004, 2008) study discrete-time return processes with jumps and provide analytical filters but do not provide the risk-neutral process, which is necessary for option valuation. Duan, Ritchken, and Sun (2006, 2007) propose return processes similar to (1), and provide risk-neutralization arguments, but their setup differs from ours in important ways. First, we use a different specification of the equity premium. Our specification of the equity premium is affine in the state variables. We use this structure below to separate the jump and normal risk premiums, and it also makes it relatively straightforward to derive the risk-neutral dynamic, which is crucial for option pricing. Second, Duan, Ritchken, and Sun (2006, 2007) consider processes without time-varying jump intensities. Third, in our specification the jump and normal components carry separate risk premiums. Fourth, the conditional skewness and kurtosis in Duan, Ritchken, and Sun (2006, 2007) are constant through time, whereas they are dynamic in our specifications.

2.2. The structure of the jump innovation

The compound Poisson structure assumes that the jump size is independently drawn from a normal distribution with mean θ and variance δ^2. The number of jumps n_{t+1} arriving between times t and $t+1$ is a Poisson counting process with intensity $h_{y,t+1}$. The jump component in the period $t+1$ return is therefore given by

$$y_{t+1}=\sum_{j=0}^{n_{t+1}} x_{t+1}^j,$$

where $x_{t+1}^j, j=0,1,2,\ldots n_{t+1}$ is a sequence of independent
and identically distributed normal random variables, $x_{t+1}^j \overset{iid}{\sim} N(\theta,\delta^2)$. The conditional expectation of the number of jumps arriving over the time interval $(t,t+1)$ equals the jump intensity, $E_t[n_{t+1}]=h_{y,t+1}$.

The mean and variance of the jump component, y_{t+1}, are given by $\theta h_{y,t+1}$ and $(\delta^2+\theta^2)h_{y,t+1}$, respectively. Intuitively, $h_{y,t+1}$ should be time-varying as the number of jumps occurring at any time period will depend on market conditions. Unfortunately, jump models with time-varying jump intensity are difficult to estimate and implement in continuous-time models with latent factors, because the likelihood function is typically not available in closed form. The filtration procedure for the latent jump and stochastic volatility processes is also far from straightforward. The filtration problem in our setup can be solved analytically. We will test several specifications for $h_{y,t+1}$, ranging from the simple case of a constant arrival rate to modeling it using a separate GARCH dynamic.

2.3. Conditional moments of returns

Conditional moments of returns for our proposed models can be readily derived using the moment generating function of the normal and compound Poisson processes. The first four conditional moments are

$$E_t(R_{t+1})\equiv\mu_{t+1}=r+(\lambda_z-\tfrac{1}{2})h_{z,t+1}+(\lambda_y-\xi)h_{y,t+1}, \tag{3}$$

$$Var_t(R_{t+1})=h_{z,t+1}+(\delta^2+\theta^2)h_{y,t+1}, \tag{4}$$

$$Skew_t(R_{t+1})=\frac{\theta(3\delta^2+\theta^2)h_{y,t+1}}{(h_{z,t+1}+(\delta^2+\theta^2)h_{y,t+1})^{3/2}}, \tag{5}$$

$$Kurt_t(R_{t+1})=3+\frac{(3\delta^4+6\delta^2\theta^2+\theta^4)h_{y,t+1}}{(h_{z,t+1}+(\delta^2+\theta^2)h_{y,t+1})^2}, \tag{6}$$

where $Skew_t(R_{t+1})$ is the conditional skewness of returns, and $Kurt_t(R_{t+1})$ is the conditional kurtosis. The sign of the conditional skewness depends on the sign of the mean jump size θ. For positive $h_{y,t+1}$, which means in the presence of jumps, the dynamics of conditional skewness and kurtosis are driven by the conditional variance of the normal component as well as the jump intensity. Harvey and Siddique (1999, 2000) have shown the importance of time-varying skewness in asset pricing. The time-variation in conditional skewness and kurtosis is a key difference between our approach and the framework in Duan, Ritchken, and Sun (2006, 2007) in which conditional skewness and kurtosis are constant through time.

2.4. The Heston and Nandi GARCH(1,1) benchmark model

Heston and Nandi (2000) propose an affine GARCH model for option valuation. The GARCH(1,1) version of this model is given by

$$R_{t+1} = r + (\lambda_z - \tfrac{1}{2})h_{z,t+1} + \sqrt{h_{z,t+1}}\varepsilon_{t+1},$$

$$h_{z,t+1} = w_z + b_z h_{z,t} + a_z\left(\varepsilon_t - c_z\sqrt{h_{z,t}}\right)^2, \tag{7}$$

where r is the risk free rate, and ε_{t+1} is the innovation term distributed i.i.d. $N(0,1)$. The conditional variance of return is $h_{z,t+1}$. The asymmetric variance response, or the leverage effect, is captured by the parameter c_z. This model is based on conditional normality and thus cannot generate one-period-ahead conditional skewness and excess kurtosis.

The GARCH dynamic in (7) is different from the more conventional nonlinear NGARCH model used by Engle and Ng (1993) and Hentschel (1995), which is used by Duan (1995) to price options. We choose the Heston and Nandi GARCH dynamic in (7) as our benchmark since it ensures a closer correspondence of our jump parameter estimates with the continuous-time literature that typically relies on affine models. However, our jump modeling approach can be applied to any type of GARCH dynamic.

Before we extend the Heston–Nandi framework to the return process in (1), we rewrite the GARCH(1,1) dynamic in (7). Letting $z_{t+1} = \sqrt{h_{z,t+1}}\varepsilon_{t+1}$, we get

$$R_{t+1} = r + (\lambda_z - \tfrac{1}{2})h_{z,t+1} + z_{t+1},$$

$$h_{z,t+1} = w_z + b_z h_{z,t} + \frac{a_z}{h_{z,t}}(z_t - c_z h_{z,t})^2.$$

The unconditional variance is given by $E[h_{z,t+1}] = (w_z + a_z)/(1 - b_z - a_z c_z^2)$, where $b_z + a_z c_z^2$ is the variance persistence. We will use the empirical performance of the Heston–Nandi model as the main benchmark for the evaluation of our proposed jump models. Because we only use the GARCH(1,1) implementation, we henceforth refer to it simply as the GARCH model.

2.5. Four nested models

We now apply an extended Heston–Nandi GARCH dynamic to the two return innovations in (1). For the most general specification, the jump intensity and the variance of the normal innovation are governed by the following dynamic:

$$h_{z,t+1} = w_z + b_z h_{z,t} + \frac{a_z}{h_{z,t}}(z_t - c_z h_{z,t})^2 + d_z(y_t - e_z)^2, \tag{8}$$

$$h_{y,t+1} = w_y + b_y h_{y,t} + \frac{a_y}{h_{z,t}}(z_t - c_y h_{z,t})^2 + d_y(y_t - e_y)^2. \tag{9}$$

In (8)–(9), the dynamics of $h_{z,t+1}$ and $h_{y,t+1}$ are known conditional on information available at time t. The realized jump and normal innovations are identified from returns using an analytical filter that we discuss in Appendix A. Filtering z_t and y_t is relatively simple and extremely fast. It takes less than a second to filter 47 years of daily return using a package such as Matlab on a standard PC. Because the variance and the jump intensity dynamic can be updated analytically, we can conveniently estimate the model using MLE.

The specification implies jumps in volatility, which are supported in the SVJ context by the empirical findings of Broadie, Chernov, and Johannes (BCJ, 2007), Eraker (2004), and Eraker, Johannes, and Polson (2003). Furthermore, the above specification allows for jumps in the jump intensity. Notice that the normal and jump innovations, z_t and y_t, enter separately into the GARCH updating dynamic and hence allow each type of innovation to impact the variance and jump intensity differently.

We investigate four nested models based on the general dynamic in (8)–(9). We now present these four specifications.

2.5.1. The dynamic volatility and constant jumps (DVCJ) model

In this model we turn off the time-varying jump intensity dynamic in (9), while maintaining the normal component's GARCH dynamic. This amounts to the restrictions

$$b_y = 0, \quad a_y = 0, \quad c_y = 0, \quad d_y = 0, \quad e_y = 0.$$

The DVCJ model contains 11 parameters, six more than the Heston–Nandi model. In any given period, the DVCJ model implies that jumps arrive at a constant rate of w_y, regardless of the level of volatility in the market. Although this can seem counter-intuitive, it is assumed in most of the existing SVJ literature. In fact, the DVCJ specification has features similar to the stochastic volatility with correlated jumps (SVCJ) model in the continuous-time SVJ literature, in which jumps are present in both returns and volatilities.

2.5.2. The constant volatility and dynamic jumps (CVDJ) model

The CVDJ model allows for jump dynamics but turns off the dynamic in the conditional variance of the normal component. It is a special case of the general dynamic in (8)–(9), with the restrictions

$$b_z = 0, \quad a_z = 0, \quad c_z = 0, \quad d_z = 0, \quad e_z = 0.$$

In this specification, time-variation in the return distribution is driven by the jump component. The normal component of returns is homoskedastic, with a variance equal to w_z. Given that time-variation is restricted to the jump intensity, we expect to see an increase in the relative importance of the jump component for this specification.

2.5.3. The dynamic volatility and dynamic jumps (DVDJ) model

In the DVDJ specification $h_{z,t+1}$ and $h_{y,t+1}$ are both time-varying but driven by the same dynamic. The jump intensity is affine in the conditional variance of the normal component

$$h_{y,t+1} = k h_{z,t+1}, \tag{10}$$

where k is a parameter to be estimated. This linear specification of jump intensity on the conditional variance has been studied in the SVJ context in Bates (2000) and Pan (2002). The DVDJ specification can be written as a special

case of the most general specification, subject to the following restrictions on the parameters of $h_{y,t+1}$ in (9)

$$w_y = w_z k \quad b_y = b_z, \quad a_y = a_z k, \quad d_y = d_z k$$

$$c_y = c_z, \quad e_y = e_z.$$

2.5.4. The dynamic volatility and separate dynamic jumps (DVSDJ) model

We refer to the most general specification, in which the conditional variance of the normal component and the jump intensity are governed by separate processes, as the DVSDJ model. As the conditional variance and jump intensity vary over time, their relative contribution to return dynamics will also be time-varying as we will show below. Separate dynamics also allow the variance components to mean-revert at different rates.

A number of remarks are in order regarding the relationship of the newly proposed models with the SVJ literature.

First, while GARCH models are useful volatility filters, it is well-known that stochastic volatility (SV) models dominate normal GARCH models in terms of likelihood values based on daily returns. Two-shock SV models create nonnormality in the daily conditional distribution that the data clearly require. Once the GARCH model is augmented with nonnormal innovations, however, the GARCH likelihood values improve drastically (see, for example, Kim, Shephard, and Chib, 1998) and the likelihood-based comparisons between GARCH and SV become much less conclusive. From this perspective, our nonnormal GARCH-jump models help bridge the gap between the GARCH and SVJ literature.

Second, the biggest advantage of the proposed models is that they are straightforward to estimate, while existing models with time-varying jumps in returns and volatility are cumbersome to implement, even when only using returns in estimation. In a continuous-time setting, adding jumps to stochastic volatility models involves an additional set of latent state variables. The study of option pricing in stochastic volatility models with jumps therefore relies on econometric methods that can filter the unobserved state variables, or alternatively, state variables need to be inverted using a limited number of option contracts each day.[2] Presumably, the relative paucity of available estimates of time-varying jump intensities in the SVJ literature is due to the econometric complexity involved in estimating these models. In contrast, we are able to estimate even our most complex models using a long time-series and very large cross-sections of options because the state variables $h_{z,t+1}$ and $h_{y,t+1}$ are computed from returns using a simple and fast analytical filter that we describe in Appendix A.

Third, it is not necessarily straightforward to relate our estimates of jump processes to those available in the SVJ literature. The underlying volatility processes are different, and this will impact the estimates of the compound Poisson processes. The motivation for introducing jump processes into a GARCH setup is also somewhat different from the continuous-time case, because trading is not continuous in our models. In a GARCH setup, the difference between the jump and normal innovations is mainly a distributional one; the jump processes are useful because they allow for flexible modeling of the tails and higher moments, and the resulting features and parsimonious parameterization of our models allow us to analyze large samples of option prices. While the characteristics of the jump processes are of interest, because they affect the measurement of risk premiums, the ultimate criterion to judge the performance of these reduced-form models is how well they fit returns and options. Nevertheless, we provide continuous-time limits of our models in the online supplement and we relate them to the dynamic jump intensity model in Santa-Clara and Yan (2010).

2.6. A two-factor benchmark model

It could be argued that the Heston–Nandi GARCH model (7) is insufficient to benchmark the jump models in Section 2.5, because the jump models contain a second component and are more richly parameterized. We therefore provide an additional benchmark where the return dynamic contains two heteroskedastic normal innovations each driven by Heston–Nandi dynamics. The return dynamic is given by

$$R_{t+1} = r + (\lambda_1 - \tfrac{1}{2})h_{1,t+1} + (\lambda_2 - \tfrac{1}{2})h_{2,t+1} + z_{1,t+1} + z_{2,t+1}, \tag{11}$$

where $z_{1,t+1}$ is distributed $N(0,h_{1,t+1})$ and $z_{2,t+1}$ is distributed $N(0,h_{2,t+1})$, and the variance dynamics are given by

$$h_{1,t+1} = w_1 + b_1 h_{1,t} + \frac{a_{1,t}}{h_{1,t}}(z_{1,t+1} + z_{2,t+1} - c_1 h_{1,t})^2,$$

$$h_{2,t+1} = w_2 + b_2 h_{2,t} + \frac{a_{2,t}}{h_{2,t}}(z_{1,t+1} + z_{2,t+1} - c_2 h_{2,t})^2. \tag{12}$$

We refer to this model as the two-factor GARCH. Because the two return innovations are normally distributed, we can write the conditional return density analytically, and the construction of its likelihood is straightforward.

2.7. The likelihood function

A key benefit of our modeling approach is that the likelihood function is easily derived and easily optimized. The likelihood function for returns depends on the normal and Poisson distributions. First, notice that conditional on $n_{t+1} = j$ jumps occurring between time t and $t+1$, the

[2] Chernov, Gallant, Ghysels, and Tauchen (2003) use an efficient method of moments method, while Eraker, Johannes, and Polson (EJP) (2003), Eraker (2004), and Li, Wells, and Yu (2007) employ Markov Chain Monte Carlo techniques. Bates (2006) develops a characteristic function-based filter for affine models with time-varying jump intensity using returns data, but the filter has not yet been extended to models with jumps in volatility. Santa-Clara and Yan (2010) invert state variables from option prices and estimate their model using implied-state MLE. Pan (2002) uses an implied-state generalized method of moments technique to estimate her model on returns and option prices which also involves inverting state variables from option prices.

conditional density of returns is normal

$$f_t(R_{t+1}|n_{t+1}=j)=\frac{1}{\sqrt{2\pi(h_{z,t+1}+j\delta^2)}}\exp\left(-\frac{(R_{t+1}-\mu_{t+1}-j\theta)^2}{2(h_{z,t+1}+j\delta^2)}\right), \tag{13}$$

where μ_{t+1} is the conditional mean return given by (3). Because the number of jumps is finite in the compound Poisson process, the conditional probability density of returns can be derived by summing over the number of jumps, which is distributed as a Poisson counting process

$$\Pr_t(n_{t+1}=j)=\frac{(h_{y,t+1})^j}{j!}\exp(-h_{y,t+1}). \tag{14}$$

This yields the conditional density

$$f_t(R_{t+1})=\sum_{j=0}^{\infty} f_t(R_{t+1}|n_{t+1}=j)\Pr_t(n_{t+1}=j) \tag{15}$$

and the likelihood function can now be constructed as the product of the conditional distributions across the sample. The log-likelihood is given by

$$L_{returns}=\sum_{t=1}^{T-1}\ln(f_t(R_{t+1})). \tag{16}$$

When implementing maximum likelihood estimation, the summation in (15) must be truncated. Most existing studies assume that jumps of the Poisson type are rare. However, we relax this assumption and allow for the possibility of clustering of several jumps on a day. We truncate the summation at 50 jumps per day. This truncation limit is twice the value used in Maheu and McCurdy (2004). Through experimentation, we have found that our estimation results are robust when increasing the truncation limit beyond 50.

