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

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

March 7, 2016

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. An evaluation of marginal likelihoods for the proposed model relative to a large number of alternative models, including some that have featured in the literature, is provided. An extensive empirical investigation is undertaken using data on the S&P500 market index over the 1996 to 2014 period, 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.*

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, and where the contemporaneous occurrence of the two types of jumps (i.e. the occurrence of 'co-jumps') is imposed. Under this specification, large fluctuations in price tend to occur in successive periods following a (contemporaneous) jump in price and volatility, again due to persistence in the volatility process. This impact is exacerbated by the fact that the expected price jump size is assumed to be conditionally (positively) dependent on the magnitude of the latent variance jump. Broadie, Chernov and Johannes (2007) also specify co-jumps, but impose independence between the sizes of the two different types of jump. Eraker, Johannes and Polson (2003), Chernov, Gallant, Ghysels and Tauchen (2003) and Eraker (2004) use more general specifications, in which both non-contemporaneous jumps and correlated jump sizes are accommodated, although insignificant correlation between the price and variance jump sizes is documented in all cases.

More recently, volatility and jump measures constructed from high frequency data have been used to investigate price and variance jumps, including the relationship between them. 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 a certain proportion of those jumps occurring simultaneously for the S&P500 market index. Jacod, Klüppelberg and Müller (2013) use high frequency data to explore the correlation between (imposed) co-jumps for several series, but fail to reject the null hypothesis of zero correlation in the majority of cases considered.

As highlighted clearly by Bandi and Reno (2016), however, the use (or otherwise) of op-

tion price data (and the associated risk premia specifications) in past analyses, plus the very nature of the volatility filter adopted (and data frequency exploited in the measurement of volatility), is likely to have had an impact on conclusions drawn regarding the joint evolution of a price and its variance, including discontinuities therein; with such considerations possibly underlying the inconclusive results recorded. We speculate that the rather restricted manner in which the *dynamics* in jumps have been modelled may also have played a role here.

With this background in mind, we propose a very general model for the joint evolution of price and volatility in which both processes are permitted to jump, co-jumps are possible (but not imposed), and both jump processes are allowed to be dynamic. To this end, we adopt a bivariate Hawkes process (Hawkes, 1971a,b) for the *intensity* of price and variance (accordingly volatility, defined as the square root of the variance) jumps, with both jump processes being (potentially) self-exciting as a consequence; that is, the intensity of each jump process is functionally dependent on the realized past increments of that process. We allow the variance jump intensity to depend on past price jumps, enabling extreme price movements to influence the occurrence of extreme movements in volatility. Possible leverage effects operating at the level of extreme price and volatility movements are also 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. Three measures constructed from high frequency data, in addition to the daily return measure, are used to define the multiple measurement equations. The high frequency measures represent observed (price) jump occurrences and size, plus (logarithmic) bipower variation. A Bayesian analysis of the model is undertaken using a Markov chain Monte Carlo (MCMC) algorithm that accommodates the numerous sources of non-linearity in the state space model, and that samples the latent diffusion variances efficiently in blocks. The conditionally deterministic (Hawkes) specification for the jump intensities is computationally convenient, with the posterior distribution of both intensities at any time point - including future time points - able to be estimated from the MCMC draws of the parameters and latent variables to which the intensities are functionally related.

Application of the methodology to data on the S&P500 index for the period January 1996 to June 2014 is documented in detail. The empirical analysis includes the calculation of marginal likelihoods for evaluating the proposed specification against multiple alternatives, most of which are nested within our general state space model and many of which share features with (or, indeed, coincide with) models that have featured prominently in the literature. Predictive distributions are also computed, for the purpose of out-of-sample assessments. The comparative models include those in which the price and variance jump intensities are dynamic as a consequence of a functional dependence (either linear or non-linear) on the level of volatility. Two realized generalized autoregressive conditional heteroscedasticity (RGARCH) specifications of Hansen, Huang and Shek (2012) are also entertained, as

alternatives to the state space form.

As in Bandi and Reno (2016) spot price data only is used to analyse all models, with the results unaffected as a consequence by the nature of - and potential dynamics in - volatility and jump risk premia (see Bollerslev, Gibson and Zhou, 2011, and Maneesoonthorn, Martin, Forbes and Grose, 2012, for analyses in which such specifications do feature). However, and in contrast with Bandi and Reno, data measured at the daily frequency (including that which aggregates to the daily level over intraday observations) underpins the analysis. In common with the large part of the relevant literature (Bollerslev, Kretschmer, Pigorsch and Tauchen, 2009, and Liu, Patton and Sheppard, 2015, amongst many others) we also choose to construct all measures using within-day observations only, thereby avoiding the need to model close-to-open movements in the index (as in, for example, Ahoniemi, Fuertes and Olmo, 2015, and Andersen, Bollerslev and Huang, 2011) and any specific dynamic movements therein. (See Hansen and Lunde, 2005, and Takahashi, Omori and Watanabe, 2009, for earlier discussions on the role played by non-trading periods in the construction of high frequency measures).

The remainder of the paper is organized as follows. Section 2 describes our proposed asset price model and its main properties. The continuous time representation is presented first, followed by the discrete time state space structure adopted for inference. Details are given of the high frequency measures of volatility and price jumps that are used to supplement daily returns in defining the state space model. The Bayesian inferential approach is then outlined in Section 3, including the way in which the alternative specifications are to be assessed, relative to the most general model, both in terms of marginal likelihoods and cumulative log scores. Results from the extensive empirical analysis of the S&P500 index are presented and discussed in Section 4. The benefits of allowing for a very flexible dynamic specification for price and variance jumps are confirmed by both the within-sample and predictive assessments, with the bivariate Hawkes specification given strong support by the data, relative to other more restrictive models. The empirical results also indicate that two jump intensity processes differ in terms of their time series behaviour. Most notably, the variance jump intensity is much more closely aligned with market conditions, exhibiting its most dramatic increase at the peak of the global financial crisis in late 2008. Section 5 provides some conclusions. Certain technical results, including algorithmic and prior specification details, are included in appendices to the paper.

4

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 $\sigma_v^2 \leq 2\kappa\theta$ 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 a differential impact (as compared with a positive price jump) 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. The parameters δ_∞^i, $i = \{p, v\}$, are the steady state levels of the respective intensity processes to which the intensities revert once the impact of excitation dissipates. See Hawkes (1971a,b)

for seminal discussions regarding self-exciting point processes, and Aït-Sahalia, Cacho-Diaz and Laeven (2015) 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 assumption of contemporaneous price and volatility jumps as imposed, for example, by Duffie *et al.* (2000), Broadie *et al.* (2007) and Bandi and Reno (2016). Instead, price and volatility jumps are governed by separate, but dependent, dynamic random processes, such that the two types of jumps may or may not coincide. As detailed below, the probability of co-jumps can be readily computed from the MCMC output, as can the posterior distributions for the magnitude of both types of jumps (whether coincident or not). The specification can also be viewed as an extension of the stochastic volatility model of Aït-Sahalia *et al.* (2015), 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 *et al.* (2014), in which price jump intensity (only) is characterized by a Hawkes process, along with the 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 state space representation of the continuous time model for the asset price, applying an Euler discretization to (1) through (6) with $\Delta t = 1/252$ (equivalent to one trading day). Given the complexity of the proposed model, and the multiple features on which we wish to draw inference, we supplement the daily return measure, defined as

$$r_t = p_{t+1} - p_t,$$

where p_t denotes the logarithm of the asset price at the end of day t, with three additional measures computed from high frequency (intraday) returns. For expositional clarity we begin, in this section, by focusing on the measurement equation for the return only, describing in detail the latent components that feature therein. In Section 2.3 we then introduce the high frequency quantities that are used to define the three additional measurement equations, making clear the assumed link between observed and latent quantities. In Section 2.4, we collect all components of the model together, introducing appropriate labelling to facilitate subsequent referencing.

We begin then with the measurement equation based on the daily return,

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

where the variation in r_t is driven by the latent diffusive volatility process V_t and the latent price jump component $Z_t^p \Delta N_t^p$. The (daily) evolution of diffusive volatility is given by

$$V_{t+1} = \kappa\theta + (1-\kappa) V_t + \sigma_v \rho \left(r_t - Z_t^p \Delta N_t^p - \mu - \gamma V_t \right) + \sigma_v \sqrt{(1-\rho^2) V_t}\xi_t^v + Z_t^v \Delta N_t^v, \tag{8}$$

with the leverage parameter, ρ, taken into account explicitly. The error components ξ_t^p and ξ_t^v, in (7) and (8), respectively, are defined as marginally serially independent $N(0,1)$ sequences, with $corr\left(\xi_t^p, \xi_t^v\right) = 0$ for each t.

The latent occurrences of price and volatility jumps on day t are expressed as

$$\Delta N_t^p \sim Bernoulli(\delta_t^p) \tag{9}$$

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

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,

$$\delta_t^p = \alpha_p \delta_\infty^p + (1 - \alpha_p)\,\delta_{t-1}^p + \beta_{pp}\Delta N_{t-1}^p \tag{11}$$

$$\delta_t^v = \alpha_v \delta_\infty^v + (1 - \alpha_v)\,\delta_{t-1}^v + \beta_{vv}\Delta N_{t-1}^v + \beta_{vp}\Delta N_{t-1}^p + \beta_{vp}^{(-)}\Delta N_{t-1}^{p(-)}. \tag{12}$$

The discretized 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 (11) 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 $\delta_0^p = E\left(\delta_t^p\right) = E\left(\delta_{t-1}^p\right) = E\left(\Delta N_{t-1}^p\right)$ as

$$\delta_0^p = \frac{\alpha_p \delta_\infty^p}{\alpha_p - \beta_{pp}}. \tag{13}$$

Similarly, the unconditional mean of the variance jump intensity process in (12) is given by $\delta_0^v = E\left(\delta_t^v\right) = E\left(\delta_{t-1}^v\right) = E\left(\Delta N_{t-1}^v\right)$, 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

$$\delta_0^v = \frac{\alpha_v \delta_\infty^v + \beta_{vp}\delta_0^p + \beta_{vp}^{(-)}\pi_p\delta_0^p}{\alpha_v - \beta_{vv}}, \tag{14}$$

where $\pi_p = \Pr\left(Z_t^p < 0\right)$ denotes the probability that the price jump is negative. By substituting into equation (14) the expression for δ_0^p in (13), δ_0^v may be re-expressed as the following function of static parameters,

$$\delta_0^v = \frac{\alpha_v \delta_\infty^v}{\alpha_v - \beta_{vv}} + \frac{\beta_{vp}\alpha_p \delta_\infty^p + \beta_{vp}^{(-)}\pi_p\alpha_p\delta_\infty^p}{(\alpha_v - \beta_{vv})(\alpha_p - \beta_{pp})}.$$

To ensure that $\delta_0^p \in (0,1)$ and $\delta_0^v \in (0,1)$, the restrictions $0 < \delta_\infty^p < \frac{\alpha_p - \beta_{pp}}{\alpha_p}$, $0 < \delta_\infty^v < \frac{\alpha_v - \beta_{vv} - \beta_{vp}\delta_0^p - \beta_{vp}^{(-)}\pi_p\delta_0^p}{\alpha_v}$, $0 < \beta_{pp} < \alpha_p < 1$ and $0 < \beta_{vv} < \alpha_v < 1$ are required. Note that the

parameters used in (7)-(12) are the *discrete time* versions of the corresponding parameters in the continuous time model (1)-(6), but with the same symbols used for notational simplicity.

