MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Inference on Self-Exciting Jumps in Prices and Volatility using High Frequency Measures

Worapree Maneesoonthorn, Catherine S. Forbes and Gael M. Martin

December 2014

Inference on Self-Exciting Jumps in Prices and Volatility using High Frequency Measures

Worapree Maneesoonthorn,[*] Catherine S. Forbes[†]and Gael M. Martin[‡]

December 21, 2014

Abstract

Dynamic jumps in the price and volatility of an asset are modelled using a joint Hawkes process in conjunction with a bivariate jump diffusion. A state space representation is used to link observed returns, plus nonparametric measures of integrated volatility and price jumps, to the specified model components; with Bayesian inference conducted using a Markov chain Monte Carlo algorithm. The calculation of marginal likelihoods for the proposed and related models is discussed. An extensive empirical investigation is undertaken using the S&P500 market index, with substantial support for dynamic jump intensities - including in terms of predictive accuracy - documented.

Keywords: *Dynamic price and volatility jumps; Stochastic volatility; Hawkes process; Nonlinear state space model; Bayesian Markov chain Monte Carlo; Global financial crisis. JEL Classifications: C11, C58, G01.*

[*]Melbourne Business School, The University of Melbourne.

[†]Department of Econometrics and Business Statistics, Monash University.

[‡]Corresponding author: gael.martin@monash.edu. Department of Econometrics and Business Statistics, Monash University.

1

1 Introduction

Planning for unexpected and large movements in asset prices is central to the management of financial risk. Key to this planning is the ability to distinguish extreme price changes arising from a persistent shift in the asset's underlying volatility from idiosyncratic movements that occur due to random shocks in the market environment. Making this task more difficult is the fact that volatility itself exhibits discontinuous behaviour which, via the stylized occurrence of feedback from volatility to current and future returns (e.g. Bollerslev, Sizova and Tauchen, 2012), has the potential to cause seemingly discontinuous behaviour in the asset price. Moreover, it is unclear whether the apparent clustering behaviour of asset price jumps during times of market turbulence is evidence of dynamics in the jump *intensity* of either process (or both), or simply a result of the propagation through time of (independent) variance jumps due to persistence in the level of volatility.

Traditionally, parametric jump diffusion models have been used to capture the discontinuous behaviour in prices and, potentially, in their underlying volatility. Notable in this literature are the studies of Bates (2000) and Pan (2002), which propose models that characterize the intensity of a jump in price as proportional to the level of the underlying (diffusive) variance. In these models, the (price) jump intensity will be high in periods with high volatility and dependent over time as a consequence of the dynamic specification adopted for volatility itself. Duffie, Pan and Singleton (2000), on the other hand, introduce a model with both price and volatility jumps, where the two jump processes always occur contemporaneously. Under this specification, large fluctuations in price would tend to occur in the successive periods following a contemporaneous jump in price and volatility due to persistence in the volatility process. This impact is further strengthened by the fact that the average magnitude of price jumps is (positively) dependent on the average magnitude of variance jumps. See Eraker, Johannes and Polson (2003) and Broadie, Chernov and Johannes (2007) for related specifications; and also Eraker (2004), who proposes a model that specifies contemporaneous price and variance jumps as well as allowing the (price) jump intensity to depend on the volatility level.

More recently, nonparametric methods, based on volatility and jump measures constructed from high frequency data, have been used to investigate the nature of price and/or variance jumps. For example, the empirical findings of Todorov and Tauchen (2011) indicate the presence of jumps in volatility, whilst those of Jacod and Todorov (2010) provide evidence of both price and variance jumps, with less than 50% of those jumps occurring

simultaneously for the S&P500 market index. Most notably, Tauchen and Zhou (2011) conclude in favour of a *dynamic* price jump process, and further link the clusters of price jumps to changes in the US credit market. The dynamic behaviour of price jump processes has also been investigated by Aït-Sahalia, Cacho-Diaz and Laeven (2014), who introduce the Hawkes process (Hawkes, 1971a,b) to model price jumps (only) in multivariate asset prices subject to stochastic volatility. These authors analyse the dynamic transmission, or 'contagion', of price jumps across equity markets in several countries, concluding that price jump intensities increase dramatically during a crisis period, across all connected markets. See also Liao, Anderson and Vahid (2010) for an investigation of common factors in price jumps across assets in the Chinese stock market.

In the spirit of this more recent empirical work, we propose a general model in which both the price and volatility of a single asset are permitted to jump, with the two jump processes being dynamic. Adopting a bivariate Hawkes process for the intensity process associated with jumps in price and its latent variance, jumps in both price and variance (accordingly volatility, defined as the square root of the variance) are (potentially) self-exciting; that is, the intensity of each jump process is functionally dependent on the realised past increments of that process. We allow the variance jump intensity to depend also on past price jumps, in an attempt to capture any potential transmission of information in extreme price movements to extreme movements in its underlying volatility. Possible leverage effects operating at the level of extreme price and volatility movements are accommodated via the modeling of the differential impacts of negative and positive price jumps on the variance jump intensity.

A multivariate nonlinear state space framework, based on a discrete time representation of the proposed model, is specified. Volatility and price jump measures constructed from high frequency data, in addition to the daily return measure, are used to define the multiple measurement equations. A Markov chain Monte Carlo (MCMC) methodology is then developed to undertake a Bayesian analysis of the model. The analysis includes the calculation of marginal likelihoods for evaluating the proposed specification against alternatives from the literature, several of which are nested within our general model; plus the determination of predictive distributions for out-of-sample assessments.

Application of the methodology using measures constructed from intraday data from the S&P500 index for the period January 1996 to May 2012 is documented in detail. Marginal posterior inference for the general model is broadly consistent with certain comparable results from the existing literature. However, differences are noted, including the fact that the inclusion of a dynamic process for volatility jump intensity reduces estimated persistence in

3

the diffusive component of the volatility model. We also explore the use of the estimated dynamic jump intensities as a barometer of financial crises by reviewing their behaviour during a five year period encompassing the global financial crisis. Empirical results indicate that the variance jump intensity is much more informative, in this sense, than the price jump intensity, with the former closely tracking key events over this period, and exhibiting its most dramatic increase at the peak of the global financial crisis in late 2008.

The benefits of including a flexible dynamic specification for price and variance jumps are investigated via an out-of-sample forecasting exercise over May 2012 to June 2014; a period that includes market instabilities induced by the US 'fiscal cliff' and the subsequent debate surrounding its remedies. The superiority of the general model's predictive accuracy is evident in its (relatively) high cumulative predictive log density score over the evaluation period, as well as in the accuracy of the empirical coverage of 95% prediction intervals for the daily settlement price of S&P500 index futures. Support for the proposed model is also found in the closer match between the empirical coverage of extreme quantiles of returns and the nominal coverage, when compared to alternative models. A comparison of the predictive higher-order moments of returns over the out-of-sample period also highlights the different tail behaviour invoked by the alternative jump specifications. An understanding of these properties is, in turn, crucial for achieving effective hedging strategies and for constructing optimal portfolios; see, for example, Cvitanic, Polimenis and Zapatero (2008) and Harvey, Liechty, Liechty and Müller (2010). Whilst it is beyond the scope of this paper to investigate such financial applications, we believe that documenting the link between the structure of this particular type of dynamic model and the properties of returns, provides an important foundation for future investigations of this type.

The remainder of the paper is organized as follows. Section 2 describes our proposed asset price model and its main properties. It then further outlines the state space representation used to analyze the proposed model, followed by a discussion of the relevant models to be investigated. Details of the Bayesian inferential methodology are then given in Section 3. Results from the empirical analysis of the S&P500 index are presented and discussed in Section 4, and Section 5 provides some conclusions. Certain technical results, including algorithmic details, are included in appendices to the paper.

2 An asset price process with stochastic volatility and self-exciting jumps

2.1 The continuous time representation

Let $p_t = \ln(P_t)$ be the natural log of the asset price, P_t at time $t > 0$, whose evolution over time is described by the following bivariate jump diffusion process,

$$dp_t = (\mu + \gamma V_t)\, dt + \sqrt{V_t}\, dB_t^p + dJ_t^p \tag{1}$$

$$dV_t = \kappa\,(\theta - V_t)\, dt + \sigma_v \sqrt{V_t}\, dB_t^v + dJ_t^v, \tag{2}$$

with B_t^p and B_t^v denoting standard Brownian motion processes, $corr(dB_t^p, dB_t^v) = \rho dt$ and $dJ_t^i = Z_t^i dN_t^i$, for $i = \{p, v\}$. Without the discontinuous sample paths dJ_t^p and dJ_t^v this form of asset pricing process replicates that of the Heston (1993) square root stochastic volatility model, where the parameter restriction $\sigma_v^2 \leq 2\mu\theta$ ensures the positivity of the variance process, denoted by V_t, for $t > 0$. The drift component of (1) contains the additional component γV_t, allowing for a volatility feedback effect (that is, the impact of volatility on future returns) to be captured, while $corr(dB_t^p, dB_t^v) = \rho dt$ in (2) captures the leverage effect (that is, the impact of (negative) returns on future volatility). See Bollerslev, Livitnova and Tauchen (2006) who also propose a model that separates volatility feedback from leverage effects. The J_t^i, $i = \{p, v\}$, are dependent random jump processes that permit occasional jumps in either p_t or V_t, or both, and have random sizes Z_t^p and Z_t^v, respectively.

A novel contribution of this paper is the specification of a bivariate Hawkes process for the point processes, N_t^i, $i = \{p, v\}$, which feeds into the bivariate jump process, J_t^i, $i = \{p, v\}$. Specifically, we assume that

$$\Pr(dN_t^p = 1) = \delta_t^p dt + o(dt), \quad \text{with} \tag{3}$$

$$d\delta_t^p = \alpha_p\,(\delta_\infty^p - \delta_t^p)\, dt + \beta_{pp} dN_t^p, \tag{4}$$

and that

$$\Pr(dN_t^v = 1) = \delta_t^v dt + o(dt), \quad \text{with} \tag{5}$$

$$d\delta_t^v = \alpha_v\,(\delta_\infty^v - \delta_t^v)\, dt + \beta_{vv} dN_t^v + \beta_{vp} dN_t^p + \beta_{vp}^{(-)} dN_t^{p(-)}, \tag{6}$$

where $dN_t^{p(-)} = dN_t^p \mathbf{1}(Z_t^p < 0)$ denotes the occurrence of a *negative* price jump, corresponding to a value of one for the indicator function $\mathbf{1}(\cdot)$. Due to the inclusion of the terms dN_t^p and $dN_t^{p(-)}$ in (6), the process dN_t^v defined by (5) is not only 'self-exciting', but is

also excited by a concurrent price jump. The additional threshold component, $\beta_{vp}^{(-)} dN_t^{p(-)}$, allows a contemporaneous negative price jump to have an impact on $d\delta_t^v$, thereby serving as an additional channel for leverage, over and above the non-zero correlation between the Brownian motion increments, dB_t^p and dB_t^v. See Hawkes (1971a,b) for seminal discussions regarding self-exciting point processes, and Aït-Sahalia *et al.* (2014) for the introduction of the Hawkes process into asset pricing models.

Our proposed specification can be viewed as a natural extension of the various models in the literature that accommodate both stochastic volatility and jumps. Most notably we relax the strict assumptions of the well-known stochastic volatility with contemporaneous jumps (SVCJ) model (Duffie *et al.*, 2000) to allow jumps in both the price and volatility to be governed by separate, though dependent, dynamic random processes. The specification can also be viewed as an extension of the stochastic volatility model of Aït-Sahalia *et al.* (2014), in which a Hawkes process is used to characterize multivariate price jump occurrences, but with variance jumps absent from the model. Similarly, it extends the model proposed by Fulop, Li and Yu (2014), in which price jump intensity (only) is characterized by a Hawkes process, along with a restrictive assumption that variance jumps occur contemporaneously with negative price jumps.

2.2 A discrete time model for returns

In common with the literature, we undertake inference in the context of a discrete time representation of the model, as described in detail in this section. Applying an Euler discretization to (1) through (6) with $\Delta t = 1/252$ (equivalent to one trading day), leads to

$$r_t = \mu + \gamma V_t + \sqrt{V_t}\xi_t^p + Z_t^p \Delta N_t^p \tag{7}$$

$$V_{t+1} = \kappa\theta + (1-\kappa)V_t + \sigma_v\sqrt{V_t}\xi_t^v + Z_t^v \Delta N_t^v, \tag{8}$$

where $r_t = p_{t+1} - p_t$, and ξ_t^p and ξ_t^v are defined as marginally serially independent $N(0,1)$ sequences, but with $corr\left(\xi_t^p, \xi_t^v\right) = \rho$ for each t. Recognizing the nonlinear state space structure of the model, we re-write (8) as

$$V_{t+1} = \kappa\theta + (1-\kappa)V_t + \sigma_v\rho\left(r_t - Z_t^p \Delta N_t^p - \mu - \gamma V_t\right) + \sigma_v\sqrt{(1-\rho^2)V_t}\xi_t^{ind} + Z_t^v \Delta N_t^v, \tag{9}$$

to explicitly take into account the leverage parameter, ρ, leaving $corr\left(\xi_t^p, \xi_t^{ind}\right) = 0$. In addition, we write

$$\Delta N_t^p \sim Bernoulli(\delta_t^p) \tag{10}$$

$$\Delta N_t^v \sim Bernoulli(\delta_t^v), \tag{11}$$

with $\Delta N_t^p = N_{t+1}^p - N_t^p$, $\Delta N_t^v = N_{t+1}^v - N_t^v$, and where the probabilities of success are driven (respectively) by the discretized intensity processes,

$$\delta_t^p = \alpha_p \delta_\infty^p + (1 - \alpha_p) \delta_{t-1}^p + \beta_{pp} \Delta N_{t-1}^p \tag{12}$$

$$\delta_t^v = \alpha_v \delta_\infty^v + (1 - \alpha_v) \delta_{t-1}^v + \beta_{vv} \Delta N_{t-1}^v + \beta_{vp} \Delta N_{t-1}^p + \beta_{vp}^{(-)} \Delta N_{t-1}^{p(-)}. \tag{13}$$

The price and variance jump sizes, Z_t^p and Z_t^v, are assumed to be distributed as

$$Z_t^p | V_t \sim N\left(\mu_p + \gamma_p V_t, \sigma_p^2\right) \tag{14}$$

and

$$Z_t^v \sim Exp\left(\mu_v\right), \tag{15}$$

respectively. The price jump size Z_t^p in (14), being characterized as conditionally Gaussian, allows for the possibility of both positive and negative extreme movements in the asset price, and with the relative magnitude of the (conditional) expected price jump being proportional to V_t. The variance jump, on the other hand, is restricted to be positive, via the exponential distributional assumption in (15). Note that the parameters used in (7)-(13) are the discrete time versions of the corresponding parameters in the continuous time model (1)-(6), with the same symbols used for notational simplicity.