The conventional interpretation of jumps is that they capture large and rare events. In line with this interpretation, we restrict the daily jump intensity, $h_{y,t+1}$, to be less than one. This means that the conditionally expected number of jumps the next day is less than one. The actual number of jumps, n_{t+1}, can be larger than one but is restricted to be less than 50 as mentioned above.

As indicated by (8)–(9), we need to separately identify the two unobserved shocks z_{t+1} and y_{t+1} to filter the conditional variance $h_{z,t+1}$ and the conditional jump intensity $h_{y,t+1}$ which enter the likelihood. The structure of the model allows us to do this in a straightforward way by using the analytical filter in Appendix A.

3. Daily return empirics

In this section we present the data and maximum likelihood estimates on returns.

3.1. Data and method

We estimate our models using the time-series of S&P500 returns from June 1, 1962 through December 31, 2009. The data are obtained from the Center for Research in Security Prices (CRSP). We use the daily time-series of 3-month Treasury bills as our proxy for the risk free rate. The top panel of Fig. 1 shows the daily logarithmic return on the S&P500 for our sample. Several large or "jump-like" movements in returns are apparent. The largest daily price change is the crash of October 1987, when the index fell by almost 25% in a single day. Several large price changes also occur during the financial crisis of 2008. We use a long sample of returns because it is well-known that it is difficult to estimate the dynamic volatility model parameters precisely using short samples. Even more importantly, jumps are rare events, and given an average occurrence of one or two jumps per year estimated in the existing literature, it is difficult to get sensible estimates on jump parameters with daily return data that span less than 20 years. We estimate the models using standard maximum likelihood. The optimization converges quickly and the estimates are robust to a wide range of starting values.

3.2. Maximum likelihood estimates

Table 1 presents maximum likelihood estimates (MLEs) for the new jump models, obtained using returns data for 1962–2009. For reference, Table 2 presents MLEs for four benchmark models: the Heston and Nandi (2000) GARCH model in (7), the Black and Scholes (1973) model, the Merton (1976) jump model, and the two-factor GARCH model in (11). For each jump model, we separate the parameters into two columns, one representing the normal component, and the other representing the jump component. For example, the parameter λ in the "Jump" column refers to the λ_y parameter. Beneath each parameter estimate, we report its robust standard error calculated from the outer product of the gradient at the optimum parameter value. For the two-factor GARCH model, we report its parameters in two columns corresponding to the first and second normal innovations.

The DVDJ model is a distinct case because the jump intensity is linear in the conditional variance by a constant multiple k. Following (2), the conditional equity premium is $(\lambda_z+k\lambda_y)h_{z,t+1}$ and thus, λ_z and λ_y are not separately identified when estimating using returns data. We therefore assume that $\lambda_z=0$ when estimating the DVDJ model using only returns data. This assumption is similar to the one used in Pan (2002) who estimates options and returns of the S&P500 jointly using a linear jump intensity specification similar to the DVDJ model. She finds that neither the diffusion nor the volatility risk premium in her model are statistically significant, and constrains the volatility risk premium to be zero. We will show in Section 6 that the market prices of jump and normal risks are identified in the DVDJ model when it is estimated using options and returns jointly. The robustness of our option valuation results to different risk premium specifications will be analyzed in detail in Section 5.4.

The log-likelihood values show that three of the four jump models in Table 1 significantly improve on the fit of the GARCH model in Table 2. Perhaps not surprisingly, the CVDJ model, which has constant conditional variance of the normal component, has the worst performance among all jump models. It also underperforms the GARCH model, but it must be noted that the log-likelihood comparison of

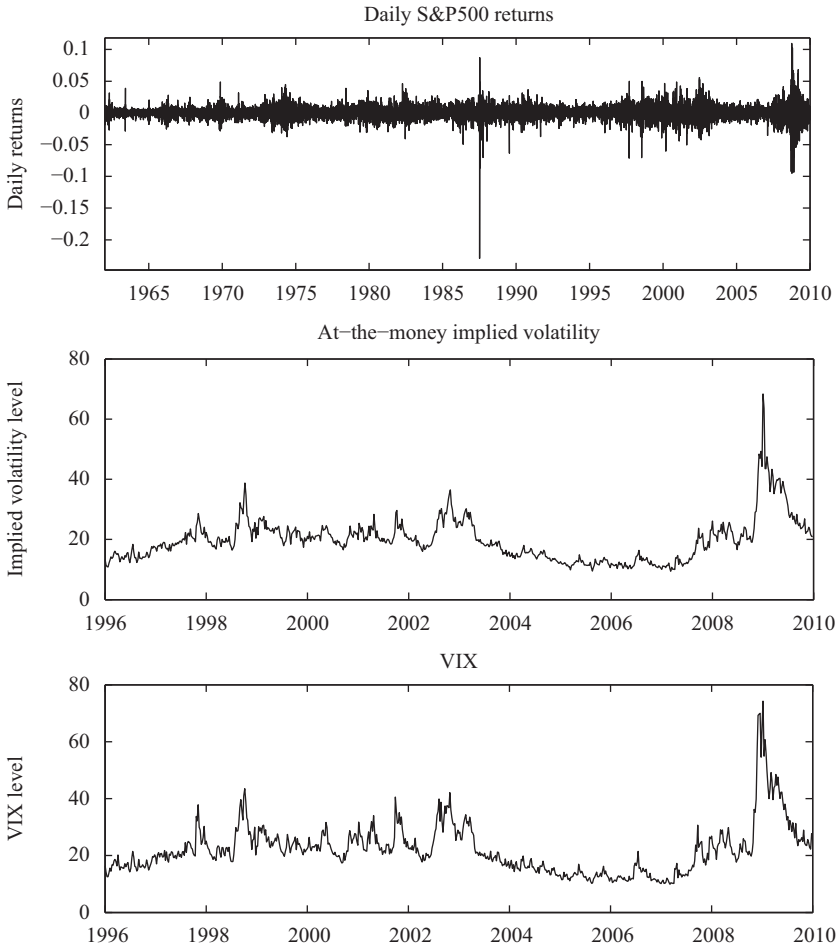

Fig. 1. Daily return and implied volatility on the S&P500 index. The top panel plots the daily S&P500 returns from July 1962 to December 2009. The middle panel plots the average weekly implied Black–Scholes volatility for at-the-money S&P500 options, and the bottom panel plots the VIX index from January 1996 through October 2009. The implied Black–Scholes volatility and VIX index levels are expressed in annualized percentages.

the GARCH model with the CVDJ model is heuristic because these two models are not nested. Note also that the jump intensity dynamic in the CVDJ model is fully persistent.

Fig. 2 plots the autocorrelation function for the squared returns in the top left panel and the autocorrelation of the normalized squared normal model residuals, $z_t^2/h_{z,t}$, in the other five panels. Fig. 2 clearly shows that the DVDJ and DVSDJ models absorb the heteroskedasticity in the S&P500 returns. The CVDJ model, on the other hand, shows substantial autocorrelations in $z_t^2/h_{z,t}$, strongly suggesting that the i.i.d. assumption is violated in that model. The online supplement provides further diagnostics on the models.

The left column of panels in Fig. 3 shows the annualized conditional return volatility for the four proposed jump models. We plot the total conditional return volatility, which is the square root of the conditional return variance given by Eq. (4), from June 1, 1962 through December 31, 2009. To aid the comparison with the conditional volatility path of the GARCH model without jumps, we plot the difference between the conditional volatility of each jump model and the Heston–Nandi model in the right column of Fig. 3. The conditional volatility paths look fairly similar for three of the models in the left column. The conditional volatility for the CVDJ model is very different, and the model has great difficulty capturing large spikes in volatility, again suggesting severe misspecification. For the three models that have more similar volatility paths, the right column shows that important events such as the 1987 crash and the 2008 financial crisis are captured quite differently across models.

The left panels in Fig. 4 report the jump intensity $h_{y,t+1}$ over time for each model. It shows that the CVDJ specification has the largest jump intensity of the proposed jump models. This is perhaps not surprising as jumps are the only source of heteroskedasticity in this model. Table 1 indicates that the size of the jumps arriving in the CVDJ model are also the smallest across models, with a mean jump size $\theta = -0.062\%$.

Consider now the DVCJ jump specification. Adding a simple constant jump component to a GARCH dynamic can significantly improve model fit, even though the variance parameters of the normal component in the

Table 1
MLE estimates of jump models on S&P500 returns, 1962-2009.

We apply MLE to daily returns on the S&P500 index from June 1962 to December 2009. Columns labeled "Normal" contain estimates of parameters governing the normal component, and columns labeled "Jump" contain parameters governing the jump component. Reported in parentheses are standard errors computed using the outer product of the gradient. Under Properties we report Persistence, which refers to variance and jump intensity persistence, respectively, the Percent of annual variance, which refers to the contribution to overall return variation from the normal and jump components, respectively, the Average annual volatility (standard deviation), the Log-likelihood value, and the option-implied volatility root mean squared error (IVRMSE) from using the models in option valuation.

	DVCJ		CVDJ		DVDJ		DVSDJ	
Parameters	Normal	Jump	Normal	Jump	Normal	Jump	Normal	Jump
λ	2.11E+00	−2.57E−02	7.840E+00	−2.10E−04		4.29E−03	5.06E−01	5.88E−03
	(1.08E+00)	(1.67E−02)	(1.68E−01)	(4.45E−04)		(2.19E−03)	(1.48E+00)	(2.4E−03)
w	−1.17E−05	4.85E−03	4.28E−05	−3.25E−03	−1.40E−04		−1.16E−04	−5.29E−02
	(4.75E−06)	(9.57E−04)	(9.364E−07)	(7.02E−09)	(9.74E−05)		(2.52E−06)	(1.90E−06)
b	9.58E−01			9.96E−01	9.44E−01		9.26E−01	9.73E−01
	(3.40E−03)			(2.80E−08)	(3.85E−03)		(4.14E−03)	(1.65E−06)
a	2.28E−06			1.09E−03	2.18E−06		2.23E−06	8.11E−04
	(1.13E−07)			(4.96E−08)	(1.47E−07)		(1.41E−07)	(2.27E−07)
c	1.12E+02			−7.67E+01	1.06E+02		1.32E+02	2.86E−03
	(7.43E+00)			(3.77E+01)	(9.39E+00)		(3.92E−03)	(2.77E+00)
d	1.51E−02			7.26E+00	9.49E−03		1.19E−02	4.74E+00
	(2.56E−03)			(7.46E−01)	(4.92E−03)		(2.56E−04)	(5.16E−01)
e	2.62E−02			1.90E−02	1.21E−01		9.84E−02	1.05E−01
	(7.65E−03)			(1.42E−03)	(7.49E−02)		(7.35E−06)	(5.76E−03)
θ		−2.70E−02		−6.25E−04		−1.66E−02		−1.75E−02
		(8.34E−03)		(4.53E−04)		(2.25E−03)		(8.37E−07)
δ		4.21E−02		1.84E−02		1.03E−02		9.78E−03
		(3.49E−03)		(4.35E−04)		(9.14E−04)		(1.35E−03)
k						4.53E+02		
						(1.01E+02)		
Properties								
Persistence	0.996			1.000	0.996		0.994	0.997
Percent of annual variance	87.25%	12.75%	22.60%	77.40%	85.21%	14.79%	85.01%	14.99%
Average annual volatility	15.48%		16.19%		15.37%		15.31%	
Log-likelihood	40,597		39,671		40,642		40,672	
Option IVRMSE	6.38		7.48		4.38		3.61	

DVCJ model are similar to those of the GARCH model. Table 1 reports a mean jump size of $\theta = -2.703\%$, and jump volatility of $\delta = 4.208\%$. As for the jump intensity $E[h_{y,t+1}] = w_y$, the model implies that jumps arrive at a frequency of $252 \times w_y \simeq 1.22$ jumps per year. These estimates are similar to the ones obtained by Eraker (2004) and EJP (2003) using the SVCJ model. Eraker, using data for 1970–1990, gets approximately one jump per year, $\theta = -2.84\%$ and $\delta = 4.91\%$. EJP (2003) use 1980–2000 data and get 1.52 jumps per year, $\theta = -2.63\%$ and $\delta = 2.89\%$.

The results for the DVDJ model indicate that allowing for state-dependent jump intensities can further improve model performance. The estimate of k is statistically significant, confirming that the arrival rate of jumps depends on the level of volatility. We also estimate the DVDJ model with a linear jump intensity specification $h_{y,t+1} = k_0 + k_1 h_{zt+1}$ (not reported), and find that k_0 is not significant. The mean jump size in the DVDJ model is somewhat smaller than in the DVCJ model. However, jumps arrive more frequently, with on average $E[h_{y,t+1}] = kE[h_{z,t+1}] = 0.0342$, or 8.62 jumps per year. When allowing for time-varying jump intensities, smaller jumps can occur at a frequency that depends on the level of risk in the market. Fig. 4 indeed shows considerable variation over time in the jump intensity of the DVDJ model.

The likelihood for the most general specification, DVSDJ, further improves on DVDJ. This illustrates the benefit of allowing the jump intensity to be driven by its own independent dynamic. The persistence of the jump intensity is slightly larger than the persistence of the variance of the normal component. The MLE estimates of the DVSDJ model indicate that the average jump intensity is $E[h_{y,t+1}] = 0.0336$, which translates into 8.47 jumps per year, very similar to the DVDJ model. The average jump size is also similar to that in the DVDJ model.

Although Fig. 4 shows that the jump intensity dynamics of the DVSDJ model are similar to those of the more parsimonious DVDJ model, the models differ with respect to the response of the jump intensity to shocks. Table 1 shows that the parameter c_y is not significant in the DVSDJ model, while for the DVDJ model, c_y is restricted to equal c_z, which is positive and significant. The parameter c_y captures the asymmetric response in the jump intensity dynamic to negative and positive normal shocks. The more richly parameterized DVSDJ model therefore indicates that the normal return innovation does not induce asymmetry in the news impact curve of the jump intensity. On the other hand, the leverage parameter e_y is positive and highly significant, suggesting that jump intensity responds asymmetrically to positive

Table 2
MLE estimates of benchmark models on S&P500 returns, 1969–2009.

We apply MLE to daily returns on the S&P500 index from June 1962 to December 2009. Columns labeled "Normal" contain estimates of parameters governing the normal component, and columns labeled "Jump" contain parameters governing the jump component. For the Two-factor GARCH model, "Normal-1" and "Normal-2" refer to the parameters that govern the first and the second normally distributed shocks. BSM refers to the Black and Scholes (1973) model, and Merton refers to the jump-diffusion model of Merton (1976). Reported in parentheses are standard errors computed using the outer product of the gradient. Under Properties we report Persistence, which refers to variance persistence, the Percent of annual variance, which refers to the contribution to overall return variation from the normal and jump components respectively, the Average annual volatility (standard deviation), the Log-likelihood value, and the option-implied volatility root mean squared error (IVRMSE) from using the models in option valuation.