The size of the latent volatility jump is assumed to be exponentially distributed,

$$Z_t^v \sim Exponential\,(\mu_v)\,,$$

with only positive volatility jumps allowed as a consequence. The latent price jump size, on the other hand, is assumed to be composed of two parts: magnitude and sign, ensuring adequate characterisation of the empirically observed bimodal feature of the measured price jump distribution (see further discussion of this point in Section 4). Specifically, we assume,

$$Z_t^p = S_t^{Z^p} \exp\,(M_t^p)\,,$$

where, with π_p as defined earlier, the random variable,

$$S_t^{Z^p} = \left\{ \begin{array}{ll} -1 & \text{with probability } \pi_p \\ +1 & \text{with probability } (1-\pi_p) \end{array} \right.$$

determines the sign of the price jump, and

$$M_t^p \sim N\left(\mu_p + \gamma_p V_t, \sigma_p^2\right) \tag{15}$$

determines the logarithmic magnitude, with a mean value that is proportional to the underlying volatility V_t.

The factors that drive the return in (7) can be interpreted as follows. Consistent with the empirical finance literature (see, for example, 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 (11)), as well as provoking (or exciting) an increase in the future intensity of variance jumps (via (12)). The threshold parameter $\beta_{vp}^{(-)}$ in (12) 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 (11) that drives

8

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 (12), 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. We return to this point in Section 4.

2.3 Incorporating high frequency measurements of volatility and price jumps

In the spirit of Barndorff-Nielsen and Shephard (2002), Creal (2008), Takahashi *et al.* (2009), Dobrev and Szerszen (2010), Jacquier and Miller (2010), Hansen *et al.* (2012), Maneesoonthorn *et al.* (2012), and Koopman and Scharth (2013), amongst others, we exploit 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. As is now standard knowledge, realized variance, defined by

$$RV_t = \sum_{t < t_i \leq t+1}^{M} r_{t_i}^2, \tag{16}$$

where $r_{t_i} = p_{t_{i+1}} - p_{t_i}$ denotes the ith observed return the over the horizon t to $t + 1$, and there being M such returns, is a consistent estimator of quadratic variation under the assumption of no microstructure noise. (See, for example, Barndorff-Nielsen and Shephard, 2002 and Andersen, Bollerslev, Diebold and Labys, 2003). Quadratic variation, $QV_{t,t+1}$, in turn, captures the variation in the return over the horizon t to $t + 1$ due to both the stochastic volatility and price jump components, with $QV_{t,t+1} = V_{t,t+1} + \mathcal{J}_{t,t+1}^2$, where $V_{t,t+1} = \int_t^{t+1} V_s ds$ 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. With bipower variation,

$$BV_t = \frac{\pi}{2} \sum_{t < t_i \leq t+1}^{M} |r_{t_i}| |r_{t_{i-1}}|, \tag{17}$$

being a consistent measure of $V_{t,t+1}$ (again, in the absence of microstructure noise), the discrepancy between RV_t in (16) and BV_t in (17) serves as a measure of price jump variation, and has, as a consequence, formed the basis of various tests of the significance of jump variation on any particular day; see, for example, Barnorff-Nielsen and Shephard (2004, 2006) and Huang and Tauchen (2005).

Three measurement equations that exploit the information content in RV_t and BV_t are constructed as follows. First we define a measure that indicates the *occurrence* or otherwise

of a price jump on day t,

$$I_t^p = \mathbf{1}\left(Z_{RJ,t} > c_a\right), \qquad (18)$$

where

$$Z_{RJ,t} = \frac{RJ_t}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right) M^{-1} \max\left(1, \frac{TQ_t}{BV_t^2}\right)}}, \qquad (19)$$

$RJ_t = (RV_t - BV_t)/RV_t$ and $c_a = \Phi^{-1}(1-a)$ is the critical value in a standard normal distribution, associated with significance level a. The term TQ_t in the denominator of (19) denotes an estimate of the integrated quarticity, with $Z_{RJ,t}$ having a limiting standard normal distribution under the assumption of no jumps, as a result of the particular standardization used in its definition; see Huang and Tauchen (2005) for details. The indicator function in (18) is then viewed as a noisy measure of the latent price jump occurrence in (9). That is, we specify the measurement equation,

$$I_t^p = \begin{cases} Bernoulli\,(\beta) & \text{if } \Delta N_t^p = 1 \\ Bernoulli\,(\alpha) & \text{if } \Delta N_t^p = 0 \end{cases}.$$

with constant probabilities α and β to be estimated from the data.

Second, we assume that the latent (logarithmic) price jump size in (15) is measured with error by

$$\widetilde{M}_t^p = \ln\left(\widetilde{Z}_t^p\right), \qquad (20)$$

where

$$\widetilde{Z}_t^p = \sqrt{\max\left(RV_t - BV_t, 0\right)}, \qquad (21)$$

by specifying the measurement equation,

$$\widetilde{M}_t^p = M_t^p + \sigma_{M_p}\xi_t^{M_p} \qquad \text{for } \widetilde{Z}_t^p \neq 0,$$

with $\xi_t^{M_p} \sim i.i.d.N\,(0,1)$.[1]

Third, as a direct measure of integrated volatility, BV_t is assumed to bring noisy information about the diffusive volatility process, including any jumps in such a process. Hence (and utilizing BV_t in logarithmic form in order to better justify the assumption of a Gaussian measurement error), we specify a final measurement equation as

$$\ln BV_t = \psi_0 + \psi_1 \ln V_t + \sigma_{BV}\xi_t^{BV},$$

where $\xi_t^{BV} \sim i.i.d.N\,(0,1)$, where $\ln V_t$ is a discretization of $\ln \mathcal{V}_{t,t+1}$ and the estimation of ψ_0 and ψ_1 as free parameters allows for $\ln BV_t$ to be a biased measure of $\ln V_t$.

---

[1]Note that when $\widetilde{Z}_t^p = 0$, which, from (21), occurs when $RV_t - BV_t \leq 0$, we do not view the data as providing any information about price jump size, and with $\widetilde{M}_t^p$ being undefined in this case.

2.4 The full discrete time state space model

For expositional clarity, we collect together here all components of the model, beginning with the four measurement equations:

$$Daily\ return: \quad r_t = \mu + \gamma V_t + \sqrt{V_t}\xi_t^p + Z_t^p \Delta N_t^p \tag{22}$$

$$Price\ jump\ indicator: \quad I_t^p = \begin{cases} Bernoulli\,(\beta) & \text{if } \Delta N_t^p = 1 \\ Bernoulli\,(\alpha) & \text{if } \Delta N_t^p = 0 \end{cases} \tag{23}$$

$$Log\ price\ jump\ size: \quad \widetilde{M}_t^p = M_t^p + \sigma_{M_p}\xi_t^{M_p} \quad (\text{for } \widetilde{Z}_t^p \neq 0) \tag{24}$$

$$Log\ bipower\ variation: \quad \ln BV_t = \psi_0 + \psi_1 \ln V_t + \sigma_{BV}\xi_t^{BV}. \tag{25}$$

The stochastic state processes comprise:

$$Latent\ volatility: \quad V_{t+1} = \kappa\theta + (1-\kappa)V_t + \sigma_v \rho \left(r_t - Z_t^p \Delta N_t^p - \mu - \gamma V_t\right)$$
$$+ \sigma_v \sqrt{(1-\rho^2)V_t}\xi_t^v + Z_t^v \Delta N_t^v \tag{26}$$

$$Latent\ price\ jump\ occurrence: \quad \Delta N_t^p \sim Bernoulli(\delta_t^p) \tag{27}$$

$$Latent\ volatility\ jump\ occurence: \quad \Delta N_t^v \sim Bernoulli(\delta_t^v) \tag{28}$$

$$Latent\ price\ jump\ size: \quad Z_t^p = S_t^{Z^p} \exp\left(M_t^p\right) \tag{29}$$

$$Latent\ volatility\ jump\ size: \quad Z_t^v \sim Exponential\,(\mu_v), \tag{30}$$

where the specification of the latent price jump Z_t^p in (29) is given by the product of two random components, with one relating to the jump direction:

$$S_t^{Z^p} = \begin{cases} -1 & \text{with probability } \pi_p \\ +1 & \text{with probability } (1-\pi_p) \end{cases}$$

and the other relating to the log price jump magnitude:

$$M_t^p \sim N\left(\mu_p + \gamma_p V_t, \sigma_p^2\right).$$

Finally, the two conditionally deterministic states are given by:

$$Price\ jump\ intensity: \quad \delta_t^p = \alpha_p \delta_\infty^p + (1-\alpha_p)\delta_{t-1}^p + \beta_{pp}\Delta N_{t-1}^p \tag{31}$$

$$Volatility\ jump\ intensity: \quad \delta_t^v = \alpha_v \delta_\infty^v + (1-\alpha_v)\delta_{t-1}^v + \beta_{vv}\Delta N_{t-1}^v$$
$$+ \beta_{vp}\Delta N_{t-1}^p + \beta_{vp}^{(-)}\Delta N_{t-1}^{p(-)}. \tag{32}$$

All subsequent referencing of the model makes use of the equation numbering in this section.

3 Bayesian inference

3.1 Overview

For notational convenience, we collectively denote, at time point t, the measurement vector as $Y_t = \left(r_t, I_t^p, \widetilde{M}_t^p, \ln BV_t\right)'$, and the latent state vector as $X_t = \left(V_t, \Delta N_t^p, \Delta N_t^v, S_t^{Z_p}, M_t^p, Z_t^v\right)'$,

with the static parameters also collectively denoted by the vector

$$\phi = (\mu, \gamma, \rho, \mu_p, \gamma_p, \sigma_p, \pi_p, \alpha, \beta, \sigma_{M_p}, \psi_0, \psi_1, \sigma_{BV}, \kappa, \theta, \sigma_v, \mu_v, \delta_0^p, \alpha_p, \beta_{pp}, \delta_0^v, \alpha_v, \beta_{vv}, \beta_{vp}, \beta_{vp}^{(-)})'.$$
(33)

In addition, we denote time-indexed variables generically as, for example, $W_{1:t} = (W_1, ..., W_t)'$ for $t = 1, ..., T$, where $W_{1:0}$ is empty. The joint posterior density associated with the full model in (22)-(32), denoted subsequently by $\mathcal{M}_F$, satisfies

$$p(X_{1:T}, \phi | Y_{1:T}) \propto p(Y_1 | X_1, \phi) p(X_1 | \phi) p(\phi) \left[\prod_{t=2}^{T} p(Y_t | X_{1:t-1}, \phi) \times p(X_t | X_{1:t-1}, \phi) \right]. \quad (34)$$

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 the prior $p(\phi)$ in (34) we use a combination of noninformative and weakly informative distributions for the various elements of ϕ. Other than exploiting the natural groupings of parameters that arise from the regression structures embedded within the model, we adopt *a priori* independence for the individual parameters. The detailed specifications of each component of $p(\phi)$ are documented in Appendix A.

Given the complexity of the state space representation, and the high dimension of the set of unknowns, the posterior indicated by (34) is not available in closed form. Hence, a hybrid of the Gibbs and Metropolis-Hastings (MH) MCMC algorithms is developed to obtain draws of the static parameters and latent variables of interest from the joint posterior distribution, with inference - including the construction of posterior predictive distributions - conducted using those draws. Details of this algorithm, including a reference made to Maneesoonthorn *et al.* (2012) for a full description of the multi-move algorithm adopted for sampling the variance state vector, $V_{1:T}$, are given in Appendix B.1.

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 simpler specifications.