The factors that drive the asset price process can be interpreted as follows. Consistent with the empirical finance literature (see, for examples, Engle and Ng, 1993, Maheu and McCurdy, 2004, and Malik, 2011), the diffusive price shock, $\sqrt{V_t}\xi_t^p$, and the price jump occurrence, $Z_t^p \Delta N_t^p$, are collectively viewed as 'news'. Regular modest movements in price, as driven by $\sqrt{V_t}\xi_t^p$, are assumed to result from typical daily information flows, with (all else equal) the typical direction of the impact of $\sqrt{V_t}\xi_t^p$ on the variance of the subsequent period, V_{t+1}, captured by the sign of ρ. The occurrence of a price jump however, indicated by $\Delta N_t^p = 1$, can be viewed as a sizably larger than expected shock that may signal a shift in market conditions, with the probability of subsequent price and/or variance jumps adjusted accordingly, through the model adopted here for the jump intensities. That is, the process ΔN_t^p can be viewed as being potentially self-exciting: provoking an increase in the future intensity (and thus occurrence) of price jumps (via (12)), as well as provoking (or exciting) an increase in the future intensity of variance jumps (via (13)). The threshold parameter $\beta_{vp}^{(-)}$ in (13) allows for a possible additional impact of a negative price jump on the variance jump intensity (and, hence, the level of volatility), providing an additional channel for leverage, as noted above.

From the form of (7) and (8), the implications for returns of the occurrence of the two types of jumps are clear. From (7), the direct impact of a given price jump at time t, $\Delta J_t^p = Z_t^p \Delta N_t^p$, is felt only at time t. However, clusters of price jumps and, hence, successive extreme values in returns, can occur via the dynamic intensity process in (12) that drives subsequent realizations of ΔN_t^p. The impact of a given (positive) variance jump at time t will carry forward through time via the persistence of the V_{t+1} process, as governed by κ. That is, if the return variance jumps in any period, it will tend to remain higher in subsequent periods and, thus, be expected to cause larger movements in successive prices than would otherwise have occurred. In addition, any clustering of variance jumps, driven by the dynamic intensity process in (13), would simply cause an exaggeration of the resultant clustering of extreme returns. Arguably, clusters of jumps in the latent variance would typically be associated with sustained market instability, with the variance jump intensity expected to increase and remain high during periods of heightened market stress. Accordingly, the quantity δ_t^v is seen as a potential barometer of a crisis, a characteristic that is investigated empirically in Section 4.3.

The jump intensities, δ_t^p and δ_t^v, possess a conditionally deterministic structure that is analogous to that of a generalized autoregressive conditional heteroskedastic (GARCH) model for latent volatility, with the lagged jump occurrences playing a similar role to the lagged (squared) returns in a GARCH model (Bollerslev, 1986). Assuming stationarity, the unconditional mean for the price intensity process is determined by taking expectations through (12) as follows,

$$E\left(\delta_t^p\right) = E\left(\alpha_p \delta_\infty^p + (1 - \alpha_p)\,\delta_{t-1}^p + \beta_{pp}\Delta N_{t-1}^p\right)$$

and solving for the common value $\delta_0^p = E\left(\delta_t^p\right) = E\left(\delta_{t-1}^p\right)$ as

$$\delta_0^p = \frac{\alpha_p \delta_\infty^p}{\alpha_p - \beta_{pp}}. \tag{16}$$

Similarly, the unconditional mean of the variance jump intensity process in (13) is given by $\delta_0^v = E\left(\delta_t^v\right) = E\left(\delta_{t-1}^v\right)$, with

$$E\left(\delta_t^v\right) = E\left(\alpha_v \delta_\infty^v + (1 - \alpha_v)\,\delta_{t-1}^v + \beta_{vv}\Delta N_{t-1}^v + \beta_{vp}\Delta N_{t-1}^p + \beta_{vp}^{(-)}\Delta N_{t-1}^{p(-)}\right),$$

resulting in

$$\delta_0^v = \frac{\alpha_v \delta_\infty^v + \beta_{vp}\delta_0^p + \beta_{vp}^{(-)} F_{Z^p}(0)\,\delta_0^p}{\alpha_v - \beta_{vv}}, \tag{17}$$

8

where $F_{Z^p}(0) = E_{V_t}\left[\Phi\left(\frac{-(\mu_p + \gamma_p V_t)}{\sigma_p}\right)\right]$, with $\Phi(\cdot)$ denoting the cumulative distribution function (cdf) of a standard normal random variable. By substituting into equation (17) the expression for δ_0^p in (16), δ_0^v may be re-expressed as the following function of static parameters,

$$\delta_0^v = \frac{\alpha_v \delta_\infty^v}{\alpha_v - \beta_{vv}} + \frac{\beta_{vp}\alpha_p\delta_\infty^p + \beta_{vp}^{(-)}F_{Z^p}(0)\alpha_p\delta_\infty^p}{(\alpha_v - \beta_{vv})(\alpha_p - \beta_{pp})}. \tag{18}$$

To ensure that the quantities defined in (16) and (18) are finite and positive, the restrictions $\delta_\infty^p > 0$, $0 < \alpha_p < 1$, $\beta_{pp} > 0$, $\alpha_p - \beta_{pp} > 0$, $\delta_\infty^v > 0$, $0 < \alpha_v < 1$, $\beta_{vv} > 0$ and $\alpha_v - \beta_{vv} > 0$ are required. In addition, the restrictions $\delta_\infty^p < \frac{\alpha_p - \beta_{pp}}{\alpha_p}$ and $\delta_\infty^v < \frac{\alpha_v - \beta_{vv} - \beta_{vp}\delta_0^p - \beta_{vp}^{(-)}F_{Z^p}(0)\delta_0^p}{\alpha_v}$ ensure that $\delta_0^p \in (0,1)$ and $\delta_0^v \in (0,1)$, respectively.

2.3 Incorporating high-frequency measurements of volatility and price jumps via additional measurements

In the spirit of Barndorff-Nielsen and Shephard (2002), Creal (2008), Jacquier and Miller (2010), Maneesoonthorn, Martin, Forbes and Grose (2012) and Koopman and Scharth (2013), amongst others, we exploit the existence of high-frequency data to supplement the measurement equation in (7) with additional equations based on nonparametric measures of return variation: both its diffusive and jump components. In particular, in order to reduce the number of unobservable components to be estimated, we assume that both the price jump occurrence and price jump size are measured without error.

We begin with the most well-known nonparametric measure of variation, realized variance, defined by

$$RV_t = \sum_{t < t_i \leq t+1}^{M} r_{t_i}^2, \tag{19}$$

where $r_{t_i} = p_{t_{i+1}} - p_{t_i}$ denotes the ith observed return the over the horizon t to $t+1$, with there being M such returns. Under the assumption of no microstructure noise,

$$RV_t \xrightarrow{p} Q\mathcal{V}_{t,t+1} \quad \text{as } M \to \infty.$$

The quadratic variation, $Q\mathcal{V}_{t,t+1}$, captures the variation in returns due to both the stochastic volatility and price jump components, with

$$Q\mathcal{V}_{t,t+1} = \mathcal{V}_{t,t+1} + \mathcal{J}_{t,t+1}^2,$$

where $\mathcal{V}_{t,t+1} = \int_t^{t+1} V_s ds$ denotes the integrated variance, and $\mathcal{J}_{t,t+1}^2 = \sum_{t < s \leq t+1}^{N_{t+1}^p} (Z_s^p)^2$ denotes the price jump variation over the horizon t to $t+1$. Inherent in the measure of

integrated variance is information about both the diffusive and jump components of the variance process itself.

Barndorff-Nielsen and Shephard (2004) have proposed the quantity referred to as bipower variation,

$$BV_t = \frac{\pi}{2} \sum_{t < t_i \leq t+1}^{M} |r_{t_i}| \left| r_{t_{i-1}} \right|, \tag{20}$$

as a consistent measure of $\mathcal{V}_{t,t+1}$ as $M \to \infty$ (again, in the absence of microstructure noise). The discrepancy between RV_t in (19) and BV_t in (20) then serves as a measure of price jump variation, forming the basis of various statistical tests for price jumps (Barndorff-Nielsen and Shephard, 2004 and 2006, and Huang and Tauchen, 2005). For example, the asymptotic properties of RV_t and BV_t are used to define the distribution of the relative price jump statistic,

$$Z_{RJ,t} = \frac{RJ_t}{\sqrt{(v_{bb} - v_{qq}) M^{-1} \max\left(1, \frac{TQ_t}{BV_t^2}\right)}}, \tag{21}$$

where $RJ_t = \frac{RV_t - BV_t}{RV_t}$, TQ_t denotes an estimate of the integrated quarticity, and $Z_{RJ,t}$ has a limiting standard normal distribution under the assumption of no jumps. It is this price jump test statistic that we use to extract the price jump measurements ΔN_t^p and $Z_t^p|(\Delta N_t^p = 1)$, following the price jump extraction method of Tauchen and Zhou (2011). Assuming that a maximum of one price jump occurs over a day, the price jump occurrence, ΔN_t^p, is taken as

$$\Delta N_t^p = \mathbf{1}\left(Z_{RJ,t} > \Phi^{-1}(1 - \alpha)\right), \tag{22}$$

where $\Phi^{-1}(\cdot)$ is the inverse cdf of a standard normal distribution and α is the significance level. The price jump size is then extracted using

$$(Z_t^p|\Delta N_t^p = 1) = sign(r_t) \times \sqrt{\max(RV_t - BV_t, 0)}, \tag{23}$$

under the assumption that the price jump takes the same sign of the return on that particular day, and that it is observable only when $\Delta N_t^p = 1$. Measures of price jump occurrence and size, constructed using (22) and (23), are used as observable versions of the model components defined in (10) and (14), respectively, as motivated by Tauchen and Zhou, who investigate the time series dynamics of price jumps via these measures.

The bipower variation in (20), on the other hand, is viewed as a measure of the latent variance, V_t, in which measurement error is modelled explicitly. Specifically, and in order to keep the form of the measurement error as close to Gaussian as possible, we utilize BV_t in

log form, giving the measurement equation,

$$\ln BV_t = \ln V_t + \sigma_{BV}\xi_t^{BV}. \tag{24}$$

2.4 The full state space model

Collectively, (and with some repetition for ease of exposition) the measurement equations for the state space representation of our model can be expressed as

$$r_t = \mu + \gamma V_t + \sqrt{V_t}\xi_t^p + Z_t^p \Delta N_t^p \tag{25}$$

$$Z_t^p = \mu_p + \gamma_p V_t + \sigma_p \xi_t^{Z^p} \tag{26}$$

$$\Delta N_t^p \sim Bernoulli\left(\delta_t^p\right) \tag{27}$$

$$\ln BV_t = \ln V_t + \sigma_{BV}\xi_t^{BV}, \tag{28}$$

with the evolution of (unobserved) state variables given by,

$$V_{t+1} = \kappa\theta + (1-\kappa)V_t + \sigma_v\rho\left(r_t - Z_t^p\Delta N_t^p - \mu - \gamma V_t\right) + \sigma_v\sqrt{(1-\rho^2)V_t}\xi_t^{ind} + Z_t^v\Delta N_t^v \tag{29}$$

$$Z_t^v \sim Exponential\left(\mu_v\right) \tag{30}$$

$$\Delta N_t^v \sim Bernoulli\left(\delta_t^v\right). \tag{31}$$

In addition, the two conditionally deterministic states representing the price and variance jumps intensities, and contributing respectively to (27) and (31), are specified (once again) as

$$\delta_t^p = \alpha_p\delta_\infty^p + (1-\alpha_p)\delta_{t-1}^p + \beta_{pp}\Delta N_{t-1}^p \tag{32}$$

$$\delta_t^v = \alpha_v\delta_\infty^v + (1-\alpha_v)\delta_{t-1}^v + \beta_{vv}\Delta N_{t-1}^v + \beta_{vp}\Delta N_{t-1}^p + \beta_{vp}^{(-)}\Delta N_{t-1}^{p(-)}, \tag{33}$$

where $\Delta N_{t-1}^{p(-)} = \Delta N_{t-1}^p \mathbf{1}\left(Z_{t-1}^p < 0\right)$. As noted earlier, the restrictions $\delta_\infty^p > 0$, $0 < \alpha_p < 1$, $\beta_{pp} > 0$, $\alpha_p - \beta_{pp} > 0$, $\delta_\infty^v > 0$, $0 < \alpha_v < 1$, $\beta_{vv} > 0$, and $\alpha_v - \beta_{vv} > 0$ are required. The error components $\left(\xi_t^p, \xi_t^{Z^p}, \xi_t^{BV}, \xi_t^{ind}\right)'$ are collectively assumed to be multivariate standard normal (truncated in the case of ξ_t^{ind} to ensure the positivity of V_{t+1}) with an identity variance-covariance matrix. Assumptions are also required for the initial stochastic variance, V_1, the initial variance jump occurrence ΔN_1^v, and the initial bivariate jump intensities, δ_1^i, for $i = v, p$. Here we impose $V_1 = \theta + \frac{\mu_v\delta_0^v}{\kappa}$, $\delta_1^p = \delta_0^p$, $\delta_1^v = \delta_0^v$ and $\Delta N_1^v = \Delta N_1^p$; however the methodology outlined in the next section may be adapted to accommodate alternative assumptions about initial conditions.

