ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

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

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

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

December 2013

Working Paper 28/13

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

Worapree Maneesoonthorn
Melbourne Business School, University of Melbourne
(O.Maneesoonthorn@mbs.edu)

Catherine S. Forbes and Gael M. Martin
Department of Econometrics & Business Statistics, Monash University

Abstract

This paper investigates the dynamic behaviour of jumps in financial prices and volatility. The proposed model is based on a standard jump diffusion process for price and volatility augmented by a bivariate Hawkes process for the two jump components. The latter process specifies a joint dynamic structure for the price and volatility jump intensities, with the intensity of a volatility jump also directly affected by a jump in the price. The impact of certain aspects of the model on the higher-order conditional moments for returns is investigated. In particular, the differential effects of the jump intensities and the random process for latent volatility itself, are measured and documented. A state space representation of the model is constructed using both financial returns and non-parametric measures of integrated volatility and price jumps as the observable quantities. Bayesian inference, based on a Markov chain Monte Carlo algorithm, is used to obtain a posterior distribution for the relevant model parameters and latent variables, and to analyze various hypotheses about the dynamics in, and the relationship between, the jump intensities. An extensive empirical investigation using data based on the S&P500 market index over a period ending in early-2013 is conducted. Substantial empirical support for dynamic jump intensities is documented, with predictive accuracy enhanced by the inclusion of this type of specification. In addition, movements in the intensity parameter for volatility jumps are found to track key market events closely over this period.

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) volatility 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 allowing both price and volatility to jump, and for jumps in the two different processes to (always) occur contemporaneously. Under this type of specification, it is the (positive) volatility jumps that occur concurrently with price jumps that lead to further large fluctuations in the price over subsequent periods, via the persistence in the volatility process. This impact is exacerbated further by specifying the average magnitude of price jumps to be (positively) dependent on the average magnitude of volatility jumps. See Eraker, Johannes and Polson (2003) and Broadie, Chernov and Johannes (2007) for related investigations; and also Eraker (2004), who specifies a model allowing for *both* contemporaneous price and volatility jumps *and* dependence of the (price) jump intensity on the volatility level.

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

ously for the S&P500 market index. Most notably, Tauchen and Zhou (2011) conclude in favour of a *dynamic* price jump process, with clusters of price jumps linked to dynamic changes in credit risk in the US market. The dynamic behaviour of price jump processes has also been investigated by Aït-Sahalia, Cacho-Diaz and Laeven (2013), who introduce the Hawkes process (Hawkes 1971a,b) to model jumps in multivariate asset prices. These authors analyze the dynamic transmission, or 'contagion', of price jumps across equity markets in several countries, concluding that price jump intensities increase dramatically during a crisis period, across all connected markets. See also Liao, Anderson and Vahid (2010) for an investigation of common factors in price jumps across assets in the Chinese stock market.

In the spirit of this more recent empirical work, we propose a model in which both the price and volatility of a single asset are permitted to jump, with the two jump processes being dynamic. Adopting a bivariate Hawkes process for the intensity process associated with jumps in price and its latent variance, jumps in both price and variance (accordingly volatility, defined as the square root of the variance) are (potentially) self-exciting; that is, the intensity of each jump process is functionally dependent on the realized past increments of that process. We allow the variance jump intensity to depend on past price jumps, in an attempt to capture any potential transmission of information in extreme price movements to extreme movements in its underlying volatility. Possible leverage effects operating at the level of extreme price and volatility movements are 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, incorporating volatility and price jump measures constructed from high frequency data, in addition to the daily return measure, is used to specify the proposed dynamic components. The state space model is analyzed using Bayesian methods, implemented via a Markov chain Monte Carlo (MCMC) sampling algorithm. Posterior odds ratios, constructed using the marginal likelihood computation methods of Chib (1995) and Chib and Jeliazkov (2001), are used to assess various hypotheses associated with the price and volatility jump dynamics. The benefits of including a very flexible dynamic specification for price and volatility jumps are also investigated via an out-of-sample forecasting exercise. The implications of the proposed model for the conditional higher-order moments of the return process are explored in some detail. An understanding of these properties is, in turn, crucial for achieving effective hedging strategies and for constructing optimal portfolios; see, for example, Cvitanic, Polimenis and Zapatero (2008) and Harvey, Liechty, Liechty and Müller (2010). Whilst it is beyond the scope of this paper to investigate such financial applications, we believe that documenting the link between the

structure of this particular type of dynamic model and the properties of returns, provides an important foundation for future investigations of this type.

The proposed methodology is applied to measures constructed from intraday data for the S&P500 index for the period beginning in January 1996 and ending in early-May 2013. The existence of dynamics in both types of intensity processes is supported by the posterior odds analysis, while price jumps (positive or negative) are found to have a negligible direct impact on the variance jump intensity. Incorporating dynamic structures for both intensities is found to lead to more accurate predictive return distributions over the period used for out-of-sample analysis, namely mid-May 2012 to early-May 2013; a period that includes the market instabilities induced by emerging evidence of the US 'fiscal cliff' and the subsequent debates surrounding its remedies. This increased accuracy is evident in the positive cumulative difference in the predictive log density score, computed using a comparison model in which constant jump intensities are imposed. Support for the proposed model is also found in the closer match between the empirical coverage of extreme quantiles of returns and the nominal coverage, when compared to alternative models that either impose constant jump intensities or exclude jumps altogether. We also view estimates of the dynamic jump intensities as possible indicators of impending financial crises, and review their behaviour during a recent five year period. Empirical results indicate that the variance jump intensity is much more informative, in this sense, than the price jump intensity, with the former closely tracking key events over this period, and exhibiting its most dramatic increase at the peak of the global financial crisis in late 2008.

The remainder of the paper is organized as follows. Section 2 describes our proposed asset price model and its main properties, including the implications of the model for the conditional moments of returns. Section 3 outlines the state space representation used to analyze the model, followed by a discussion of the relevant hypotheses to be investigated. Details of the Bayesian inferential methodology are then given in Section 4. Results from the empirical analysis of the S&P500 index are presented and discussed in Section 5, and Section 6 provides some conclusions. Certain technical details, including algorithmic details, are included in appendices to the paper.

2 An asset price process with stochastic volatility and 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,

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

$$dV_t = \tilde{\kappa}\left(\tilde{\theta} - V_t\right) dt + \tilde{\sigma}_v \sqrt{V_t} dB_t^v + dJ_t^v, \tag{2}$$

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 $\tilde{\sigma}_v^2 \leq 2\tilde{\kappa}\tilde{\theta}$ ensures the positivity of the variance process, denoted by V_t, for $t > 0$. The drift component of (1) contains the additional component $\tilde{\gamma} V_t$, allowing the 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 p_t and V_t of 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 two 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

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

$$d\tilde{\delta}_t^p = \tilde{\alpha}_p\left(\tilde{\delta}_\infty^p - \tilde{\delta}_t^p\right) dt + \tilde{\beta}_{pp} dN_t^p, \tag{4}$$

and that

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

$$d\tilde{\delta}_t^v = \tilde{\alpha}_v\left(\tilde{\delta}_\infty^v - \tilde{\delta}_t^v\right) dt + \tilde{\beta}_{vv} dN_t^v + \tilde{\beta}_{vp} dN_t^p + \tilde{\beta}_{vp}^{(-)} dN_t^{p(-)}, \tag{6}$$

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, $\tilde{\beta}_{vp}^{(-)} dN_t^{p(-)}$, allows a contemporaneous negative price jump to have an impact on $d\tilde{\delta}_t^v$, thereby serving as an additional channel for leverage, over and above the non-zero correlation between the Brownian motion increments, dB_t^p and dB_t^v. See Hawkes (1971a,b) for seminal discussions regarding self-exciting point processes, and Aït-Sahalia, Cacho-Diaz and Laeven (2013) for the introduction of the Hawkes process into asset pricing models.

Our proposed specification can be viewed as a natural extension of the various models in the literature that accommodate both stochastic volatility and jumps. Most notably we relax the strict assumptions of the well-known stochastic volatility with contemporaneous jumps (SVCJ) model (Duffie *et al.* 2000) to allow jumps in both the price and volatility to be governed by separate, though dependent, dynamic random processes. The specification can also be viewed as an extension of the stochastic volatility model of Aït-Sahalia, Cacho-Diaz and Laeven (2013), in which a Hawkes process is used to characterize the occurrence of price jumps, but with volatility jumps absent from the model. Similarly, it extends the model proposed by Fulop, Li and Yu (2012), in which price jump intensity (only) is also characterized by a Hawkes process, along with a restrictive assumption that volatility jumps occur contemporaneously with negative price jumps.

2.2 A discrete time model for returns

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

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

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

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

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

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

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

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

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

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

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

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

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

and

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

respectively. The price jump size Z_t^p in (14), being characterized as conditionally Gaussian, allows for the possibility of both positive and negative extreme movements in the asset price, and with the relative magnitude of the expected price jump being proportional to V_t. The variance jump, on the other hand, is restricted to be positive, via the exponential distributional assumption in (15). Note that the parameters used in (7)-(13) are the discrete time versions of the corresponding parameters (with super-imposed tilda) in the continuous time model (1)-(6). For example, κ, being the counterpart of $\tilde{\kappa}$, denotes the parameter that measures the daily persistence in the annualized stochastic variance process, V_t.