To this end, we consider several competing models, summarized in Table 1, most of which are nested in the full model specification $\mathcal{M}_F$ in (22) to (32), and all of which are to be evaluated empirically in Section 4. First (and with reference to the labelling of models in the left-most column of the table), a model without a threshold component (that is, without the differential impact on variance jump intensity due to the occurrence of *negative* price jumps) is specified as $\mathcal{M}_1$. Next, model $\mathcal{M}_2$ specifies that the occurrence of price jumps has no impact at all on the variance jump intensity, via the removal of both price jump feedback terms from δ_t^v. That price and variance jumps occur contemporaneously, or that the variance does not jump at all, each correspond to further restrictions specified in $\mathcal{M}_3$ and $\mathcal{M}_4$, respectively. Note that $\mathcal{M}_4$ shares some features with the model proposed by

Aït-Sahalia *et al.* (2015), albeit in a single asset setting here. In models $\mathcal{M}_5$ to $\mathcal{M}_7$, and as an alternative to the use of the bivariate Hawkes process, the jump intensities are specified as various functions (both linear and non-linear) of the latent variance, with $\mathcal{M}_5$ sharing some common features with the models adopted in Bates (1996), Pan (2002), and Eraker (2004). In $\mathcal{M}_8$ we then specify constant jump intensities, yielding the stochastic volatility with the independent jumps (SVIJ) model of Duffie *et al.* (2000), whilst in $\mathcal{M}_9$ we consider the absence of both price and variance jumps, with the resultant latent process thereby coinciding with that of the conventional Heston (1993) square root model.

Finally, we provide an alternative to the state space form, entertaining the conditionally deterministic RGARCH(1,1) model of Hansen *et al.* (2012), specified as

$$r_t = \sqrt{h_t} z_t$$

$$h_t = v\left(h_{t-1}, BV_{t-1}\right) \tag{35}$$

$$BV_t = m\left(h_t, z_t, u_t\right), \tag{36}$$

where $z_t \sim N(0,1)$ and $u_t \sim N(0, \sigma_u^2)$, with $v\left(.\right)$ and $m\left(.\right)$ defining the evolution of the deterministic variance and the bipower variation, respectively. In our comparison, we entertain both the linear and the log-linear specifications of RGARCH, denoted by $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$, respectively, with details given in Table 1. These two models are, of course, not nested in the general state space framework, and details of the separate MCMC algorithm used to estimate the relevant posterior densities are provided in Appendix B.2.

To examine the relative merits of these twelve models of interest, computation of their corresponding marginal likelihood values,

$$p\left(Y_{1:T} | \mathcal{M}_i\right), \tag{37}$$

for $i = F$ and $i = 1, ..., 11$, is required. Under the assumption that each of the models is *a priori* equally likely, the posterior odds ratio for the full state space model $\mathcal{M}_F$ relative to any restricted model $\mathcal{M}_i$ is equivalent to the Bayes factor BF_i, given in turn by

$$BF_i = \frac{p\left(Y_{1:T} | \mathcal{M}_F\right)}{p\left(Y_{1:T} | \mathcal{M}_i\right)}. \tag{38}$$

Given Bayes factors BF_i and BF_j, the Bayes factor for model $\mathcal{M}_i$ relative to model $\mathcal{M}_j$ is obtained simply as $BF_{i,j} = BF_j / BF_i$.

Note that both $\mathcal{M}_F$ and the first eight comparator models are estimated using the full set of measurements. However, models $\mathcal{M}_9$, $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$ do not exploit observed price jump information, given the absence of any jumps specified in either the price or latent volatility component, with all three models estimated using only observations on r_t and BV_t as a consequence. This gives us two possibilities regarding the computation of the Bayes factors for these models. Firstly, we can compute the marginal likelihoods using only observations on r_t and BV_t and compare the three marginal likelihoods to each other only. Alternatively, we can augment these marginal likelihoods with an additional factor that caters for the

price jump measurements, computed using priors that are consistent with the imposition of no jumps within the models. These latter (expanded) quantities can then be used in a comparison, via the computation of Bayes factors, with the other eight specifications. We record both forms of results in Section 4.

Table 1: Specification of the full set of models used in the comparative evaluation. All parametric restrictions described herein relate to the parameters of either equations (38) and (39) or equations (43) and (44).

Model	Restrictions	Description
$\mathcal{M}_F$	None	Full state space model: (22) to (32)
$\mathcal{M}_1$	$\beta_{vp}^{(-)} = 0$	Full model without threshold component
$\mathcal{M}_2$	$\beta_{vp}^{(-)} = \beta_{vp} = 0$	Full model without price jump feedback
$\mathcal{M}_3$	$\Delta N_t^p = \Delta N_t^v$	Full model with contemporaneous jumps
$\mathcal{M}_4$	$\Delta N_t^v = 0$	Hawkes model without volatility jumps
$\mathcal{M}_5$	$\begin{cases} \delta_t^p = \alpha_{p_0} + \alpha_{p_1} V_t \\ \delta_t^v = \alpha_{v_0} + \alpha_{v_1} V_t \end{cases}$	State dependent jump intensity: linear
$\mathcal{M}_6$	$\begin{cases} \delta_t^p = \alpha_{p_0} + \alpha_{p_1} V_t + \alpha_{p_2} V_t^2 \\ \delta_t^p = \alpha_{v_0} + \alpha_{v_1} V_t + \alpha_{v_2} V_t^2 \end{cases}$	State dependent jump intensity: quadratic
$\mathcal{M}_7$	$\begin{cases} \delta_t^p = \frac{\exp\left(\alpha_{p_0} + \alpha_{p_1} V_t\right)}{1 + \exp\left(\alpha_{p_0} + \alpha_{p_1} V_t\right)} \\ \delta_t^v = \frac{\exp\left(\alpha_{v_0} + \alpha_{v_1} V_t\right)}{1 + \exp\left(\alpha_{v_0} + \alpha_{v_1} V_t\right)} \end{cases}$	State dependent jump intensity: logistic
$\mathcal{M}_8$	$\delta_t^p = \delta_0^p, \, \delta_t^v = \delta_0^v$	Constant jump intensity
$\mathcal{M}_9$	$\delta_t^p = 0$ and $\delta_t^v = 0$	Stochastic volatility model without jumps
$\mathcal{M}_{10}$	$\begin{cases} h_t = \omega + \beta h_{t-1} + \gamma BV_{t-1} \\ BV_t = \xi + \varphi h_t + \tau_1 z_t \\ \quad + \tau_2 \left(z_t^2 - 1\right) + u_t \end{cases}$	RGARCH model: linear
$\mathcal{M}_{11}$	$\begin{cases} \ln h_t = \omega + \beta \ln h_{t-1} + \gamma \ln BV_{t-1} \\ \ln BV_t = \xi + \varphi \ln h_t + \tau_1 z_t \\ \quad + \tau_2 \left(z_t^2 - 1\right) + u_t \end{cases}$	RGARCH model: log-linear

The marginal likelihood for model $\mathcal{M}_i$ in (37) is challenging to compute, in particular for the state space specifications, which require the calculation of an integral over a very large dimension due to the number of latent variables present. As per Chib (1995) and Chib and Jeliazkov (2001), we estimate the marginal likelihood of each model using the output of a series of auxiliary MCMC algorithms, in addition to the full MCMC algorithm associated with estimation of the given model. Specific details of this computation are provided in Appendix C. A brief explanation of the computation of the marginal likelihoods for the three restricted models ($\mathcal{M}_9$, $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$) is also provided in this appendix.

3.3 Predictive performance

With reference to the *joint* measurement vector at time t, and model $\mathcal{M}_i$, $i = F$ and $i = 1, ..., 11$, the one-step-ahead predictive distribution as based on information up to time $t - 1$ is given by

$$p\left(Y_t|Y_{1:t-1}, \mathcal{M}_i\right) \tag{39}$$
$$= \int p\left(Y_t|Y_{1:t-1}, X_{1:t}, \phi_i, \mathcal{M}_i\right) p\left(X_{1:t}|Y_{1:t-1}, \phi_i, \mathcal{M}_i\right) p\left(\phi_i|Y_{1:t-1}, \mathcal{M}_i\right) dX_{1:t} d\phi_i,$$

where ϕ_i denotes the vector of static parameters associated with model $\mathcal{M}_i$, and $X_{1:t}$ represents the full set of latent variables that feature therein. For the RGARCH models, of course, $X_{1:t}$ is an empty set, so that the integration occurs over ϕ_i only. As is well known (see, for example, Geweke, 2001) the log marginal likelihood for any model $\mathcal{M}_i$, $i = F$ and $i = 1, ..., 11$, computed over the entire sample period 1 to T, may be expressed as the sum of T log marginal *predictive* densities, each associated with $\mathcal{M}_i$, and evaluated *ex-post* at the corresponding observed values:

$$\ln p\left(Y_{1:T}|\mathcal{M}_i\right) = \sum_{t=1}^{T} \ln p\left(Y_t|Y_{1:t-1}, \mathcal{M}_i\right). \tag{40}$$

It follows that the Bayes factor BF_i in (38) may actually be interpreted as providing a measure of predictive accuracy for model $\mathcal{M}_F$, relative to that of model $\mathcal{M}_i$, over the entire sample period. Importantly, the component predictive distributions in (40) do not rely upon any unknown parameters, and reflect the evaluation of predictions made without reference to any future information.

What is absent from the computation of the full sample Bayes factor, however, is any information on the *change* over the sample period in the predictive performance $\mathcal{M}_F$ relative to $\mathcal{M}_i$. To better capture this dynamic predictive behaviour, we compute the cumulative difference in log score (CLS) over an evaluation (sub-) period of $(T - T_0)$ trading days, according to

$$CLS_i\left(n\right) = \sum_{t=T_0+1}^{n} \ln \left[\frac{p\left(Y_t|Y_{1:t-1}, \mathcal{M}_F\right)}{p\left(Y_t|Y_{1:t-1}, \mathcal{M}_i\right)}\right], \tag{41}$$

for $n = T_0 + 1, ..., T$. An increase in the value of $CLS_i(n)$, relative to $CLS_i(n-1)$, indicates an improvement in the performance of the reference model, $\mathcal{M}_F$, relative to $\mathcal{M}_i$, in terms of predicting all four elements of Y_n, with a *persistent* positive level in CLS_i indicating sustained predictive superiority of $\mathcal{M}_F$ relative to $\mathcal{M}_i$, over that period. Note that for the early values in this set of sequential $CLS_i(n)$ calculations to be reliable, an initial sample consisting of T_0 observations, $Y_{1:T_0}$, is used to initialise the computation. (See also Geweke and Amisano, 2010). Estimation of (39) and, subsequently, computation of (41), occurs via a combination of MCMC and particle filtering algorithms, with further details provided in Section 4.4.

To complement the CLS results as they pertain to the joint measurement vector Y_t, in Section 4.4 we also report predictive performance as it relates to certain individual sub-vectors of Y_t. Denote any sub-vector of Y_t by g_t. Then we define the *marginal CLS* for g_t as

$$[g]\,CLS_i\,(n) = \sum_{t=T_0+1}^{n} \ln\left[\frac{p\,(g_t|Y_{1:t-1}, \mathcal{M}_F)}{p\,(g_t|Y_{1:t-1}, \mathcal{M}_i)}\right], \tag{42}$$

again for $n = T_0 + 1, ..., T$. Specifically, we report predictive results for the return, where $g_t = r_t$, and the log bipower variation, where $g_t = \ln BV_t$, as well as for the bi-variate subvector containing the price jump indicator and size components, where

$$g_t = \left(\widetilde{M}_t^p, I_t^p\right)'. \tag{43}$$

Computation of the terms comprising the marginal CLS in (42) for a given g_t requires a relatively small modification of the methodology used to compute the joint quantities in (39) and (41), with details provided in Section 4.4. Isolation of the predictives for individual elements of Y_t also enables us to directly compare the predictive performance of models $\mathcal{M}_9$, $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$ with $\mathcal{M}_F$, in terms of the accuracy with which these two models forecast future values of $g_t = r_t$ and $g_t = \ln BV_t$ specifically. In the case of $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$, given the absence of stochastic latent variables, computation of the predictive quantities using the MCMC draws from the joint posterior is standard, without there being any need for additional filtering steps.