11

3 Bayesian inference

Given the complexity of the state space representation, and the high dimension of the set of unknowns, the joint posterior required for Bayesian inference is not available in closed form. Hence, an MCMC algorithm is developed to obtain draws of the parameters and latent variables of interest from the joint posterior distribution, with inference - including the construction of posterior predictive distributions - conducted using those draws. The marginal likelihoods used to conduct the in-sample comparison of the various models of interest (as specified in Section 3.2) are computed from the simulation output using the computation methods of Chib (1995) and Chib and Jeliazkov (2001).

3.1 MCMC algorithm and priors

For notational convenience, we denote time-indexed variables generically as, for example, $X_{1:t} = (X_1, ..., X_t)'$ for $t = 1, ..., T$ and where $X_{1:0}$ is empty. The static parameters are collectively denoted by the vector $\phi = (\mu, \gamma, \mu_p, \gamma_p, \sigma_p, \sigma_{BV}, \kappa, \theta, \sigma_v, \rho, \delta_0^p, \alpha_p, \beta_{pp}, \delta_0^v, \alpha_v, \beta_{vv}, \beta_{vp}, \beta_{vp}^{(-)}, \mu_v)'$. The joint posterior density associated with the full model in (25)-(33) satisfies

$$
\begin{aligned}
& p\left(V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}^v, \phi | r_{1:T}, \ln BV_{1:T}, Z_{1:T}^p, \Delta N_{1:T}^p\right) \\
& \propto p\left(r_1 | V_1, Z_1^p, \Delta N_1^p, \mu, \gamma\right) \times p\left(\ln BV_1 | V_1, \sigma_{BV}\right) \times p\left(Z_1^p | \Delta N_1^p, V_1, \mu_p, \gamma_p, \sigma_p\right) \\
& \times p\left(\Delta N_1^p | \delta_1^p\right) \times p\left(Z_1^v | \mu_v\right) \times p\left(V_1 | \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right) \times p\left(\Delta N_1^v | \Delta N_1^p\right) \times p(\phi) \\
& \times \left[\prod_{t=2}^T p\left(r_t | V_t, Z_t^p, \Delta N_t^p, \mu, \gamma\right) \times p\left(\ln BV_t | V_t, \sigma_{BV}\right) \times p\left(Z_t^p | \Delta N_t^p, V_t, \mu_p, \gamma_p, \sigma_p\right) \right. \\
& \times p\left(\Delta N_t^p | \Delta N_{1:t-1}^p, \alpha_p, \beta_{pp}, \delta_0^p\right) \times p\left(V_t | V_{t-1}, Z_{t-1}^v, \Delta N_{t-1}^v, r_{t-1}, Z_{t-1}^p, \Delta N_{t-1}^p, \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right) \\
& \left. \times p\left(Z_t^v | \mu_v\right) \times p\left(\Delta N_t^v | \Delta N_{1:t-1}^v, \Delta N_{1:t-1}^p, Z_{1:t-1}^p, \alpha_v, \beta_{vv}, \beta_{vp}, \beta_{vp}^{(-)}, \delta_0^v\right)\right].
\end{aligned}
\tag{34}
$$

Note that this joint posterior assumes that $\delta_1^p = \delta_0^p$, $\delta_1^v = \delta_0^v$ and $\Delta N_1^v = \Delta N_1^p$.

In the specification of $p(\phi)$ in (34) we use a combination of noninformative and weakly informative priors for the various elements of ϕ. Uniform priors are assumed for the parameters κ and θ, truncated from below at zero, while the parameter σ_v^2 is blocked with the leverage parameter, ρ, via the reparameterization: $\psi = \rho \sigma_v$ and $\omega = \sigma_v^2 - \psi^2$; see Jacquier, Polson and Rossi (2004). This reparameterization is convenient as, given $V_{1:T}$, it allows for ψ and ω to be treated respectively as the slope and error variance coefficients in a normal

linear regression model. Direct sampling of ψ and ω is then conducted using standard posterior results, based on conjugate prior specifications in the form of conditional normal and inverse gamma distributions, respectively, given by $p\left(\psi|\omega\right) \sim N\left(\psi_0 = -0.005, \sigma_0^2 = \omega/5.0\right)$ and $p\left(\omega\right) \sim IG\left(\alpha = 10, \beta = 0.001\right)$, where β denotes the scale parameter in the context of the inverse gamma distributions discussed here. The prior specifications for ψ and ω are chosen such that the implied prior distributions for ρ and σ_v have means of -0.34 and 0.012, and standard deviations of 0.33 and 0.0027, respectively. As such, the prior distributions for ρ and σ_v are relatively diffuse, with the ranges being broadly in line with the range of the empirical values of these parameters reported in the literature.

Truncated uniform priors are specified for the parameters μ, γ, μ_p and γ_p, with the priors for μ and μ_p both defined on the interval $(-10, 10)$ and γ_p defined on the interval $(-100, 100)$. Very wide ranges of values for these parameters, over both the negative and positive regions of the real line, are thus specified *a priori*. The volatility feedback parameter γ is assumed *a priori* to be bounded from above at zero, and to be uniformly distributed over the interval $(-100, 0)$. This is consistent with recent findings of negative volatility feedback in the high frequency literature. (See, for example, Bollerslev *et al.* 2006, and Jensen and Maheu 2013). Conjugate inverse gamma priors, $IG\left(\alpha = 3, \beta = 1\right)$, are applied to the parameters σ_p^2 and σ_{BV}^2, with both prior distributions being centred around a mean of 0.5, and with (a relatively large) standard deviation of 0.5, as a consequence.

Conjugate beta priors are employed for the unconditional jump intensities, δ_0^p and δ_0^v. Specifically, a $Beta(a = 10, b = 90)$ prior is specified for δ_0^p, such that the prior mean of 0.1 matches the sample mean of the observed $\Delta N_{1:T}^p$. The prior distribution of δ_0^v is, in turn, equated with that of δ_0^p, stemming from the prior belief that if there is a price jump $(\Delta N_t^p = 1)$, then it is *likely* (albeit not strictly necessary) that the variance process also contains a jump (that is, $\Delta N_t^v = 1$). A conjugate inverse gamma prior, $IG\left(\alpha = 20, \beta = 7.2\right)$, is employed for μ_v, implying a prior mean of 0.007 and prior standard deviation of 0.002, where this prior mean is a proportion of the average of $\max\left(RV_t - BV_t, 0\right)$. The initial stochastic variance is assumed to be degenerate, with $V_1 = \theta + \frac{\mu_v \delta_0^v}{\kappa}$. Uniform priors are employed for the jump intensity parameters, $\alpha_p, \beta_{pp}, \beta_{vp}, \beta_{vp}^{(-)}, \alpha_v$ and β_{vv}, conforming to the theoretical restrictions listed at the end of Section 2.4, and the prior belief that $\beta_{vp} > 0$ and $\beta_{vp}^{(-)} > 0$.

A hybrid of the Gibbs and Metropolis-Hastings (MH) algorithms is used for sampling the latent variables and the elements of the static parameter vector. Details of the algorithm, including a reference made to Maneesoonthorn *et al.* (2012) for a description of the multi-

move algorithm adopted for sampling the state vector, $V_{1:T}$, are given in Appendix A.

3.2 Models of interest and their marginal likelihoods

As has been highlighted, a novel aspect of our specification is that it allows for dynamic behaviour in both price and variance jumps, as well as various types of dependencies between those extreme movements. It is of interest then to explore whether or not this rich dynamic structure is warranted empirically, through an investigation of various models associated with the jump intensity processes in (32) and (33).

We consider several competing models, all but one of which are nested in the general specification, and all of which are to be evaluated empirically in Section 4. First, we determine whether the occurrence of negative price jumps (per se) influences the variance jump intensity by testing the model

$$\mathcal{M}_1 : \beta_{vp}^{(-)} = 0; \tag{35}$$

and then whether the occurrence of price jumps has any impact at all on the variance jump intensity, by considering

$$\mathcal{M}_2 : \beta_{vp}^{(-)} = \beta_{vp} = 0. \tag{36}$$

Another issue of interest is whether the price and variance jumps occur contemporaneously, and whether variance jumps are a necessary part of the model, with consideration of

$$\mathcal{M}_3 : \Delta N_t^p = \Delta N_t^v \text{ for all } t = 1, ... T \tag{37}$$

and

$$\mathcal{M}_4 : \Delta N_t^v = 0 \text{ for all } t = 1, ... T \ , \tag{38}$$

respectively. Note that the model under restriction (38) corresponds to the model proposed by Aït-Sahalia *et al.* (2014), albeit in a single asset setting here.

Next, we specify the jump intensities to be proportional to the latent variance, with this specification sharing some common features with the models adopted in Bates (1996), Pan (2002), and Eraker (2004),

$$\mathcal{M}_5 : \delta_t^p = \alpha_{p_0} + \alpha_p V_t \text{ and } \delta_t^v = \alpha_{v_0} + \alpha_v V_t \text{ for all } t = 1, ... T. \tag{39}$$

We then specify constant jump intensities, yielding the stochastic volatility with the independent jumps (SVIJ) model of Duffie *et al.* (2000):

$$\mathcal{M}_6 : \delta_t^p = \delta_0^p, \ \delta_t^v = \delta_0^v \text{ for all } t = 1, ... T. \tag{40}$$

14

Finally, we consider the absence of both price and variance jumps:

$$\mathcal{M}_7 : \delta_t^p = 0 \text{ and } \delta_t^v = 0 \text{ for all } t = 1, ...T, \tag{41}$$

coinciding with the conventional Heston (1993) square root model.

To examine these seven models of interest, computation of their corresponding marginal likelihood values,

$$p\left(r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p | \mathcal{M}_i\right), \tag{42}$$

for $i = 1, ..., 7$, is required. In addition to the specific models, $\mathcal{M}_i$, $i = 1, ..., 7$ listed above, we denote the state space model with the complete specification - collectively defined by (25)-(33) - by $\mathcal{M}_F$. Under the assumption that each of the models is *a priori* equally likely, the posterior odds ratio for any restricted model $\mathcal{M}_i$ relative to the full state space model $\mathcal{M}_F$, is equivalent to the Bayes factor $BF_{i,F}$, given in turn by

$$BF_{i,F} = \frac{p\left(r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p | \mathcal{M}_i\right)}{p\left(r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p | \mathcal{M}_F\right)}. \tag{43}$$

Note that given Bayes factors $BF_{i,F}$ and $BF_{j,F}$, the Bayes factor for model $\mathcal{M}_i$ against $\mathcal{M}_j$ is obtained simply as $BF_{i,j} = BF_{i,F}/BF_{j,F}$. As such, the information about the relative performance of the competing models can be obtained from the relative ranking of the models' marginal likelihoods. The marginal likelihood for model $\mathcal{M}_i$ in (42) and that appears in the numerator of (43) is challenging to compute, requiring the calculation of an integral over a very large dimension due to the number of latent variables present. We estimate the marginal likelihood of each model using the output of a series MCMC algorithms, details of which are provided in Appendix B.

4   Empirical analysis

4.1   Data description

In this section we report the results of the application of the proposed method to daily annualized returns, plus annualized bipower variation and annualized price jump measures, constructed from data on the S&P500 index, for the period January 3, 1996 to June 23, 2014. The period of January 3, 1996 to May 15, 2012 (comprising 4070 trading days) is used for in-sample analysis, while the period of May 16, 2012 to June 23, 2014 (comprising 528 days) is preserved for evaluating out-of-sample predictive return distributions, associated value at risk (VaR) estimates, as well as the predictive distributions of daily settlement prices of a

particular S&P500 futures contract. The index data has been supplied by the Securities Industries Research Centre of Asia Pacific (SIRCA) on behalf of Reuters, with the raw intraday index data having been cleaned using methods similar to those of Brownlees and Gallo (2006). The numerical results reported in this empirical section have been produced using the JAVA programming language.

The in-sample daily annualized return, r_t, is computed as the difference between the log of the closing index value and that of the opening value on a particular trading day, scaled by an annualization factor of 252, and is plotted in Panel A of Figure 1 for the combined in-sample and out-of-sample periods. Figure 1 also contain plots of the corresponding daily annualized quantities of the bipower variation measure (BV_t) and its log $(\ln BV_t)$ in Panel B and the price jump measure $Z_t^p \Delta N_t^p$ in Panel C, with the latter plot simultaneously indicating the price jump timing and the size and direction of such jumps. Both measures, BV_t and $Z_t^p \Delta N_t^p$, are based on fixed five minute sampling, with a 'nearest price' method (Andersen, Bollerslev and Diebold, 2007) used to construct the relevant returns five minutes apart. With reference to (22), the price jump occurrence measure ΔN_t^p is extracted using the significance level of $\alpha = 0.001$, as recommended by Tauchen and Zhou (2011), for the price jump test statistic in (21). The descriptive summary statistics of the in-sample period data (only) are tabulated in Table 1. As is evident in Panels A and B, and as is completely expected in this setting, volatility clustering in returns exists, with the most extreme variation in returns, and BV_t values of unprecedented magnitude, observed towards the end of 2008. The plot of the price jump measure, $Z_t^p \Delta N_t^p$, in Panel C of Figure 1 shows clear evidence of price jump clustering, with clusters appearing intermittently over the sample period. The marginal intensity of price jumps is around 10%, with the price jump size being approximately Gaussian, as evidenced by the large p-value for the corresponding Jarque-Bera statistic (see Table 1). The most intense volatility clustering, and the clusters of price jumps that are largest in magnitude, occur during three of the most volatile market periods: late 2001 and throughout 2002 following the September 11 terrorist attacks, the global financial crisis period in 2008 and 2009, and the culmination of the period of Euro-zone debt crises, in 2011.