The factors that drive the asset price process 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 volatility jumps adjusted accordingly, through the model adopted here for the jump intensities. That is, the process

ΔN_t^p can be viewed as being potentially self-exciting: provoking an increase in the future intensity (and thus occurrence) of price jumps (via (12)), as well as provoking (or exciting) an increase in the future intensity of variance jumps (via (13)). The threshold parameter $\beta_{vp}^{(-)}$ in (13) allows for a possible additional impact of a negative price jump on the variance jump intensity (and, hence, the level of volatility), providing an additional channel for leverage, as noted above.

From the form of (7) and (8), the implications for returns of the occurrence of the two types of jumps are clear. From (7), the impact of a given price jump at time t, $\Delta J_t^p = Z_t^p \Delta N_t^p$, is felt only at time t. However, clusters of price jumps and, hence, successive extreme values in returns, can occur via the dynamic intensity process in (12) that drives ΔN_t^p. The impact of a given (positive) variance jump at time t persists through time via the persistence of the V_{t+1} process, as governed by κ. That is, if the return volatility 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. Any clustering of variance jumps, driven by the dynamic intensity process in (13), simply causes 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. Hence, the quantity δ_t^v is seen as a potential indicator of a crisis, a characteristic that is investigated empirically in Section 5.3.

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

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

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

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

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

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

resulting in

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

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

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

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

2.3 Implied higher-order moments

Rare and extreme movements in the asset price contribute to the higher-order moments of the conditional return distribution (see also Eraker *et al.* 2003, for relevant discussion). In this section, we derive the third- and fourth-order conditional moments for returns implied by the discrete time model in (7) to (15), with particular focus given to the impact on these moments of the dynamic jump intensities in (12) and (13), and the various components of the jump size specifications in (14) and (15). Specifically, we derive expressions for the standardized (conditional) skewness,

$$sk(r_t|\cdot) = \frac{E\left(\left(r_t - E(r_t|\cdot)\right)^3|\cdot\right)}{\left[Var(r_t|.)\right]^{3/2}}$$

and standardized (conditional) kurtosis

$$ku(r_t|\cdot) = \frac{E\left(\left(r_t - E(r_t|\cdot)\right)^4|\cdot\right)}{\left[Var(r_t|\cdot)\right]^2}$$

of r_t, under two different conditioning settings. First, we condition on the current latent stochastic variance, V_t. This enables us to derive the moments of r_t (conditional on V_t) as explicit functions of the current price jump intensity, δ_t^p, and various other *static* parameters in the model, including those that determine the price jump size in (14). This, in turn, enables us to identify the incremental impact on the higher-order moments of the price

jump specification, over and above any effect induced by the stochastic variance V_t. Next, to measure the incremental impact of stochastic volatility itself (including jumps in the latent variance process), beyond the effect induced by random price jumps, we condition on the lagged stochastic variance, V_{t-1} and the current price jump, ΔJ_t^p. We then derive the (conditional) moments of r_t as explicit functions of δ_{t-1}^v and other static parameters, including μ_v; the relevant relationship here being between (the conditional moments of) r_t and δ_{t-1}^v (rather than δ_t^v) due to the nature of the discretization adopted in (7) and (8).

Conditioning on V_t, we obtain the moment generating function for r_t, as an explicit function of δ_t^p, as

$$M_{r_t|V_t}(h) = (1-\delta_t^p)\exp\left((\mu+\gamma V_t)\,h + \frac{1}{2}V_t h^2\right) + \delta_t^p \exp\left(\left(\mu+\gamma V_t+\mu_p+\gamma_p V_t\right)h + \frac{1}{2}\left(V_t+\sigma_p^2\right)h^2\right),$$

from which the first four central conditional moments are derived and tabulated in Panel A of Table 1. The contribution of the dynamic price jump intensity to conditional skewness and kurtosis can be gauged by noting the role played by δ_t^p in determining the third and fourth central conditional moments. Specifically, as the price jump size, Z_t^p, is assumed to be conditionally Gaussian, the distribution of $r_t|V_t$ can be expressed as a mixture of two normal densities, with the first component being the Gaussian density of the diffusive element only, weighted by $1-\delta_t^p$, and the second component being the Gaussian density of sum of the diffusive element and the random price jump size, weighted by δ_t^p. Thus, if either $\delta_t^p = 0$ or $\delta_t^p = 1$, the distribution of $r_t|V_t$ is Gaussian, with differing means and variances in the two cases, but with (standardized) skewness and kurtosis coinciding with the Gaussian values of zero and three. In the presence of random price jumps ($0 < \delta_t^p < 1$), the distribution of $r_t|V_t$ is non-Gaussian and the direction of skewness, positive or negative, depends entirely on the sign of the conditional mean of the price jump size, $\left(\mu_p+\gamma_p V_t\right)$, with the magnitude of the third and fourth central moments depending on the magnitude of both $\left(\mu_p+\gamma_p V_t\right)$ and the variance of the price jump size, σ_p^2.

Table 1: The model-implied conditional moments

This table summarizes the conditional moments of the returns based on the model defined in (7) to (15). The conditional skewness and kurtosis are computed by $sk\left(r_t|\cdot\right)=\frac{E\left(\left(r_t-E(r_t|\cdot)\right)^3|\cdot\right)}{\left[Var(r_t|\cdot)\right]^{3/2}}$ and $ku\left(r_t|\cdot\right)=\frac{E\left(\left(r_t-E(r_t|\cdot)\right)^4|\cdot\right)}{\left[Var(r_t|\cdot)\right]^2}$, respectively. The functions $g_1\left(\cdot,\cdot\right)$, $g_2\left(\cdot,\cdot\right)$, $g_3\left(\cdot,\cdot\right)$ and $g_4\left(\cdot,\cdot\right)$ are detailed in Appendix A.

Panel A: Central Moments of $r_t|V_t$

$$
\begin{aligned}
E\left(r_t|V_t\right) &= \mu+\gamma V_t+\delta_t^p\left(\mu_p+\gamma_p V_t\right)\\
Var\left(r_t|V_t\right) &= V_t+\delta_t^p\sigma_p^2+\delta_t^p\left(1-\delta_t^p\right)\left(\mu_p+\gamma_p V_t\right)^2\\
E\left(\left(r_t-E\left(r_t|V_t\right)\right)^3|V_t\right) &= \delta_t^p\left(1-3\delta_t^p+2\left(\delta_t^p\right)^2\right)\left(\mu_p+\gamma_p V_t\right)^3+3\delta_t^p\left(1-\delta_t^p\right)\sigma_p^2\left(\mu_p+\gamma_p V_t\right)\\
E\left(\left(r_t-E\left(r_t|V_t\right)\right)^4|V_t\right) &= 3\left(V_t^2+2\delta_t^p\sigma_p^2V_t+\delta_t^p\sigma_p^4\right)+\left(\delta_t^p-4\left(\delta_t^p\right)^2+6\left(\delta_t^p\right)^3-3\left(\delta_t^p\right)^4\right)\left(\mu_p+\gamma_p V_t\right)^4\\
&\quad+\left(6\delta_t^p-12\left(\delta_t^p\right)^2+6\left(\delta_t^p\right)^3\right)\sigma_p^2\left(\mu_p+\gamma_p V_t\right)^2+6\delta_t^p\left(1-\delta_t^p\right)V_t\left(\mu_p+\gamma_p V_t\right)^2
\end{aligned}
$$

Panel B: Central Moments of $r_t|V_{t-1},\Delta J_t^p$

$$
\begin{aligned}
E\left(r_t|V_{t-1},\Delta J_t^p\right) &= \mu+\gamma\left(\overline{V}_t^d+\delta_{t-1}^v\mu_v\right)+\Delta J_t^p+g_1\left(\phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right),\Phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right)\right)\\
Var\left(r_t|V_{t-1},\Delta J_t^p\right) &= \overline{V}_t^d+\gamma^2\sigma_v^2V_{t-1}\left(1-\rho^2\right)+\delta_{t-1}^v\mu_v+\left(2-\delta_{t-1}^v\right)\gamma^2\delta_{t-1}^v\mu_v^2\\
&\quad+g_2\left(\phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right),\Phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right)\right)\\
E\left(\left(r_t-E\left(r_t|V_{t-1},\Delta J_t^p\right)\right)^3|V_{t-1},\Delta J_t^p\right) &= 3\gamma\sigma_v^2V_{t-1}\left(1-\rho^2\right)+3\gamma\left(2-\delta_{t-1}^v\right)\delta_{t-1}^v\mu_v^2+2\gamma^3\left(3-3\delta_{t-1}^v+\left(\delta_{t-1}^v\right)^2\right)\delta_{t-1}^v\mu_v^3\\
&\quad+g_3\left(\phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right),\Phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right)\right)\\
E\left(\left(r_t-E\left(r_t|V_{t-1},\Delta J_t^p\right)\right)^4|V_{t-1},\Delta J_t^p\right) &= 3\left(Var\left(r_t|V_{t-1},\Delta J_t^p,\delta_{t-1}^v\right)\right)^2+3\sigma_v^2\left(1-\rho^2\right)V_{t-1}\\
&\quad+\left(6-5\delta_{t-1}^v\right)\delta_{t-1}^v\mu_v^2+\left(24-48\delta_{t-1}^v+32\left(\delta_{t-1}^v\right)^2-6\left(\delta_{t-1}^v\right)^3\right)\delta_{t-1}^v\mu_v^4\gamma^4\\
&\quad+\left(16\left(\delta_{t-1}^v\right)^2-54\delta_{t-1}^v+36\right)\delta_{t-1}^v\mu_v^3\gamma^2\\
&\quad+g_4\left(\phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right),\Phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right)\right)\\
&\quad\text{where } \overline{V}_t^d=\kappa\theta+\left(1-\kappa\right)V_{t-1}+\sigma_v\rho\left(r_{t-1}-Z_{t-1}^p\Delta N_{t-1}^p-\mu-\gamma V_{t-1}\right)
\end{aligned}
$$