	BSM	Merton		GARCH		Two-factor GARCH	
Parameters	Normal	Normal	Jump	Normal	Jump	Normal-1	Normal-2
λ	1.46E+00	4.48E+00		1.66E+00		5.20E+00	5.02E−01
	(9.12E−01)	(2.42E+00)		(9.05E−01)		(2.26E+00)	(1.93E+00)
w	1.04E−04	3.31E−05	3.63E−01	−1.27E−06		−1.12E−06	4.25E−07
	(3.44E−07)	(1.12E−06)	(2.12E−02)	(8.55E−08)		(8.54E−08)	(1.47E−07)
b				9.45E−01		9.82E−01	3.45E−01
				(3.90E−03)		(3.69E−03)	(7.68E−02)
a				2.81E−06		1.10E−06	3.41E−07
				(1.16E−07)		(1.12E−07)	(5.90E−08)
c				1.16E+02		8.89E+01	1.30E+03
				(8.36E+00)		(1.50E+01)	(1.86E+02)
θ			−6.64E−04				
			(2.45E−04)				
δ			1.31E−02				
			(2.63E−04)				
Properties							
Persistence				0.983		0.997	09.592
Percent of annual variance	100%	34.48%	65.52%	100%		69.56%	30.44%
Average annual volatility	16.20%	15.56%		15.34%		15.07%	
Log- likelihood	37,853	39,341		40,347		40,393	
Option IVRMSE	9.83	7.82		6.61		7.11	

and negative jump innovations in the current period return. That is, negative jumps increase future jump intensities by more than positive jumps.

Table 2 also contains estimation results for the two-factor GARCH model in (11). The number of parameters in the two-factor GARCH model is more similar to that in the jump models, which themselves can be interpreted as two-factor models, with one normal and one jump component. However, the two-factor model is significantly outperformed by the DVCJ, DVDJ, and DVSDJ models. The results in Table 2 therefore indicate that the jump models in Table 1 perform well in terms of in-sample fit, also when compared with more stringent benchmarks than the Heston–Nandi GARCH(1,1) model.

3.3. Variance decomposition and conditional higher moments

At the bottom of Table 1 we report the decomposition of the total unconditional return variance into the normal and jump components. Taking the expectation of (4), we can write the total unconditional return variance, σ^2, as

$$\sigma^2 \equiv \sigma_z^2 + (\delta^2 + \theta^2)\sigma_y^2, \tag{17}$$

where σ_z^2 and σ_y^2 are computed as the time-series averages of $h_{z,t+1}$ and $h_{y,t+1}$. Clearly, the jump contribution to the total return variance in (17) is $(\delta^2+\theta^2)\sigma_y^2$, while the contribution of the normal component to the total return variance is σ_z^2. We report the jump contribution to the total return variance in percentages. Overall, the contribution of return jumps to the total equity return variance is roughly similar for the DVCJ, DVDJ, and DVSDJ models, between 12% and 15%. In the CVDJ model, which does not allow for volatility dynamics, the contribution of return jumps to the overall variance is much higher at 77.40%. In the nonparametric jump literature, Andersen, Bollerslev, and Diebold (2007) find that the ratio of mean realized jump variance over the mean total realized variance is 14.6%, which is very close to our finding for the three models that allow for volatility dynamics. Huang and Tauchen (2005) find that about 7% of the overall equity return volatility is due to jumps.

The right-side panels of Fig. 4 show the percentage of total variance captured by the jump component over time, that is, we plot the time path of $Var_t(y_{t+1})/Var_t\ (R_{t+1}) \times 100$, where $Var_t(y_{t+1})$ is given by $(\delta^2+\theta^2)h_{y,t+1}$ and $Var_t(R_{t+1})$ is given in (4). These plots show another interesting difference between the DVDJ and DVSDJ models. In the DVDJ model, we assume that $h_{y,t} = kh_{z,t}$, which in turn implies that the proportion of total variance explained by jumps is constant and equal to $k(\delta^2+\theta^2)/\ (1+k(\delta^2+\theta^2))$. On the other hand, the DVSDJ model allows for additional flexibility by allowing for the jump contribution to change over time.

Fig. 5 plots the conditional one-day-ahead skewness in (5) and excess kurtosis (6) from the four models. Note that the results for DVCJ are reported using a different scale. The DVCJ model, which allows for dynamic variance and constant-intensity jumps, implies conditional

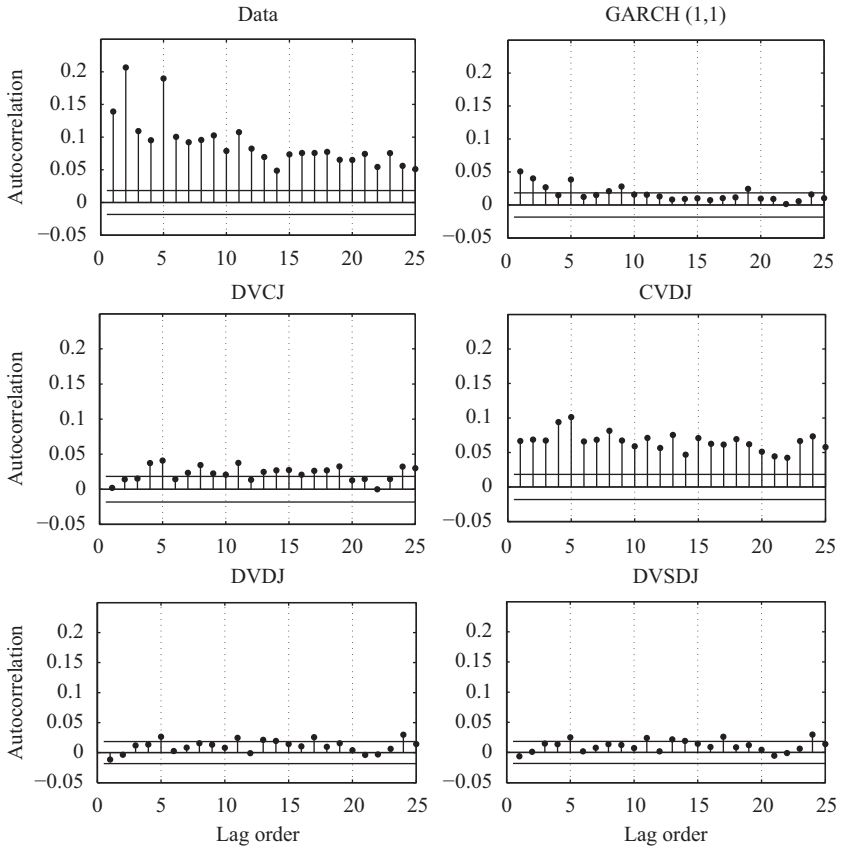

Fig. 2. Autocorrelations of squared S&P500 returns and squared model residuals. The top left panel shows the autocorrelations of daily squared S&P500 returns. The other five panels plot the autocorrelation of the squared standardized return residuals from fitting various models to daily S&P500 index returns. For jump models, we plot the autocorrelation of the squared standardized normal component of return residuals, $z_t^2/h_{z,t}$. The underlying parameter estimates are from Tables 1 and 2.

skewness as low as -15 and conditional excess kurtosis as high as 400. The CVDJ model, in which the normal component has a constant variance, implies almost constant conditional skewness and kurtosis. The DVDJ model, which allows for dynamic jump intensities to be a linear function of the variance of the normal component, implies conditional skewness as low as -2 and conditional excess kurtosis as high as 40. The general DVSDJ model in turn exhibits skewness and kurtosis patterns that are of the same order of magnitude as the DVDJ model but with higher variance in both moments. The DVCJ, DVDJ, and DVSDJ models all display substantially more negative skewness and excess kurtosis in the early part of the sample.

In summary, the four models have very different implications for the level as well as the dynamics of conditional skewness and kurtosis.

3.4. Decomposing daily returns

Fig. 6 presents the results from applying the filtering density in Appendix A to the proposed jump models. Using the MLE estimates in Table 1, we apply the filter to each jump model and back out the time-series of three unobservables: the number of jumps n_{t+1}, the jump component of the return y_{t+1}, and the normal component of the return z_{t+1}. Because the filter is analytic, the filtering process is extremely fast and precise.

The filtered number of jumps occurring each day is shown in the left panels. The filtered jump, y_t, and the standardized normal component, $z_t/\sqrt{h_{z,t}}$, are presented in the middle and right panels, respectively. We find strong evidence for multiple jumps per day in all the proposed jump models. This is especially true for the October 19, 1987 crash when the CVDJ, DVDJ, and DVSDJ models indicate five jumps. The number of jumps for the CVDJ model is nearly always higher than for the other models, because the model is severely misspecified, and the jump component needs to capture the dynamics that are captured by time-varying volatility in the three other models. The average jump size θ in Table 1 is negative for all four models, and Fig. 6 confirms that the estimated jumps induce negative skewness in the return distribution for all models.

The left column confirms that when jump intensities are time-varying (in the DVDJ and DVSDJ models), jumps arrive at much higher frequencies compared to the DVCJ model. The filtered state variables for the DVDJ and DVSDJ models indicate the importance of time-varying jump intensities with clustering effects. The highest jump

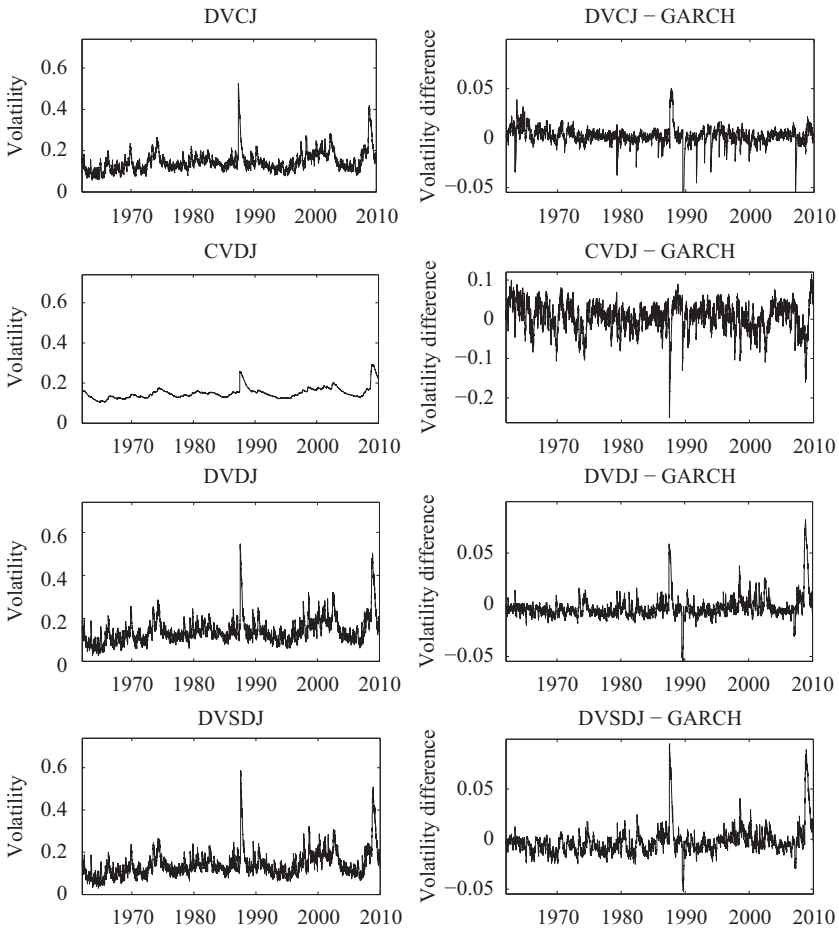

Fig. 3. Conditional volatility: levels and differences with benchmark GARCH. In the left column we plot the annualized conditional volatility for each of the four jump models. In the right column we plot the difference between the annualized conditional volatility of the jump model and that of the benchmark conditionally normal GARCH model. The underlying parameter estimates are from Tables 1 and 2.

arrival frequencies are observed in 1987, the early period of the dot-com collapse, and the 2008 financial crisis. Because more jumps can arrive when the level of risk rises, the DVDJ and the DVSDJ models use a smaller negative jump mean (see θ in Table 1) than the DVCJ model to produce the 1987 crash, the volatility of the late 1990s, and the 2008 financial crisis.

The standardized normal residuals suggest that the CVDJ model misses some volatility dynamics which therefore remain in the shocks. For the DVDJ and DVSDJ models, the shocks appear homoskedastic as desired. The constant jump intensity (DVCJ) model appears to have more medium-sized outliers in the left and right tails than the dynamic jump models (DVDJ and DVSDJ). This indicates that the constant intensity jumps are not fully able to capture the nonnormality in returns.

When jumps occur, they are usually the dominating shock to returns. It can be seen from the middle column that during the crash of October 1987, the jump component accounts for most of the 22.9% drop in the S&P500 index for all four models. This finding is similar to EJP (2003), who find that the jump component accounts for most of the drop in the index on this date.

In summary, the filtering results confirm the importance of time-varying jump intensities.

4. Option valuation theory

We now derive results that allow us to value derivatives using the four proposed jump models. In our framework, the stock price can jump to an infinite set of values in a single period, and the equivalent martingale measure is therefore not unique. We proceed by establishing the existence of a risk-neutral probability density which is of a similar form as the physical density but under which the returns on all assets are equal to the risk free rate. First, we explain our risk-neutralization procedure with an emphasis on the importance of the assumption on the equity premium in (1).