4.2 The parameter estimates

Estimates of the static parameters of the full state space model, $\mathcal{M}_F$, for the sample period between January 3, 1996 and May 15, 2012, are recorded in Table 2. Marginal posterior means (MPMs) and 95% higher posterior density (HPD) interval estimates are calculated from $G = 100{,}000$ MCMC draws, following a 100,000 draw burn-in period of which every
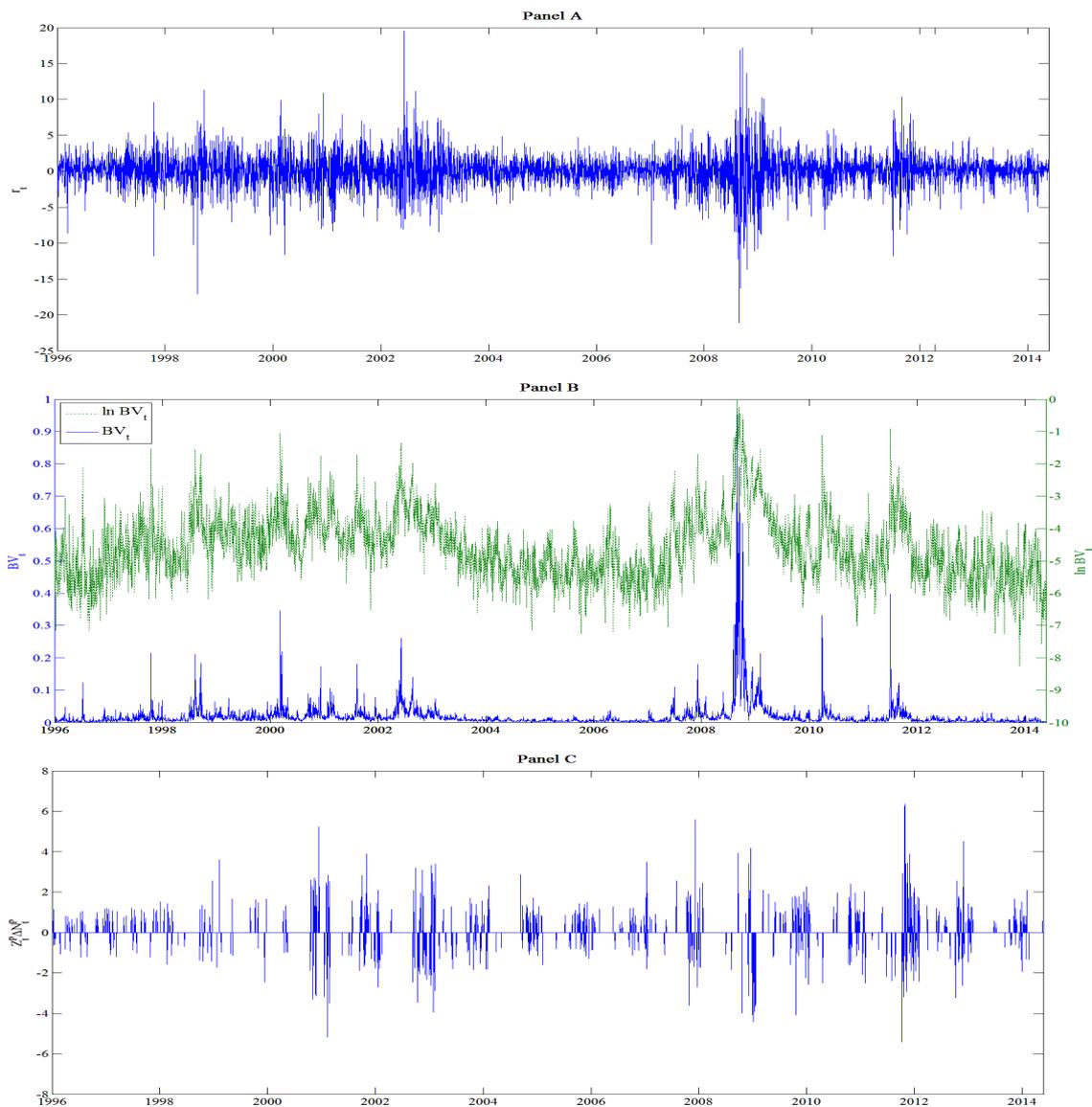
Figure 1: The daily annualized daily return, r_t, of the S&P500 index (Panel A), the measure of bipower variation, BV_t, and its log, $\ln BV_t$, (Panel B) and price jumps, $Z_t^p \Delta N_t^p$, (Panel C). Note that the left vertical axis in Panel B provides the scale for BV_t, whereas the right vertical axis proves the relevant scale for $\ln BV_t$. The non-parametric measures in Panel B and C are constructed using intraday sampling at the 5 minute frequency. The data ranges from January 3, 1996 to June 23, 2014, inclusive, comprising both the in-sample period, ending on May 15, 2012, and the subsequent out-of-sample period.

Table 1: Summary statistics for r_t, BV_t, $\ln(BV_t)$, ΔN_t^p and $Z_t^p | \Delta N_t^p = 1$ for the S&P 500 stock index for January 3, 1996 to May 15, 2012, inclusive.

	Annualized r_t	BV_t	$\ln(BV_t)$	ΔN_t^p	$Z_t^p | \Delta N_t^p = 1$
Mean	0.0421	0.0191	-4.5913	0.1070	0.1494
Standard Deviation	2.4689	0.0407	1.0187	0.3092	1.7601
Minimum	-21.070	0.0003	-8.2593	0	-5.4288
Maximum	19.511	0.9671	-0.0334	1	6.3586
25th percentile	-1.1080	0.0049	-5.3249	0	-1.1264
50th percentile	0.1660	0.0095	-4.6561	0	0.6511
75th percentile	1.2711	0.0192	-3.9513	0	1.2505
Skewness	-0.2793	10.143	0.4859	2.5427	-0.0679
Kurtosis	9.3022	156.40	3.5736	7.4654	3.3325
Jarque-Bera p-value	0.00	0.00	0.00	0.00	0.24

10th draw is saved. Inefficiency factors computed from the retained posterior draws are also reported in the table, estimated as the ratio of the variance of the sample mean of a set of MCMC draws of a given unknown, to the variance of the sample mean from a hypothetical independent sample. All parameter estimates are reported in annualized terms where appropriate. For example, estimates of θ have a scale that accords with an annualized variance quantity, whilst estimates of κ reflect the daily persistence in annualized variance.

The parameters associated with the measurement equations (25) and (28), namely μ, γ and σ_{BV}, are all within ranges that are consistent with those obtained in the study reported in Maneesoonthorn *et al.* (2012), where the integrated variance measure BV_t, supplemented by a model-free option-implied volatility measure and subsequently augmented by daily returns, is used to conduct inference on the latent variance process over a sample period from July 1999 and December 2008. The reported MPM of σ_v in Table 2 is broadly consistent with point estimates that have been reported earlier for similar models (see Broadie *et al.*, 2007, for a summary). The value of κ, on the other hand, is rather high compared to other estimates reported in the literature, with a possible explanation being that the degree of persistence in the latent variance process is partially captured by the dynamic model for the variance jump intensity in our specification.[1] The estimates of θ and μ_v are also consistent with

---
[1]This observation is further supported by the MCMC estimation results for the various alternative dynamic jump specifications explored in Section 4.4. In brief, diffusive volatility under those specifications with restrictive assumptions about the dynamics in volatility jumps is more persistent than otherwise. The unconditional diffusive variance is also larger in magnitude in these cases. The estimation results for the alternative models are available in a supplementary online appendix, and are discusssed in more detail in

those reported in Maneesoonthorn *et al.* (2012), although higher in magnitude than those reported by Broadie *et al.*, based on a less volatile sample of data. The MPM of the leverage parameter, ρ, is negative and the magnitude of ρ, ranging from 0.25 to 0.40 according to the HPD interval, is consistent with figures previously reported in the literature (see Broadie *et al.* and Aït-Sahalia, Fan and Li, 2013, for examples).

The parameters associated with the two jump intensity processes are, of course, our primary interest. As the price jump variables, Z_t^p and ΔN_t^p, are treated as observable, and certain prior specifications are linked to these observed variables, it is not surprising that the reported MPM of the unconditional price jump intensity, δ_0^p, is close to the sample mean of the ΔN_t^p measure. The dynamic price jump intensity process, δ_t^p, possesses a strong degree of persistence, as indicated by the relatively low MPM of α_p, and an 95% HPD interval for β_{pp} that is well above zero, consistent with the presence of self-excitation. The magnitudes of α_p and β_{pp} reported here, once annualized, are consistent with the parameters reported in Aït-Sahalia *et al.* (2014), who (as noted earlier) propose a Hawkes process for price jumps, but omit variance jumps in their stochastic volatility specification.

The MPM of the long-run variance jump intensity, δ_0^v, is relatively high compared with previously reported (comparable) quantities (Eraker *et al.*, 2003, Eraker, 2004 and Broadie *et al.*, 2007). The variance jump intensity process is also more persistent than the price jump intensity process, with the MPM of α_v being lower in magnitude than that of α_p. In addition there is evidence of self-exciting dynamics, as indicated by the non-zero MPM of β_{vv}. The self-exciting dynamics in δ_t^v, measured by β_{vv}, are much stronger than the feedback from the previous price jump occurrence, measured by β_{vp}, and its threshold component, measured by $\beta_{vp}^{(-)}$, with the marginal posterior densities for both β_{vp} and $\beta_{vp}^{(-)}$ being highly concentrated around mean values very close to zero. Further assessment of the importance of these feedback effects, plus the importance of the dynamic structures specified for price and variance jumps, and of the presence of jumps per se, is conducted in Section 4.4, via a comparison of marginal likelihoods.

The inefficiency factors reported in Table 2 for the static parameters range from 1 to 49, with the parameters associated with the dynamic variance jump intensity and size producing the highest inefficiency factors. The factors for the latent variance, $V_{1:T}$, computed at selected time points, range from 3 to 4. The acceptance rates for all parameters drawn using MH schemes range from 15-30%, with the acceptance rate for drawing $V_{1:T}$ (in blocks) - computed as the proportion of times that at least one block of $V_{1:T}$ is updated over the entire MCMC

Section 4.4.

19

Table 2: Empirical results for the S&P 500 stock index for January 3, 1996 to May 15, 2012, inclusive, for the full dynamic model, $\mathcal{M}_F$.

Parameter	MPM (95% HPD) Inefficiency Factor	Parameter	MPM (95% HPD) Inefficiency Factor
μ	0.179 (0.118, 0.234) 1.18	δ_0^p	0.104 (0.095, 0.114) 1.04
γ	-7.683 (-9.903, -3.716) 1.00	α_p	0.095 (0.070, 0.127) 3.76
μ_p	0.353 (0.118, 0.591) 1.45	β_{pp}	0.059 (0.045, 0.076) 4.65
γ_p	-13.43 (-24.58, -2.115) 2.41	δ_0^v	0.058 (0.035, 0.087) 49.18
σ_p	1.773 (1.658, 1.897) 1.05	α_v	0.041 (0.024, 0.064) 27.56
σ_{BV}	0.468 (0.454, 0.481) 3.08	β_{vv}	0.032 (0.019, 0.048) 32.03
κ	0.101 (0.080, 0.125) 66.83	β_{vp}	$4.79e^{-4}$ $(1.19e^{-5}, 1.73e^{-3})$ 1.29
θ	0.0078 (0.0069, 0.0087) 18.36	$\beta_{vp}^{(-)}$	$1.04e^{-3}$ $(2.64e^{-5}, 3.67e^{-3})$ 1.43
ρ	-0.328 (-0.401, -0.254) 5.37	μ_v	0.021 (0.016, 0.027) 49.28
σ_v	0.016 (0.014, 0.017) 18.04		

chain - being approximately 99%. The convergence of the MCMC chains for all model unknowns is also confirmed via inspection of graphical CUSUM plots (Yu and Mykland, 1998).

4.3   The latent processes

Time series plots of the MPMs and the 95% HPD intervals of the latent variance, V_t, and the variance jump, $\Delta J_t^v = Z_t^v \Delta N_t^v$, computed at every time point over the estimation period for the full model, $\mathcal{M}_F$, are shown in Panels A and B, respectively, of Figure 2. Similarly, Figure 3 displays the time series plots of the MPMs and the 95% HPD intervals of the jump intensity processes, δ_t^p and δ_t^v, in Panels A and B respectively.
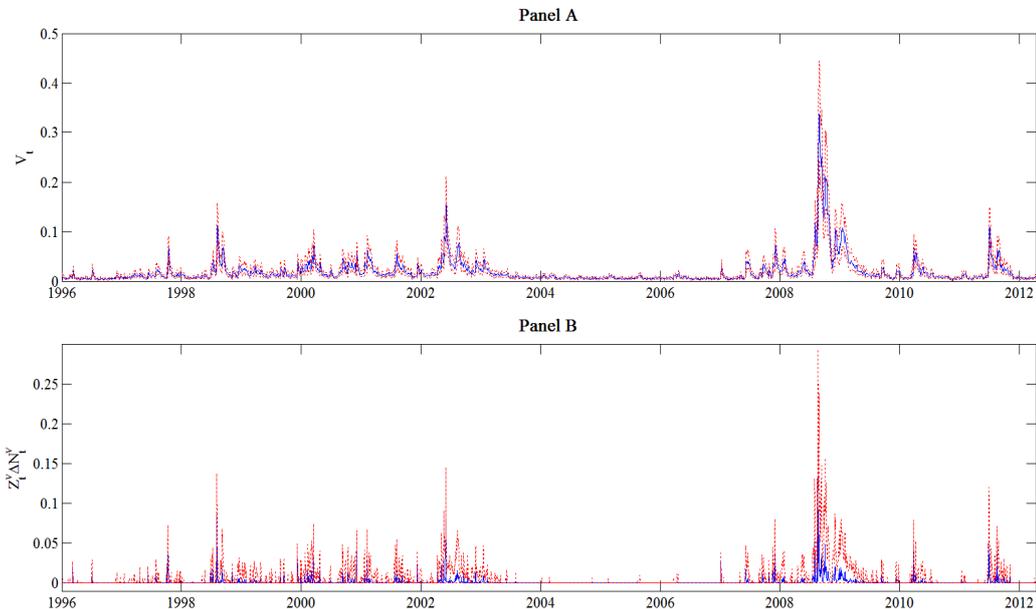
Figure 2: Time series plots of the MPMs and (pointwise) 95% HPD intervals for the latent variance, V_t, (Panel A) and variance jump, $\Delta J_t^v = Z_t^v \Delta N_t^v$, (Panel B) over the sample period January 3, 1996 to May 15, 2012, inclusive.