In Figure 1 we plot the (conditional) third central moment (Panel A), skewness (Panel B), fourth central moment in log scale (Panel C) and kurtosis (Panel D) over the range $\delta_t^p \in [0,1]$. We condition on three representative levels of the latent variance, V_t, and use values for the static parameters that match the marginal posterior means produced in the empirical analysis associated with the S&P500 index detailed in Section 5. The levels of V_t are given by the square of the annualized standard deviation of returns, with a low volatility of 6% corresponding to August 1996, a medium volatility level of 18% corresponding to December 2002 and a high volatility of 68% corresponding to October 2008. The values employed for the static parameters are $\mu = 0.09$, $\gamma = -3.7$, $\mu_p = -0.03$, $\gamma = -3.8$, $\sigma_p = 1.3$, $\rho = -0.6$, $\sigma_v = 0.02$, and $\mu_v = 0.025$. At the high level of V_t, the relative magnitude of $\left(\mu_p + \gamma_p V_t\right)$ is high and, hence, the third central moment (Panel A) and the fourth central moment in log scale (Panel C) exceed the corresponding values for the medium and low volatility cases, for any value of $\delta_t^p \in (0,1)$. However, as the high level of volatility also translates into a higher level for the conditional variance of r_t, the overall impact on the magnitude of skewness (Panel B) and kurtosis (Panel D) is reduced by the process of standardization. As a consequence, the greatest impact of price jumps on standardized skewness and kurtosis is observed at the lowest level of V_t, for any given value of $\delta_t^p \in (0,1)$. The given values for the (static) parameters are shown to induce negative skewness and kurtosis in excess of three, for $0 < \delta_t^p < 1$, reverting to the Gaussian values of zero and three as δ_t^p approaches either bound. At the lowest level of V_t, the magnitudes of skewness and kurtosis both peak at about $\delta_t^p = 0.25$, whilst at the medium and high levels of V_t, the peaks are observed at larger values of δ_t^p, and with both quantities assuming values closer to the Gaussian quantities, for all $\delta_t^p \in (0,1)$.

We next investigate the characteristics of the distribution of r_t, conditional on V_{t-1} and ΔJ_t^p, with a view to isolating the impact of the (lagged) intensity of variance jumps, δ_{t-1}^v, on the higher-order moments of the return. The relevant moment generating function in this

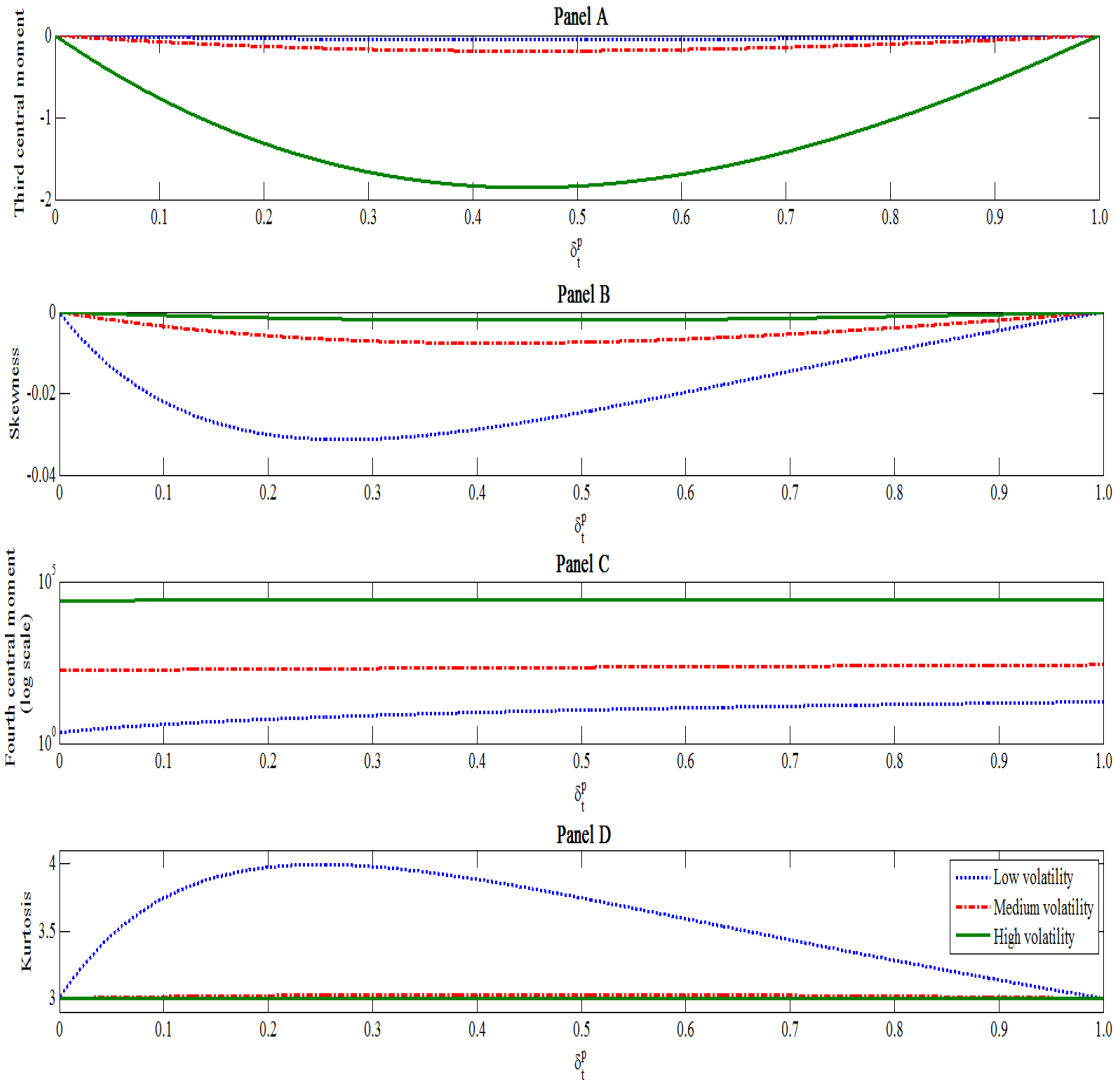

Figure 1: Conditional third central moment (Panel A), skewness (Panel B), fourth central moment, in log scale (Panel C) and kurtosis (Panel D) of the distribution $r_t|V_t$, evaluated at a set of selected values of the relevant parameters. All four statistics are computed over a grid of $\delta_t^p \in [0, 1]$, for three specific cases of volatility. The specific values are 6% (low volatility: August 1996), 18% (medium volatility: December 2002) and 68% (high volatility: October 2008).

case is

$$M_{r_t|V_{t-1},\Delta J_t^p}(h) = \left[\left(1-\delta_{t-1}^v\right) + \delta_{t-1}^v\left(1-\mu_v\left(\gamma h+\frac{1}{2}h^2\right)\right)^{-1}\right] \times \exp\left(\left(\mu+\Delta J_t^p\right)h + \overline{V}_t^d\left(\gamma h+\frac{1}{2}h^2\right) + \frac{1}{2}\sigma_v^2\left(1-\rho^2\right)V_{t-1}\left(\gamma h+\frac{1}{2}h^2\right)^2\right) \times \left[\frac{1-\Phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}} - \sigma_v\sqrt{\left(1-\rho^2\right)V_{t-1}}\left(\gamma h+\frac{1}{2}h^2\right)\right)}{1-\Phi\left(\frac{-\left(\overline{V}_t^d\right)}{\sigma_v\sqrt{(1-\rho^2)V_{t-1}}}\right)}\right], \tag{19}$$

where $\overline{V}_t^d = \kappa\theta + (1-\kappa)V_{t-1} + \sigma_v\rho\left(r_{t-1} - Z_{t-1}^p\Delta N_{t-1}^p - \mu - \gamma V_{t-1}\right)$ denotes the conditional mean of the diffusive component of the stochastic variance process. The final term on the right hand side of (19) arises from the positivity bound imposed on the (diffusive) variance process, implying a truncated normal density for the conditional distribution $V_t|V_{t-1}, \Delta J_{t-1}^v, r_{t-1}$.

The first four central moments of the conditional distribution, $r_t|V_{t-1}, \Delta J_t^p$, are tabulated in Panel B of Table 1, with the relevant adjustment terms of the central moments that result from the truncation (denoted by $g_1(\cdot,\cdot)$, $g_2(\cdot,\cdot)$, $g_3(\cdot,\cdot)$ and $g_4(\cdot,\cdot)$) documented in Appendix A. Focussing first on the (standardized) skewness, we note that it is non-zero if and only if γ - the parameter governing volatility feedback to the return in (7) - is non-zero, with a negative value for γ implying negative conditional skewness. Characteristics of the variance jumps - namely the variance jump intensity, δ_{t-1}^v, and the (strictly positive) expected variance jump size, μ_v - then play a direct role in determining the magnitude of the third central moment, and hence skewness, of the conditional return distribution. Non-zero values for the variance jump intensity, δ_{t-1}^v, the expected variance jump size, μ_v, and γ, also serve to produce fatter tails over and above the impact of marginalization over the random variance V_t, as indicated by their contributions to the fourth central moment. Even without variance jumps (i.e. for $\delta_{t-1}^v = 0$), the implied (standardized) kurtosis exceeds three, due to the randomness in the variance process (captured by $\sigma_v^2 > 0$). (For further elaboration of this point, see Appendix A.)