4.1. The pricing kernel

Following the affine asset pricing literature (see Bates, 2006, 2008; Ho, Perraudin, and Sorensen, 1996, for examples), we assume that the pricing kernel π_{t+1} follows the

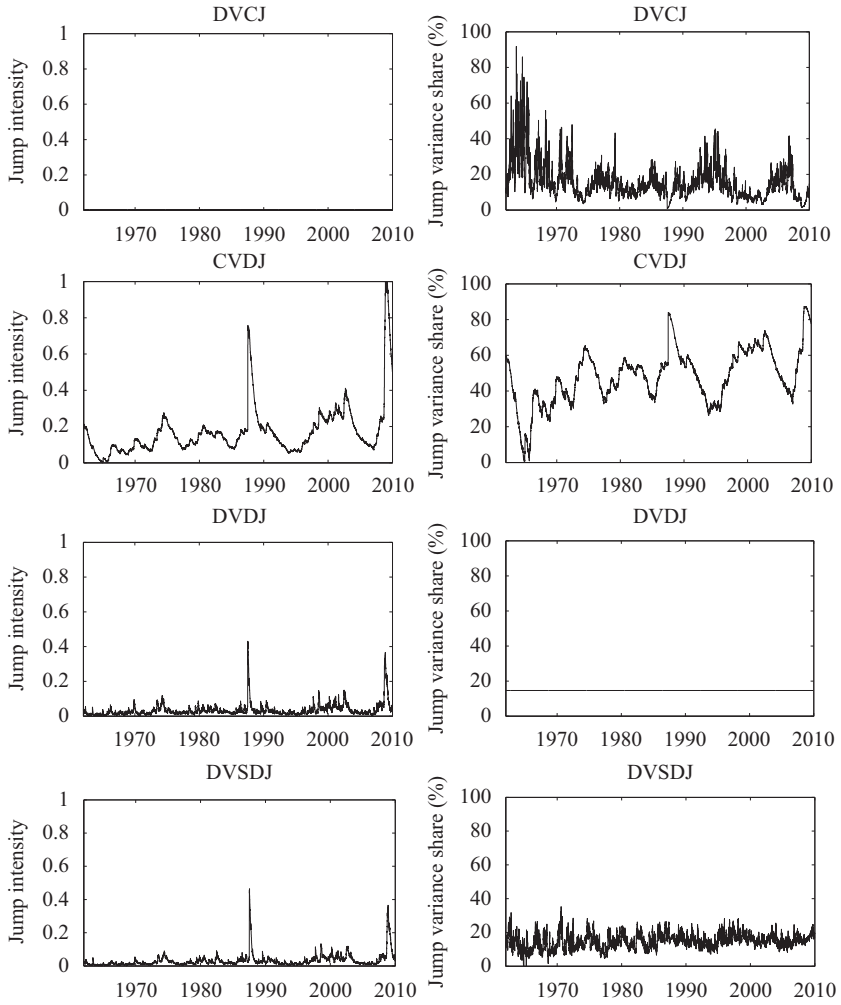

Fig. 4. Conditional jump intensity and variance decomposition. Using the parameter estimates in Table 1, we plot in the left panels the daily conditional jump intensity, $h_{y,t}$, for each of the four jump models, from July 1, 1962 to December 31, 2009. In the right panels we plot the daily percentage jump share of total variance defined as $Var_t(y_{t+1})/Var_t(R_{t+1}) \times 100$.

affine dynamic

$$\log\left(\frac{\pi_{t+1}}{\pi_t}\right) = \alpha_{t+1} - \gamma R_{t+1} - \gamma_J y_{t+1}, \tag{18}$$

where R_{t+1} is the logarithm of the return on the index, and y_{t+1} is the jump component in the index return. The pricing kernel in Eq. (18) is a discrete-time analog to the affine pricing kernel used in various continuous-time studies, including Bates (1991, 2006), Liu, Pan, and Wang (2005), and Eraker (2008). The coefficient α_{t+1} is a normalizing factor such that $E_t[\pi_{t+1}/\pi_t]$ equals the risk free rate. We get

$$\alpha_{t+1} = \log E_t[\exp(r + \gamma R_{t+1} + \gamma_J y_{t+1})]. \tag{19}$$

When $\gamma_J = 0$, and assuming that the index return R_{t+1} is a good proxy for the return on aggregate wealth, the pricing kernel in (18) is consistent with the power utility function where γ is relative risk aversion. The parameter γ_J determines the additional risk premium on jumps in returns beyond the wealth effect already captured by γR_{t+1}.

When $\gamma_J \neq 0$, Liu, Pan, and Wang (2005) and Bates (2008) show that the pricing kernel in (18) is consistent with an economy in which investors are averse to jumps. Eraker (2008) shows that the affine pricing kernel (18) is an equilibrium outcome for a representative agent with Epstein–Zin recursive utility, when the consumption growth rate is affine in the state variable, i.e., the jump component in our case.

Substituting the dynamic of R_{t+1} from Eq. (1) into Eqs. (18) and (19), we get the following dynamic:

$$\frac{\pi_{t+1}}{\pi_t} = \frac{\exp(r - \gamma z_{t+1} - (\gamma + \gamma_J) y_{t+1})}{E_t[\exp(-\gamma z_{t+1} - (\gamma + \gamma_J) y_{t+1})]} \tag{20}$$

for the pricing kernel.

4.2. The equivalent martingale measure

Motivated by the affine structure of the pricing kernel in (20), we proceed by specifying the following

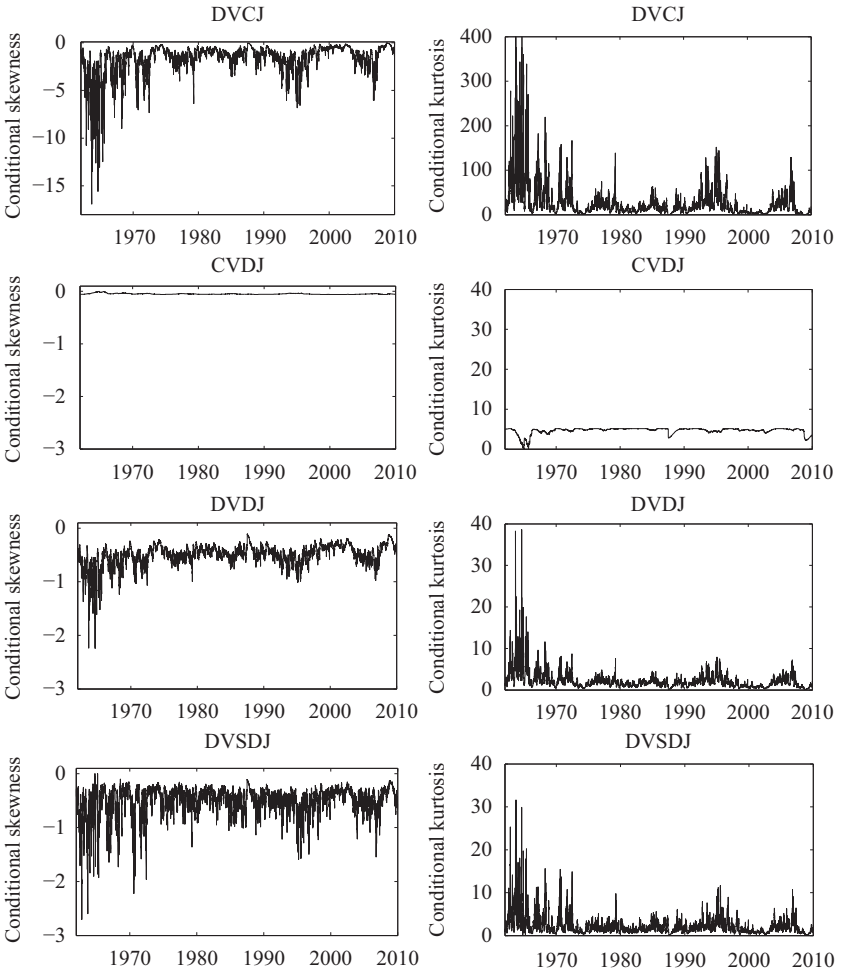

Fig. 5. Conditional skewness and kurtosis. Using the parameter estimates in Table 1, we plot daily conditional skewness and excess kurtosis for the four jump models, from July 1, 1962 to December 31, 2009. Note that the scales are different for the DVCJ model in the top row.

conditional Radon–Nikodym derivative

$$\frac{\frac{dQ_{t+1}}{dP_{t+1}}}{\frac{dQ_t}{dP_t}} = \frac{\exp(\Lambda_z z_{t+1} + \Lambda_y y_{t+1})}{E_t[\exp(\Lambda_z z_{t+1} + \Lambda_y y_{t+1})]}, \tag{21}$$

where z_{t+1} and y_{t+1} are the normal and jump shocks to returns. We define Λ_z and Λ_y as the equivalent martingale measure (EMM) coefficients that capture the wedge between the physical and the risk-neutral measure. Comparing (21) and (20), we see that when $\Lambda_z = -\gamma$ and $\Lambda_y = -\gamma - \gamma_J$, then $(dQ_{t+1}/dP_{t+1})/(dQ_t/dP_t) = \exp(r)\pi_{t+1}/\pi_t$.[3] The Radon–Nikodym derivative we use can therefore be motivated by a representative agent with power utility. The following result puts the conditions on the Radon–Nikodym derivative required to make it an equivalent martingale measure.

Proposition 1. If the dynamic of returns under the physical measure P is given by (1)*, the risk-neutral probability measure Q defined by the Radon–Nikodym derivative in* (21) *is an equivalent martingale measure* (*EMM*) *if and only if*

$$\lambda_y - (e^{(\delta^2/2)+\theta} - 1) - e^{(\Lambda_y^2\delta^2/2)+\Lambda_y\theta}(1 - e^{(1/2+\Lambda_y)\delta^2+\theta}) = 0, \tag{22}$$

$$\Lambda_z + \lambda_z = 0. \tag{23}$$

Proof. See Appendix B.

Eq. (23) of course implies that $\Lambda_z = -\lambda_z$ so that $\gamma = \lambda_z$. This result is identical to the one obtained using Duan's (1995) local risk-neutral valuation method, which builds on Brennan (1979). It is not possible to solve for the second EMM coefficient Λ_y in (22) analytically. However, it is well-behaved and can easily be solved numerically. Note that due to the structure of the equity premium, the market prices of risk λ_z and λ_y enter separately into the above two Eqs. (22) and (23). An estimate of the physical market price of risk λ_z is

[3] A univariate variant of the Radon–Nikodym derivative in (21) has also been used by Christoffersen, Elkamhi, Feunou, and Jacobs (2010), and by Gerber and Shiu (1994), who motivate it using the Esscher (1932) transform. See also Buhlmann, Delbaen, Embrechts, and Shiryaev (1998) and Bakshi, Kapadia, and Madan (2003). See Carr and Wu (2004) for applications to option pricing, and Dai, Le, and Singleton (2010) for applications in discrete-time dynamic term structure modeling.

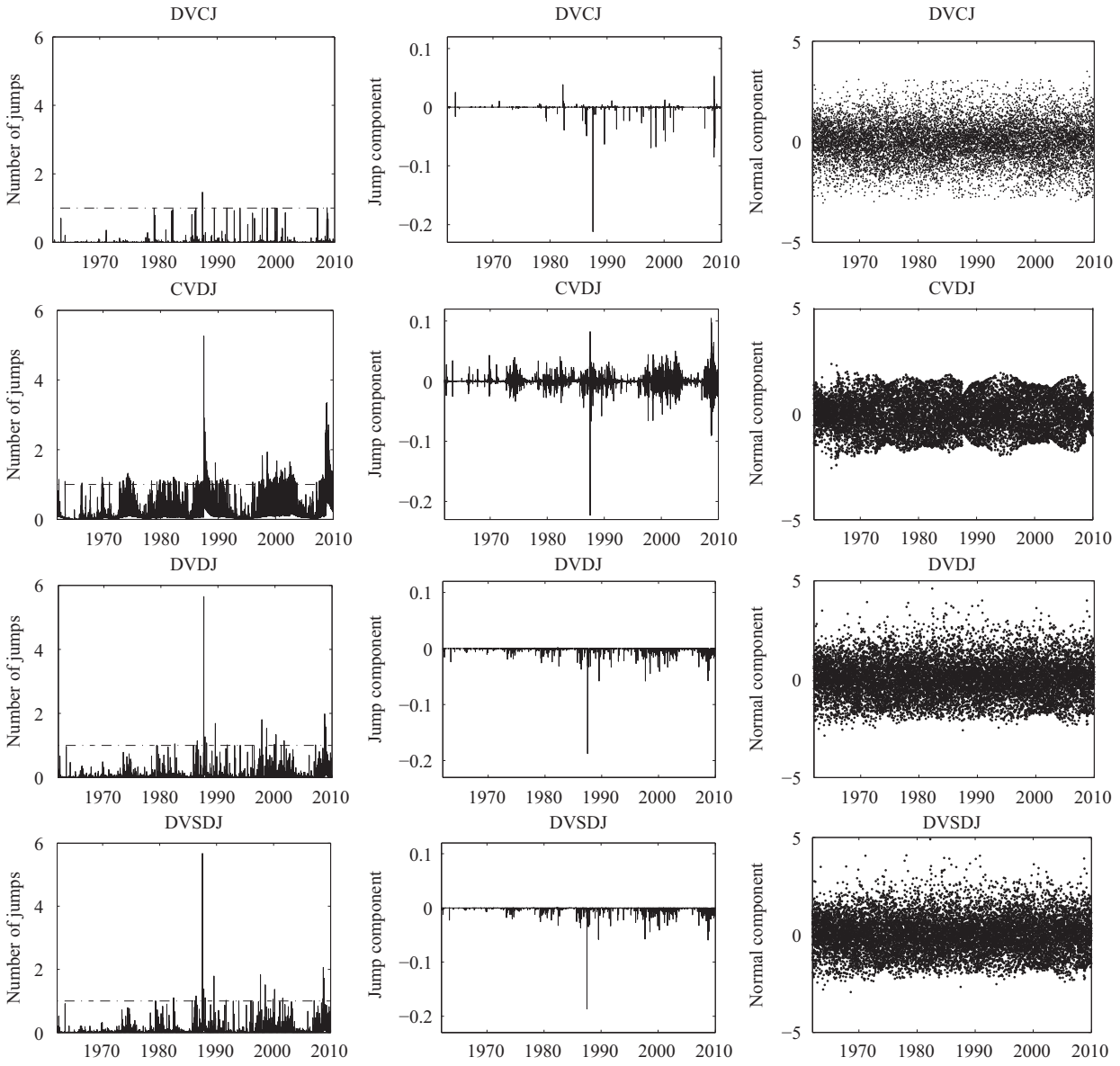

Fig. 6. Decomposition of daily returns. We plot the number of jumps, n_t (left), the jump component, y_t (middle), and the standardized normal component, $z_t/\sqrt{h_{z,t}}$ (right) using the analytical filter in Appendix A. Results are obtained using the MLE estimates from Table 1. The horizontal dashed lines in the left panels are drawn at one jump per day.

sufficient to determine the EMM coefficient Λ_z, and hence the wedge that links the two measures for the normal innovation. Similarly, estimates of the physical jump parameters λ_y, θ, and δ provide the change of measure parameter for the jump innovation, Λ_y. We discuss the implications of these two different types of risk premiums further in Section 5.4.

4.3. Risk-neutral dynamics

We have now characterized an equivalent martingale measure. We can therefore derive the corresponding risk-neutral probability measure for the normal and jump components of returns using a simple change of measure.

Proposition 2. Consider a stochastic process that is the sum of two contemporaneously independent random variables $z_{t+1}+y_{t+1}$, with each component distributed as

$$z_{t+1} \sim N(0,h_{z,t+1}), \quad y_{t+1} \sim J(h_{y,t+1},\theta,\delta^2) \tag{24}$$

under the physical measure P. $N(\cdot)$ and $J(\cdot)$ refer to the normal and compound Poisson distribution, respectively. According to the Radon–Nikodym derivative in (21)*, under the risk-neutral measure Q, the stochastic process can be written as $z^*_{t+1}+y^*_{t+1}$,*

where

$$z^*_{t+1} \sim N(\Lambda_z h_{z,t+1}, h_{z,t+1}) \quad \textit{and} \quad y^*_{t+1} \sim J(h^*_{y,t+1}, \theta^*, \delta^2), \tag{25}$$

with

$$h^*_{y,t+1} = h_{y,t+1} \exp\left(\frac{\Lambda_y^2 \delta^2}{2} + \Lambda_y \theta\right) \quad \textit{and} \quad \theta^* = \theta + \Lambda_y \delta^2. \tag{26}$$

Proof. See Appendix C.

The change of measure shifts the mean of the normal component to the left by $\Lambda_z h_{z,t+1}$, which amounts to $z^*_{t+1} = z_{t+1} - \lambda_z h_{z,t+1}$. This result is identical to Duan (1995), who also motivates the risk-neutralization using the power utility function. The compound Poisson process under the Q measure differs from its distribution under the physical measure in terms of the jump intensity $h^*_{y,t+1}$ and the mean jump size θ^*. This finding is consistent with existing studies, including Naik and Lee (1990), Pan (2002), Bates (2006, 2008), and Liu, Pan, and Wang (2005).