Reconciling the patterns in Panel A of Figure 2 with those in Panel B of Figure 1, it is clear that the overall dynamic behaviour in the MPM of V_t closely tracks that of its direct measure, BV_t, even in the most extreme period towards the end of 2008. In Panel B of Figure 2, there is also distinct evidence of clustering of variance jumps in high volatility periods associated, in turn, with an increase in the point and interval estimates of δ_t^v, as shown in Panel B of Figure 3. Once a period of multiple variance jumps has passed, the
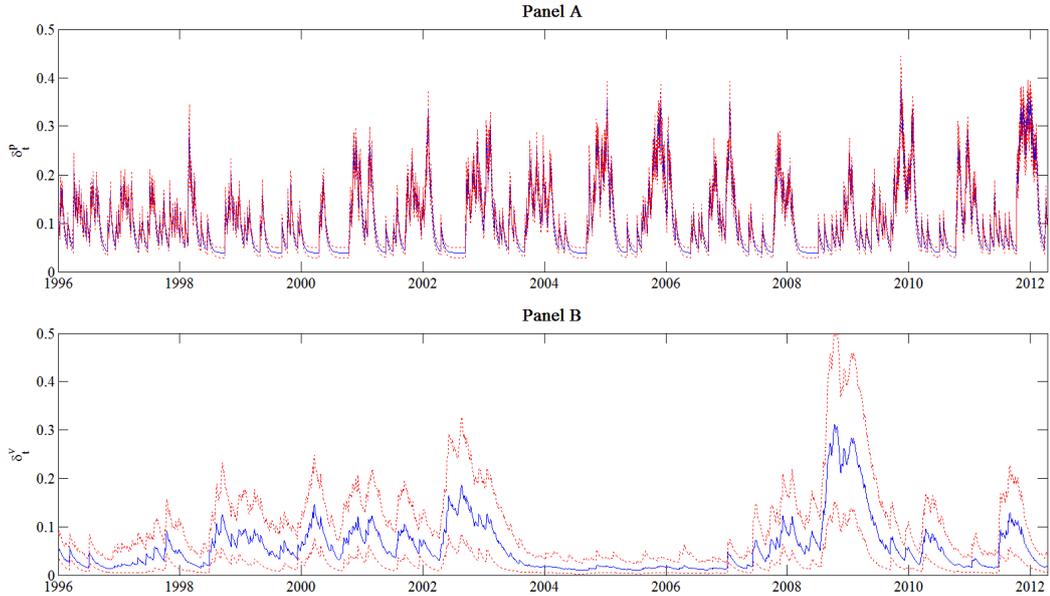
21

Figure 3: Time series plots of the MPMs and (pointwise) 95% HPD intervals for the price jump intensity, δ_t^p, (Panel A) and variance jump intensity, δ_t^v, (Panel B), over the sample period January 3, 1996 to May 15, 2012, inclusive.

value of δ_t^v declines rather slowly, with this high level of persistence being consistent with the low estimates (point and interval) of α_v recorded in Table 2.

In order to identify any relationship between the variance jump intensity and general market conditions, and to assess whether or not δ_t^v may be useful as a barometer of extreme market events, Figure 4 reproduces the MPM of δ_t^v, over the period between January 2007 and May 2012, with many well-known market events indicated. As can be observed, some of the sharpest rises in δ_t^v are either synchronous with, or occur soon after, certain key events. In particular, the collapse of the Lehman Brothers (September, 2008) and the subsequent intervention by the US Federal Reserve (December, 2008) are followed closely by the largest variance jump intensity levels observed throughout the entire sample period (reaching a peak of almost 0.37 on December 5th, 2008). During the various phases of the recent Euro-zone debt crisis (starting from late 2009), sharp increases in the MPM of δ_t^v are also evident, albeit with the magnitude of these being less than the rises observed during the global financial crisis.

Panel A of Figure 3 depicts the MPM and 95% HPD interval of the price jump intensity variable, δ_t^p, over the sample period. In contrast to the evident link between the magnitude of the variance jump intensity and the characteristics of the market, the magnitude and
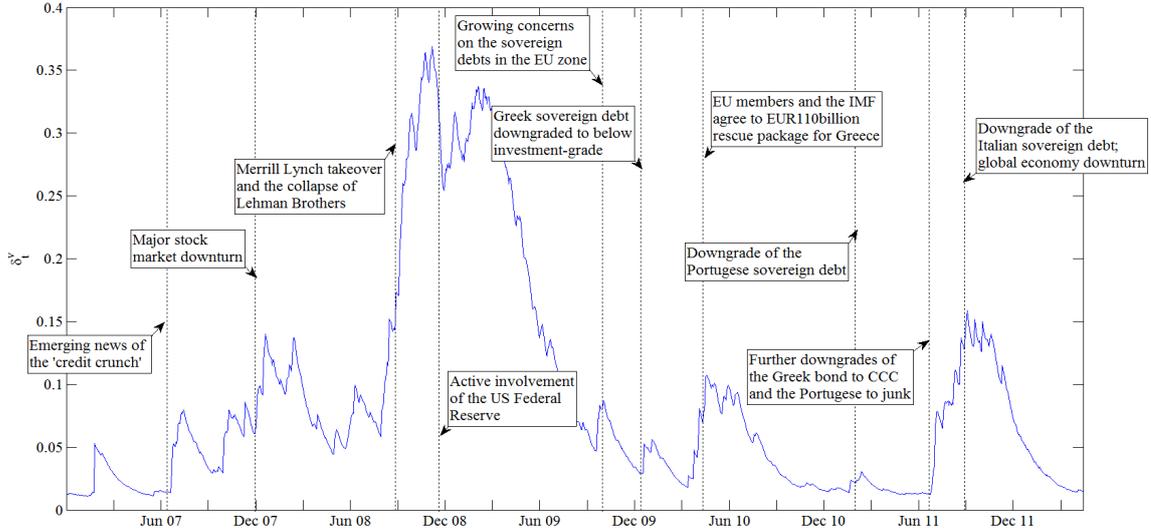
Figure 4: Time series plot of the variance jump intensity process, δ_t^v, over the period January 1, 2007 to May 15, 2012, inclusive, with the timing of various important market events noted, including the recent global financial crisis, as well as the Euro-zone debt crisis.

dynamic behaviour of δ_t^p do not track market conditions in any obvious way. Specifically, an increase in the intensity of price jumps is both relatively short lived (compared to that of variance jumps) and associated with periods in which the magnitude of the observed jumps (Figure 1, Panel C) may either be large or small. That is, an increase in price jump intensity does not appear to correlate with a period of *large* price jumps only.

4.4   Model comparisons

Evaluation of models $\mathcal{M}_1$ to $\mathcal{M}_7$, described in (35) to (41) respectively in Section 3.2, in addition to the full dynamic model described in Section 2.4, is conducted by ranking the corresponding marginal likelihoods. Table 3 reports the logarithm of the marginal likelihoods for all eight models. Note that the MPMs and 95% HPD intervals for the relevant parameter sets of the alternative models investigated, $\mathcal{M}_1$ to $\mathcal{M}_7$, are not reported here to conserve space, but are (as noted in an earlier footnote) available in a supplementary online appendix. In summary, the estimates (point and interval) of the static parameters under the alternative models are relatively robust to model specification, with the exception to this being certain parameters in the models in which restrictive assumptions are imposed on the variance jumps. In particular, the model imposing the contemporaneous occurrence of price and variance jumps, the model that excludes variance jumps, and the model that imposes

23

constant variance jump intensities ($\mathcal{M}_3$, $\mathcal{M}_4$ and $\mathcal{M}_6$, respectively) each imply a higher degree of persistence in the diffusive volatility component - as measured by κ - and a higher overall level of volatility - as measured by θ - than do the remaining models considered.

A key message from the results recorded in Table 3 is that models in which both types of jumps are present, and are allowed to occur independently, are strongly supported by the data. The model that imposes contemporaneous price and variance jumps ($\mathcal{M}_3$) performs more poorly, according to marginal likelihood, than does the model with no variance jumps at all ($\mathcal{M}_4$). In addition, models in which some kind of jump dynamics are allowed are more favoured, with models $\mathcal{M}_F$, $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_5$ being ranked in the top four. Models $\mathcal{M}_1$ and $\mathcal{M}_2$, which impose restrictions on the impact of price jumps on variance jump intensity, produce higher marginal likelihoods than the full dynamic model, $\mathcal{M}_F$. Indeed, the model $\mathcal{M}_2$, in which feedback from a price jump (positive or negative) to the subsequent variance jump intensity is omitted, is assigned the top ranking amongst the set of eight models. The support for $\mathcal{M}_2$ is also consistent with the extremely small magnitudes of the point and interval estimates of β_{vp} and $\beta_{vp}^{(-)}$ from the full dynamic model, $\mathcal{M}_F$, reported in Table 2. Not surprisingly, the Heston specification in $\mathcal{M}_7$, with both types of jumps excluded, performs the poorest overall, with model $\mathcal{M}_6$, which assumes no dynamics in the jump intensities, ranking second last.

Table 3: Log marginal likelihoods and model rankings, computed using the data from January 3, 1996 to May 15, 2012, inclusive.

Model	ln (marginal likelihood)	Ranking
$\mathcal{M}_F$: Full Hawkes	-11853	3
$\mathcal{M}_1$: Hawkes with $\beta_{vp}^{(-)} = 0$	-11750	2
$\mathcal{M}_2$: Hawkes with $\beta_{vp}^{(-)} = \beta_{vp} = 0$	-11738	1
$\mathcal{M}_3$: Hawkes with $\Delta N_t^p = \Delta N_t^v$	-12376	6
$\mathcal{M}_4$: Hawkes with $\Delta N_t^v = 0$	-12280	5
$\mathcal{M}_5$: Model with intensity specification $\delta_t^p = \alpha_{p_0} + \alpha_p V_t$ and $\delta_t^v = \alpha_{v_0} + \alpha_v V_t$	-12153	4
$\mathcal{M}_6$: Constant intensity specification $\delta_t^p = \delta_0^p$ and $\delta_t^v = \delta_0^v$	-12437	7
$\mathcal{M}_7$: Heston model specification $\delta_t^p = 0$ and $\delta_t^v = 0$	-18353	8

4.5 Evaluation of predictive densities

The model comparison exercise conducted in the previous section focuses on in-sample performance. It is also of interest to evaluate the forecasting value of the dynamic structures proposed in the paper. For this purpose we choose to assess the accuracy of the predictive returns distribution associated the most favoured Hawkes model, $\mathcal{M}_2$, compared with the models that assume the most distinct specifications for jump intensities, namely: $\mathcal{M}_5$, in which a non-Hawkes dynamic structure is adopted for the intensities; $\mathcal{M}_6$, in which the jump intensities are constant; and $\mathcal{M}_7$, in which no jumps at all are modelled.[2] Using the posterior draws of the static parameters based on the data from January 3, 1996 to May 15, 2012, inclusive, 528 one-step-ahead predictive densities are produced over the out-of-sample period of May 16, 2012 to June 23, 2014, inclusive. Forecasts of the latent processes, $V_{T+1}, Z_T^v, \Delta N_T^v, \delta_{T+1}^p$ and δ_{T+1}^v, are produced using a fixed window of $n = 4070$ daily observations, with the forecasts updated recursively over the forecast evaluation period. The predictive ability of the four models under investigation is evaluated in two ways: in terms of the accuracy of density forecasts of S&P500 index returns themselves; and in terms of the accuracy of density forecasts of the prices of futures contracts written on the index.

4.5.1 The predictive return distribution

The one-step-ahead predictive return distribution from model $\mathcal{M}_i$ at time T is given by

$$
p\left(r_{T+1}|r_{T-n+1:T}, \ln BV_{T-n+1:T}, \Delta N_{T-n+1:T}^p, Z_{T-n+1:T}^p, \mathcal{M}_i\right)
$$
$$
= \int p\left(r_{T+1}|V_{T+1}, Z_{T+1}^p, \Delta N_{T+1}^p, \phi_i\right) p\left(V_{T+1}|V_T, Z_T^v, \Delta N_T^v, r_T, Z_T^p, \Delta N_T^p, \phi_i\right)
$$
$$
\times p\left(V_T, Z_T^v, \Delta N_T^v|r_{T-n+1:T}, \ln BV_{T-n+1:T}, \Delta N_{T-n+1:T}^p, Z_{T-n+1:T}^p, \phi_i\right)
$$
$$
\times p\left(Z_{T+1}^p|\Delta N_{T+1}^p, V_{T+1}, \phi_i\right) p\left(\Delta N_{T+1}^p|\Delta N_{T-n+1:T}^p, \phi_i\right) d\left(Z_{T+1}^p, \Delta N_{T+1}^p, V_{T+1}, V_T, Z_T^v, \Delta N_T^v, \phi_i\right).
$$
$$(44)$$

where ϕ_i denotes the vector of static parameters associated with model $\mathcal{M}_i$. As is now standard, the integration in (44) is computed by posterior simulation, using the output of the MCMC algorithm associated with the relevant model, $\mathcal{M}_i$. In addition, note that the predictive distribution in (44) relates to the return on day $T+1$, and is computed marginal of a possible price jump occurrence, ΔN_{T+1}^p, and its corresponding price jump size, Z_{T+1}^p.