In Figure 2, we display graphically the impact of δ_{t-1}^v on the (conditional) third central moment (Panel A), skewness (Panel B), fourth central moment in log scale (Panel C) and kurtosis (Panel D) of the distribution of $r_t|V_{t-1}, \Delta J_t^p$. These conditional moments are plotted over the range $\delta_{t-1}^v \in [0,1]$, with the same representative values as previously used for V_t in the conditional moments of $r_t|V_t$, imposed for V_{t-1}, and the empirically estimated

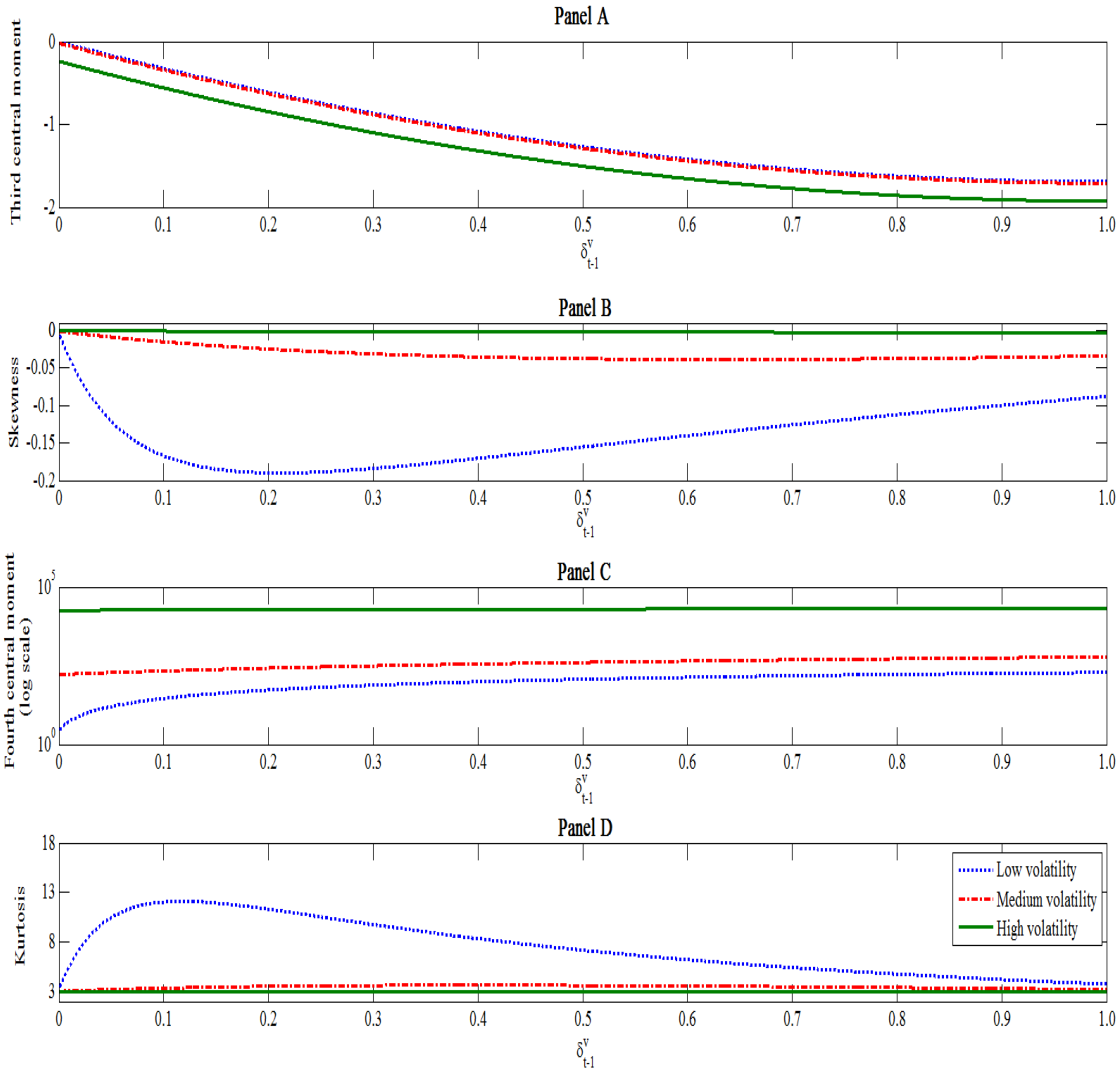

Figure 2: Conditional third central moment (Panel A), skewness (Panel B), fourth central moment, in log scale (Panel C) and kurtosis (Panel D) of the distribution $r_t|V_{t-1}, \Delta J_t^p$, evaluated at a set of selected values of the relevant parameters. All four statistics are computed over a grid of $\delta_t^v \in [0, 1]$, for three specific cases of volatility. The specific values are 6% (low volatility: August 1996), 18% (medium volatility: December 2002) and 68% (high volatility: October 2008).

values (reported in Table 2) used to represent the unknown static parameters in the relevant expressions. Although the behaviour of the central third and fourth moments in Figure 2 is different from that of the corresponding quantities in Figure 1, the qualitative nature of the relationship between δ_{t-1}^{v} and the skewness and kurtosis of $r_t|V_{t-1}, \Delta J_t^p$ is very similar to that between the intensity of price jumps, δ_t^p, and the skewness and kurtosis of $r_t|V_t$. Both the skewness and kurtosis in Figure 2, however, are of a much greater magnitude than the corresponding quantities in Figure 1, indicating that the marginal impact of volatility jumps on the tail behaviour of returns is greater than that of price jumps. Similar to the impact of the price jump intensity, the variance jump intensity has the most marked impact on skewness and kurtosis at the lowest value of V_{t-1}, with the discontinuous movements in the latent variance process inducing the greatest deviation from (conditional) normality at approximately $\delta_{t-1}^{v} = 0.2$. While the skewness and kurtosis approach the values associated with a normal distribution as the variance jump intensity approaches 1, there remains evidence of both fat tails and negative skewness in the conditional distribution, resulting from the diffusive component of the model at time t and the (negative) volatility feedback effect captured by γ. At the medium and high levels of V_{t-1}, the impact of volatility jumps on the tail behaviour of $r_t|V_{t-1}, \Delta J_t^p$, over and above that produced by the diffusive variance, is minimal.

Whilst it has been beyond the scope of the analysis to directly map the dynamic behaviour in the jump intensities into dynamic behaviour in the higher-order conditional return moments, these results linking the magnitude of the moments to the magnitude of the intensity parameters provide some insight into how the conditional moments of returns would change over time were the intensity parameters to change in particular ways. In addition, we see that the impact of jumps would differ according to the level of volatility itself and, thus, the state of the market environment in general. Jumps of either type would exert the greatest influence on the tail behaviour of returns if the market were stable, and have little additional impact if the market environment were already volatile. We explore these characteristics in an empirical setting in Section 5.

3 An augmented state space model

In the spirit of Barndorff-Nielsen and Shephard (2002), Creal (2008), Jacquier and Miller (2012) and Maneesoonthorn, Martin, Forbes and Grose (2012), we exploit the existence of high-frequency data to augment the measurement equation in (7) with additional equations

based on nonparametric measures of return variation: both its diffusive and jump components. In particular, in order to reduce the number of unobservable components to be estimated, we assume that both the price jump occurrence and price jump size are measured without error. Details of the additional measurements are given in the following section, whilst in Section 3.2 a complete outline of the full state space model underlying the empirical analysis in Section 5 is provided.

3.1 The additional measurements

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

$$RV_t = \sum_{t<t_i\leq t+1}^{M} r_{t_i}^2, \tag{20}$$

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

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

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

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

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

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

$$BV_t = \frac{\pi}{2} \sum_{t<t_i\leq t+1}^{M} |r_{t_i}| \left|r_{t_{i-1}}\right|, \tag{21}$$

as a consistent measure of $\mathcal{V}_{t,t+1}$ as $M \to \infty$ (again, in the absence of microstructure noise). The discrepancy between RV_t in (20) and BV_t in (21) then serves as a measure of price jump variation, forming the basis of various statistical tests for price jumps (Barndorff-Nielsen and Shephard 2004 and 2006, and Huang and Tauchen 2005). For example, the asymptotic properties of RV_t and BV_t are used to define the distribution of the relative price jump

statistic,

$$Z_{RJa,t} = \frac{RJ_t}{\sqrt{\left(v_{bb} - v_{qq}\right) M^{-1} \max\left(1, \frac{TQ_t}{BV_t^2}\right)}}, \tag{22}$$

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

$$\Delta N_t^p = \mathbf{1}\left(Z_{RJa,t} > \Phi^{-1}\left(1 - \alpha\right)\right), \tag{23}$$

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

$$\left(Z_t^p | \Delta N_t^p = 1\right) = sign\left(r_t\right) \times \sqrt{\max\left(RV_t - BV_t, 0\right)}, \tag{24}$$

under the assumption that the price jump takes the same sign of the return on that particular day, and that it is observable only when $\Delta N_t^p = 1$. Measures of price jump occurrence and size, constructed using (23) and (24), are used as observable versions of the model components defined in (10) and (14), respectively, as motivated by Tauchen and Zhou, who investigate the time series dynamics of price jumps via these measures.