Recall from Eq. (2) that the conditional equity premium for the proposed return dynamic (1) is $\lambda_z h_{z,t+1} + \lambda_y h_{y,t+1}$. The conditional equity premium consists of two components: the normal risk premium and the jump risk premium. The above propositions show that λ_z is related to Λ_z and exclusively determines the shift in the risk-neutral distribution of the normal innovation. We thus refer to $\lambda_z h_{z,t+1}$ as the normal risk premium component. Similarly, λ_y is related to the EMM coefficient Λ_y which exclusively determines the change in jump distribution under the risk-neutral measure. We therefore refer to $\lambda_y h_{y,t+1}$ as the jump risk premium component in the total equity premium. The online supplement provides further discussion of the risk premium structure of our models.

We are now ready to derive the risk-neutral process required for option valuation.

Proposition 3. The stock return process under the risk-neutral dynamic takes the form

$$\log\left(\frac{S_{t+1}}{S_t}\right) = r - \frac{1}{2} h_{z,t+1} - \xi^* h^*_{y,t+1} + z_{t+1} + y^*_{t+1}, \tag{27}$$

with the following variance dynamics:

$$h_{z,t+1} = w_z + b_z h_{z,t} + \frac{a_z}{h_{z,t}}(z_t - c^*_z h_{z,t})^2 + d_z (y^*_t - e_z)^2,$$

$$h^*_{y,t+1} = w^*_y + b_y h^*_{y,t} + \frac{a^*_y}{h_{z,t}}(z_t - c^*_y h_{z,t})^2 + d^*_y (y^*_t - e_y)^2, \tag{28}$$

where

$$h^*_{y,t+1} = h_{y,t+1} \Pi, \quad \xi^* = e^{(\delta^2/2)+\theta^*} - 1, \quad w^*_y = w_y \Pi,$$

$$a^*_y = a_y \Pi, \quad c^*_y = c_y + \lambda_z, \quad c^*_z = c_z + \lambda_z,$$

$$d^*_y = d_y \Pi; \quad \Pi = \exp\left(\frac{\Lambda_y^2 \delta^2}{2} + \Lambda_y \theta\right). \tag{29}$$

Proof. See Appendix D.

The discounted stock price process in (27) is a martingale where $-\frac{1}{2} h_{z,t+1}$ and $\xi^* h^*_{y,t+1}$ are the compensating terms for the normal and jump components, respectively. The risk-neutral dynamic for the Heston–Nandi model is a special case of (27) and (28) with $h^*_{y,t+1} = y^*_{t+1} = 0$.

Closed-form option valuation results are not available. This is due to the jump innovation which does not yield an exponentially affine moment generating function. However, the discrete-time GARCH structure of the model renders option valuation straightforward via Monte Carlo simulation implemented using the antithetic variance reduction technique as well as the empirical martingale simulation of Duan and Simonato (1998).

5. Option valuation empirics using return estimates

When estimating the models on returns data, the evidence in favor of jumps is overwhelming, and the results point to the importance of time-varying jump intensities. We now use the return-based parameter estimates to discuss the importance of jump dynamics when using the models for option valuation.

5.1. Option data

We evaluate the option pricing performance of our models using a rich sample of out-of-the-money S&P500 put and call option data for the period January 1, 1996 to October 31, 2009, which was the time period available in OptionMetrics at the time of writing of the paper. We retrieve European option quotes from OptionMetrics that have maturities between two weeks and one year. We eliminate quotes with zero trading volume, and we keep only out-of-the-money option quotes because they tend to be more liquid. We further eliminate illiquid quotes by keeping only the six most liquid strikes at each maturity. We define moneyness as F/K, where $F = Se^{-r\tau}$ is the implied index futures price at maturity τ, and K is the strike price. We only keep Wednesday options because it is the least likely day to be a holiday and it is least likely to be affected by weekend effects. For further discussion of the advantages of using Wednesday data, we refer to Dumas, Fleming, and Whaley (1998). After, we apply the filters proposed by Bakshi, Cao, and Chen (1997) to the data. Finally, we remove quotes that have inconsistent option-implied volatilities with respect to the other inputs reported by the database provider.[4]

Table 3 presents descriptive statistics for the option quotes by moneyness and maturity. The shape of the volatility smirk is evident from Panel C across all maturities, with short-term options exhibiting the steepest volatility smirk. The middle panel of Fig. 1 plots the Black-Scholes implied volatility using at-the-money options. The bottom

[4] We obtain our index option data from the SAS dataset, midxopprcd, provided by Wharton Research Data Services (WRDS). Using the information (e.g. price, strike, interest rates, dividend) reported in this database, we compute the implied volatilities (IV) for each contract. We then remove quotes where the computed IV differ from the IV reported in the database by more than 5%.

Table 3
S&P500 index option data (1996–2009).

We use out-of-the-money European put and call options on the S&P500 index. The data are obtained from OptionMetrics. The prices are taken from non-zero trading volume quotes on each Wednesday during the January 1, 1996 to October 31, 2009 period. We eliminate illiquid quotes by keeping the six most liquid strikes at each maturity. We also apply the moneyness and maturity filters used by Bakshi, Cao, and Chen (1997) to the data. The implied volatilities are calculated using the Black–Scholes formula. DTM refers to the number of days-to-maturity, and F/K refers to moneyness defined as the ratio of the implied index futures price divided by the strike price.

Panel A: Number of option contracts

	DTM ≤ 30	30 < DTM ≤ 90	90 < DTM ≤ 180	180 < DTM ≤ 250	DTM > 250	All
$F/K \le 0.96$	38	652	685	617	1,040	3,032
$0.96 < F/K \le 0.98$	96	821	372	198	291	1,778
$0.98 < F/K \le 1.02$	491	3,531	1,052	541	672	6,287
$1.02 < F/K \le 1.04$	150	1,219	417	213	316	2,315
$1.04 < F/K \le 1.06$	85	811	468	167	213	1,744
$1.06 < F/K$	102	1,504	1,845	1,451	1,651	6,553
All	962	8,538	4,839	3,187	4,183	21,709

Panel B: Average option prices

	DTM < 30	30 < DTM < 90	90 < DTM < 180	180 < DTM < 250	DTM > 250	All
$F/K \le 0.96$	6.45	13.81	23.98	30.32	38.66	27.90
$0.96 < F/K \le 0.98$	9.56	19.29	34.94	48.35	64.25	32.63
$0.98 < F/K \le 1.02$	16.36	30.08	50.40	66.75	79.74	40.87
$1.02 < F/K \le 1.04$	12.09	22.81	48.12	74.34	94.87	41.25
$1.04 < F/K \le 1.06$	9.60	19.04	35.54	63.27	98.66	36.97
$1.06 < F/K$	5.31	13.31	22.78	28.95	39.30	25.86
All	12.85	22.76	33.31	41.67	54.60	33.58

Panel C: Average implied volatility

	DTM < 30	30 < DTM < 90	90 < DTM < 180	180 < DTM < 250	DTM > 250	All
$F/K \le 0.96$	0.2227	0.2134	0.2011	0.1916	0.1894	0.1981
$0.96 < F/K \le 0.98$	0.1865	0.1877	0.1859	0.1837	0.1910	0.1874
$0.98 < F/K \le 1.02$	0.1785	0.1835	0.1966	0.1932	0.1923	0.1871
$1.02 < F/K \le 1.04$	0.2162	0.2097	0.2205	0.2052	0.1889	0.2088
$1.04 < F/K \le 1.06$	0.2595	0.2306	0.2190	0.2145	0.2071	0.2245
$1.06 < F/K$	0.3247	0.2785	0.2639	0.2527	0.2538	0.2632
All	0.2096	0.2111	0.2263	0.2213	0.2163	0.2169

panel of Fig. 1 contains a time-series for the volatility index, VIX, from the Chicago Board Options Exchange for the same period. Clearly, the data in our sample are representative of prevailing market conditions.

5.2. Overall option valuation performance

We use the MLE estimates from Tables 1 and 2, risk-neutralize them using Proposition 3, and then compute option prices for 1996–2009. Note that due to the GARCH structure of the models, all parameters needed to value options can be estimated using MLE on the returns of the underlying asset. Even though we use a long time-series to estimate the models, Tables 1 and 2 show that the estimates of the market prices of risks λ_z and λ_y are not always statistically significant. When valuing options, we set the market price of risk parameter equal to zero if it is not statistically significant at the 5% level. We investigate the robustness of this assumption in Section 5.4.

We report model fit using implied volatility root mean squared error (IVRMSE). We refer to BCJ (2007) for a discussion on the benefits of using the IVRMSE metric for comparing option pricing models. For the computation of the IVRMSE, we invert each computed option price O_j from the model using the Black-Scholes formula to get the implied volatilities $IV(O_j,K_j,\tau_j,S_j,r_j)$. The IVRMSE is then computed as

$$IVRMSE \equiv \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\sigma_j^{BS}-IV(O_j,K_j,\tau_j,S_j,r_j))^2}, \tag{30}$$

where σ_j^{BS} is the Black–Scholes implied volatility of the jth observed option price, and $N=21{,}709$ is the total number of option contracts used in the analysis. The variables K_j,τ_j, S_j, and r_j are the strike, maturity, the underlying index level, and the risk free rate associated with each option. The models' overall option valuation performance is indicated in the last row of Tables 1 and 2. Panel A in Table 4 expresses the jump models' option valuation performance relative to that of the benchmark GARCH model. The IVRMSE for the benchmark GARCH model is 6.61%. For the other models, Table 4 reports the IVRMSE ratio relative to the GARCH model for ease of comparison.

The most important finding is that the DVDJ and DVSDJ models provide a substantial improvement in option valuation compared to the Heston–Nandi GARCH model. Panel A of Table 4 indicates that the DVDJ model outperforms the GARCH model by almost 34%. The most

Table 4
IVRMSE and ratio IVRMSE by moneyness, maturity, and VIX level.

We use the MLE estimates from Tables 1 and 2 to compute the implied volatility root mean squared error (IVRMSE) for various moneyness, maturity, and VIX level bins. The IVRMSE is reported in levels for the GARCH model. For the jump models, we report the IVRMSE ratio relative to the GARCH model. DTM refers to the number of days-to-maturity, and F/K refers to moneyness defined as the ratio of the implied index futures price divided by the strike price. VIX is the implied volatility index level obtained from the Chicago Board of Options Exchange (CBOE).

Panel A: Overall IVRMSE and ratios

	GARCH	DVCJ	CVDJ	DVDJ	DVSDJ
	IVRMSE(%)	IVRMSE ratio	IVRMSE ratio	IVRMSE ratio	IVRMSE ratio
	6.607	0.966	1.132	0.662	0.546

Panel B: Sorting by moneyness

	GARCH	DVCJ	CVDJ	DVDJ	DVSDJ
	IVRMSE(%)	IVRMSE ratio	IVRMSE ratio	IVRMSE ratio	IVRMSE ratio
$F/K \leq 0.96$	4.849	0.956	0.928	0.671	0.619
$0.96 < F/K \leq 0.98$	4.580	0.966	1.113	0.702	0.607
$0.98 < F/K \leq 1.02$	4.874	0.965	1.153	0.701	0.574
$1.02 < F/K \leq 1.04$	6.067	0.964	1.161	0.669	0.553
$1.04 < F/K \leq 1.06$	6.606	0.959	1.169	0.650	0.529
$1.06 < F/K$	9.227	0.960	1.149	0.643	0.520

Panel C: Sorting by maturity

	GARCH	DVCJ	CVDJ	DVDJ	DVSDJ
	IVRMSE(%)	IVRMSE ratio	IVRMSE ratio	IVRMSE ratio	IVRMSE ratio
$DTM \leq 30$	6.085	0.996	1.380	0.766	0.709
$30 < DTM \leq 90$	5.993	0.975	1.252	0.727	0.633
$90 < DTM \leq 180$	7.097	0.962	1.109	0.648	0.508
$180 < DTM \leq 250$	7.088	0.958	1.042	0.609	0.457
$DTM > 250$	6.938	0.959	0.974	0.593	0.474

Panel D: Sorting by VIX level

	GARCH	DVCJ	CVDJ	DVDJ	DVSDJ
	IVRMSE(%)	IVRMSE ratio	IVRMSE ratio	IVRMSE ratio	IVRMSE ratio
$VIX \leq 14$	1.704	0.814	1.479	1.276	1.477
$14 < VIX \leq 18$	2.659	0.953	1.058	0.864	0.887
$18 < VIX \leq 22$	4.374	0.979	1.089	0.714	0.582
$22 < VIX \leq 26$	6.707	0.987	1.068	0.728	0.558
$26 < VIX \leq 30$	7.713	0.997	1.130	0.706	0.562
$30 < VIX$	12.583	0.950	1.162	0.575	0.465

flexible model, DVSDJ, outperforms the GARCH model by 45%. The jump structure in the DVCJ model is similar to that of the continuous-time SVCJ model, thus, we can compare our findings on its option pricing performance to the existing literature. Based on the IVRMSE metric, the DVCJ outperforms the GARCH(1,1) model by 3.3%. This finding is similar to the 2.3% improvement in fit shown by Eraker (2004) using a different loss function. Our finding is very different from BCJ (2007) who show that the SVCJ model can improve on the IVRMSE over the Heston (1993) model by a striking 50%. However, their implementation is very different. In their setup, the spot volatility is estimated from options rather than filtered from returns and they estimate the risk premiums by minimizing the option pricing error.

Just as with the likelihood comparison in Section 3.2, it could be argued that the jump models have an unfair advantage, because they contain two components, whereas the GARCH model has only one component. However, note that because model parameters are estimated on returns, and the models are evaluated using option data, this is effectively an out-of-sample exercise. In this setup it is not necessarily surprising for a more richly parameterized model to be outperformed by a more parsimonious model. Indeed, the IVRMSE in Tables 1 and 2 indicate that the GARCH model, with an IVRMSE of 6.61%, outperforms the two-factor GARCH model, with an IVRMSE of 7.11%.

Although we do not tabulate the results in the paper, the online supplement reports on the related but non-nested two-component model of Christoffersen, Jacobs, Ornthanalai, and Wang (2008). It has an IVRMSE of 5.15% and thus outperforms the GARCH model. It also outperforms the DVCJ model, but is itself outperformed by the DVDJ and DVSDJ models.

Once again, we conclude that the jump models perform well. Jump models with time-varying intensities perform exceptionally well.

5.3. Valuation errors by moneyness, maturity, and volatility

Table 4 provides additional evidence on the option fit of the four proposed jump models. We report option IVRMSE for the simple GARCH model and the IVRMSE ratio of the four proposed jump models versus the simple GARCH model. We report the results by moneyness, maturity, and VIX index level. The performance of the DVDJ and DVSDJ models is very robust across moneyness and maturity (see Panels B and C). This finding confirms the importance of time-varying jump intensities. The DVCJ outperforms the GARCH model for all moneyness and maturity categories. The CVDJ model performs poorly across the board, except for deep out-of-the-money call options and long maturities.

Panel D of Table 4 reports IVRMSE and IVRMSE ratios sorted by the level of the VIX index. The DVDJ and DVSDJ models perform exceptionally well in medium and high volatility periods (when VIX ≥ 14). In contrast, in the lowest volatility period, which contains less than 2.5% of the sample, only the DVCJ model can significantly improve on the simple GARCH model. These findings suggest that time-varying jumps provide no benefit for option pricing in low volatility periods.

Because of space constraints, we do not report a more detailed comparison between the two-factor models and jump models by moneyness, maturity, and VIX levels, as in Table 4, but the DVDJ and DVSDJ models robustly outperform the two-factor models across different categories. These results are available on request.