---

[2]The marginal likelihoods reported in the previous section imply that the posterior model probability of $\mathcal{M}_2$ is virtually equal to one, while the posterior probabilities for the other seven models are negligible. This indicates that there would be no meaningful difference between the predictive results conditional on $\mathcal{M}_2$ and those produced via Bayesian model averaging. Hence, the latter is not entertained in this section.

We begin our assessment by highlighting the features of the predictive return distributions implied by the (four) competing models. Figure 5 depicts the predictive variance (Panel A), predictive skewness (Panel B) and predictive kurtosis (Panel C) for all four models over the out-of-sample period. The predictive moments are computed as the sample moments of the posterior draws obtained from the predictive distribution in (44). Panel A also plots the short-term (fortnight maturity) option-implied volatility index (VXST), expressed as an annualized variance quantity.

The predictive variance associated with the best fitting Hawkes model, $\mathcal{M}_2$, is seen to be larger overall than that of the three alterative models. Further comparing the predictive variance from $\mathcal{M}_2$ with the VXST, the former is seen to track the dynamics in the latter relatively well, with correlation of 0.568 over the evaluation period. Perhaps not surprisingly, due to the fact that implied volatility indices can be viewed as a market indicator, there are also significant cross-autocorrelations between the two quantities, with the VXST tending to lead the predictive variance of the returns.[3] The predictive skewness implied by $\mathcal{M}_2$ is negative at most time points, while the other three models produce predictive skewness values that tend to fluctuate more around zero, and with the Heston model, $\mathcal{M}_7$, exhibiting the smallest absolute values of skewness overall. The predictive kurtosis of the Heston model also fluctuates closely around the Gaussian value of three, whilst the remaining models generate substantially larger kurtosis values, illustrating the impact of including jumps in the predictive model. Although model $\mathcal{M}_6$ assumes constant jump intensities, it still generates notable non-Gaussian tail behaviour (evident from both Panel B and Panel C), albeit behaviour that exhibits less persistence over time than that of the two models ($\mathcal{M}_2$ and $\mathcal{M}_5$) which specify dynamic models for jumps. For model $\mathcal{M}_2$ in particular, we observe that the deviation from Gaussianity is greater when the predictive variance is low, as also indicated by the high negative correlation between the predictive variance and the predictive skewness (in absolute terms) and kurtosis (-0.23 and -0.72, respectively).

We evaluate the accuracy of (44) - for the four alternative models - using three different criteria. First, the predictive performance of model $\mathcal{M}_2$ relative to a competing model $\mathcal{M}_i$, $i = 5, 6, 7$, is evaluated using the cumulative difference in log score (CLS), computed over

---

[3]Note that we have opted to use the VXST index as the option-implied comparator rather than the VIX index due to the short(er)-term nature of the former. With the VIX index computed from options with a month-long maturity, the effect of jumps (of either type) may be smoothed out. The discrepency between (all four) returns-based variance measures and the option-implied variance measure can be viewed as evidence of variance-related risk premia being factored into option prices (see Bollerslev *et al.*, 2011, and Maneesoonthorn *et al.*, 2012, for relevant illustrations). For more information on the VXST, see www.cboe.com/VXST.
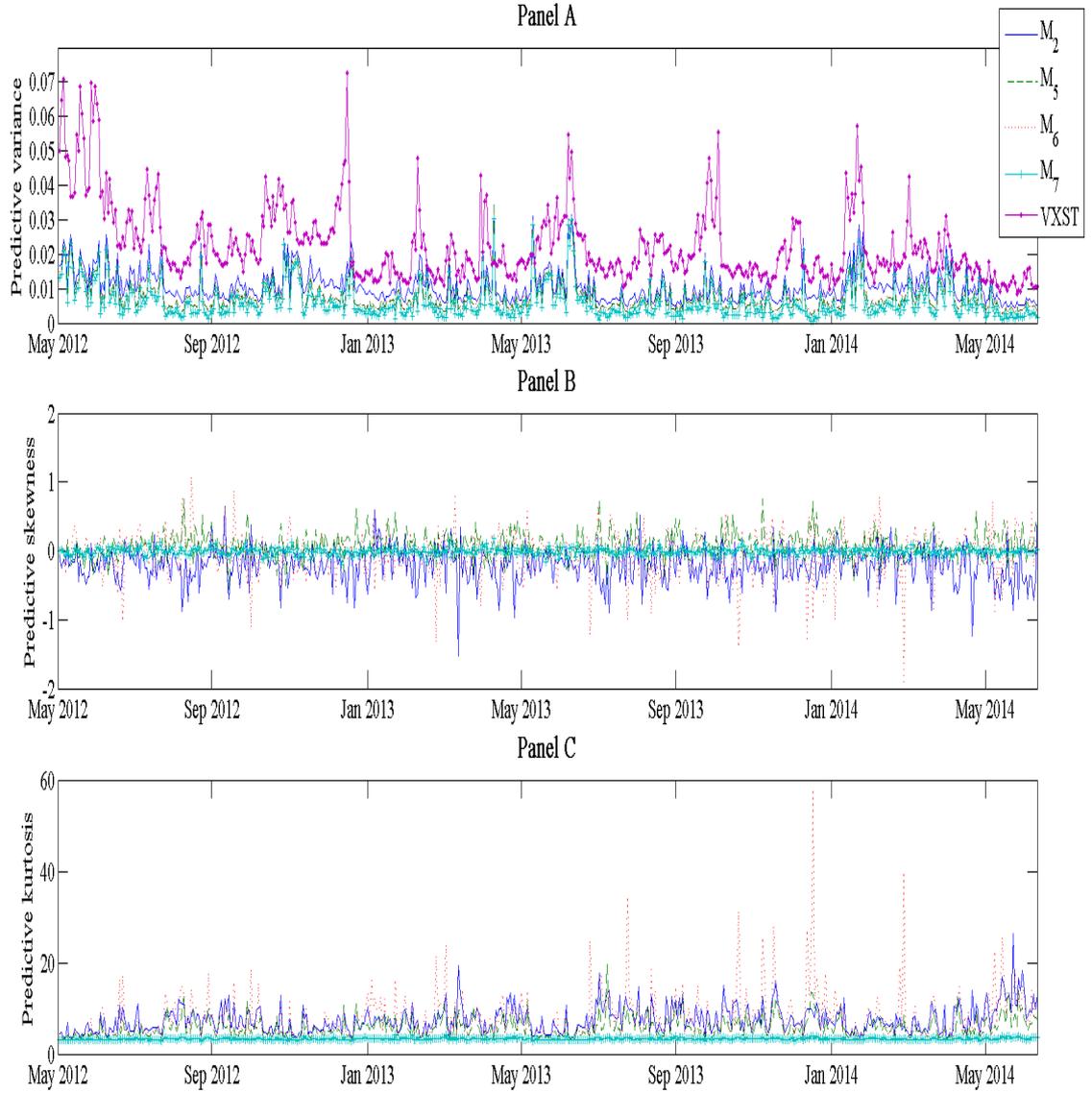
Figure 5: Predictive variance (Panel A), predictive skewness (Panel B) and predictive kurtosis (Panel C) implied by the four competing models over the period of May 16, 2012 to June 23, 2014. The predictive statistics are plotted in solid lines for model $\mathcal{M}_2$ (best fitting Hawkes), dashed lines for model $\mathcal{M}_5$ (jump intensities dependent on V_t), dotted lines for model $\mathcal{M}_6$ (constant jump intensities) and solid lines with cross markers for model $\mathcal{M}_7$ (Heston). Panel A also contains the VXST index, an implied variance measure from short-term option prices.

27

the evaluation period as

$$CLS_i(k) = \sum_{t=T+1}^{T+k} \ln \left[\frac{p\left(r_t^o | r_{k:t-1}, \ln BV_{k:t-1}, \Delta N_{k:t-1}^p, Z_{k:t-1}^p, \mathcal{M}_2\right)}{p\left(r_t^o | r_{k:t-1}, \ln BV_{k:t-1}, \Delta N_{k:t-1}^p, Z_{k:t-1}^p, \mathcal{M}_i\right)} \right], \qquad (45)$$

for $k = 1, ..., 528$ (see also Geweke and Amisano 2010, for discussion of the use of CLS in predictive performance evaluation). The notation $p(r_t^o | r_{k:t-1}, \ln BV_{k:t-1}, \Delta N_{k:t-1}^p, Z_{k:t-1}^p, \mathcal{M}_i)$ denotes the (estimated) predictive density - conditional on model $\mathcal{M}_i$ - evaluated at the observed value of r_t^o, with the conditioning dataset of time subscript $k : t - 1$ always based on the sample size $n = 4070$. A positive value of CLS_i indicates that the reference model, $\mathcal{M}_2$, with the Hawkes jump structure, outperforms the alternative model, $\mathcal{M}_i$.
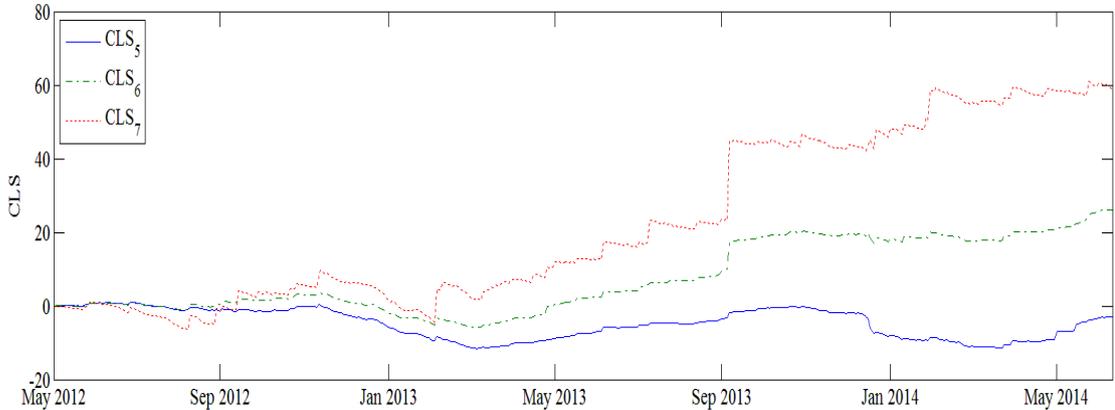

Figure 6: Cumulative log score for the Hawkes model ($\mathcal{M}_2$), relative to models $\mathcal{M}_5$ (CLS_5, solid line), $\mathcal{M}_6$ (CLS_6, dash line) and $\mathcal{M}_7$ (CLS_7, dotted line). The out-of-sample period is May 16, 2012 to June 23, 2014 inclusive.

In Figure 6, the CLS_i of the predictive return distribution over the out-of-sample period is depicted. As can be observed, both CLS_6 and CLS_7 are positive at the final time point, at 26.02 and 58.81, respectively. This provides solid evidence that the best fitting Hawkes model $\mathcal{M}_2$ outperforms the model with constant jump intensities, $\mathcal{M}_6$, and the Heston specification, $\mathcal{M}_7$. Even though the final CLS_5 value of -3.04 indicates that the Hawkes model performs more poorly over the entire period than the model in which the jump intensities are expressed as linear functions of volatility, there are, nevertheless, periods when the Hawkes model is superior, as indicated by periods with a positive slope for the CLS_5 path. These periods are between March 2013 and December 2013, and from May 2014 to the end of the evaluation period.

Secondly, we investigate the coverage of the predictive return distributions. Figure 7 (Panels A and B, respectively) depicts the 95% highest posterior predictive (HPP) intervals

28

for r_{T+1} implied by models $\mathcal{M}_2$ and $\mathcal{M}_7$, along with the observed returns for the forecasting period. Given that the plots for models $\mathcal{M}_2$ and $\mathcal{M}_7$ represent the two extreme cases, the predictive interval plots for models $\mathcal{M}_5$ and $\mathcal{M}_6$ are omitted to conserve space. As to be expected, the prediction intervals for $\mathcal{M}_2$, which models jumps under the Hawkes specification, tend to be wider than those of $\mathcal{M}_7$, a model that excludes the possibility of any jumps. The coverage statistics for the HPP intervals for all four competing models, along with the corresponding unconditional coverage test results (Christoffersen, 1998), are reported in the last column of Table 4. The test results indicate that $\mathcal{M}_2$ is the only model that generates empirical coverage that is statistically insignificantly different from the nominal level of 95%. The empirical coverage of $\mathcal{M}_6$ is statistically different from the nominal coverage at the 10% significance level, while that of models $\mathcal{M}_5$ and $\mathcal{M}_7$ is statistically different at the 5% significance level.

The varying widths of the HPP intervals across models also sheds some light on the previously discussed CLS results. In particular, while the path of CLS_5 indicates that model $\mathcal{M}_2$ is inferior to $\mathcal{M}_5$ overall, in the sense that the predictive density for $\mathcal{M}_5$ is more concentrated around the observed value than is the density for $\mathcal{M}_2$, the coverage statistics illustrate that the predictive density of $\mathcal{M}_5$ is too narrow, and hence does not adequately capture the more extreme values of the future return.

Lastly, and picking up on the previous point made regarding tail coverage, we assess the performance of the four competing models in predicting the one-step-ahead 5% and 1% VaR for the market portfolio associated with the S&P500 market index. In Table 4, we report the p-values of the Christoffersen (1998) tests of correct unconditional coverage and independence of exceedances of the (predicted) VaR. Models that produce forecasts that fail to reject both of these tests are deemed adequate in predicting VaR. Only $\mathcal{M}_2$, a model where both price and variance jumps follow the Hawkes process, provides satisfactory VaR predictions at both the 5% and 1% levels, with the relevant null hypotheses not rejected in all cases. As indicated by the empirical tail coverage statistics, models $\mathcal{M}_5$, $\mathcal{M}_6$, and $\mathcal{M}_7$ all *underestimate* downside risk at both the 5% and 1% VaR levels across this out-of-sample period, with the worst performance given by model $\mathcal{M}_7$.