4   Empirical application

4.1   Data description and preliminary analysis

For the empirical analysis documented below, 4598 observations on the open-to-close logarithmic S&P500 return (r_t), price jump indicator (I_t^p), logarithmic price jump size $\left(\widetilde{M}_t^p\right)$ and logarithmic bipower variation ($\ln BV_t$) were analyzed, over the period January 3, 1996 to June 23, 2014. 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 measures constructed from high-frequency data are based on fixed five minute sampling, with a 'nearest price' method (Andersen, Bollerslev and Diebold, 2007) applied to construct the relevant returns five minutes apart, and only index values recorded within the New York Stock Exchange market trading hours. The numerical results reported in this empirical section have been produced using a combination of the JAVA and MATLAB programming languages. Marginal posterior point and interval summaries for the static parameters are reported in Section 4.2 for the full model $\mathcal{M}_F$ only, with results pertaining to all competing models recorded in Table 1 provided in the on-line supplementary appendix.

In Figure 1 we provide a graphical representation of two of the four measures, r_t and BV_t (the latter in both raw and logarithmic form) for the entire sample period, recorded in
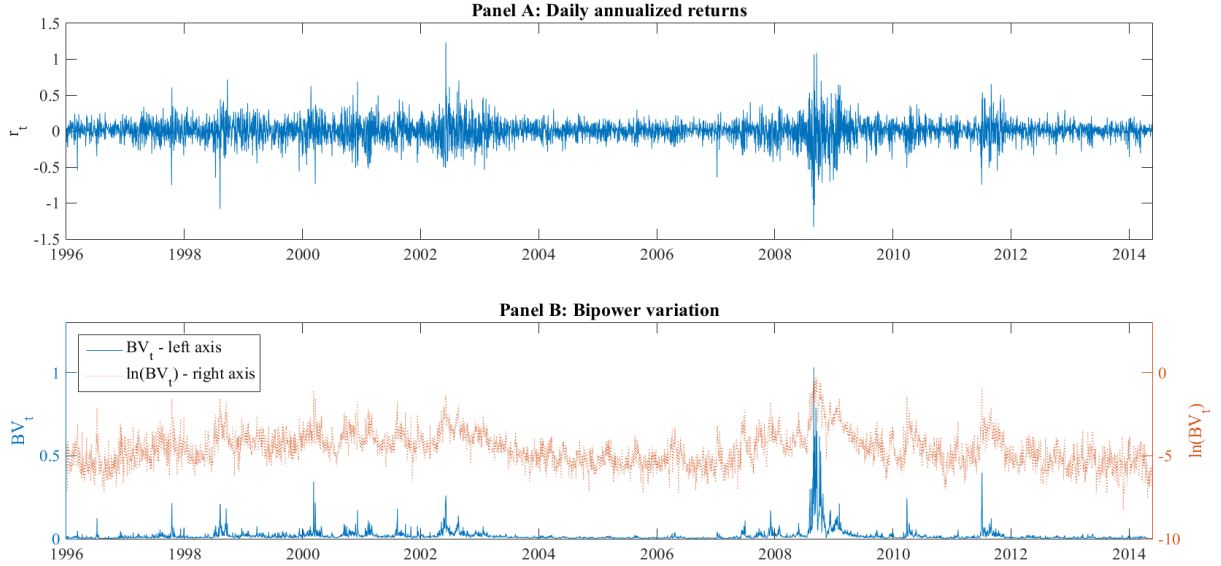
Figure 1: Plots of the S&P500 logarithmic returns (r_t) (Panel A); bipower variation (BV_t) and its logarithm $(\ln BV_t)$ (Panel B) for January 3, 1996 to June 23, 2014.

annualized form. As is evident in both panels of Figure 1, and as is completely expected in this setting, volatility clustering is a marked feature. The most extreme variation in returns, along with the occurrence of BV_t values of unprecedented magnitude, is observed towards the end of 2008. The large jumps observed periodically in BV_t, in addition to the jumps in evidence in the return series itself, plus the tendency for both types of jumps to cluster, all provide motivation for the specification of a dynamic model for both price and volatility jumps.

In Panel A of Figure 2 we plot the time series of the *signed* jump size measure $I_t^p \times \widetilde{Z}_t^p \times sign(r_t)$, with I_t^p and $\widetilde{Z}_t^p$ as defined in (18) and (21) respectively, with the data indicating that price jump intensity is 10.64% on average.[2] Values of the combined measure are indicated on the left-hand-side axis.[3] Distinct variation in the observed price jump size over the sample period, including clusterings of both small and large jumps, is evident, with there being no particular tendency for *negative* price jumps (as identified here simply by the occurrence of a negative return) to predominate over this extended period. Clusters of large jumps appear intermittently over the sample period; however, the clusters 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. The logarithmic measure of price jump magnitude, $\widetilde{M}_t^p$, (as defined in (20)), is also included in Panel A,

---

[2]With reference to (18), I_t^p is defined using a significance level of 0.001, as recommended by Tauchen and Zhou (2011).

[3]We reiterate that the sign of the price jump is modelled as a latent process only (in (29)), and is estimated along with all other unknowns in the model. We do not assume that the sign of the price jump coincides exactly with the sign of the return on that day. We represent the direction of the price jump by the sign of the return for the purpose of this preliminary diagnostic exercise only.

Panel A: Price jump measures

Panel B: Empirical Distribution of I_t^p x $\widetilde{Z}_t^{-p}$ x sign(r_t)
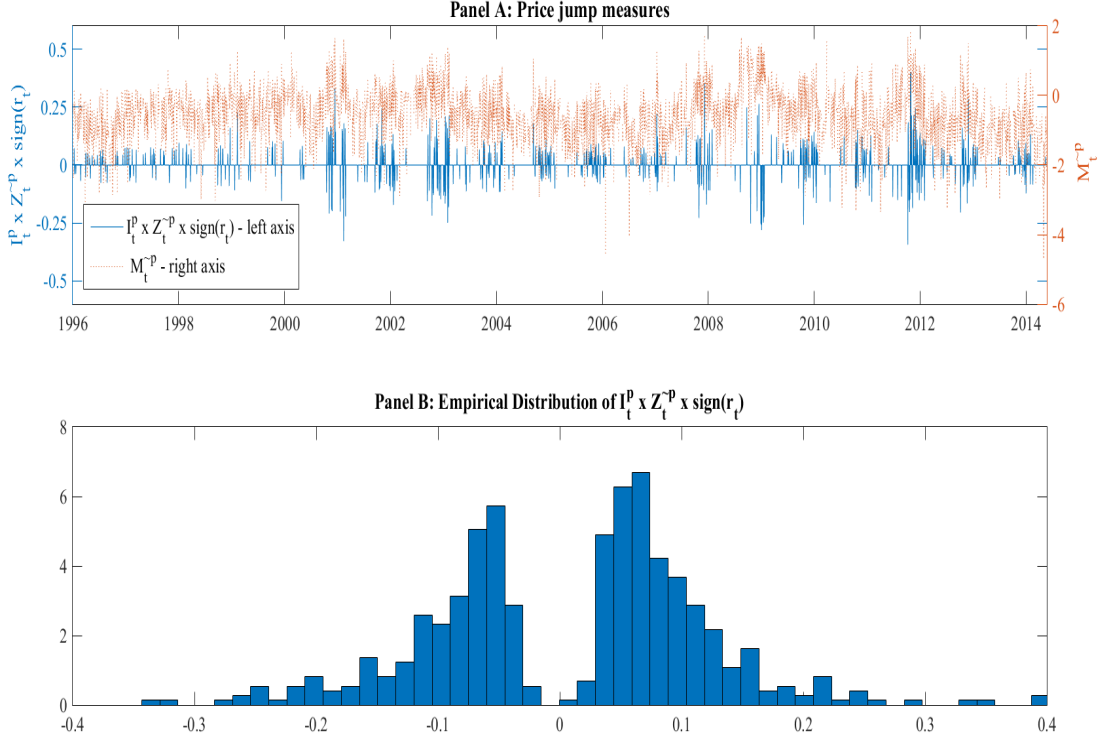
Figure 2: Panel A superimposes two time series plots for January 3, 1996 to June 23, 2014: i) the solid line (and left-hand-side axis) depicts the product of the measure of price jump occurrence (I_t^p), the measure of price jump size $\left(\widetilde{Z}_t^p\right)$, and the sign of the return (r_t); ii) the dotted line (and right-hand-side axis) depicts the logarithmic price jump size measurement $\left(\widetilde{M_t^p}\right)$. Panel B plots the histogram of the empirical distribution of the signed price jump measure $\left(\widetilde{Z}_t^p \times sign\left(r_t\right)\right)$ for days when the price jump indicator signals the presence of a jump.

with values indicated on the right-hand-side axis. The fluctuations in this variable reflect (via the logarithmic transformation of $\widetilde{Z}_t^p$) the changes in the observed price jump size, with changes that are large in magnitude producing large positive values for $\widetilde{M_t^p}$, and very small magnitude changes in $\widetilde{Z}_t^p$ yielding negative values for $\widetilde{M_t^p}$.

Panel B of Figure 2 depicts the histogram of the signed jump size measure, $I_t^p \times \widetilde{Z}_t^p \times sign(r_t)$. As is consistent with the time series plot in Panel A, there is no evidence of negative jumps occurring more frequently than positive jumps throughout the entire sample period. In addition, the empirical distribution is seen to be bimodal, with the very small probability mass in the neighbourhood of zero reflecting the fact that, *conditional* on a significant jump occurring, the size of that jump is, necessarily, bounded away from zero. This bimodal feature of the observed price jump magnitude does not appear to have been recognized in the literature, with a Gaussian distribution typically adopted for the latent variable Z_t^p. (See Eraker *et al.*, 2003, and Tauchen and Zhou, 2011, for example). In contrast, our approach

18

adopts $\widetilde{M}_t^p$ as a (noisy) measure of the latent (log) jump size, M_t^p, only when $\widetilde{Z}_t^p$ is non-zero, and thereby both accommodates this observed bimodality and avoids a Gaussian assumption for Z_t^p itself.[4]

4.2 The implied Hawkes dynamics

To illustrate the dynamic structure implied by our full state space model $\mathcal{M}_F$, we provide here posterior summary information relating to the static parameters, including the parameters of the two jump intensity processes, corresponding to the full sample period. Reported in Table 2 are the marginal posterior means (MPMs) and 95% highest posterior density (HPD) intervals for the static parameters in (33), calculated from 30,000 MCMC draws (following a 30,000 draw burn-in period) of which every 5th 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 summaries are reported in annualized terms where appropriate. For example, the magnitude of the parameter θ accords with an annualized variance quantity, whilst κ reflects the daily persistence in that annualized variance. We also record point and interval estimates of the probability of simultaneous and sequential price and volatility jumps, in the last two lines in the table.