5.4. The jump risk and normal risk premiums

The option valuation results are based on parameter estimates obtained from physical returns. However, it is well-known that it is difficult to reliably estimate conditional means, and consequently risk premiums, from return data. For the results in Table 4, we set these statistically insignificant risk premiums equal to zero. It is therefore worthwhile to further explore the implications of alternative risk premium estimates for option pricing performance.

Table 5 reports option pricing performance based on the IVRMSE metric for different levels of the equity premium, keeping all the other parameters at their values from Tables 1 and 2. Panel A shows the GARCH IVRMSE. Panels B, C, and D show the IVRMSE ratios for three jump models relative to the GARCH model. In each panel, the columns represent various levels of the equity premium ranging from 0% to 10%. The rows represent various levels of the risk premium associated with the normal component. For example, in Panel C, if the total equity premium is 8% and the risk premium for the normal innovation is 2%, which means that the jump risk premium is 6%, the IVRMSE ratio is 0.57. For this combination of risk premiums, the DVDJ model thus improves on the GARCH model by 43%. Note that for the GARCH case, we only have entries on the diagonal since the normal innovation is the only source of risk in this model. To save space, we do not report the results of the CVDJ model in Table 5, as it performs poorly relative to GARCH at all combinations of the risk premiums in this exercise.

Table 5
Effects of risk premiums on IVRMSE option valuation performance.

We compute S&P500 Wednesday out-of-the money put and call option prices and implied volatilities for 1996–2009 using MLE estimates from Tables 1 and 2, together with various assumptions on the long-run equity risk premium. Reported are the IVRMSEs of the benchmark Heston–Nandi (2000) GARCH model and IVRMSE ratios of three jump models relative to the GARCH model. The columns represent relative IVRMSEs as the total equity premium increases, and the rows represent the relative IVRMSEs as the normal risk premium increases. For example, when the total equity premium is 6% and the normal risk premium is 2%, the jump risk premium is 4%. Boldfaced font on the diagonal indicates that normal risk accounts for all of the equity premium. Boldface font in the first row indicates that jump risk accounts for all of the equity risk premium.

Panel A: GARCH IVRMSE

Normal risk premium (%)	Total equity premium (%)					
	0.0	2.0	4.0	6.0	8.0	10.0
0.0	**7.01**					
2.0		**6.85**				
4.0			**6.70**			
6.0				**6.53**		
8.0					**6.36**	
10.0						**6.19**

Panel B: DVCJ over GARCH IVRMSE ratio

Normal risk premium (%)	Total equity premium (%)					
	0.0	2.0	4.0	6.0	8.0	10.0
0.0	**0.97**	**0.93**	**0.88**	**0.84**	**0.81**	**0.81**
2.0		**0.97**	0.93	0.88	0.84	0.82
4.0			**0.97**	0.93	0.88	0.85
6.0				**0.97**	0.93	0.89
8.0					**0.97**	0.93
10.0						**0.97**

Panel C: DVDJ over GARCH IVRMSE ratio

Normal risk premium (%)	Total equity premium (%)					
	0.0	2.0	4.0	6.0	8.0	10.0
0.0	**0.83**	**0.73**	**0.64**	**0.57**	**0.60**	**0.78**
2.0		**0.83**	0.72	0.63	0.57	0.63
4.0			**0.81**	0.71	0.62	0.59
6.0				**0.81**	0.70	0.61
8.0					**0.80**	0.69
10.0						**0.79**

Panel D: DVSDJ over GARCH IVRMSE ratio

Normal risk premium (%)	Total equity premium (%)					
	0.0	2.0	4.0	6.0	8.0	10.0
0.0	**0.79**	**0.67**	**0.57**	**0.55**	**0.68**	**0.99**
2.0		**0.79**	0.68	0.58	0.56	0.71
4.0			**0.80**	0.69	0.59	0.58
6.0				**0.80**	0.68	0.60
8.0					**0.81**	0.69
10.0						**0.81**

Table 5 allows us to address the impact of the overall equity premium and the impact of different combinations of the jump risk premium and the risk premium associated with the normal component. It yields some important conclusions. Compare the case when the equity premium is entirely due to non-jump risk, which corresponds to the diagonal entries, with the case when the equity premium is entirely due to jump risk, which corresponds to the bolded cells in the first rows of Panels B–D. Two important conclusions obtain. First, the jump models improve much more dramatically on the GARCH model when the entire premium is due to jump risk, except for the case of the DVSDJ model with a 10% equity premium. Second, when the premium is entirely due to jump risk, the proposed jump models first improve on the GARCH model as the total equity premium increases, but the performance of the DVDJ and DVSDJ models eventually deteriorates as the premium exceeds 6%. To ensure that our finding is not due to a specific choice of loss function, we have repeated this analysis using $RMSE instead of IVRMSE (not reported), and identical conclusions obtain.

The overall option pricing performance of the DVDJ and DVSDJ models is very good, and it is therefore of special interest to see if this performance is robust to the specification of risk premiums. Panels C and D of Table 5

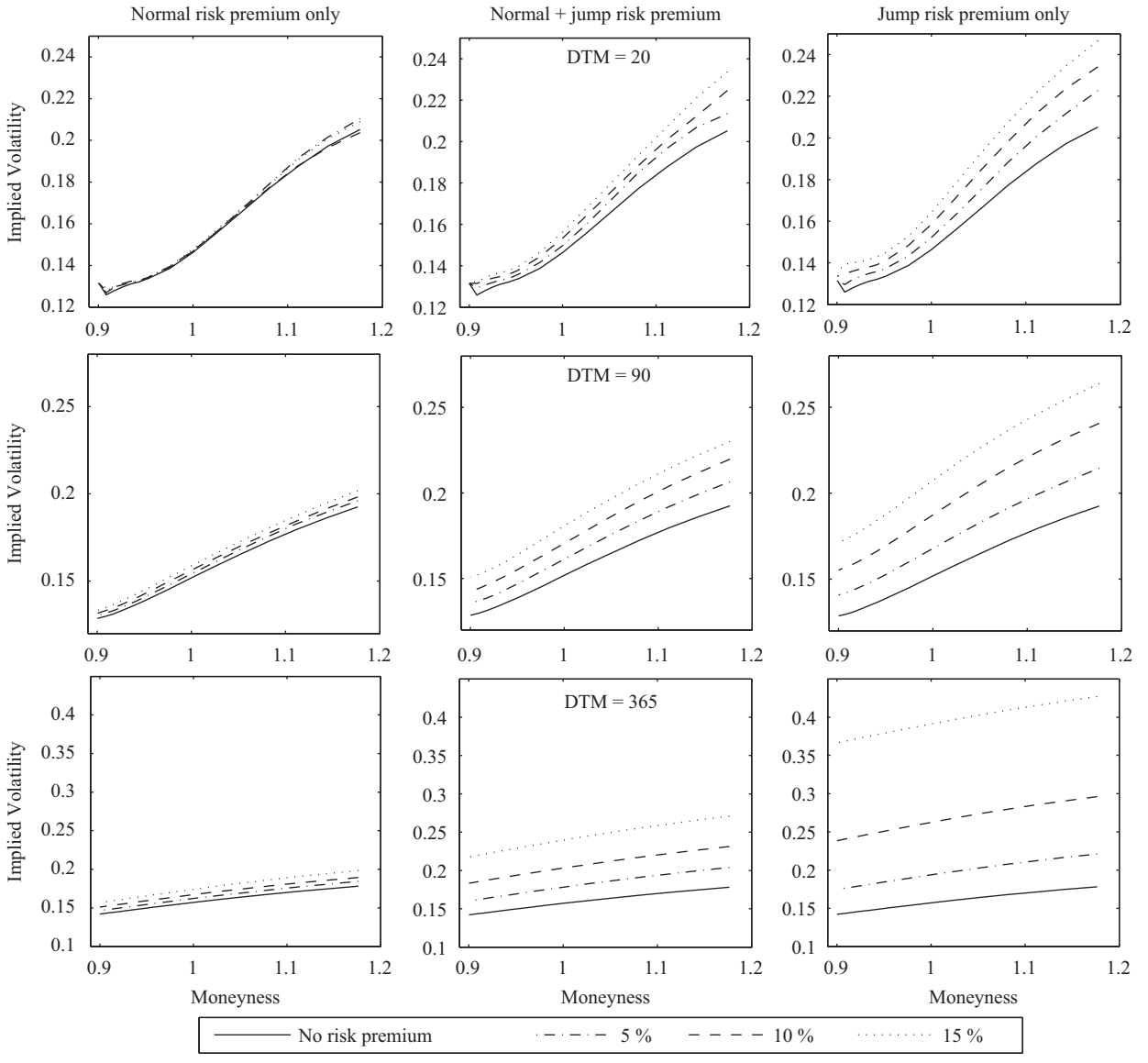

Fig. 7. The impact of risk premiums on the IV smirk in the DVDJ model. We plot implied volatility smirks in annualized terms at different days-to-maturity (DTM) and according to different levels and sources of risk premiums. In the left column, the equity premium stems entirely from normal risk. In the middle column, jump and normal risk premiums contribute equally to the risk premium. In the right column, the equity premium stems entirely from jump risk. The top, middle, and bottom rows correspond to the plots of implied volatility smirks at 20, 90, and 365 days-to-maturity, respectively. The conditional variance of the normal component and the conditional jump intensity are set equal to the model's implied long-run mean.

indicate that the DVDJ and DVSDJ models perform well compared to the GARCH model at all levels and combinations of the risk premiums. They perform particularly well when a large part of the equity risk premium is attributed to jump risk.

To further study the role of jump risk premiums, Fig. 7 presents implied volatility (IV) smirks generated by the DVDJ model at three different maturities, for different levels and sources of risk premiums. The left column presents results without a jump risk premium ($\lambda_y = 0$), and with total equity premiums of zero, 5%, 10%, and 15%. The right column presents results without a normal risk premium ($\lambda_z = 0$), and the middle column represents the mixed case in which each component delivers half the equity risk premium. The conditional jump intensity and conditional variance of the normal component are set to their long-run values.

The importance of the jump risk premium is clearly evident in Fig. 7. The shape and especially the level of the implied volatility smirks are highly sensitive to the jump risk premium, but not to the risk premium associated with the normal innovation. We have repeated this analysis with the other proposed jump models, and obtain identical conclusions (not reported). For options with 20 days to maturity, increasing the jump risk premium results in steeper slopes. For longer maturities, we see that high implied volatilities can be generated with a small jump risk premium. In the absence of jump risk, this is not possible.

6. Joint estimation using returns and options

The option valuation empirics in Section 5 were based on risk-neutralized MLE estimates from daily returns. This method has the advantage that it does not calibrate the models directly to option prices, thus highlighting the performance of well-specified parsimonious models. However, increasing importance has been placed in the literature on models' ability to reconcile the price of the options and the underlying returns, for example, in Chernov and Ghysels (2000), Pan (2002), Eraker (2004), and Santa-Clara and Yan (2010). To learn more about the models' performance along this dimension, we re-estimate the four jump models as well as the benchmark GARCH and the two-factor GARCH model using information from options and returns.

We estimate the model parameters using the joint MLE of index options quoted on each Wednesday from January 1996 to October 2009, as well as daily S&P500 index returns for 1995–2009. We use one year of daily returns prior to the option sample to initialize volatility. The log-likelihood for fitting daily index returns, $L_{returns}$, can be computed as in Section 2.7. However, we need to make further assumptions on the option pricing errors to write the log-likelihood for options fitting, $L_{options}$. We use the relative implied volatility loss function, and assume the following error structure:

$$u_j = \frac{\sigma_j^{BS} - IV(O_j, K_j, \tau_j, S_j, r_j)}{\sigma_j^{BS}}, \tag{31}$$

where σ_j^{BS} and IV are as defined in Section 5.2. We use the relative implied volatility loss function to avoid assigning excessive weighting to option data from the high volatility period 2008–2009. Fig. 1 shows that at-the-money implied volatility during 2008–2009 is up to three times higher than the observed level in 2006. We assume that the relative implied volatility error $u_j \sim Normal(0, \sigma_u^2)$ is uncorrelated with shocks in returns, which facilitates the construction of $L_{options}$ as

$$L_{options} = -\frac{N}{2}\ln(2\pi\sigma_u^2) - \frac{1}{2}\sum_{j=1}^{N}\frac{u_j^2}{\sigma_u^2}. \tag{32}$$

Option prices are available in the time-series as well as in the cross-section, thus, the number of data points available for computing $L_{options}$ is significantly larger than the number of data points available for computing $L_{returns}$. To ensure that the parameter estimates from the joint estimation are not dominated by the option prices, we assign an equal weight to each data point when computing the total log-likelihood

$$L_{joint} = \frac{T+N}{2}\frac{L_{returns}}{T} + \frac{T+N}{2}\frac{L_{options}}{N}, \tag{33}$$

where T is the number of days in the return time-series for 1995–2009, and N is the total number of option contracts. Other weighting schemes could of course equally well be used. Hu and Zidek (2002) study the properties of the weighted likelihood version of the standard likelihood and show that the key asymptotic results continue to hold.

The task of fitting jump models to option prices is more numerically demanding than estimating from returns, because option valuation under this class of models relies on Monte Carlo simulation. We alleviate some of the numerical challenges by focusing our joint estimation on option contracts with short to medium maturities. That is, we eliminate contracts with maturities of more than 250 days, which leads to a sample with 17,529 option contracts.

Table 6 presents joint maximum likelihood estimates for the four jump models as well as for the benchmark GARCH and the two-factor GARCH models. Note that the market prices of jump and normal risks are more precisely estimated in the joint MLE using options and returns, as compared to the MLE on returns in Table 1. This is not surprising as market incompleteness implies that options are non-redundant securities. Option prices therefore contain valuable information about the pricing kernel that is not available from the underlying return dynamics. Note that when estimating using options and returns, both market prices of risk are identified in the DVDJ model, as can be seen from (22) and (23).

Based on the log-likelihood values, jump models perform significantly better than the simple GARCH model and the two-factor GARCH model. The DVSDJ model performs very well compared to the other jump models. This is not necessarily surprising, because this model is more richly parameterized. For the other three jump models, which have the same number of parameters, the log-likelihood values show that the DVDJ model produces the best fit for both options and returns. Results

Table 6
Joint MLE estimates of selected models on S&P500 returns and options.

We apply MLE to jointly fit daily S&P500 returns from January 1995 to October 2009, and weekly out-of-the money S&P500 options with maturity of 250 days or less from January 1996 to October 2009. Columns labeled "Normal" contain estimates of parameters governing the normal component, and columns labeled "Jump" contain parameters governing the jump component. Reported in brackets are the standard errors computed using the outer product of the gradient. Under Properties we report the long-run equity risk premium contributions from the normal and jump components. The average annual volatility is reported in standard deviations. The "Total log-likelihood" row reports the joint log-likelihood values of each model. The total log-likelihood values are decomposed into the contributions from returns, "Returns log-likelihood", and from options, "Options log-likelihood".