The results of the three different assessment criteria discussed in this section thus highlight the importance of including both price and variance jumps in this empirical setting and, moreover, confirm the added value of augmenting the basic stochastic volatility structure with the particular dynamic structure for the price and volatility jumps as represented by the Hawkes process.
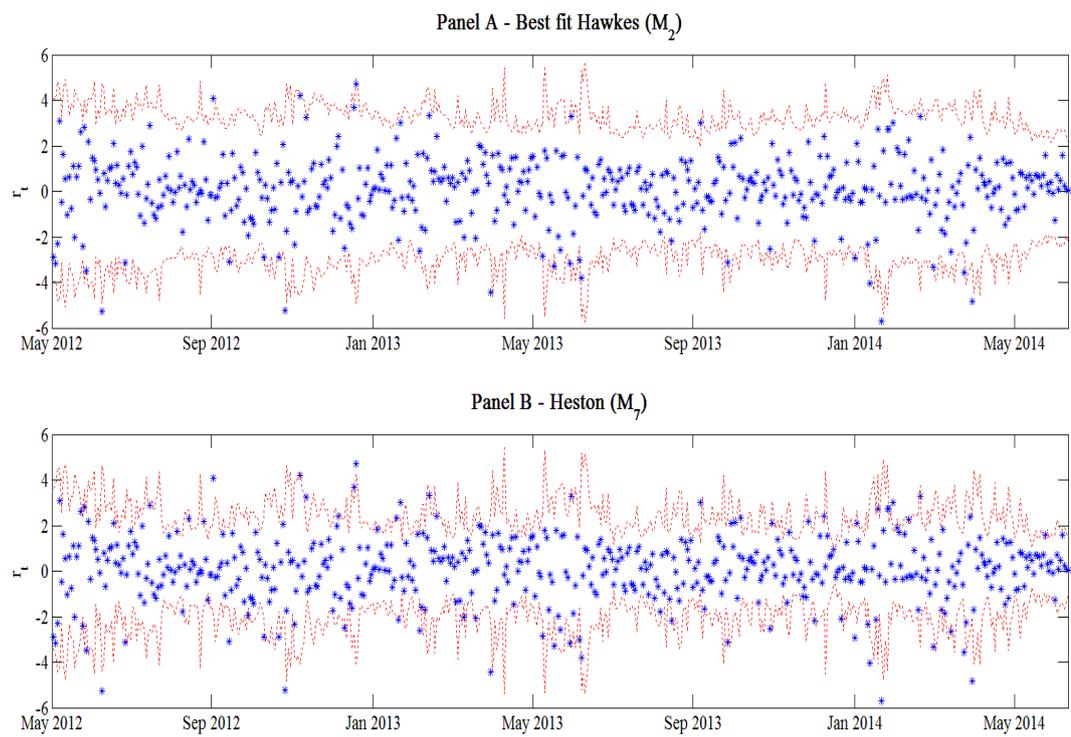
Figure 7: Time series plots of the observed return with the corresponding 95% HPP interval from model $\mathcal{M}_2$ (Panel A) and model $\mathcal{M}_7$ (Panel B) between May 16, 2012 and June 23, 2014.

Table 4: Empirical tail coverage, computed as the proportion of observed returns that are lower than the 5% and 1% VaR predictions, respectively (with p-values for the corresponding unconditional coverage (UC) and independence tests), is given in Columns 3 and 4. The empirical coverage of the 95% HPP interval of the predictive returns distribution is given in Column 5. The subscripts * and ** denote empirical coverage that is statistically different from the nominal coverage at the 10% and 5% significance levels, respectively. All statistics are computed over the out-of-sample period between May 16, 2012 and June 23, 2014, inclusive.

		5% VaR	1% VaR	95% HPP interval
$\mathcal{M}_2$	Empirical coverage	6.06%	0.76%	96.02%
	UC (p-value)	0.28	0.56	
	Independence (p-value)	0.43	0.80	
$\mathcal{M}_5$	Empirical coverage	7.77%	3.03%	92.80%**
	UC (p-value)	0.01	0.00	
	Independence (p-value)	0.13	0.32	
$\mathcal{M}_6$	Empirical coverage	7.95%	2.27%	93.18%*
	UC (p-value)	0.00	0.01	
	Independence (p-value)	0.11	0.45	
$\mathcal{M}_7$	Empirical coverage	9.09%	5.11%	89.58%**
	UC (p-value)	0.00	0.00	
	Independence (p-value)	0.48	0.72	

4.5.2 Prediction of S&P500 index daily futures settlement prices

Here we exploit the relationship between the predictive distribution for the return on the S&P500 market index and the daily closing settlement price of futures contracts written on the index, in order to investigate the performance of the alternative models in predicting the futures price. Applying the cost of carry model, the futures settlement price at time $T + 1$, denoted by F_{T+1}, can be written as a function of the corresponding asset (spot) price P_{T+1} according to

$$F_{T+1} = P_{T+1}e^{r_{f,T}(T_M - (T+1))},$$

where $r_{f,T}$ denotes the risk-free rate of return on day T and T_M denotes the maturity date. Since r_{T+1} denotes the change in the logarithm of the price from the beginning to the end of day $T + 1$, draws from the posterior predictive distribution of the futures settlement price, $F_{T+1}^{(g)}$, can be obtained using draws of the predictive return $r_{T+1}^{(g)}$, for any model $\mathcal{M}_i$, according to deterministic relationship

$$\begin{aligned} F_{T+1}^{(g)} &= P_{T+1}^{(g)}e^{r_{f,T}(T_M - (T+1))} \\ &= P_T e^{r_{T+1}^{(g)}}e^{r_{f,T}(T_M - (T+1))} \\ &= P_T e^{\left\{r_{T+1}^{(g)} + r_{f,T}(T_M - (T+1))\right\}}. \end{aligned}$$

For the purpose of illustration, we obtain the posterior predictive distributions for $F_{T+1}^{(g)}$ associated with a December 2014 maturity. The daily data on US 3-month Treasury Bill rates obtained from the St Louis Federal Reserve Economic Data (FRED) database over the evaluation period is used to represent $r_{f,T}$. With spot returns constructed from market open to market close, we take the opening spot price of the day indexed by $T + 1$ to correspond to P_T. The accuracy of the posterior predictive distribution for $F_{T+1}^{(g)}$ is assessed using the empirical coverage of the corresponding 95% HPP interval, for each of the four alternative models. The resulting coverage statistics, as reported in Table 5, indicate that all four models tend to under-cover the observed daily settlement price. Nevertheless, $\mathcal{M}_2$ produces a coverage level that is closest to the nominal 95% level, indicating the benefit to be derived from the proposed dynamic jump specification, even in this derivative price setting.

5   Conclusions

In this paper, a stochastic volatility model incorporating dynamic behaviour in price and variance (and, hence, volatility) jumps is proposed. The model allows price and variance

32

Table 5: The empirical coverage of the 95% HPP interval of the predictive distribution of the December 2014 S&P500 index futures, with ** denoting empirical coverage that is statistically different from the nominal coverage at the 5% significance level, between May 16, 2012 and June 23, 2014, inclusive.

Model	Coverage of 95% HPP intervals
$\mathcal{M}_2$	89.20%**
$\mathcal{M}_5$	85.04%**
$\mathcal{M}_6$	84.85%**
$\mathcal{M}_7$	77.65%**

jumps to cluster over time, and for the occurrence of a price jump to impact on the likelihood of a subsequent variance jump, via a bivariate Hawkes process. A nonlinear state space model using daily return measurements along with nonparametric measures of volatility and price jumps for the S&P500 market index is constructed, with a hybrid Gibbs-MH MCMC algorithm used to analyze the model. Perhaps not surprisingly, our investigation suggests that the price jump intensity possesses qualitatively different time series behaviour from that of the variance jump intensity. Clusters of inflated price jump intensities are relatively short-lived and scattered throughout the sample period, whilst clusters of high variance jump intensities occur less frequently but persist for longer when they do occur. Furthermore, rises in the intensity of variance jumps are very closely associated with negative market events, whereas as no corresponding link is evident for the price jump intensity.

Several different models that impose various restrictions on the price and variance jump processes are explored using marginal likelihoods, with the overall conclusion being in favour of the model that specifies dynamics in both price and variance jump intensity, but that does not allow for feedback from past price jumps to the variance jump intensity. The dynamic structures imposed on the occurrences of price and variance jumps are also shown to add value to the predictions of both returns on the index and the price of index futures, in an out-of-sample context. Having thus quantified the importance of dynamic jumps - and of respecting the particular nature of the interaction between price and volatility jumps - in the modelling of index returns, such features would appear to deserve more careful attention in future risk management strategies.

Appendix A: MCMC algorithm The MCMC algorithm for sampling from the joint posterior in (34) can be broken down into four main steps:

1. Generating $V_{1:T}$ from

$$
\begin{aligned}
&p\left(V_{1:T}|Z_{1:T}^v, \Delta N_{1:T}^v, r_{1:T}, \ln BV_{1:T}, Z_{1:T}^p, \Delta N_{1:T}^p, \phi\right) \\
&\propto p\left(r_1|V_1, Z_1^p, \Delta N_1^p, \mu, \gamma\right) \times p\left(\ln BV_1|V_1, \sigma_{BV}\right) \\
&\times p\left(Z_1^p|\Delta N_1^p, V_1, \mu_p, \gamma_p, \sigma_p\right) \times p\left(V_1|\kappa, \theta, \sigma_v, \mu, \gamma, \rho\right) \\
&\times \left[\prod_{t=2}^{T} p\left(r_t|V_t, Z_t^p, \Delta N_t^p, \mu, \gamma\right) \times p\left(\ln BV_t|V_t, \sigma_{BV}\right) \times p\left(Z_t^p|\Delta N_t^p, V_t, \mu_p, \gamma_p, \sigma_p\right)\right. \\
&\left.\times p\left(V_t|V_{t-1}, Z_{t-1}^v, \Delta N_{t-1}^v, r_{t-1}, Z_{t-1}^p, \Delta N_{t-1}^p, \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right)\right],
\end{aligned}
$$

2. Generating $\Delta N_{1:T}^v$ (T Bernoulli random variables) from

$$
\begin{aligned}
&p\left(\Delta N_{1:T}^v|V_{1:T}, Z_{1:T}^v, r_{1:T}, \ln BV_{1:T}, Z_{1:T}^p, \Delta N_{1:T}^p, \phi\right) \\
&\propto p\left(\Delta N_1^v|\Delta N_1^p\right) \times \left[\prod_{t=2}^{T} p\left(V_t|V_{t-1}, Z_{t-1}^v, \Delta N_{t-1}^v, r_{t-1}, Z_{t-1}^p, \Delta N_{t-1}^p, \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right)\right. \\
&\left.\times p\left(\Delta N_t^v|\Delta N_{1:t-1}^v, \Delta N_{1:t-1}^p, Z_{1:t-1}^p, \alpha_v, \beta_{vv}, \beta_{vp}, \beta_{vp}^{(-)}, \delta_0^v\right)\right],
\end{aligned}
$$

3. Generating $Z_{1:T}^v$ (T truncated normal random variables) from

$$
\begin{aligned}
&p\left(Z_{1:T}^v|V_{1:T}, \Delta N_{1:T}^v, r_{1:T}, BV_{1:T}, Z_{1:T}^p, \Delta N_{1:T}^p, \phi\right) \\
&\propto p\left(Z_1^v|\mu_v\right) \left[\prod_{t=2}^{T} p\left(Z_t^v|\mu_v\right)\right. \\
&\left.\times p\left(V_t|V_{t-1}, Z_{t-1}^v, \Delta N_{t-1}^v, r_{t-1}, Z_{t-1}^p, \Delta N_{t-1}^p, \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right)\right],
\end{aligned}
$$

4. Generating elements of ϕ from

$$
\begin{aligned}
&p\left(\phi|V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}^v, r_{1:T}, \ln BV_{1:T}, Z_{1:T}^p, \Delta N_{1:T}^p\right) \\
&\propto p\left(V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}^v, \phi|r_{1:T}, \ln BV_{1:T}, Z_{1:T}^p, \Delta N_{1:T}^p\right).
\end{aligned}
$$

The most challenging part of the algorithm is the generation of the state vector $V_{1:T}$, due to the nonlinear functions of V_t that feature in the measurement equations (25) and (28), and in the state equation (29). As in Maneesoonthorn *et al.* (2012) - in which a

nonlinear state space model is specified for both option- and spot-price based measures, and forecasting risk premia is the primary focus - we adopt a multi-move algorithm for the latent volatility that extends an approach suggested by Stroud, Müller and Polson (2003). In the current context this involves augmenting the state space model with mixture indicator vectors corresponding to the latent variance vector $V_{1:T}$ and the two observation vectors $r_{1:T}$ and $\ln BV_{1:T}$. Conditionally, the mixture indicators define suitable linearizations of the relevant state or observation equation and are used to establish a linear Gaussian candidate model for use within an MH subchain. Candidate vectors of $V_{1:T}$ are sampled and evaluated in blocks. With due consideration taken of the different model structure and data types, Appendix A of Maneesoonthorn *et al.* provides sufficient information for the details of this component of the algorithm applied herein to be extracted.

The elements of ϕ are sampled using MH subchains wherever necessary. Given observations on $\Delta N_{1:T}^p$ and $Z_{1:T}^p$, plus draws of $V_{1:T}$ and of all other unknowns that appear in (25), the parameters μ and γ can be treated as regression coefficients, with exact draws produced in the standard manner from a (truncated) Gaussian joint conditional posterior distribution, as a consequence of the previously specified priors. The sampling of σ_{BV} is also standard, given the inverse gamma form of its conditional posterior. As described in Section 3.1, the parameters ρ and σ_v are sampled indirectly via the conditionals of $\psi = \rho\sigma_v$ and $\omega = \sigma_v^2 - \psi^2$, which take the form of normal and inverse gamma distributions, respectively. Conditional upon the draws of $V_{1:T}$, $\Delta N_{1:T}^v$ and $Z_{1:T}^v$, the parameters κ, θ, ψ and ω are drawn in blocks, taking advantage of the (conditionally) linear regression structure with truncated Gaussian errors, and with the constraint $\sigma_v^2 \leq 2\kappa\theta$ imposed.