The inefficiency factors reported in Table 2 (for the static parameters) range from 1 to 150, with certain parameters associated with the variance jump intensity producing the highest values. The inefficiency factors for all latent variables, computed at selected time points (and not reported here), range from 3 to 5. 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 chain - being approximately 99%. The convergence of the MCMC chains for all unknowns is also confirmed via inspection of graphical CUSUM plots (Yu and Mykland, 1998), and using the convergence diagnostics prescribed by Heidelberger and Welch (1983) and Geweke (1992).

The parameters associated with the two jump intensity processes are, of course, our primary interest. The dynamic price jump intensity, δ_t^p, possesses a reasonably 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 by Aït-Sahalia *et al.* (2015), 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

---

[4]We are grateful to an anonymous referee who highlighted the need to accommodate this non-Gaussianity in our modelling of the price jump size.

Table 2: Empirical results for the S&P 500 stock index for January 3, 1996 to June 23, 2014, inclusive, for the full state space model, $\mathcal{M}_F$.

Parameter	MPM	95% HPD interval	Inefficiency Factor
μ	0.199	(0.139,0.256)	1.59
γ	-8.628	(-9.955,-5.679)	1.10
ρ	-0.357	(-0.421,-0.289)	6.54
μ_p	-0.419	(-0.435,-0.403)	6.43
γ_p	10.955	(9.967,11.970)	22.83
σ_p	0.207	(0.187,0.226)	13.84
π_p	0.382	(0.297,0.470)	12.76
α	$8.99e^{-4}$	$(2.39e^{-5},3.33e^{-3})$	1.82
β	0.814	(0.633,0.956)	17.17
σ_{M_p}	0.183	(0.162,0.203)	13.47
ψ_0	0.970	(0.796,1.142)	116.60
ψ_1	1.290	(1.255,1.325)	81.16
σ_{BV}	0.436	(0.423,0.450)	5.75
κ	0.116	(0.092,0.167)	45.97
θ	$8.19e^{-3}$	$(7.41e^{-3},9.11e^{-3})$	15.79
σ_v	0.016	(0.014,0.017)	19.25
μ_v	$9.66e^{-3}$	$(8.21e^{-3},0.012)$	48.48
δ_0^p	0.132	(0.108,0.170)	14.96
α_p	0.097	(0.072,0.127)	9.18
β_{pp}	0.062	(0.047,0.079)	12.12
δ_0^v	0.121	(0.082,0.158)	41.23
α_v	0.035	(0.024,0.050)	149.68
β_{vv}	0.030	(0.021,0.043)	134.58
β_{vp}	$5.51e^{-4}$	$(1.33e^{-5},2.00e^{-3})$	1.77
$\beta_{vp}^{(-)}$	$1.14e^{-3}$	$(3.11e^{-5},3.84e^{-3})$	2.34
$\Pr\left(\Delta N_t^v = 1 | \Delta N_t^p = 1\right)$	0.097	(0.059,0.139)	20.35
$\Pr\left(\Delta N_{t+1}^v = 1 | \Delta N_t^p = 1\right)$	0.107	(0.066,0.149)	30.12

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 close to zero. The probability of instantaneous co-jumps, measured by the MCMC-based estimate of $\Pr\left(\Delta N_t^v = 1 | \Delta N_t^p = 1\right)$, is 9.7%, whilst the probability that a volatility jump will follow in the period subsequent to a price jump is 10.7%. Thus, whilst the estimated model discounts the importance of feedback from observed price jumps to volatility jump intensity, it remains flexible enough to capture the phenomenon of both simultaneous - and close to simultaneous - price and volatility jumps, with such events estimated to happen with nearly 20% probability. Further assessment of the importance of the dynamic structures specified for price and variance jumps, and of the presence of jumps *per se*, is conducted in Section 4.3, via a comparison of marginal likelihoods.

It is interesting to note that the value of κ 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[5]. The MPM of the other parameters associated with stochastic volatility, for examples ρ, σ_v, θ and μ_v, are broadly consistent with those reported in the literature (see, for example, Broadie *et al.,* 2007 Maneesoonthorn *et al.,* 2012 and Aït-Sahalia, Fan and Li, 2013), albeit differing slightly in magnitude presumably due to the varying sample periods.

Time series plots of the MPMs and the 95% HPD intervals of both jump intensity processes, δ_t^p and δ_t^v, computed at every time point over the estimation period, are displayed in Panels A and B, respectively, of Figure 3. As is evident from a comparison of the two panels, the dynamics of the price and volatility jumps are quite distinct. Price jump clustering - associated with sustained periods of high values for δ_t^p - occurs intermittently throughout the sample period, and without any obvious tracking of market conditions. 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 2, Panel A) is either large or small. That is, an increase in price jump intensity does not appear to correlate with a period of large price jumps only. The magnitude of price jumps, however, is found to be associated with the level of volatility, with the MPM and 95% HPD interval of the parameter γ_p being in the highly positive region.

In contrast, the variance jumps tend to cluster during high volatility periods specifically, with an increase in marginal posterior mean and 95% HPD intervals associated with δ_t^v

---

[5]This observation is further supported by the posterior results (recorded in the on-line supplementary appendix) for the alternative models listed in Table 1. In brief, diffusive volatility under those specifications with restrictive assumptions about the dynamics in volatility jumps ($\mathcal{M}_3, \mathcal{M}_4, \mathcal{M}_8$ and $\mathcal{M}_9$) is more persistent than otherwise. The unconditional diffusive variance is also larger in magnitude in these cases.
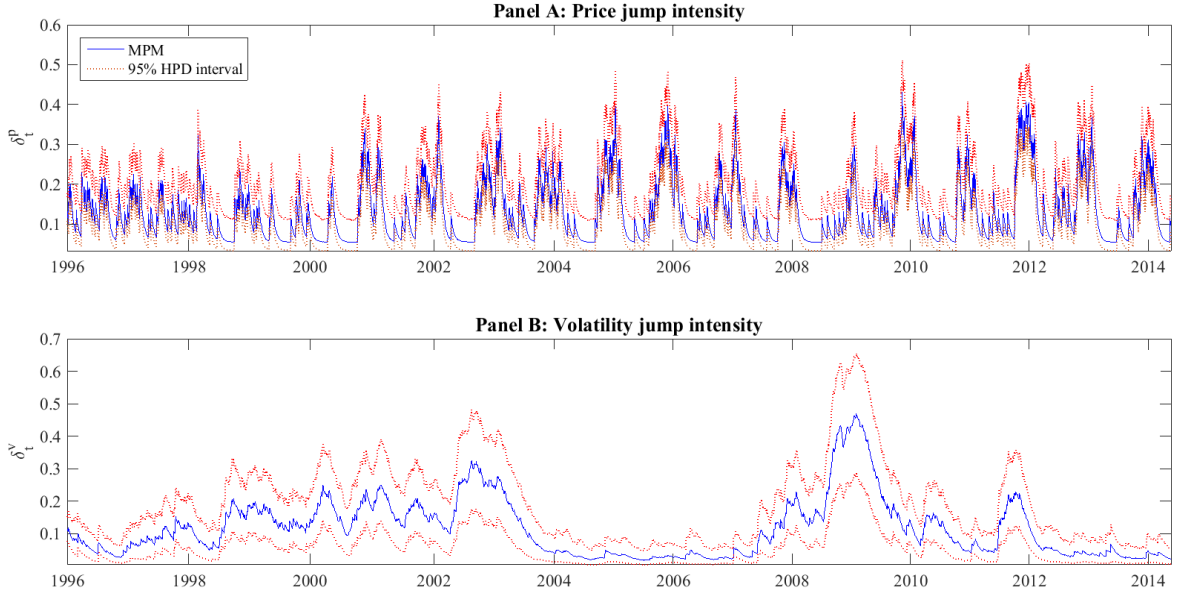
Figure 3: Posterior results for the price jump intensity, δ_t^p, (Panel A) and volatility jump intensity, δ_t^v, (Panel B) over the period of January 3, 1996 to June 23, 2014. The solid blue lines represent the marginal posterior means (MPM), while the 95% HPD intervals are depicted by the dotted red lines.

coinciding with the rises in the observed volatility measure BV_t, as recorded in Panel B of Figure 1. Some of the sharpest rises in δ_t^v are either synchronous with, or occur soon after, certain key events, as illustrated in Figure 4, in which the MPM of δ_t^v is plotted over the 2007-2014 period. 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 (the MPM reaching a peak of 47% on March 18th, 2009). During the various phases of the recent US debt ceiling concerns and the 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. Once a period of multiple variance jumps has passed, the value of δ_t^v declines rather slowly, with this high level of persistence being consistent with the point and interval estimates of α_v recorded in Table 2.[6]

4.3 Model ranking

Table 3 reports the log marginal likelihood of each of the eleven models, $\mathcal{M}_1$ to $\mathcal{M}_{11}$, as well as that of the full model $\mathcal{M}_F$, and as computed over the entire sample period. The Bayes

---

[6]The dynamics of volatility jump intensity implied by models $\mathcal{M}_5$ to $\mathcal{M}_7$ are not dissimilar to those presented here, as all three models assume that the jump intensity is driven by the latent volatility process. The key difference is in the dynamics of the price jump intensity, with the MPMs and 95% HPD intervals of δ_t^p implied by these three models (and as reported in the on-line supplementary document) indicating that the price jump intensity is roughly constant. Such a model-implied feature is obviously inconsistent with the empirical characteristics of the price jump indicator evident in Figure 2 (Panel A).
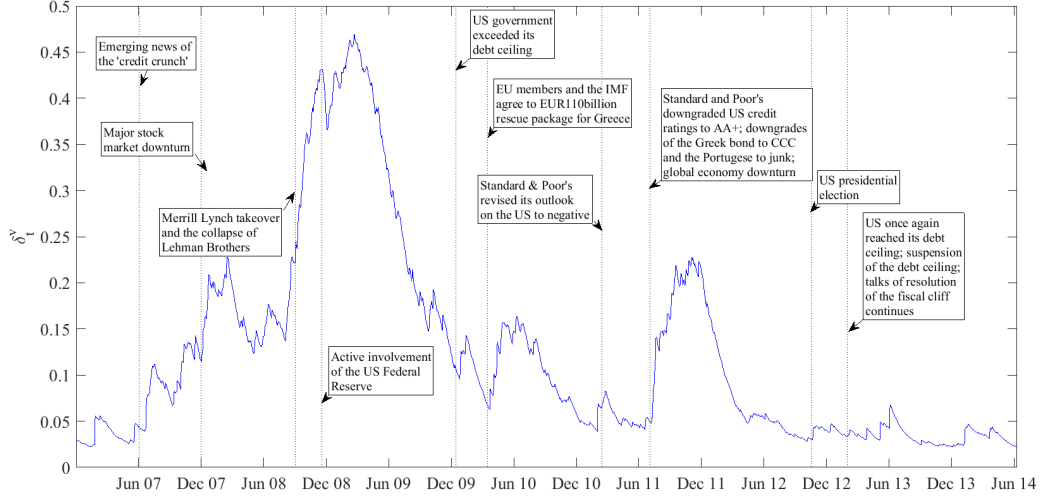
Figure 4: Time series plot of the variance jump intensity process, δ_t^v, over a sub-period of January 3, 2007 to June 23, 2014, inclusive, with the timing of various important market events noted, including the recent global financial crisis, as well as the events related to the US debt ceiling and Euro-zone debt crises.

factor for each of $\mathcal{M}_1$ to $\mathcal{M}_{11}$ relative to $\mathcal{M}_F$ are also computed, as per (38), as the ratio of the marginal likelihood of $\mathcal{M}_F$ to that of $\mathcal{M}_i$, $i = 1, 2, ..., 11$, and are recorded in logarithmic form. We also report the ranking (from one to twelve) of all of these models, as based on their marginal likelihoods values.[7] As noted earlier, the marginal likelihoods of $\mathcal{M}_9$ to $\mathcal{M}_{11}$, are directly comparable to those of the other models only if an extra component (based on the two jump measures) is used to supplement the marginal likelihoods computed directly from the r_t and BV_t measures. These augmented figures are recorded in the middle panel of Table 3. For completeness, we also record in the bottom panel of the table the marginal likelihood based on the r_t and BV_t measures only, with these figures not allowing for a direct comparison with the remaining nine models.