	GARCH	DVCJ		DVDJ		DVSDJ		Two-factor GARCH	
Parameters	Normal	Normal	Jump	Normal	Jump	Normal	Jump	Normal-1	Normal-2
λ	5.78E+00	1.71E+00	7.67E−02	1.67E+00	1.87E−03	8.18E−01	2.50E−03	7.73E+00	6.15E−01
	(3.50E−01)	(1.2E−01)	(1.9E−02)	(4.6E−01)	(1.7E−07)	(3.8E−01)	(3.0E−04)	(2.0E−02)	(8.80E−03)
w	−2.57E−07	−8.00E−05	1.14E−03	−1.38E−04		−1.12E−04	−5.51E−02	−1.03E−06	1.81E−07
	(9.04E−09)	(3.13E−09)	(4.70E−08)	(1.24E−09)		(3.76E−06)	(5.28E−11)	(9.14E−10)	(6.431E−10)
b	8.46E−01	9.44E−01		9.53E−01		9.20E−01	9.71E−01	9.47E−01	1.10E−01
	(1.14E−03)	(4.24E−06)		(6.11E−06)		(1.90E−03)	(6.52E−06)	(3.91E−05)	(9.51E−05)
a	1.13E−06	2.78E−06		1.80E−06		2.34E−06	1.15E−03	1.12E−06	2.60E−07
	(9.41E−09)	(1.49E−09)		(2.25E−13)		(6.84E−08)	(2.43E−12)	(3.24E−10)	(3.30E−11)
c	3.58E+02	1.35E+02		1.23E+02		1.44E+02	2.71E−03	2.03E+02	1.80E+03
	(1.70E−01)	(5.11E−02)		(6.86E−01)		(1.43E−07)	(4.49E+00)	(9.51E−02)	(2.14E−01)
d		8.57E−03		9.28E−03		9.85E−03	4.66E+00		
		(4.79E−07)		(1.08E−09)		(3.35E−04)	(1.85E−02)		
e		9.53E−02		1.21E−01		1.06E−01	1.08E−01		
		(1.33E−06)		(1.08E−08)		(7.52E−08)	(2.15E−04)		
θ			−4.80E−02		−1.67E−02		−2.26E−02		
			(2.61E−03)		(4.85E−04)		(1.55E−06)		
δ			2.28E−02		1.03E−02		1.09E−02		
			(2.31E−03)		(5.79E−04)		(5.14E−04)		
k					3.63E+02				
					(2.48E+00)				
Properties									
Long-run risk premium	22.15%	6.48%	2.19%	5.67%	2.30%	2.50%	3.12%	22.30%	0.66%
Average annual volatility	19.57%	19.70%		19.64%		19.62%		19.88%	
Percent of annual var	100%	97.92%	2.08%	87.72%	12.28%	79.55%	20.45%	72.48%	27.52%
Total log-likelihood	36,985	38,725		38,861		39,455		37,902	
Options log-likelihood	3,500	4,966		5,006		5,521		4,215	
Returns log-likelihood	33,486	33,759		33,855		33,933		33,687	

for the CVDJ model indicated serious misspecification, and we therefore do not report its joint estimation results. Overall, the results from joint estimation point to the importance of time-varying jump intensities.

The row labeled "Long-run risk premium" in Table 6 decomposes the unconditional equity premium in each model into its jump and normal components. The unconditional jump and normal risk premiums are computed as the sample averages of $\lambda_y h_{y,t+1}$ and $\lambda_z h_{z,t+1}$, respectively. For instance, under the DVCJ model, we find a 6.48% normal risk premium and a 2.19% jump risk premium. We find that the jump risk premium is roughly similar across jump specifications. Table 6 also reports the average variance decomposition between the normal and jump components, which shows that the normal component in the DVCJ model captures 98% of the total return variance whereas it only captures 80% of the variance in the DVSDJ model in which jumps are dynamic.

Table 6 shows that the unconditional equity premium in the GARCH model is 22.15%, which is unrealistically large, indicating substantial misspecification. Similarly, the two-factor GARCH model yields an unrealistically large equity premium. The unconditional equity premium for the first and second factors of the two-factor GARCH model are 22.30% and 0.66%, respectively. This finding further demonstrates the importance of jump risk premiums for fitting option prices. Overall, the results from Table 6 once again confirm the importance of allowing for jumps with time-varying intensity.

7. Conclusion and directions for future work

This paper presents a new framework for modeling dynamic fat tails in the conditional distribution of returns. Our specification of jump models is inspired by a popular class of jump-diffusion models in the continuous-time literature. However, we specify the models using a discrete-time GARCH setting, and as a result, models with time-varying jump intensities and jumps in volatility can be easily estimated from long time-series of return data using maximum likelihood procedures. It is also possible to estimate these models using large option data sets, or using options and returns jointly. This enables us to analyze complex jump models, with features that are difficult to study in a continuous-time setup because of their computational complexity. Our general model encompasses four nested specifications which exhaust all possible sources of heteroskedasticity. The time-varying properties of the proposed jump models are driven by the dynamics of the variance of the normal shock and the dynamics of the jump intensity. Our model shares some features with Maheu and McCurdy (2004), who study equity returns but not options, and Duan, Ritchken, and Sun (2006) who study option pricing. However, our model has different implications for option valuation, because our assumptions enable us to characterize the risk-neutral dynamic with a separate risk premium for the jump and normal components.

Our empirical analysis on S&P500 index return data and option prices leads to important conclusions. We find that GARCH jump models should allow for heteroskedasticity both in the conditional variance of the normal innovation and the jump intensity. Although jumps can complement a heteroskedastic normal innovation by improving the modeling of the tails of the distribution, when restricted to at most one expected jump per day, they cannot fully replace the normal innovations. We also find that the intensity of jump arrivals varies through time, and depends on the level of risk in the market. Option data also indicate support for time-varying jump intensities, but we find that jumps are not useful in low volatility regimes. Therefore, future specifications should allow for frequent jump arrivals in high volatility regimes, and little or no possibility of jump arrivals in low volatility regimes.

Our option valuation results also demonstrate the importance of the jump risk premium. For reasonable levels of the equity premium, we find that the risk premium associated with the conditional normal innovation has little impact on the implied volatility term structure, and that realistic shapes of the implied volatility term structure can only be generated with a sizeable jump risk premium.

Our results can be extended in a number of ways. First, we could incorporate more general types of jumps. Although we analyze GARCH jumps of finite activity and finite variation, the framework that we introduce can be extended to incorporate the infinite-activity Lévy processes considered in the SVJ context by Huang and Wu (2004). Second, while the GARCH structure provides a convenient filter that allows us to analyze complex jump structures, it is somewhat different from the stochastic volatility models studied in many existing papers. A detailed investigation of the implications of these differences for the estimated jump structure would be valuable. Third, our jump approach to dynamic fat tails could be contrasted with the dynamic density models developed in Hansen (1994) and Huang and Tauchen (2005). Finally, applying the models to intraday returns data would be interesting and feasible given the computational advantages of our approach.

Appendix A. Filtering return innovations

The filtering process is extremely fast and precise because the filtering density for the ex post number of jumps as well as for the jump size are known analytically. We are grateful to David Bates for pointing out that analytical filtering of the jump component from the total return innovation is feasible in our class of models. Maheu and McCurdy (2004) provide an analytical filter for the number of jumps in their discrete-time model and we build on their work below.

We discuss the filtering of three time-series of unobservables: the number of jumps n_t, the normal component of the return z_t, the jump component of the return y_t. Given filtered time-series of these unobservables, we can determine the variance of the normal component $h_{z,t+1}$, and the jump intensity $h_{y,t+1}$. To distinguish the filtered observables from the non-filtered observables, we use $\tilde{n}_t$ to denote the filtered version of n_t, etc.

First, we describe the filtering density for the number of jumps at time t, n_t. Applying Bayes' rule, the filtering

density is given by

$$\Pr_t(n_t=j)\equiv\Pr_{t-1}(n_t=j|R_t)=\frac{f_{t-1}(R_t|n_t=j)\Pr_{t-1}(n_t=j)}{f_{t-1}(R_t)},\tag{34}$$

where the expressions on the right-hand side of (34) are given by (13), (14), and (15). $\Pr_t(n_t=j)$ represents the ex post inference on n_t, or the probability that j jumps have arrived between time $t-1$ and t conditional on the information available at time t. The filtered number of jumps is then given by

$$\tilde{n}_t=\sum_{j=0}^{\infty}j\Pr_t(n_t=j).\tag{35}$$

Next, we discuss the analytical filter for the normal component of the return z_t. The filtration of z_t involves solving the expectation $\tilde{z}_t=E_t[z_t]$. Notice that if the return R_t and the number of jumps $n_t=j$ at time t are known, we can express z_t as

$$z_t(R_t,n_t=j)=\sqrt{\frac{\tilde{h}_{z,t}}{\tilde{h}_{z,t}+j\delta^2}}(R_t-\mu_t-j\theta),\tag{36}$$

where $\tilde{h}_{z,t}$ denotes the filtered $h_{z,t}$, and recall that μ_t is the first conditional return moment given in (3). Since $z_t(R_t,n_t=j)$ depends on the discrete number of jumps $n_t=j$, the expectation $E_t[z_t]$ can be solved using the following summation:

$$\tilde{z}_t=E_t[z_t]=\sum_{j=0}^{\infty}z_t(R_t,n_t=j)\Pr_t(z_t,n_t=j),\tag{37}$$

where $\Pr_t(z_t,n_t=j)$ is the joint probability of z_t and $n_t=j$ given that R_t is known. Using Bayes' rule, we can write the filtering density $\Pr_t(z_t,n_t=j)$ up to the normalizing constant as

$$\begin{aligned}\Pr_t(z_t,n_t=j)&\equiv\Pr_{t-1}(z_t,n_t=j|R_t)\\&\propto\Pr_{t-1}(z_t|R_t,n_t=j)\Pr_t(n_t=j).\end{aligned}\tag{38}$$

This represents the ex post inference on z_t given time t information. The second term on the right-hand side of (38) is given by (34), while the first term is the probability of z_t given that R_t and $n_t=j$ are known. To compute the first term on the right-hand side, we express z_t as $z_t(R_t,n_t=j)$ following (36), and use the change of variable technique. This allows us to write

$$\Pr_{t-1}(z_t|R_t,n_t=j)=\Pr_{t-1}(R_t,n_t=j|R_t,n_t=j)\sqrt{\frac{\tilde{h}_{z,t}}{\tilde{h}_{z,t}+j\delta^2}},\tag{39}$$

where the first term on the right-hand side of (39) is trivially one. Using (36), (37), (38), and (39), we can write the expected ex post normal component of the return as

$$\begin{aligned}\tilde{z}_t&=\sum_{j=0}^{\infty}z_t(R_t,n_t=j)\Pr_{t-1}(z_t|\ R_t,n_t=j)\Pr_t(n_t=j)\\&=\sum_{j=0}^{\infty}\frac{\tilde{h}_{z,t}}{\tilde{h}_{z,t}+j\delta^2}(R_t-\mu_t-j\theta)\Pr_t(n_t=j).\end{aligned}\tag{40}$$

Once $\tilde{z}_t$ is known, we can directly infer $\tilde{y}_t$ given time t information by noting that from (1), $y_t=R_t-\mu_t-z_t$. Therefore, the filtered jump innovation is given by

$$\tilde{y}_t=R_t-\mu_t-\tilde{z}_t.\tag{41}$$

After we have obtained $\tilde{z}_t$ and $\tilde{y}_t$, the filtered variance, $\tilde{h}_{z,t+1}$, and the jump intensity, $\tilde{h}_{y,t+1}$ can be computed according to

$$\begin{aligned}\tilde{h}_{z,t+1}&=w_z+b_z\tilde{h}_{z,t}+\frac{a_z}{\tilde{h}_{z,t}}(\tilde{z}_t-c_z\tilde{h}_{z,t})^2+d_z(\tilde{y}_t-e_z)^2,\\\tilde{h}_{y,t+1}&=w_y+b_y\tilde{h}_{y,t}+\frac{a_y}{\tilde{h}_{z,t}}(\tilde{z}_t-c_y\tilde{h}_{z,t})^2+d_y(\tilde{y}_t-e_y)^2.\end{aligned}\tag{42}$$

Note that the above equations are identical to the GARCH dynamics (8)–(9) but with filtered innovations.

At time $t=0$, we need to make an assumption on $\tilde{h}_{z,0}$ and $\tilde{h}_{y,0}$ to propagate the GARCH dynamics into the next period. We assume that $\tilde{h}_{z,0}$ and $\tilde{h}_{y,0}$ are equal to the mean of their filtered time-series from the last optimization iteration.

Once we have filtered $\tilde{h}_{z,t+1}$ and $\tilde{h}_{y,t+1}$, we can evaluate the log-likelihood built from the conditional distribution in (15), namely,

$$L_{returns}=\sum_{t=1}^{T-1}\ln(f_t(R_{t+1}))=\sum_{t=1}^{T-1}\ln\left(\sum_{j=0}^{\infty}f_t(R_{t+1}|n_{t+1}=j)\Pr_t(n_{t+1}=j)\right),\tag{43}$$

where

$$f_t(R_{t+1}|n_{t+1}=j)=\frac{1}{\sqrt{2\pi(\tilde{h}_{z,t+1}+j\delta^2)}}\exp\left(-\frac{(R_{t+1}-\mu_{t+1}-j\theta)^2}{2(\tilde{h}_{z,t+1}+j\delta^2)}\right)\tag{44}$$

and

$$\Pr_t(n_{t+1}=j)=\frac{(\tilde{h}_{y,t+1})^j}{j!}\exp(-\tilde{h}_{y,t+1}).\tag{45}$$

Appendix B. Proof of Proposition 1

For an EMM to exist, the expected return of S_t from time t to $t+1$ must equal the risk free rate

$$E_t\left[\left(\frac{\frac{dQ_{t+1}}{dP_{t+1}}}{\frac{dQ_t}{dP_t}}\right)\frac{S_{t+1}}{S_t}\right]=e^r.\tag{46}$$

Substituting the Radon–Nikodym derivative in (21) and the return dynamic in (1), and taking expectations, we have

$$\frac{e^{r+(\lambda_z-1/2)h_{z,t+1}+(\lambda_y-\xi)h_{y,t+1}}E_t[\exp((\Lambda_z+1)z_{t+1}+(\Lambda_y+1)y_{t+1})]}{E_t[\exp(\Lambda_z z_{t+1}+\Lambda_y y_{t+1})]}=e^r\tag{47}$$

and taking logs yields

$$\log\frac{E_t[\exp((\Lambda_z+1)z_{t+1}+(\Lambda_y+1)y_{t+1})]}{E_t[\exp(\Lambda_z z_{t+1}+\Lambda_y y_{t+1})]}$$

$$+\left(\lambda_z-\frac{1}{2}\right)h_{z,t+1}+(\lambda_y-\xi)h_{y,t+1}=0, \tag{48}$$

where $\xi=(e^{(\delta^2/2)+\theta}-1)$.