The parameters associated with the price and variance jump processes are dealt with as follows. The mean of the variance jump size, μ_v, is sampled directly from an inverse gamma distribution, and the unconditional jump intensities, δ_0^p and δ_0^v are sampled directly from beta posteriors. Each of the parameters, $\alpha_p, \beta_{pp}, \alpha_v, \beta_{vv}, \beta_{vp}, \beta_{vp}^{(-)}$, is sampled using an appropriate candidate beta distribution in an MH algorithm, subject to restrictions that ensure that (32) and (33) define stationary processes, and that (16) and (17) are defined on the $[0, 1]$ interval. The intensity parameters δ_∞^p and δ_∞^v are then computed using the explicit relationships in (16) and (17), and the vectors $\delta_{1:T}^v$ and $\delta_{1:T}^p$ updated deterministically based on (32) and (33).

The algorithms for all sub-models described in Section 3.2, $\mathcal{M}_i$, for $i = 1, ..., 7$, proceed in an analogous way.

Appendix B: Marginal likelihood computation The basic idea underlying the evaluation of (42) is the recognition that it can be re-expressed as

$$p\left(r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p | \mathcal{M}_i\right) = \frac{p\left(r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p | \phi_i, \mathcal{M}_i\right) p\left(\phi_i | \mathcal{M}_i\right)}{p\left(\phi_i | r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p, \mathcal{M}_i\right)}, \quad (46)$$

for any point ϕ_i in the posterior support of model M_i, where ϕ_i denotes the vector of static parameters associated with model $\mathcal{M}_i$. The first component of the numerator on the right-hand-side of (46) is the likelihood, conditional on $\mathcal{M}_i$, marginal of the latent variables. That

is,

$$p\left(r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^{p}, Z_{1:T}^{p} | \phi_i, \mathcal{M}_i\right)$$
$$= \int p\left(r_{1:T} | V_{1:T}^{(i)}, Z_{1:T}^{p}, \Delta N_{1:T}^{p}, Z_{1:T}^{v(i)}, \Delta N_{1:T}^{v(i)}, \phi_i, \mathcal{M}_i\right) p\left(\ln BV_{1:T} | V_{1:T}^{(i)}, \phi_i, \mathcal{M}_i\right)$$
$$p\left(Z_{1:T}^{p} | \Delta N_{1:T}^{p}, V_{1:T}^{(i)}, \phi_i, \mathcal{M}_i\right) p\left(\Delta N_{1:T}^{p} | \phi_i, \mathcal{M}_i\right) p\left(V_{1:T}^{(i)} | Z_{1:T}^{v(i)}, \Delta N_{1:T}^{v(i)}, \phi_i, \mathcal{M}_i\right)$$
$$p\left(Z_{1:T}^{v} | \Delta N_{1:T}^{v(i)}, \phi_i, \mathcal{M}_i\right) p\left(\Delta N_{1:T}^{v(i)} | Z_{1:T}^{p}, \Delta N_{1:T}^{p}, \phi_i, \mathcal{M}_i\right) d\left(V_{1:T}^{(i)}, Z_{1:T}^{v(i)}, \Delta N_{1:T}^{v(i)}\right). \qquad (47)$$

The denominator on the right-hand-side of (46) is simply the conditional posterior density of the (static) parameter vector, also marginalized over the latent variables,

$$p\left(\phi_i | r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^{p}, Z_{1:T}^{p}, \mathcal{M}_i\right)$$
$$= \int p\left(\phi_i | r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^{p}, Z_{1:T}^{p}, V_{1:T}^{(i)}, Z_{1:T}^{v(i)}, \Delta N_{1:T}^{v(i)}, \mathcal{M}_i\right)$$
$$p\left(V_{1:T}^{(i)} | Z_{1:T}^{v(i)}, \Delta N_{1:T}^{v(i)}, \phi_i, \mathcal{M}_i\right) p\left(Z_{1:T}^{v(i)} | \Delta N_{1:T}^{v(i)}, \phi_i, \mathcal{M}_i\right)$$
$$p\left(\Delta N_{1:T}^{v(i)} | \Delta N_{1:T}^{p}, \phi_i, \mathcal{M}_i\right) d\left(V_{1:T}^{(i)}, Z_{1:T}^{v(i)}, \Delta N_{1:T}^{v(i)}\right). \qquad (48)$$

The evaluation of (47) at a high density posterior point ϕ_i^* (say, the vector of marginal posterior means for the elements of ϕ_i) is a straightforward use of the output of a full MCMC run for model $\mathcal{M}_i$; namely, the product of the closed form representations of $p\left(r_{1:T} | V_{1:T}^{(i)}, Z_{1:T}^{p}, \Delta N_{1:T}^{p}, Z_{1:T}^{v(i)}, \Delta N_{1:T}^{v(i)}, \phi_i, \mathcal{M}_i\right)$, $p\left(\ln BV_{1:T} | V_{1:T}^{(i)}, \phi_i, \mathcal{M}_i\right)$, $p\left(Z_{1:T}^{p} | \Delta N_{1:T}^{p}, V_{1:T}^{(i)}, \phi_i, \mathcal{M}_i\right)$ and $p\left(\Delta N_{1:T}^{p} | \phi_i, \mathcal{M}_i\right)$ is averaged over the draws of the latent states, $V_{1:T}^{(i)}$, $Z_{1:T}^{v(i)}$ and $\Delta N_{1:T}^{v(i)}$, and computed at the given point ϕ_i^*. Evaluation of (48) is more difficult, in particular when a combination of Gibbs and MH algorithms needs to be employed in the production of draws of ϕ_i. Defining $y = (r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^{p}, Z_{1:T}^{p})'$, and exploiting the structure of posterior density, we decompose $p\left(\phi_i^* | y, \mathcal{M}_i\right)$ into five constituent densities as:

$$p\left(\phi_i^* | y, \mathcal{M}_i\right) = p\left(\phi_{1i}^* | y, \mathcal{M}_i\right) p\left(\phi_{2i}^* | \phi_{1i}^*, y, \mathcal{M}_i\right) ... p\left(\phi_{5i}^* | \phi_{1i}^*, \phi_{2i}^*, ..., \phi_{4i}^*, y, \mathcal{M}_i\right), \qquad (49)$$

where $\phi_{1i} = (\sigma_{BV}, \mu_v, \delta_0^p, \delta_0^v, \rho, \sigma_v, \mu_p)$, $\phi_{2i} = (\alpha_p, \alpha_v, \kappa, \gamma_p)$, $\phi_{3i} = (\beta_{pp}, \beta_{vv}, \theta, \sigma_p)$, $\phi_{4i} = (\beta_{vp}, \mu)$, and $\phi_{5i} = \left(\beta_{vp}^{(-)}, \gamma\right)$. Following the methods outlined by Chib (1995) and Chib and Jeliazcov (2001), five additional MCMC chains, each of which involves a different level of conditioning and, hence, a reduced number of free parameters, are then run to estimate each of the last five components of (49), in turn evaluated at ϕ_{ji}^*, $j = 2, ..., 5$. The first component on the right hand side of (49), involving no such conditioning, is estimated from the output of the full MCMC chain, in the usual way. More details of the relevant reduced algorithms are available from the authors on request.

References

[1] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R.J.A. (2014), "Modeling Financial Contagion Using Mutually Exciting Jump Processes," *Journal of Financial Economics*, forthcoming.

[2] Aït-Sahalia, Y., Fan, J. and Li, Y. (2013), "The Leverage Effect Puzzle: Disentangling Sources of Bias at High Frequency," *Journal of Financial Economics*, 109, 224-249.

[3] Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2007), "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," *The Review of Economics and Statistics*, 89, 701-720.

[4] Barndorff-Nielsen, O.E. and Shephard, N. (2002), "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society B*, 64, 253-280.

[5] ———— (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1-37.

[6] ———— (2006) "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4, 1-30.

[7] Bates, D.S. (1996), "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69-107.

[8] ———— (2000), "Post-87 Crash Fears in the S&P 500 Futures Option Market," *Journal of Econometrics*, 94, 181-238.

[9] Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.

[10] Bollerslev, T., Gibson, M. and Zhou, H. (2011), "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities," *Journal of Econometrics*, 160, 235-245.

[11] Bollerslev, T., Litvinova, J. and Tauchen, G. (2006), "Leverage and Volatility Feedback Effects in High-Frequency Data," *Journal of Financial Econometrics*, 4, 353-384.

[12] Bollerslev, T., Sizova, N. and Tauchen, G. (2012), "Volatility in Equilibrium: Asymmetries and Dynamic Dependencies," *Review of Finance*, 16, 31-80.

[13] Broadie, M., Chernov, M. and Johannes, M. (2007) "Model Specification and Risk Premia: Evidence from Futures Options," *The Journal of Finance*, LXII, 1453-1490.

[14] Brownlees, C.T. and Gallo, G.M. (2006) "Financial Econometric Analysis at Ultra-High Frequency: Data Handling Concerns," *Computational Statistics and Data Analysis*, 51, 2232-2245.

[15] Chib, S. (1995), "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, 90, 1313-1321.

[16] Chib, S. and Jeliazkov, I. (2001), "Marginal Likelihood from the Metropolis-Hastings Output," *Journal of the American Statistical Association*, 96, 270-281.

[17] Christoffersen, P. F. (1998), "Evaluating Interval Forecasts," *International Economic Review*, 39, 841-862.

[18] Creal, D.D. (2008), "Analysis of Filtering and Smoothing Algorithms for Levy-driven Stochastic Volatility Models," *Computational Statistics and Data Analysis*, 52, 2863-2876.

[19] Cvitanic, J., Polimenis, V. and Zapatero, F. (2008), "Optimal Portfolio Allocation with Higher Moments," *Annals of Finance*, 4, 1-28.

[20] Duffie, D., Pan J. and Singleton, K. (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343-1376.

[21] Engle, R.F. and Ng, V.K. (1993), "Measuring and Testing the Impact of News on Volatility," *The Journal of Finance*, 48, 1749-1778.

[22] Eraker, B. (2004), "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," *The Journal of Finance*, LIX, 1367-1403.

[23] Eraker, B., Johannes, M. and Polson, N. (2003), "The Impact of Jumps in Volatility and Returns," *The Journal of Finance*, LVIII, 1269-1300.

[24] Fulop, A., Li, J. and Yu, J. (2014), "Self-Exciting Jumps, Learning, and Asset Pricing Implications," *Review of Financial Studies*, forthcoming.

[25] Geweke, J. (2005), *Contemporary Bayesian Econometrics and Statistics*, Wiley, New York.

[26] Geweke, J. and Amisano, G. (2010), "Comparing and Evaluating Bayesian Prediction Distributions of Asset Returns," *International Journal of Forecasting*, 26, 216-230.

[27] Harvey, C.R., Liechty, J.C., Liechty, M.W. and Müller, P. (2010), "Portfolio Selection with Higher Moments," *Quantitative Finance*, 10, 469-485.

[28] Hawkes, A.G. (1971a), "Spectra of Some Self-Exciting and Mutually Exciting Point Processes," *Biometrika*, 58, 83-90.

[29] ——— (1971b), "Point Spectra of Some Mutually Exciting Point Processes," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 33, 438-443.

[30] Heston, S.L. (1993), "A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *The Review of Financial Studies*, 6, 327-343.

[31] Huang, X. and Tauchen, G. (2005), "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3, 456-499.

[32] Jacod, J. and Todorov, V. (2010), "Do Price and Volatility Jump Together?," *Annals of Applied Probability*, 20, 1425-1469.

[33] Jacquier, E. and Miller, S. (2010), "The Information Content of Realized Volatility," *Working Paper*, HEC, University of Montreal.

[34] Jacquier, E., Polson, N.G. and Rossi, P.E. (2004), "Bayesian Analysis of Stochastic Volatility Model with Fat-tails and Correlated Errors," *Journal of Econometrics*, 122, 185-212.

[35] Jensen, M.J. and Maheu, J.M. (2013), "Risk, Return and Volatility Feedback: A Bayesian Nonparametric Analysis," *Working Paper*. Available at www.economics.utoronto.ca/index.php/index/research/downloadSeminarPaper/34011.

[36] Koopman, S.J. and Scharth, M. (2013), "The Anallysis of Stochastic Volatility in the Presence of Daily Realized Measure," *Journal of Financial Econometrics*, 11, 76-115.

[37] Liao, Y., Anderson, H. and Vahid, F. (2010), "Do Jumps Matter? Forecasting Multivariate Realized Volatility Allowing for Common Jumps," *Working Paper*, *Monash University Department of Econometrics and Business Statistics*. Available at http://www.buseco.monash.edu.au/ebs/pubs/wpapers/2010/wp11-10.pdf.

[38] Maheu, J.M. and McCurdy, T.H. (2004), News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns, *The Journal of Finance,* LIX, 755-793.

[39] Malik, F. (2011), "Estimating the Impact of Good News on Stock Market Volatility," *Applied Financial Economics*, 21, 545-554.

[40] Maneesoonthorn, W., Martin, G.M., Forbes, C.S. and Grose, S. (2012), "Probabilistic Forecasts of Volatility and its Risk Premia," *Journal of Econometrics*, 171, 217-236.

[41] Pan, J. (2002), "The Jump-risk Premia Implicit in Options: Evidence from an Integrated Time-series Study," *Journal of Financial Economics*, 63, 3-50.

[42] Stroud, J.R., Müller, P. and Polson, N.G. (2003), "Nonlinear State-space Models with State-Dependent Variances," *Journal of the American Statistical Association*, 98, 377-386.

[43] Tauchen G. and Zhou, H. (2011), "Realized Jumps on Financial Markets and Predicting Credit Spreads," *Journal of Econometrics*, 160, 102–118

[44] Todorov, V. and Tauchen, G. (2011), "Volatility Jumps," *Journal of Business and Economic Statistics*, 29, 356-371.

[45] Yu, B. and Mykland, P. (1998), Looking at Markov Samplers through CUSUM Path Plots: A Simple Diagnostic Idea. *Statistics and Computing*, 8, 275-286.