The key message from the results recorded in Table 3 is that the proposed Hawkes specification for both price and volatility jumps is strongly supported by the data. The log marginal likelihood of the full dynamic model $\mathcal{M}_F$ is only inferior when compared against its slightly more restrictive alternatives, $\mathcal{M}_1$ and $\mathcal{M}_2$, which assume no threshold effect and no feedback effect from price to volatility jumps, respectively. This support for $\mathcal{M}_1$ and $\mathcal{M}_2$ is consistent with the fact that most of the posterior mass associated with each of β_{vp} and $\beta_{vp}^{(-)}$ is near zero in the full dynamic model, $\mathcal{M}_F$, as indicated by the MPM and 95% HPD intervals reported in Table 2. The model that *imposes* contemporaneous price and variance jumps ($\mathcal{M}_3$) performs poorly, with the model ranked ninth overall, indeed ranked more lowly than the model in which no volatility jumps at all are allowed ($\mathcal{M}_4$, ranked sixth) and the model in which jumps have a constant intensity ($\mathcal{M}_8$, ranked eight). All three models that

---

[7]As noted in Section 3.2, a series of auxiliary MCMC algorithms is required to compute any given Bayes factor, in addition to the full MCMC algorithm associated with the two models in question. All auxiliary algorithms produce 10,000 draws, after a 10,000 draw burn-in period.

avoid the Hawkes structure in modelling the dynamic intensities ($\mathcal{M}_5$, $\mathcal{M}_6$ and $\mathcal{M}_7$) are ranked below both $\mathcal{M}_F$ and its two closest restricted versions, $\mathcal{M}_1$ and $\mathcal{M}_2$, and the Heston and RGARCH models ($\mathcal{M}_9$, $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$) are the most poorly performing models of all. Of the latter three, when considered in isolation from the remaining models, the logarithmic RGARCH specification ranks the highest but does not provide an explanation of the sample data that is close to any of the models that accommodate jumps.

Table 3: Log marginal likelihoods and model rankings, computed using the data from January 3, 1996 to June 23, 2014, inclusive.

	Model	ln (marginal likelihood)	ln BF_i	Ranking
	$\mathcal{M}_F$	-10024	0	3
	$\mathcal{M}_1$	-9861	-163	1
	$\mathcal{M}_2$	-9957	-67	2
	$\mathcal{M}_3$	-12868	2844	9
	$\mathcal{M}_4$	-10639	615	6
	$\mathcal{M}_5$	-10686	662	7
	$\mathcal{M}_6$	-10618	594	5
	$\mathcal{M}_7$	-10076	52	4
	$\mathcal{M}_8$	-10773	749	8
with	$\mathcal{M}_9$	-27793	17769	10
price jump	$\mathcal{M}_{10}$	-33486	23462	12
measures	$\mathcal{M}_{11}$	-26830	16806	11
without	$\mathcal{M}_9$	-12521	N/A	N/A
price jump	$\mathcal{M}_{10}$	-18214	N/A	N/A
measures	$\mathcal{M}_{11}$	-11558	N/A	N/A

4.4   Predictive comparison

The exercise conducted in the previous section documents the relative performance of the alternative models over the full sample period. In the current section, we compute the 'joint' CLS in (41) and the three marginal CLS values discussed in Section 3.3, over a more recent period only, with a training sample used to initialize the computation. Once again we use the full model $\mathcal{M}_F$ as the reference model, but this time conduct a comparison of it only against those alternative models that are most distinct from it, namely: $\mathcal{M}_4$, in which a Hawkes structure is adopted for price jumps but volatility jumps are omitted; $\mathcal{M}_5$, in which a linear (non-Hawkes) dynamic structure is adopted for the intensities;[8] $\mathcal{M}_8$, in which the jump intensities are constant; $\mathcal{M}_9$, in which no jumps at all are modelled within the state space structure; and $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$, which adopt conditionally deterministic specifications for the variance and also eschew jumps.

The first $T_0 = 2500$ observations in the data set are used to produce, for each model

---

[8]The relative predictive performances of models $\mathcal{M}_6$ and $\mathcal{M}_7$, in which non-linear functions of V_t were used for the jump intensities, were very similar to that of $\mathcal{M}_5$.

considered, the initial predictive distributions (for $T_0 + 1$) in both (41) and (42). To reduce the computational burden in obtaining all subsequent predictive distributions, the posterior distributions for the relevant collection of static parameters are updated only every 250 observations thereafter. For the state space models, draws of the one-step-ahead latent vector, X_{t+1}, are produced recursively for each of the 2098 trading days, from February 22, 2006 to June 23, 2014, of which the evaluation period is comprised. A particle filtering algorithm is adopted for this purpose, conditional on the draws of the static parameters. The candidate state particles are sampled from the relevant state transition density as the proposal, with the latter being prescribed by the model in Section 2.4 and the restrictions detailed in Table 1. The predictive ability of the four models under investigation is evaluated in two ways: in terms of the accuracy of the probabilistic forecasts of all relevant measurements, assessed by the joint and marginal cumulative log scores; and in terms of the accuracy of highest posterior predictive (HPP) interval coverage and Value at Risk (VaR) prediction for the return measurement alone.

4.4.1   Cumulative log score assessment

Panels A to D in Figure 5 depict, in turn, the joint CLS score associated with the full measurement vector Y_t, and the marginal $[g]$ CLS scores of $g_t = r_t$, $g_t = \ln BV_t$ and $g_t = \left(\widetilde{M}_t^p, I_t^p \right)'$, as given in (43). From Panel A it is clear that the full dynamic model, $\mathcal{M}_F$, dominates all three of the models that exploit the full set of measurements, $\mathcal{M}_4$, $\mathcal{M}_5$ and $\mathcal{M}_8$, over the assessment period. The positive CLS scores throughout are consistent with positive log Bayes factors recorded for the full sample period in Table 3. The results in Panel C are very much in line with those in Panel A, with $\mathcal{M}_F$ continuing to dominate the comparator models (now expanded to include $\mathcal{M}_9$ to $\mathcal{M}_{11}$) in terms of the accuracy with which it predicts $\ln BV_t$ alone. Somewhat in contrast with these two sets of results, in Panels B and D the relative performance of $\mathcal{M}_F$ in predicting returns and price jumps respectively is seen to fluctuate throughout the evaluation period, with $\mathcal{M}_F$ sometimes being dominated by certain alternative specifications, despite still being the best model overall (as indicated by positive final values for both CLS scores). It is interesting to note (in Panel B) that in terms of predicting returns, $\mathcal{M}_F$ performs the best, amongst all of the state space models, during high volatility periods - both over the depth of the GFC in the second half of 2008, and during the Euro-zone debt crisis in 2011 - with all four CLS_i curves seen to have strong positive slopes at those points. Clearly the dynamic specifications incorporated in $\mathcal{M}_F$ have particular predictive power (for returns) during these turbulent periods. When compared to the conditionally deterministic RGARCH specifications, $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$, the full state space model outperforms the linear specification $\mathcal{M}_{10}$ overall, but under-performs relative to the log-linear specification, $\mathcal{M}_{11}$. In predicting the measures related to price jumps alone (Panel D), $\mathcal{M}_F$ performs roughly on par with $\mathcal{M}_4$ and $\mathcal{M}_5$, both of which employ some sort of dynamic structure for price jump intensity. However, when compared to $\mathcal{M}_8$, the model

with constant jump intensity, $\mathcal{M}_F$ clearly dominates.

4.4.2 Value at risk prediction and HPP coverage

As a final exercise, we assess the ability of the five alternative models entertained in Section 4.4.1 both to accurately estimate predictive tail quantiles and to produce 95% HPP intervals with accurate empirical coverage. We focus here only on the predictive distribution for the return, with the quantile estimation coinciding with the prediction of 1% and 5% VaRs. The empirical coverage statistics associated with both the VaRs and the HPP intervals, for all five competing models, are reported Table 4. We also report the results of the Christoffersen (1998) tests of correct unconditional coverage and independence of exceedances of the (predicted) intervals. Models that produce forecasts that fail to reject both of these tests are deemed adequate in predicting VaR.

The results indicate that all seven models being assessed have empirical coverage that is significantly different from the nominal coverage of the 95% HPP intervals over the assessment period. That said, the coverages are all quite reasonable (in an absolute sense) and $\mathcal{M}_F$ performs on par with $\mathcal{M}_{11}$, with empirical coverages that are quite close to the 95% level, as well as being the only models that do not reject the null hypothesis of independent violations. In all but two cases - the 1% VaR prediction from $\mathcal{M}_5$ and the 5% VaR prediction from $\mathcal{M}_{10}$ - the competing models produce VaR predictions with independent exceedences, with $\mathcal{M}_F$ being one of the three models with the empirical tail coverage closest to the nominal quantile probabilities. Perhaps not surprisingly, the worst performance (in terms of both HPP and tail coverage) is exhibited by the (Heston) model, $\mathcal{M}_9$, in which no price or volatility jump components feature.