To solve the conditional expectation above, we will make use of the joint moment generating function (MGF) of the return innovation. Because we assume that the normal and jump components are contemporaneously independent, the conditional joint MGF of $z_{t+1}+y_{t+1}$ can be written as the product of their moment generating functions

$$E_t[\exp(\phi_z z_{t+1}+\phi_y y_{t+1})]\equiv\exp(\Psi_z(\phi_z;h_{z,t+1})+\Psi_y(\phi_y;h_{y,t+1})). \tag{49}$$

The expression $\Psi_z(\phi_z;h_{z,t+1})=\frac{1}{2}\phi_z^2 h_{z,t+1}$ is the exponent of the normal MGF with zero mean and variance $h_{z,t+1}$. For the compound Poisson process with jump intensity $h_{y,t+1}$, jump mean size θ, and jump variance δ^2, the exponent of its MGF is given by

$$\Psi_y(\phi_y;h_{y,t+1})=h_{y,t+1}(\exp(\phi_y\theta+\tfrac{1}{2}\phi_y^2\delta^2)-1). \tag{50}$$

Substituting

$$E_t[\exp(\Lambda_z z_{t+1}+\Lambda_y y_{t+1})]=\exp(\Psi_z(\Lambda_z;h_{z,t+1})+\Psi_y(\Lambda_y;h_{y,t+1})) \tag{51}$$

and

$$E_t[\exp((\Lambda_z+1)z_{t+1}+(\Lambda_y+1)y_{t+1})]$$
$$=\exp(\Psi_z(\Lambda_z+1;h_{z,t+1})+\Psi_y(\Lambda_y+1;h_{y,t+1})) \tag{52}$$

in the EMM restriction (48) gives, after simplification and collecting terms,

$$h_{y,t+1}(\lambda_y-(e^{(\delta^2/2)+\theta}-1)-e^{(\Lambda_y^2\delta^2)/2+\Lambda_y\theta}(1-e^{(1/2+\Lambda_y)\delta^2+\theta}))$$
$$+h_{z,t+1}(\lambda_z+\Lambda_z)=0. \tag{53}$$

This can be solved by equating the coefficients on $h_{y,t+1}$ and $h_{z,t+1}$ to zero, which gives (22) and (23). □

Appendix C. Proof of Proposition 2

We will prove this using the moment generating function. First, we denote the MGF of z_t+y_t under the risk-neutral measure Q,

$$E_t^Q[\exp(\phi(z_{t+1}+y_{t+1}))]=\exp(\Psi_z(\phi;h_{z,t+1})+\Psi_y(\phi;h_{y,t+1})), \tag{54}$$

where $\Psi_z^Q(\phi;h_{z,t+1})$ and $\Psi_y^Q(\phi;h_{y,t+1})$ are the exponent of the Q-measure MGF for the normal and jump components, respectively. Next, we apply the change of measure through the Radon–Nikodym derivative in (21). Subsequently, we retrieve the stochastic process that will yield such a form of the MGF. Using the notation in (49), the change of measure to (54) gives

$$E_t^Q[\exp(\phi(z_{t+1}+y_{t+1}))]=E_t\left[\left(\frac{\frac{dQ_{t+1}}{dP_{t+1}}}{\frac{dQ_t}{dP_t}}\right)\exp(\phi(z_{t+1}+y_{t+1}))\right]$$
$$=E_t\left[\frac{\exp((\phi+\Lambda_z)z_{t+1}+(\phi+\Lambda_y)y_{t+1})}{E_t[\exp(\Lambda_z z_{t+1}+\Lambda_y y_{t+1})]}\right], \tag{55}$$

where the conditional expectation on the right-hand side is now under the physical measure. Noting that $E_t[\exp(\Lambda_z z_{t+1}+\Lambda_y y_{t+1})]=\exp(\Psi_z(\Lambda_z;h_{z,t+1})+\Psi_y(\Lambda_y;h_{y,t+1}))$ is known at time t, and that z_{t+1} and y_{t+1} are conditionally independent, this gives

$$E_t^Q[\exp(\phi(z_{t+1}+y_{t+1}))]=\exp(A_{z,t+1}+A_{y,t+1}), \tag{56}$$

where

$$A_{z,t+1}=\Psi_z(\phi+\Lambda_z;h_{z,t+1})-\Psi_z(\Lambda_z;h_{z,t+1}),$$

$$A_{y,t+1}=\Psi_y(\phi+\Lambda_y;h_{y,t+1})-\Psi_y(\Lambda_y;h_{y,t+1}). \tag{57}$$

It can be seen that $A_{z,t+1}$ is the exponent of the normal MGF with mean $\Lambda_z h_{z,t+1}$ and variance $h_{z,t+1}$:

$$A_{z,t+1}=\phi\Lambda_z h_{z,t+1}+\tfrac{1}{2}\phi^2 h_{z,t+1}. \tag{58}$$

We denote the normal component of the risk-neutral measure by $z_{t+1}^*\sim N(\Lambda_z h_{z,t+1},h_{z,t+1})$. Similarly, rearranging the expression for $A_{y,t+1}$ yields

$$A_{y,t+1}=h_{y,t+1}\exp\left(\frac{\Lambda_y^2\delta^2}{2}+\Lambda_y\theta\right)(\exp^{\phi(\theta+\Lambda_y\delta^2)+(\delta^2\phi^2)/2}-1), \tag{59}$$

which is the exponent of the compound Poisson MGF with jump mean size $\theta^*=\theta+\Lambda_y\delta^2$, and jump intensity

$$h_{y,t+1}^*=h_{y,t+1}\exp\left(\frac{\Lambda_y^2\delta^2}{2}+\Lambda_y\theta\right).$$

We denote the jump component in the risk-neutral measure by $y_t^*\sim J(h_{y,t+1}^*,\theta^*,\delta^2)$. The proof is complete as we have shown that $E_t^Q[\exp(\phi(z_{t+1}+y_{t+1}))]$ is the moment generating function of the stochastic process $z_{t+1}^*+y_{t+1}^*$. □

Appendix D. Proof of Proposition 3

Using the result from Proposition 2, we see that under the risk-neutral Q measure, the return process in (1) can be written as

$$\log\left(\frac{S_{t+1}}{S_t}\right)=r+\left(\lambda_z-\frac{1}{2}\right)h_{z,t+1}+(\lambda_y-\xi)h_{y,t+1}+z_{t+1}^*+y_{t+1}^*. \tag{60}$$

The risk-neutral dynamics for (8)–(9) can be written as

$$h_{z,t+1}=w_z+b_z h_{z,t}+\frac{a_z}{h_{z,t}}(z_t^*-c_z h_{z,t})^2+d_z(y_t^*-e_z)^2,$$

$$h_{y,t+1}=w_y+b_y h_{y,t}+\frac{a_y}{h_{z,t}}(z_t^*-c_y h_{z,t})^2+d_y(y_t^*-e_y)^2. \tag{61}$$

Note that the risk-neutral distributions of the two shocks are $z_{t+1}^*\sim N(\Lambda_z h_{z,t+1},h_{z,t+1})$ and $y_{t+1}^*\sim J(h_{y,t+1}^*,\theta^*,\delta^2)$. The convention in the GARCH literature is to express the normal shock as a mean-zero innovation. We therefore use the simple transformation $z_{t+1}^*=z_{t+1}+\Lambda_z h_{z,t+1}$. Recalling that the analytical solution to Λ_z is $-\lambda_z$, we can write (60) and (61) as

$$\log\left(\frac{S_{t+1}}{S_t}\right)=r-\frac{1}{2}h_{z,t+1}+(\lambda_y-\xi)h_{y,t+1}+z_{t+1}+y_{t+1}^* \tag{62}$$

and

$$h_{z,t+1}=w_z+b_z h_{z,t}+\frac{a_z}{h_{z,t}}(z_t-(c_z+\lambda_z)h_{z,t})^2+d_z(y_t^*-e_z)^2,$$

$$h_{y,t+1} = w_y + b_y h_{y,t} + \frac{a_y}{h_{z,t}}(z_t^* - (c_y + \lambda_z)h_{z,t})^2 + d_y(y_t^* - e_y)^2. \tag{63}$$

Because the Q measure is constructed such that the discounted price process of S_t in (62) is a martingale, and we already know that $E_t[\exp(z_{t+1} - \frac{1}{2}h_{z,t+1})]$ is a martingale, we must also have

$$\exp((\lambda_y - \xi)h_{y,t+1}) = E_t[\exp(-y_{t+1}^*)] = \xi^* h_{y,t+1}^*, \tag{64}$$

where $\xi^* = (e^{(\delta^2/2)+\theta^*} - 1)$ and $h_{y,t+1}^*$ is the risk-neutral jump intensity. Consequently, the return dynamic (62) under the risk-neutral measure simplifies to

$$\log\left(\frac{S_{t+1}}{S_t}\right) = r - \frac{1}{2}h_{z,t+1} - \xi^* h_{y,t+1}^* + z_{t+1} + y_{t+1}^*. \tag{65}$$

Next, we further reparameterize the GARCH dynamics in (63) using the following transformations:

$$h_{y,t+1}^* = h_{y,t+1}\Pi, \quad w_y^* = w_y\Pi, \quad a_y^* = a_y\Pi,$$

$$c_y^* = c_y + \lambda_z, \quad c_z^* = c_z + \lambda_z, \quad d_y^* = d_y\Pi \tag{66}$$

for $\Pi = \exp((\Lambda_y^2\delta^2)/2 + \Lambda_y\theta)$, which leaves us with the following Q-dynamics:

$$h_{z,t+1} = w_z + b_z h_{z,t} + \frac{a_z}{h_{z,t}}(z_t - c_z^* h_{z,t})^2 + d_z(y_t^* - e_z)^2,$$

$$h_{y,t+1}^* = w_y^* + b_y h_{y,t}^* + \frac{a_y^*}{h_{z,t}}(z_t - c_y^* h_{z,t})^2 + d_y^*(y_t^* - e_y)^2. \quad \square \tag{67}$$

Appendix E. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org.10.1016/j.jfineco.2012.05.017.

References

Andersen, T., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. Journal of Finance 57, 1239–1284.

Andersen, T., Bollerslev, T., Diebold, F., 2007. Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility. Review of Economics and Statistics 89, 701–720.

Bakshi, C., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52, 2003–2049.

Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. Review of Financial Studies 16, 101–143.

Barone-Adesi, G., Engle, R., Mancini, L., 2008. A GARCH option pricing model with filtered historical simulation. Review of Financial Studies 21, 1223–1258.

Bates, D., 1991. The crash of '87: was it expected? The evidence from options markets. Journal of Finance 46, 1009–1044.

Bates, D., 2000. Post-87 crash fears in S&P500 futures options. Journal of Econometrics 94, 181–238.

Bates, D., 2006. Maximum likelihood estimation of latent affine processes. Review of Financial Studies 19, 909–965.

Bates, D., 2008. The market for crash risk. Journal of Economic Dynamics and Control 32, 2291–2321.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–659.

Bollerslev, T., Todorov, V., 2011. Tails, fears and risk premia. Journal of Finance 66, 2165–2211.

Brennan, M., 1979. The pricing of contingent claims in discrete-time models. Journal of Finance 34, 53–68.

Broadie, M., Chernov, M., Johannes, M., 2007. Model specification and risk premia: evidence from futures options. Journal of Finance 62, 1453–1490.

Buhlmann, H., Delbaen, F., Embrechts, P., Shiryaev, A., 1998. On Esscher transforms in discrete finance models. ASTIN Bulletin 28, 171–186.

Carr, P., Wu, L., 2004. Time-changed Lévy processes and option pricing. Journal of Financial Economics 17, 113–141.

Chernov, M., Gallant, R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. Journal of Econometrics 116, 225–257.

Chernov, M., Ghysels, E., 2000. A study towards a unified approach to the joint estimation of objective and risk-neutral measures for the purpose of option valuation. Journal of Financial Economics 56, 407–458.

Christoffersen, P., Elkamhi, R., Feunou, B., Jacobs, K., 2010. Option valuation with conditional heteroskedasticity and nonnormality. Review of Financial Studies 23, 2139–2183.

Christoffersen, P., Jacobs, K., Ornthanalai, C., Wang, Y., 2008. Option valuation with long-run and short-run volatility components. Journal of Financial Economics 90, 272–297.

Dai, Q., Le, A., Singleton, K., 2010. Discrete-time affine Q term structure models with generalized market prices of risks. Review of Financial Studies 23, 2184–2227.

Duan, J.C., 1995. The GARCH option pricing model. Mathematical Finance 5, 13–32.

Duan, J.C., Ritchken, P., Sun, Z., 2006. Approximating GARCH-jump models, jump-diffusion processes, and option pricing. Mathematical Finance 16, 21–52.

Duan, J.C., Ritchken, P., Sun, Z., 2007. Jump starting GARCH: pricing and hedging options with jumps in returns and volatilities. Unpublished working paper, National University of Singapore, Case Western Reserve University, and National City Bank.

Duan, J.C., Simonato, J.G., 1998. Empirical martingale simulation for asset prices. Management Science 44, 1218–1233.

Dumas, B., Fleming, J., Whaley, R., 1998. Implied volatility functions: empirical tests. Journal of Finance 53, 2059–2106.

Engle, R., Ng, V., 1993. Measuring and testing the impact of news on volatility. Journal of Finance 48, 1749–1778.

Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. Journal of Finance 59, 1367–1403.

Eraker, B., 2008. Affine general equilibrium models. Management Science 54, 2068–2080.

Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of Finance 58, 1269–1300.

Esscher, F., 1932. On the probability function in the collective theory of risk. Skandinavisk Aktuarietidskrift 15, 175–195.

Fleming, J., Kirby, C., 2003. A closer look at the relation between GARCH and stochastic autoregressive volatility. Journal of Financial Econometrics 1, 365–419.

Gerber, H., Shiu, E., 1994. Option pricing by Esscher transforms. Transactions of the Society of Actuaries 46, 99–140.

Hansen, B., 1994. Autoregressive conditional density estimation. International Economic Review 35, 705–730.

Harvey, C., Siddique, A., 1999. Autoregressive conditional skewness. Journal of Financial and Quantitative Analysis 34, 465–487.

Harvey, C., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 1263–1295.

Hentschel, L., 1995. All in the family: nesting symmetric and asymmetric GARCH models. Journal of Financial Economics 39, 71–104.

Heston, S., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6, 327–343.

Heston, S., Nandi, S., 2000. A closed-form GARCH option pricing model. Review of Financial Studies 13, 585–626.

Ho, M., Perraudin, W., Sorensen, B., 1996. A continuous-time arbitrage-pricing model with stochastic volatility and jumps. Journal of Business and Economic Statistics 14, 31–43.

Hu, F., Zidek, J.V., 2002. The weighted likelihood. Canadian Journal of Statistics 30, 347–371.

Huang, J.Z., Wu, L., 2004. Specification analysis of option pricing models based on time-changed Lévy processes. Journal of Finance 59, 1405–1439.

Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. Journal of Financial Econometrics 3, 456–499.

Jondeau, E., Rockinger, M., 2003. Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. Journal of Economic Dynamics and Control 27, 1699–1737.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. Review of Economic Studies 65, 361–394.

Li, H., Wells, M., Yu, C., 2007. A Bayesian analysis of return dynamics with Lévy jumps. Review of Financial Studies 21, 2345–2378.
Li, H., Wells, M., Yu, C., 2011. MCMC estimation of Lévy jump models using stock and option prices. Mathematical Finance 21, 383–422.
Liu, J., Pan, J., Wang, T., 2005. An equilibrium model of rare-event premia and its implication for option smirks. Review of Financial Studies 18, 131–164.
Maheu, J., McCurdy, T., 2004. News arrival, jump dynamics and volatility components for individual stock returns. Journal of Finance 59, 755–793.
Maheu, J., McCurdy, T., 2008. Modeling foreign exchange rates with jumps. Frontiers of Economics and Globalization 3, 449–475.
Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144.
Naik, V., Lee, M., 1990. General equilibrium pricing of options on the market portfolio with discontinuous returns. Review of Financial Studies 3, 493–521.
Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. Journal of Financial Economics 63, 3–50.
Ritchken, P., Trevor, R., 1999. Pricing options under generalized GARCH and stochastic volatility processes. Journal of Finance 54, 377–402.
Santa-Clara, P., Yan, S., 2010. Crashes volatility and the equity premium: lessons from S&P500 options. Review of Economics and Statistics 92, 435–451.