The bipower variation in (21), on the other hand, is viewed as a measure of the latent variance, V_t, in which measurement error is modelled explicitly. Specifically, and in order to keep the form of the measurement error as close to Gaussian as possible, we utilize BV_t in log form, giving the measurement equation,

$$\ln BV_t = \ln V_t + \sigma_{BV} \xi_t^{BV}. \tag{25}$$

The use of BV_t as a noisy measure of V_t is justified by the fact that BV_t is constructed using only a finite number of observations, and under the assumption of no microstructure noise.

3.2 The full state space model

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

$$r_t = \mu + \gamma V_t + \sigma_r \sqrt{V_t}\xi_t^p + Z_t^p \Delta N_t^p \tag{26}$$

$$Z_t^p = \mu_p + \gamma_p V_t + \sigma_p \xi_t^{Z^p} \tag{27}$$

$$\Delta N_t^p \sim Bernoulli\left(\delta_t^p\right) \tag{28}$$

$$\ln BV_t = \ln V_t + \sigma_{BV}\xi_t^{BV}, \tag{29}$$

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

$$V_{t+1} = \kappa\theta + (1-\kappa) V_t + \sigma_v \rho \left(r_t - Z_t^p \Delta N_t^p - \mu - \gamma V_t\right) + \sigma_v \sqrt{(1-\rho^2) V_t}\xi_t^{ind} + Z_t^v \Delta N_t^v \tag{30}$$

$$Z_t^v \sim Exponential\left(\mu_v\right) \tag{31}$$

$$\Delta N_t^v \sim Bernoulli\left(\delta_t^v\right). \tag{32}$$

In addition, the two conditionally deterministic states representing the price and variance jumps intensities, and contributing respectively to (28) and (32), are specified (once again) as

$$\delta_t^p = \alpha_p \delta_\infty^p + (1-\alpha_p)\delta_{t-1}^p + \beta_{pp}\Delta N_{t-1}^p \tag{33}$$

$$\delta_t^v = \alpha_v \delta_\infty^v + (1-\alpha_v)\delta_{t-1}^v + \beta_{vv}\Delta N_{t-1}^v + \beta_{vp}\Delta N_{t-1}^p + \beta_{vp}^{(-)}\Delta N_{t-1}^{p(-)}, \tag{34}$$

where $\Delta N_{t-1}^{p(-)} = \Delta N_{t-1}^p \mathbf{1}\left(Z_{t-1} < 0\right)$. As noted earlier, the restrictions $\delta_\infty^p > 0$, $0 < \alpha_p < 1$, $\beta_{pp} > 0$, $\alpha_p - \beta_{pp} > 0$, $\delta_\infty^v > 0$, $0 < \alpha_v < 1$, $\beta_{vv} > 0$, and $\alpha_v - \beta_{vv} > 0$ are required. The error components $\left(\xi_t^p, \xi_t^{Z^p}, \xi_t^{BV}, \xi_t^{ind}\right)'$ are collectively assumed to be multivariate standard normal (truncated in the case of ξ_t^{ind} to ensure the positivity of V_{t+1}) with an identity variance-covariance matrix. The additional scale parameter σ_r is introduced into the measurement equation (26) to accommodate the possibility of slight misspecification in the distribution of the (discretized) error term.

4 Bayesian inference

Given the complexity of the state space representation, and the high dimension of the set of unknowns, the joint posterior is not available in closed form. Hence, an MCMC algorithm is developed to obtain draws of the parameters and latent variables of interest from the

joint posterior distribution, and inference conducted using those draws. In particular, Bayes factors for the comparison of the alternative models corresponding to various hypotheses of interest, to be specified in Section 4.2, are computed from the simulation output using the marginal likelihood computation methods of Chib (1995) and Chib and Jeliazkov (2001).

4.1 MCMC algorithm and priors

For notational convenience, we denote time-indexed variables generically as, for example, $X_{1:t} = (X_1, ..., X_t)'$ for $t = 1, ..., T$ and where $X_{1:0}$ is empty. The static parameters are collectively denoted by the vector $\phi = \left(\mu, \gamma, \sigma_r, \mu_p, \gamma_p, \sigma_p, \sigma_{BV}, \kappa, \theta, \sigma_v, \rho, \delta_0^p, \alpha_p, \beta_{pp}, \delta_0^v, \alpha_v, \beta_{vv}, \beta_{vp}, \beta_{vp}^{(-)}, \mu_v\right)'$. The joint posterior density associated with the full model in (26)-(34) satisfies

$$
\begin{aligned}
& p\left(V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}^v, \phi | r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p\right) \\
& \propto p\left(r_1 | V_1, \Delta N_1^p; \mu, \gamma\right) \times p\left(\ln BV_1 | V_1; \sigma_{BV}\right) \times p\left(\Delta N_1^p | \delta_1^p\right) \\
& \times p\left(V_1 | \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right) \times p\left(Z_1^v | \Delta N_1^v, \mu_v\right) \times p\left(\Delta N_1^v | \delta_1^v\right)] \times p\left(\phi\right) \\
& \times \left[\prod_{t=2}^{T} p\left(r_t | V_t, \Delta N_t^p; \mu, \gamma\right) \times p\left(\ln BV_t | V_t; \sigma_{BV}\right) \times p\left(\Delta N_t^p | \Delta N_{1:t-1}^p; \alpha_p, \beta_{pp}, \delta_0^p\right)\right. \\
& \times p\left(V_t | V_{t-1}, Z_{t-1}^v, \Delta N_{t-1}^v, r_{t-1}, \Delta N_{t-1}^p; \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right) \\
& \left. \times p\left(Z_t^v | \Delta N_t^v, \mu_v\right) \times p\left(\Delta N_t^v | \Delta N_{1:t-1}^v, \Delta N_{1:t-1}^p, Z_{1:t-1}^p; \alpha_v, \beta_{vv}, \beta_{vp}, \beta_{vp}^{(-)}, \delta_0^v\right)\right],
\end{aligned} \tag{35}
$$

where it is assumed that $V_1 = \theta + \frac{\mu_v \delta_0^v}{\kappa}$, $\delta_1^p = \delta_0^p$, $\delta_1^v = \delta_0^v$ and $\Delta N_1^v = \Delta N_1^p$.

In the specification of $p\left(\phi\right)$ in (35) we use a combination of noninformative and weakly informative priors for the various elements of ϕ. Uniform priors are assumed for the parameters κ and θ, truncated from below at zero, while the parameter σ_v^2 is blocked with the leverage parameter, ρ, via the reparameterization: $\psi = \rho\sigma_v$ and $\omega = \sigma_v^2 - \psi^2$; see Jacquier *et al.* (2004). This reparameterization is convenient as, given $V_{1:T}$, it allows for ψ and ω to be treated respectively as the slope and error variance coefficients in a normal linear regression model. Direct sampling of ψ and ω is then conducted using standard posterior results, based on conjugate prior specifications in the form of conditional normal and inverse gamma distributions, respectively, given by $p\left(\psi|\omega\right) \sim N\left(\psi_0 = -0.005, \sigma_0^2 = \omega/5.0\right)$ and $p\left(\omega\right) \sim IG\left(v_0 = 10, v_0 t_0^2 = 0.0001\right)$, where $v_0 t_0^2$ denotes the scale parameter. The prior specifications for ψ and ω are chosen such that the implied prior distributions for ρ and σ_v (as gauged by the distributions of the draws of ψ and ω) are broadly in line with the range of empirical values for these parameters reported in the literature.

Truncated uniform priors are specified for the parameters μ, γ, μ_p and γ_p, with the prior for μ covering both the positive and negative regions over a relatively wide interval, and the volatility feedback parameter γ, and the parameters governing the conditional mean of the price jump size, μ_p and γ_p, assumed *a priori* to be bounded from above at zero. The negativity constraint imposed on both μ_p and γ_p is in line with the expectation that price jumps are negative on average, and that higher volatility induces even larger (negative) price jumps. The prior assumption of $\gamma < 0$ is consistent with recent findings of negative volatility feedback in the high frequency literature. (See, for example, Bollerslev *et al.* 2006, and Jensen and Maheu 2013). Conjugate inverse gamma priors are applied to the parameters σ_r^2, σ_p^2 and σ_{BV}^2, with hyperparameters chosen to render the prior distributions relatively diffuse.

Conjugate beta priors are employed for the unconditional jump intensities, δ_0^p and δ_0^v. The parameters of the beta prior for δ_0^p are specified such that the prior mean matches the sample mean of the observed $\Delta N_{1:T}^p$. The mean of the prior beta distribution for δ_0^v is, in turn, equated with the prior mean of δ_0^p. This stems 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, $\Delta N_t^v = 1$), generating similar unconditional jump intensities in the two processes. A conjugate prior in the form of an inverse gamma distribution is employed for μ_v, with the characteristics of the prior distribution chosen to produce a prior mean that is proportional to the average magnitude of negative price jump variation. Uniform priors are employed for the jump intensity parameters, conforming to the theoretical restrictions listed at the end of Section 3.2, and the prior belief that $\beta_{vp} > 0$ and $\beta_{vp}^{(-)} > 0$.[1]

A hybrid of the Gibbs and Metropolis-Hastings (MH) algorithms is used for sampling the latent variables and the elements of the static parameter vector. Details of the algorithm, including a reference made to Maneesoonthorn *et al.* (2012) for a description of the multi-move algorithm adopted for sampling the state vector, $V_{1:T}$, are given in Appendix B.