In summary then, these results are consistent with the rankings produced by the marginal CLS computations for the return, as reported in the previous section. They confirm the importance of including both price and variance jumps in this empirical setting and, moreover, highlight the added value of augmenting the basic stochastic volatility structure with the particular dynamic structure for the price and volatility jumps as represented by a Hawkes process.
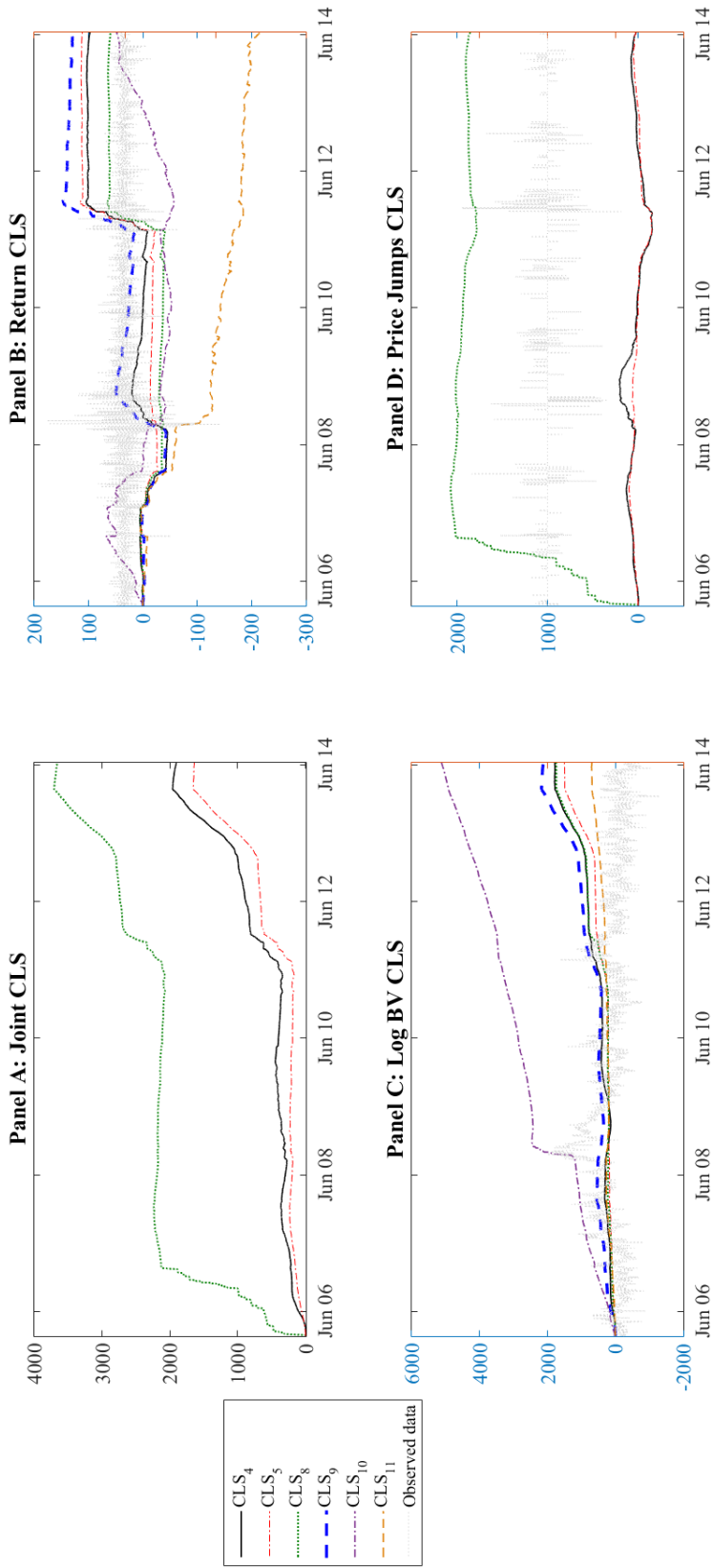
Figure 5: CLS_i of $\mathcal{M}_F$ relative to competing models: $\mathcal{M}_4$ (solid black lines), $\mathcal{M}_5$ (dash-dot red lines), $\mathcal{M}_8$ (dotted green lines), $\mathcal{M}_9$ (thick dashed blue lines), $\mathcal{M}_{10}$ (thick dashed-dot purple lines) and $\mathcal{M}_{11}$ (dashed orange lines) for February 22, 2006 to June 23, 2014. Panels A-D depict, in turn: the CLS scores for Y_t, r_t, $\ln BV_t$, and the joint price jump measure g_t given in (43). The relevant observed data over the assessment period are also plotted in the background of each of panel B-D in light dotted grey.

Table 4: Empirical tail coverage, computed as the proportion of observed returns that are lower than the 5% and 1% VaR predictions, respectively, is given in Column 2 and 3. The empirical coverage of the 95% HPP interval of the predictive returns distribution is given in Column 4. The superscripts * and + denote empirical coverage that is statistically different from the nominal level, and whose exceedences fail the independence test at the 5% significance level, respectively. All statistics are computed over the assessment period of February 22, 2006 to June 23, 2014, inclusive.

| | Empirical tail coverage | | Empirical coverage |
	5% VaR	1% VaR	95% HPP interval
$\mathcal{M}_F$	$7.34\%^*$	$2.86\%^*$	$91.94\%^*$
$\mathcal{M}_4$	$8.58\%^*$	$3.91\%^*$	$90.18\%^{*+}$
$\mathcal{M}_5$	$8.06\%^*$	$3.05\%^{*+}$	$90.75\%^{*+}$
$\mathcal{M}_8$	$7.96\%^*$	$2.81\%^*$	$91.09\%^{*+}$
$\mathcal{M}_9$	$9.01\%^*$	$4.62\%^*$	$89.37\%^{*+}$
$\mathcal{M}_{10}$	$4.62\%^+$	$1.67\%^*$	$96.38\%^{*+}$
$\mathcal{M}_{11}$	$8.91\%^*$	$4.36\%^*$	$91.28\%^*$

5   Conclusions

In this paper a very flexible stochastic volatility model is proposed, in which dynamic behaviour in price and variance (and, hence, volatility) jumps is accommodated via a bivariate Hawkes process for the two jump intensities. The model allows both price and variance jumps to cluster over time, for the two types of jump to occur simultaneously, or otherwise, and for the occurrence of a price jump to impact on the likelihood of a subsequent variance jump. A nonlinear state space model that uses daily returns on the S&P500 market index, in addition to nonparametric measures of volatility and price jumps, is constructed, with a hybrid Gibbs-MH MCMC algorithm used to estimate the model and compute marginal likelihoods and various predictive quantities. As remains standard in the literature, given that within-day index data informs the analysis, the conclusions we draw regarding the dynamics in asset prices pertain to within-day movements only, with the inclusion of overnight movements potentially requiring a modified set of assumptions to be adopted regarding the factors driving the dynamics therein.

A large number of alternative models, many of which impose restrictions on the general state space specification, are explored using Bayes factors, with the overall conclusion being in favour of the models that specify Hawkes dynamics in both price and variance jump intensity. Based on the most general specification, the probability of price and volatility jumps occurring either on the same day or on successive days is estimated to be close to 20% and the price jump size is found to be associated with the latent volatility itself. The dynamic structures imposed on the occurrences of price and variance jumps are also shown to add value to the predictions of returns on the index (including VaR predictions), as well as to the prediction of the nonparametric measures of volatility and jumps. One particular

(conditionally deterministic) alternative - the logarithmic form of RGARCH - performs the best of all models in terms of the CLS for the return, but does not dominate the more complex state space specifications in terms of predicting (logarithmic) bipower variation, and is unable to be used to predict jumps of any sort.

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.

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. Importantly though, further work is also required to ascertain the robustness of our qualitative results to the manner in which high frequency data is used to measure the occurrence and size of jumps (see Dumitru and Urga, 2012) and to the use of observed quarticity measures in the modelling of integrated variance (see, for example, Dobrev and Szerszen, 2010, and Bollerslev, Patton and Quaedvlieg, 2016). Extensive work along these lines is currently being undertaken by the authors.

Appendix A: Prior specification   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 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 $p(\psi|\omega) \sim N(\psi_0 = -0.005, \sigma_0^2 = \omega/5.0)$ and $p(\omega) \sim IG(a = 10, b = 0.001)$, where b 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 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. 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, which 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 2014). Conjugate inverse gamma priors 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.

Conjugate beta priors are employed for the unconditional jump intensities, δ_0^p and δ_0^v. The hyperparameters of these priors are chosen 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 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 $\max(RV_t - BV_t, 0)$. 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 in Section 2.2, and the prior belief that $\beta_{vp} > 0$ and $\beta_{vp}^{(-)} > 0$. The prior mean and standard deviation for each parameter is documented in Table 5.

Table 5: Prior specifications for each of the elements of the parameter vector ϕ

Parameter	Prior Spec	Mean	Stdev
μ	$U(-10, 10)$	0	5.77
γ	$U(-10, 0)$	-5	2.89
ρ	ρ, σ_v joint	-0.34	0.33
μ_p	$U(-100, 100)$	0	57.7
γ_p	$U(0, 100)$	50	28.9
σ_p^2	$IG(a = 3, b = 1)$	0.5	0.5
π_p	$\beta(a = 5, b = 5)$	0.5	0.15
α	$\beta(a = 0.01, b = 10)$	0.001	0.01
β	$\beta(a = 7, b = 3)$	0.7	0.14
$\sigma_{M_p}^2$	$IG(a = 3, b = 1)$	0.5	0.5
ψ_0	$N(0, 0.1)$	0	0.1
ψ_1	$N(1, 0.1)$	1	0.1
σ_{BV}^2	$IG(a = 3, b = 1)$	0.5	0.5
κ	$U(0, 1)$	0.5	0.29
θ	$U(0, 0.1)$	0.05	0.03
σ_v	ρ, σ_v joint	0.012	0.003
μ_v	$IG(a = 20, b = 1/7.2)$	$7e^{-3}$	$2e^{-3}$
δ_0^p	$\beta(a = 1, b = 9)$	0.1	0.03
α_p	$U(0, 1)$	0.5	0.29
β_{pp}	$U(0, 1)$	0.5	0.29
δ_0^v	$\beta(a = 1, b = 9)$	0.1	0.03
α_v	$U(0, 1)$	0.5	0.29
β_{vv}	$U(0, 1)$	0.5	0.29
β_{vp}	$U(0, 1)$	0.5	0.29
$\beta_{vp}^{(-)}$	$U(0, 1)$	0.5	0.29

Appendix B.1: MCMC algorithm for $\mathcal{M}_F$ The MCMC algorithm for sampling from the joint posterior in (34) can be broken down into seven main steps, as outlined below:

Algorithm 1 *At each iteration:*

1. *Sample $V_{1:T}$ in blocks of random length from $V_{1:T}|Z_{1:T}^v, \Delta N_{1:T}^v, M_{1:T}^p, \Delta N_{1:T}^p, S_{1:T}^{Z_p}, Y_{1:T}, \phi$ using MH sampling as described below*

2. *Sample $\Delta N_{1:T}^v$ in a single block from $\Delta N_{1:T}^v|V_{1:T}, Z_{1:T}^v, M_{1:T}^p, \Delta N_{1:T}^p, S_{1:T}^{Z_p}, Y_{1:T}, \phi$ using the conditionally independent Bernoulli structure*

3. *Sample $Z_{1:T}^v$ in a single block from $Z_{1:T}^v|V_{1:T}, \Delta N_{1:T}^v, M_{1:T}^p, \Delta N_{1:T}^p, S_{1:T}^{Z_p}, Y_{1:T}, \phi$ using the conditionally independent truncated normal structure*

4. *Sample $\Delta N_{1:T}^p$ in a single block from $\Delta N_{1:T}^p|V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}^v, M_{1:T}^p, S_{1:T}^{Z_p}, Y_{1:T}, \phi$ using the conditionally independent Bernoulli structure*

5. *Sample $M_{1:T}^p$ in a single block from $M_{1:T}^p|V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}^v, \Delta N_{1:T}^p, S_{1:T}^{Z_p}, Y_{1:T}, \phi$ using the conditionally independent normal structure*

6. *Sample $S_{1:T}^{Z_p}$ in a single block from $S_{1:T}^{Z_p}|V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}^v, M_{1:T}^p, \Delta N_{1:T}^p, Y_{1:T}, \phi$ using the conditionally independent Bernoulli structure*

7. *Sample ϕ from $\phi|X_{1:T}, Y_{1:T}$ as described below*

The most challenging part of the algorithm is step 1, namely the generation of the variance process $V_{1:T}$, due to the nonlinear functions of V_t that feature in the measurement equations (22) and (25), and in the state equation (26). 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 in step 7 using MH subchains wherever necessary. Given the draws of $V_{1:T}$ and $M_{1:T}^p$, and all of the unknowns that appear in (22) - (32), the parameters μ, γ, μ_p, γ_p, ψ_0 and ψ_1 can be treated as regression coefficients, with exact draws produced in the standard manner from Gaussian conditional posterior distributions, appropriately truncated as a consequence of the previously specified priors. The sampling schemes of the conditional variance terms σ_{BV}^2, $\sigma_{M_p}^2$ and σ_p^2 are standard, with inverse gamma conditional posteriors. Similarly, parameters π_p, α and β are sampled using Gibbs schemes, as all three have closed form conditional beta posteriors. As described in Appendix A, 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 $\sigma_v^2 \le 2\kappa\theta$ imposed.