4.2 Hypotheses of interest: Posterior odds and Bayes factors

As has been highlighted, a novel aspect of our specification is that it allows for dynamic behaviour in both price and volatility jumps, as well as various types of dependencies between

[1] We assess the sensitivity of our empirical results to the prior specification using the prior and posterior predictive analysis of Geweke (2005). While the prior predictive analysis shows that certain prior specifications better describe particular characteristics of the data than others, the posterior predictive results are nevertheless robust to our prior specification. These results can be obtained from the authors on request.

those extreme movements, with the characteristics of both jump processes influencing the higher-order properties of returns. It is of interest then to explore whether or not this rich dynamic structure is warranted empirically, through an investigation of various hypotheses associated with the jump intensity processes in (33) and (34).

With a view to determining the relative importance of the various components of the proposed model, we consider several competing hypotheses nested within the general specification, all of which are to be evaluated in Section 5. First, we determine whether the occurrence of negative price jumps (per se) influences the variance jump intensity by testing the hypothesis

$$H_1 : \beta_{vp}^{(-)} = 0; \tag{36}$$

whether the variance jump intensity is *only* influenced by past negative price jumps by testing the hypothesis

$$H_2 : \beta_{vp} = 0; \tag{37}$$

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

$$H_3 : \beta_{vp}^{(-)} = \beta_{vp} = 0. \tag{38}$$

The second issue of interest is whether the jump intensities for price and variance, respectively, are constant, with the hypotheses

$$H_4 : \delta_t^p = \delta_0^p \text{ for all } t = 1, ...T \text{ and } \beta_{pp} = 0, \tag{39}$$

and

$$H_5 : \delta_t^v = \delta_0^v \text{ for all } t = 1, ...T \text{ and } \beta_{vv} = \beta_{vp} = \beta_{vp}^{(-)} = 0 \tag{40}$$

being tested. Note that if both H_4 and H_5 hold, as the joint hypothesis

$$H_6 : \delta_t^p = \delta_0^p, \delta_t^v = \delta_0^v \text{ for all } t = 1, ...T \text{ and } \beta_{pp} = \beta_{vv} = \beta_{vp} = \beta_{vp}^{(-)} = 0, \tag{41}$$

the model corresponds to the SVIJ model of Duffie *et al.* (2000). In addition, we investigate the hypothesis

$$H_7 : \delta_t^p = 0 \text{ and } \delta_t^v = 0 \text{ for all } t = 1, ...T, \tag{42}$$

in which case both price and variance jumps are absent and the model corresponds to the conventional Heston (1993) square root model.

To examine these seven hypotheses of interest, computation of posterior odds ratios is required. Let us denote the state space model with full dynamics - collectively defined by

(26)-(34) - by $\mathcal{M}_F$, and let $\mathcal{M}_i$ denote the model specified under the restrictions imposed by hypothesis H_i, for $i = 1, ..., 7$. Under the assumption that each of the models is *a priori* equally likely, the posterior odds ratio for any restricted model $\mathcal{M}_i$ relative to the full state space model $\mathcal{M}_F$, is equivalent to the Bayes factor $BF_{i,F}$, given in turn by

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

Note that given Bayes factors $BF_{i,F}$ and $BF_{j,F}$, the Bayes factor for model $\mathcal{M}_i$ against $\mathcal{M}_j$ is obtained simply as $BF_{i,j} = BF_{i,F}/BF_{j,F}$.

The marginal likelihood for model $\mathcal{M}_i$ that appears in the numerator of (43) is defined as

$$\begin{aligned}
&p\left(r_{1:T}, \ln BV_{1:T}, \Delta N^p_{1:T}, Z^p_{1:T} | \mathcal{M}_i\right) \\
&= \int p\left(r_{1:T} | V^{(i)}_{1:T}, Z^p_{1:T}, \Delta N^p_{1:T}, \Delta N^{v(i)}_{1:T}, \Delta N^{v(i)}_{1:T}; \phi_i, \mathcal{M}_i\right) p\left(\ln BV_{1:T} | V^{(i)}_{1:T}; \phi_i, \mathcal{M}_i\right) \\
&p\left(\Delta N^p_{1:T} | \phi_i, \mathcal{M}_i\right) p\left(V^{(i)}_{1:T} | Z^{v(i)}_{1:T}, \Delta N^{v(i)}_{1:T}; \phi_i, \mathcal{M}_i\right) p\left(Z^{v(i)}_{1:T} | \Delta N^{v(i)}_{1:T}; \phi_i, \mathcal{M}_i\right) \\
&p\left(\Delta N^{v(i)}_{1:T} | \Delta N^p_{1:T}; \phi_i, \mathcal{M}_i\right) p\left(\phi_i\right) d\left(V^{(i)}_{1:T}, Z^{v(i)}_{1:T}, \Delta N^{v(i)}_{1:T}, \phi_i\right),
\end{aligned} \tag{44}$$

where ϕ_i is the static parameter vector under model $\mathcal{M}_i$, while $V^{(i)}_{1:T}$, $Z^{v(i)}_{1:T}$ and $\Delta N^{v(i)}_{1:T}$ denote respectively the vectors of latent variances, variance jump sizes and variance jump occurrences under model $\mathcal{M}_i$. Evaluating the marginal likelihoods is challenging, as the integral corresponding to (44) for each model, $\mathcal{M}_F$ and $\mathcal{M}_i$, for $i = 1, ..., 7$, is not available in closed form and is of a very large dimension due to the large number of latent variables present. We estimate the marginal likelihood of each model using the output of a series MCMC algorithms, details of which are provided in Appendix C.

5 Empirical analysis

5.1 Data description

In this section we report the results of the application of the proposed method to daily annualized returns, plus annualized bipower variation and annualized price jump measures, constructed from spot price data of the S&P500 index, for the period January 3, 1996 to May 6, 2013. The period of January 3, 1996 to May 15, 2012 (comprising 4070 trading days) is used for in-sample analysis, while the period of May 16, 2012 to May 6, 2013 (comprising 242 days) is preserved for evaluating out-of-sample predictive return distributions and associated

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

The in-sample daily annualized return, r_t, is computed as the difference between the log of the closing index value and that of the opening value on a particular trading day, and is plotted in Panel A of Figure 3 for the relevant sample period. Figure 3 also contain plots of the in-sample daily annualized quantities of the bipower variation measure (BV_t) and its log ($\ln BV_t$) in Panel B and the price jump measure ($Z_t^p \Delta N_t^p$) in Panel C, with the latter plot giving a combined picture of the price jump timing and the size and direction of such jumps. Both measures, BV_t and $Z_t^p \Delta N_t^p$, are based on fixed five minute sampling, with a 'nearest price' method (Andersen, Bollerslev and Diebold, 2007) used to construct artificial returns five minutes apart. With reference to (23), the price jump occurrence measure ΔN_t^p is extracted using the significance level of $\alpha = 0.001$ for the price jump test statistic in (22). As is evident in Panels A and B, and as is completely standard knowledge in this setting, volatility clustering in returns exists, with the most extreme variation in returns, plus BV_t values of unprecedented magnitude, observed towards the end of 2008. The plot of the price jump measure, $Z_t^p \Delta N_t^p$, in Panel C of Figure 3 shows clear evidence of price jump clustering, with clusters being scattered over the full in-sample period. The most intense volatility clustering, and the clusters of price jumps that are largest in magnitude, occur during three of the most volatile market periods: the year 2002 (and the associated market instability following the September 11 terrorist attacks), the global financial crisis in 2008, and the recent Euro-zone debt crisis in 2009.

5.2 The parameter estimates

Estimates of the static parameters of the full state space model, $\mathcal{M}_F$, for the sample period between January 3, 1996 and May 15, 2012, are recorded in Table 2. Marginal posterior means (MPMs) and 95% higher posterior density (HPD) interval estimates are calculated from $G = 100{,}000$ MCMC draws, following a 100,000 draw burn-in period. Inefficiency factors 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 sampler.

The parameters associated with the measurement equations (26) and (29), namely μ, γ, σ_r

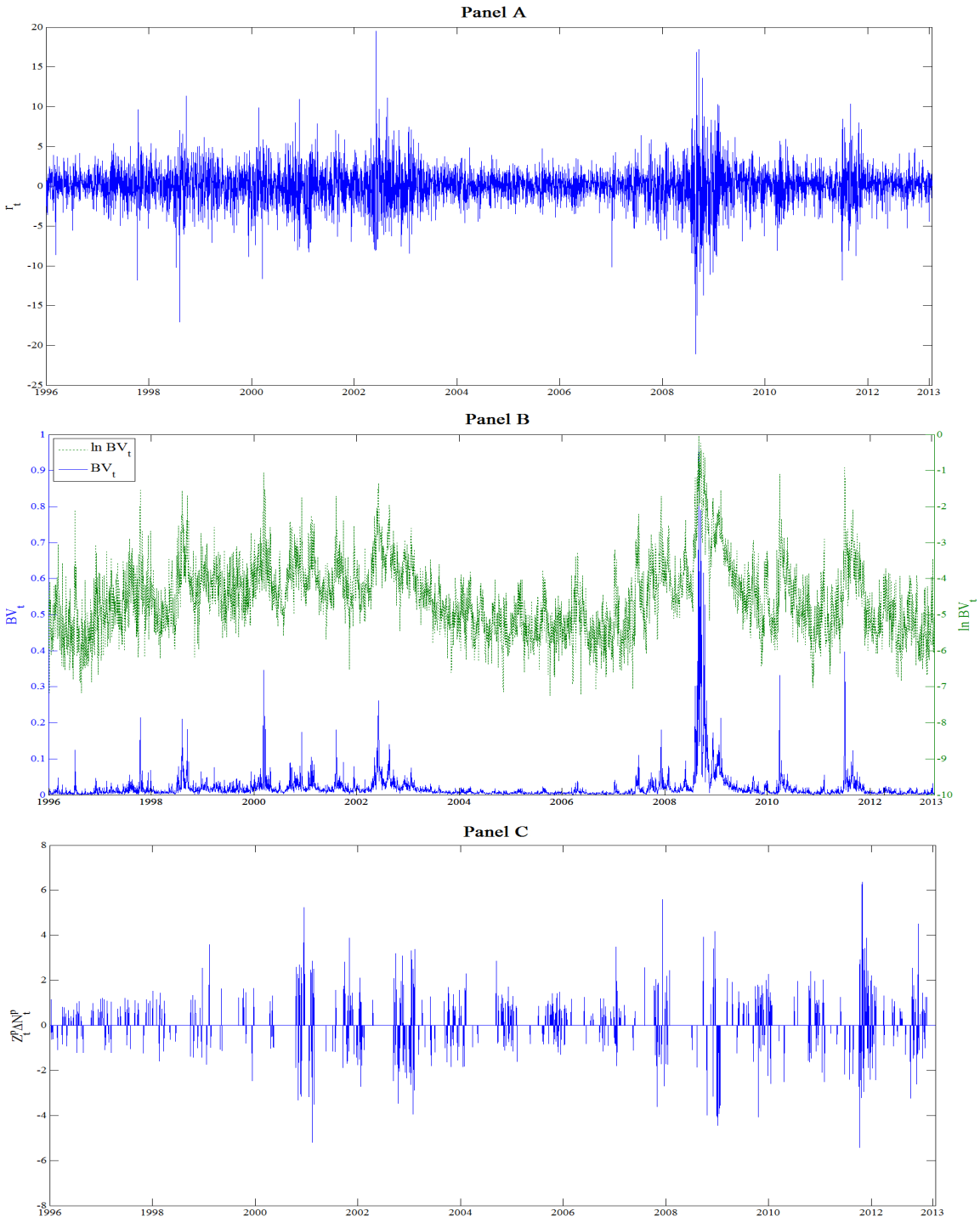

Figure 3: The daily annualized daily return, r_t, of the S&P500 index (Panel A), the measure of bipower variation, BV_t, and its log, $\ln BV_t$, (Panel B) and price jumps, $Z_t^p \Delta N_t^p$, (Panel C). The non-parametric measures in Panel B and C are constructed using intraday sampling at the 5 minute frequency. The data ranges from January 3, 1996 to May 3, 2013, inclusive.

and σ_{BV}, are all within ranges that are consistent with those obtained in the study reported in Maneesoonthorn *et al.* (2012), where the integrated variance measure BV_t, supplemented by a model-free option-implied volatility measure and subsequently augmented by daily returns, is used to conduct inference on the latent variance process over a sample period from July 1999 and December 2008. The reported MPMs of θ and σ_v in Table 2 are broadly consistent with those that have been reported earlier for similar models (see Broadie *et al.* 2007, for a summary). The value of κ, on the other hand, is rather high compared to other estimates reported in the literature, with a possible explanation being that the degree of persistence in the latent variance process is partially captured by the dynamic model for the variance jump intensity in our specification. The estimate of μ_v is also consistent with that reported in Maneesoonthorn *et al.* (2012), although higher in magnitude than that reported by Broadie *et al.*, based on a less volatile sample of data. The MPM of the leverage parameter, ρ, is negative and its magnitude, ranging from 0.51 to 0.68, albeit relatively high, is still consistent with figures previously reported in the literature (see Broadie *et al.* and Aït-Sahalia, Fan and Li 2013, for examples).

The parameters associated with the two jump intensity processes are, of course, our primary interest. As the price jump variables, Z_t^p and ΔN_t^p, are treated as observable, and certain prior specifications are linked to these observed variables, it is not surprising that the reported MPM of the unconditional price jump intensity, δ_0^p, is close to the sample mean of the ΔN_t^p measure. The dynamic price jump intensity process, δ_t^p, possesses a strong degree of persistence, as indicated by the relatively low MPM of α_p, and a marginal 95% HPD interval for β_{pp} that is well above zero, consistent with the presence of self-excitation. The magnitudes of α_p and β_{pp} reported here, once annualized, are consistent with the parameters reported in Aït-Sahalia *et al.* (2013), who propose the use of 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 quantities (Eraker *et al.* 2003, Eraker 2004 and Broadie *et al.* 2007). The variance jump intensity process is also more persistent than the price jump intensity process, with the MPM of α_v being lower in magnitude than that of α_p. In addition there is evidence of self-exciting dynamics as indicated by the non-zero MPM of β_{vv}. The self-exciting dynamics in δ_t^v, measured by β_{vv}, are much stronger than the feedback from the previous price jump occurrence, measured by β_{vp}, and its threshold component, measured by $\beta_{vp}^{(-)}$, with the marginal posterior densities for both β_{vp} and $\beta_{vp}^{(-)}$ being highly concentrated around mean values very close to zero. Further assessment of the importance of these feedback effects,

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

Parameter	MPM (95% HPD) Inefficiency Factor	Parameter	MPM (95% HPD) Inefficiency Factor
μ	0.097 (0.032, 0.166) 11.3	δ_0^p	0.104 (0.095, 0.114) 3.4
γ	-3.604 (-8.008, -0.325) 8.3	α_p	0.094 (0.069, 0.125) 34.8
σ_r	1.093 (1.068, 1.119) 4.2	β_{pp}	0.059 (0.044, 0.075) 46.5
μ_p	-0.029 (-0.102, -8.06e^{-4}) 3.4	δ_0^v	0.064 (0.042, 0.092) 114.1
γ_p	-3.276 (-9.342, -0.119) 3.7	α_v	0.046 (0.031, 0.065) 256.6
σ_p	1.795 (1.681, 1.920) 3.4	β_{vv}	0.038 (0.026, 0.051) 260.7
σ_{BV}	0.469 (0.455, 0.484) 12.4	β_{vp}	4.89e^{-4} (1.27e^{-5}, 1.76e^{-3}) 8.6
κ	0.137 (0.112, 0.164) 234.6	$\beta_{vp}^{(-)}$	1.13e^{-3} (3.06e^{-5}, 3.88e^{-3}) 10.0
θ	0.0083 (0.0075, 0.0091) 50.1	μ_v	0.025 (0.019, 0.032) 395.2
ρ	-0.599 (-0.681, -0.509) 86.4	σ_v	0.020 (0.019, 0.022) 76.3

plus the importance of the dynamic structures specified for price and variance jumps, and of the presence of jumps per se, is conducted in Section 5.4, via Bayes factors.[2]

The inefficiency factors reported in Table 2 for the static parameters range from 3 to 400, with the parameters associated with the dynamic variance jump intensity and size producing the highest inefficiency factors. The factors for the latent variance, $V_{1:T}$, computed at selected time points, range from 6 to 56. 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%.

5.3 The latent processes

Time series plots of the MPM of the latent variance, V_t, and the variance jump, $\Delta J_t^v = Z_t^v \Delta N_t^v$, computed at every time point over the estimation period for the full model, $\mathcal{M}_F$, are shown in Panels A and B, respectively, of Figure 4. Similarly, Figure 5 displays the time series plots of the MPM of the jump intensity processes, δ_t^p and δ_t^v, in panels A and B respectively.

Reconciling the patterns in Panel A of Figure 4 with those in Panel B of Figure 3, it is clear that the overall dynamics in the MPM of V_t closely track those of its direct measure, BV_t, even in the most extreme period towards the end of 2008. In Panel B of Figure 4, there is also distinct evidence of clustering of variance jumps in high volatility periods - associated, in turn, with an increase the (point) estimate of δ_t^v, as shown in Panel B of Figure 5. 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 low estimates (point and interval) of α_v recorded in Table 2. In order to identify any relationship between the variance jump intensity and general market conditions, and to assess whether or not δ_t^v may be useful as an indicator of extreme market events, Figure 6 reproduces the MPMs of δ_t^v, over the period between January 2007 and May 2012, with many well-known market events indicated. As can be observed, some of the sharpest rises in δ_t^v are either synchronous with, or occur soon

[2] The MPM and 95% HPD intervals for the relevant parameter sets of the alternative models investigated are not reported here to conserve space. In summary, the estimates (point and interval) of the static parameters under the alternative models are relatively robust to model specification, with the exception to this being certain parameters in the models in which the variance jump intensity is restricted to be constant. In particular, the models that impose constant variance jump intensities imply a higher degree of persistence in the diffusive volatility component and a higher overall level of volatility than the remaining models, in which an additional avenue for dynamic behaviour in volatility is available. These additional results are available from the authors on request.