The static 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 (31) and (32) define stationary processes, and that (13) and (14) are defined on the $[0, 1]$ interval. The intensity parameters δ_∞^p and δ_∞^v, are then computed using the explicit relationships in (13) and (14), and the vectors $\delta_{1:T}^v$ and $\delta_{1:T}^p$ updated deterministically based on (31) and (32).

The algorithms for all comparator state space models described in Section 3.2, $\mathcal{M}_i$, for $i = 1, ..., 9$, proceed in an analogous way.

Appendix B.2: MCMC algorithm for the RGARCH models The joint posterior for the RGARCH models $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$ satisfies $p\left(\phi|Y_{1:T}\right) \propto p\left(Y_1|\phi\right) p\left(\phi\right) \left[\prod_{t=2}^{T} p\left(Y_t|Y_{1:t-1},\phi\right)\right]$, with $Y_t = \left(r_t, BV_t\right)'$ for $\mathcal{M}_{10}$ and $Y_t = \left(r_t, \ln BV_t\right)'$ for $\mathcal{M}_{11}$. For the purpose of estimation, we employ the variance targeting approach, and reparameterize $\omega = \sigma_0^2\left(1 - \beta - \gamma\right)$, with σ_0^2 denoting the unconditional variance of the return. The elements of the parameter vector are identical for the two models: $\phi = (\sigma_0^2, \beta, \gamma, \xi, \varphi, \tau_1, \tau_2, \sigma_u^2)'$. We impose noninformative priors on ϕ: inverse gamma priors, $IG\left(a = 3, b = 1\right)$, are employed for both σ_0^2 and σ_u^2; uniform priors on the unit interval are employed for β and γ; and the priors for ξ, φ, τ_1 and τ_2 are uniform between -20 and $+20$. Since there are no latent variables involved in the model, the MCMC algorithm to sample from the joint posterior is quite straightforward, with MH steps required only for σ_0^2, β and γ.

Appendix C: Marginal likelihood computation The basic idea underlying the evaluation of (37) is the recognition that it can be re-expressed as

$$p\left(Y_{1:T}|\mathcal{M}_i\right) = \frac{p\left(Y_{1:T}|\phi_i, \mathcal{M}_i\right) p\left(\phi_i|\mathcal{M}_i\right)}{p\left(\phi_i|Y_{1:T}, \mathcal{M}_i\right)}, \tag{44}$$

for any point ϕ_i in the posterior support of model $\mathcal{M}_i$, 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 (44) is the likelihood, conditional on $\mathcal{M}_i$, marginal of the latent variables. That is,

$$p\left(Y_{1:T}|\phi_i, \mathcal{M}_i\right) = \int p\left(Y_{1:T}|X_{1:T}^{(i)}, \phi_i, \mathcal{M}_i\right) p\left(X_{1:T}^{(i)}|\phi_i, \mathcal{M}_i\right) dX_{1:T}^{(i)} \tag{45}$$

The denominator on the right-hand-side of (44) is simply the conditional posterior density of the (static) parameter vector, also marginalized over the latent variables,

$$p\left(\phi_i|Y_{1:T}, \mathcal{M}_i\right) = \int p\left(\phi_i|Y_{1:T}, X_{1:T}^{(i)}, \mathcal{M}_i\right) dX_{1:T}^{(i)}. \tag{46}$$

The evaluation of (45) at a high density posterior point ϕ_i^* (say, the vector of marginal posterior means for the elements of ϕ_i) is straightforward, using the output of a full MCMC run for model $\mathcal{M}_i$; namely, the closed form representation of $p\left(Y_{1:T}|X_{1:T}^{(i)}, \phi_i, \mathcal{M}_i\right)$ is averaged over the draws of the latent states, $X_{1:T}^{(i)}$, and computed at the given point ϕ_i^*. Evaluation of (46) 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. Exploiting the structure of the posterior density, we decompose $p\left(\phi_i^*|Y_{1:T}, \mathcal{M}_i\right)$ into five constituent densities as:

$$p\left(\phi_i^*|Y_{1:T}, \mathcal{M}_i\right) = p\left(\phi_{1i}^*|Y_{1:T}, \mathcal{M}_i\right) p\left(\phi_{2i}^*|\phi_{1i}^*, Y_{1:T}, \mathcal{M}_i\right) \cdots p\left(\phi_{5i}^*|\phi_{1i}^*, \phi_{2i}^*, ..., \phi_{4i}^*, Y_{1:T}, \mathcal{M}_i\right), \tag{47}$$

where $\phi_{1i} = \left(\sigma_{BV}, \mu_v, \delta_0^p, \delta_0^v, \rho, \sigma_v, \mu_p, \alpha, \beta, \pi_p, \sigma_{M_p}\right)$, $\phi_{2i} = \left(\alpha_p, \alpha_v, \kappa, \gamma_p, \psi_0\right)$, $\phi_{3i} = (\beta_{pp}, \beta_{vv}, \theta, \sigma_p, \psi_1)$, $\phi_{4i} = \left(\beta_{vp}, \mu\right)$, and $\phi_{5i} = \left(\beta_{vp}^{(-)}, \gamma\right)$. Following the methods outlined by Chib (1995) and Chib and Jeliazkov (2001), five additional auxiliary 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 (47), in turn evaluated at ϕ_{ji}^*, $j = 2, ..., 5$. The first component on the right hand side of (47), involving no such conditioning, is estimated from the output of the full MCMC chain, in the usual way.

Calculation of the marginal likelihoods of the RGARCH models follows similarly, albeit without the latent variables playing a role, and with the choice of the auxiliary chains being

determined by nature of the parameter sets for these models. The marginal likelihood for $\mathcal{M}_{10}$ also includes a Jacobian factor that accounts for the fact that $\mathcal{M}_{10}$ specifies a model for the raw measure BV_t, whereas all others are specified in terms of the transformed measure, $\ln BV_t$.

Finally, two versions of the marginal likelihood for models $\mathcal{M}_9$, $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$ are produced: one that only considers measurements that are directly used in the model, with $Y_t = (r_t, \ln BV_t)'$; and one that employs the full measurement set, $Y_t = \left(r_t, \ln BV_t, I_t^p, \widetilde{M}_t^p \right)'$. The second form of marginal likelihood allows for the comparison across all models considered in the paper. Since the possibility of price jumps is actually excluded in each of $\mathcal{M}_9$, $\mathcal{M}_{10}$ and $\mathcal{M}_{11}$, we employ the specifications: $I_t^p \sim Bernoulli\,(\alpha)$ for the price jump occurrence and $\widetilde{M}_t^p \sim N \left(-10, \sigma_{M_p}^2 \right)$ for the log price jump size, with the priors for α and $\sigma_{M_p}^2$ defined in Appendix A. The specification for I_t^p is nested in (23), associated with $\Delta N_t^p = 0$ for all t. The prior expectation of $\widetilde{M}_t^p$ is assumed to be a large negative value as this reflects a price jump magnitude that is close to zero. The marginal likelihood components related to these measures are straightforward to evaluate, with the closed form expressions of $p\left(I_t^p | \mathcal{M}_j\right)$ and $p\left(\widetilde{M}_t^p | \mathcal{M}_j\right)$ being available analytically for $j = 9, 10$ and 11.

References

[1] Ahoniemi, K., Fuertes, A. and Olmo, J. (2015), "Overnight News and Daily Equity Trading Risk Limits," *Journal of Financial Econometrics*, 14, 1-27.

[2] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R.J.A. (2015), "Modeling Financial Contagion Using Mutually Exciting Jump Processes," *Journal of Financial Economics*, 117, 585-606.

[3] 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.

[4] 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.

[5] Andersen, T.G., Bollerslev, T. and Huang X. (2011), "A Reduced Form Framework for Modeling Volatility of Speculative Prices Based on Realized Variation Measures," *Journal of Econometrics*, 160, 176-189.

[6] Bandi, F.M. and Reno, R. (2016), "Price and Volatility Co-Jumps," *Journal of Financial Economics,* 119, 107-146.

[7] 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.

[8] ———— (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1-37.

[9] ———— (2006), "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4, 1-30.

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

[11] ——— (2000), "Post-87 Crash Fears in the S&P 500 Futures Option Market," *Journal of Econometrics*, 94, 181-238.

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

[13] 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.

[14] Bollerslev, T., Kretschmer, U., Pigorsch, C. and Tauchen, G. (2009), "A Discrete-Time Model for Daily S&P 500 Returns and Realized Variations: Jumps and Leverage Effects," *Journal of Econometrics*, 150, 151-166.

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

[16] Bollerslev, T., Patton, A. and Quaedvlieg, R. (2016), "Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting," *Journal of Econometrics,* 192, 1-18.

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

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

[19] 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.

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

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

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

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

[24] Dobrev, D. and Szerszen, P. (2010), "The Information Content of High-Frequency Data for Estimating Equity Return Models and Forecasting Risk," *Working Paper*, SSRN.

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

[26] Dumitru, A.M. and Urga, G. (2012), "Identifying Jumps in Financial Assets: A Comparison Between Nonparametric Jump Tests," *Journal of Business and Economic Statistics*, 30, 242-255.

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

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

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

[30] Fulop, A., Li, J. and Yu, J. (2014), "Self-Exciting Jumps, Learning, and Asset Pricing Implications," *Review of Financial Studies*, 28, 876-912.

[31] Geweke, J. (1992), "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," in *Bayesian Statistics, Vol. 4* (eds Bernardo, J.M., Berger, J.O., Dawid, A.P. and Smith, A.F.M.), Clarendon Press, Oxford, UK.

[32] Geweke, J. (2001), "Bayesian Econometrics and Forecasting," *Journal of Econometrics*, 100, 11-15.

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

[34] Hansen, P.R., Huang, Z. and Shek, H.H. (2012), "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility," *Journal of Applied Econometrics*, 27, 877-906.

[35] Hansen, P.R. and Lunde, A. (2005), "A Realized Variance for the Whole day Based on Intermittent High-Frequency Data," *Journal of Financial Econometrics*, 3, 525–554.

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

[37] ―――― (1971b), "Point Spectra of Some Mutually Exciting Point Processes," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 33, 438-443.

[38] Heidelberger, P., and Welch, P.D. (1983), "Simulation Run Length Control in the Presence of an Initial Transient," *Operations Research*, 31, 1109-1144.

[39] 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.

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

[41] Jacod, J., Klüppelberg C. and Müller G. (2012), "Functional Relationships Between Price and Volatility Jumps and Their Consequences for Discretely Observed Data," *Journal of Applied Probability*, 49, 901-914.

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

[43] Jacod, J., Klüppelberg, C. and Müller, G. (2013), "Testing for Non-Correlation Between Price and Volatility Jumps," *Draft paper*.

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

[45] 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.

[46] Jensen, M.J. and Maheu, J.M. (2014), "Risk, Return and Volatility Feedback: A Bayesian Nonparametric Analysis," *Working Paper*. SSRN.

[47] 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.

[48] Liu, L. Y., Patton, A. J., and Sheppard, K. (2015), "Does Anything Beat 5-Minute RV? A Comparison of Realized Measures Across Multiple Asset Classes," *Journal of Econometrics*, 187, 293-311.

[49] 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.

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

[51] 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.

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

[53] 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.

[54] Takahashi, M., Omori, Y. and Watanabe, T. (2009), "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously", *Computational Statistics and Data Analysis*, 53, 2404-2426.

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

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

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