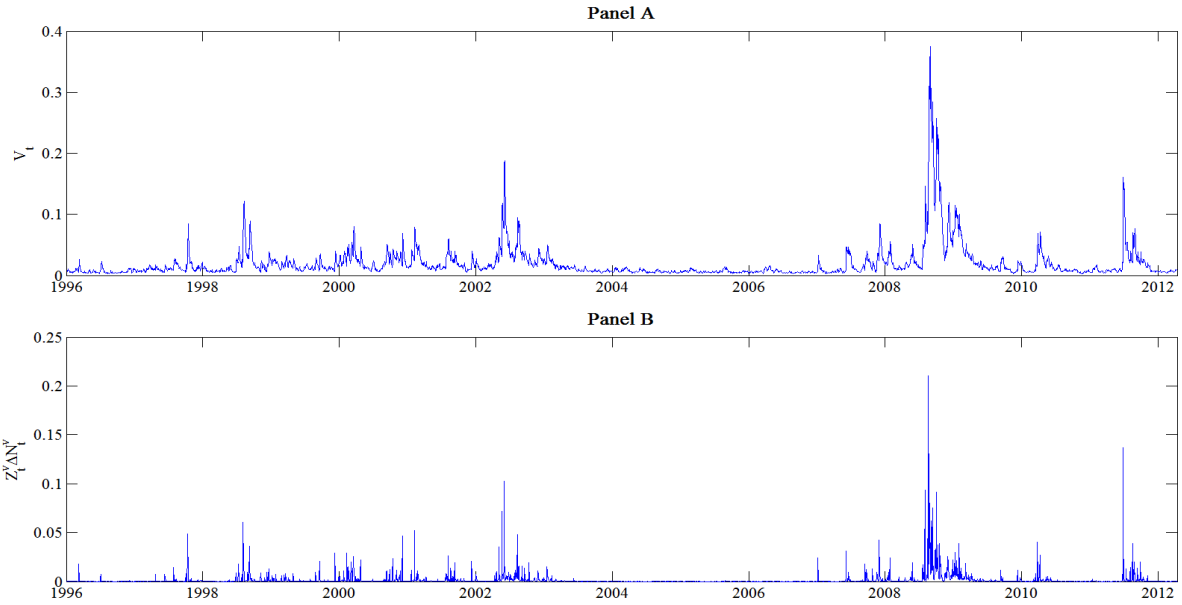

Figure 4: Time series plots of the marginal posterior mean of the latent variance, V_t, (Panel A) and variance jump, $\Delta J_t^v = Z_t^v \Delta N_t^v$, (Panel B) over the sample period January 3, 1996 to May 15, 2012, inclusive.

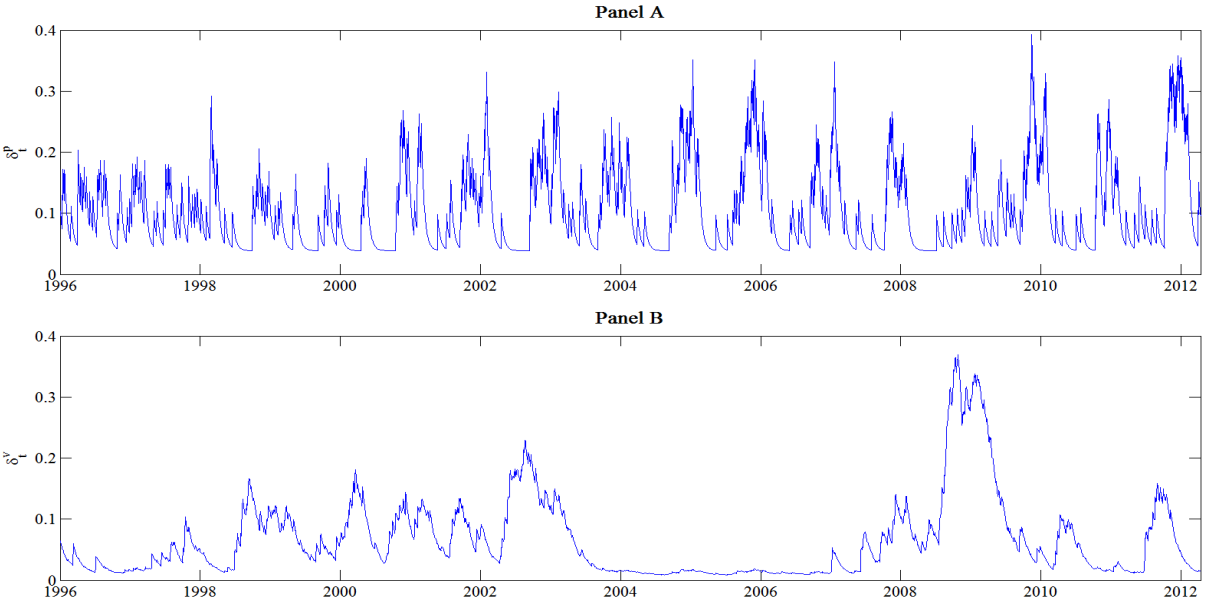

Figure 5: Time series plots of the marginal posterior mean of the price jump intensity, δ_t^p, (Panel A) and variance jump intensity, δ_t^v, (Panel B), over the sample period January 3, 1996 to May 15, 2012, inclusive.

after, certain key events. In particular, the collapse of the Lehman Brothers (September, 2008) and the subsequent intervention by the U.S. Federal Reserve (December, 2008) are followed closely by the largest variance jump intensity levels observed throughout the entire sample period (reaching a peak of almost 0.37 on December 5th, 2008). During the various phases of the recent Euro-zone debt crisis (starting from late 2009), sharp increases in the MPM of δ_t^v are also evident, albeit with the magnitude of these being less than the rises observed during the global financial crisis.

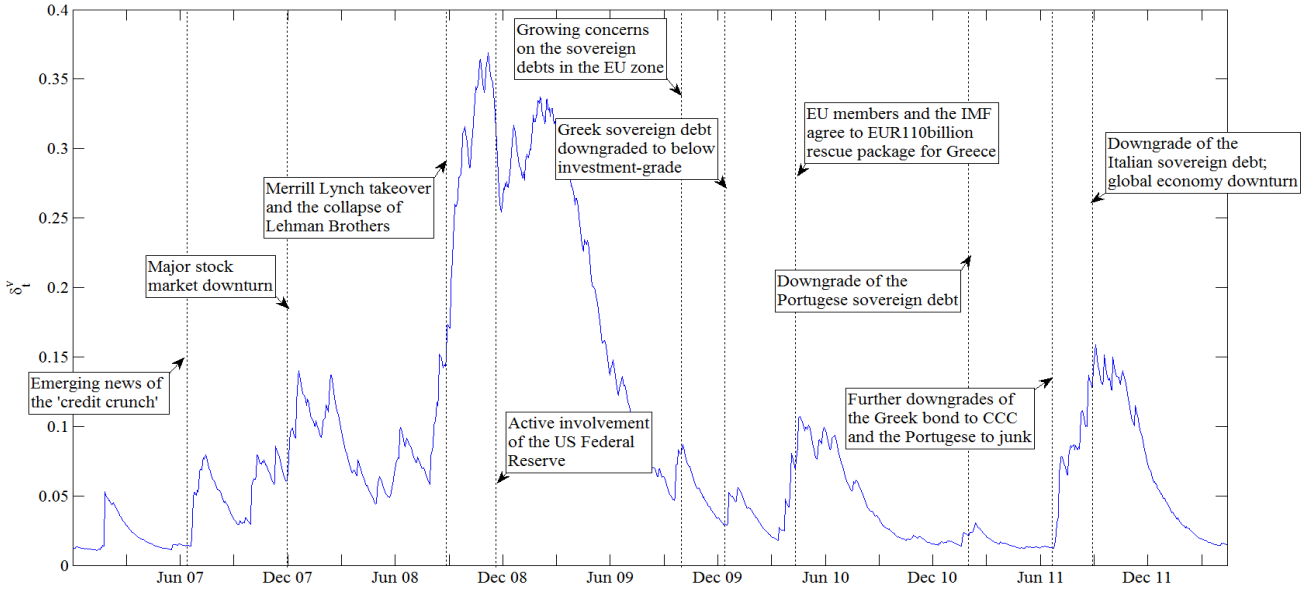

Figure 6: Time series plot of the variance jump intensity process, δ_t^v, over the period January 2007 to May 2012, inclusive, with the timing of various important market events noted, including the recent global financial crisis, as well as the Euro-zone debt crisis.

Panel A of Figure 5 depicts the MPM of the price jump intensity variable, δ_t^p, over the sample period. In contrast to the evident link between the magnitude of the variance jump intensity and the characteristics of the market, the magnitude and dynamic behaviour of δ_t^p do not track market conditions in any obvious way. Specifically, an increase in the intensity of price jumps is both relatively short lived (compared to that of variance jumps) and associated with periods in which the magnitude of the observed jumps (Figure 3, Panel C) may either be large or small. That is, an increase in price jump intensity does not appear to correlate with a period of *large* price jumps only.

Finally, as discussed in Section 2.3, the assumed jump processes contribute in particular ways to the higher-order moments of the returns. For the purpose of illustration, the MPMs

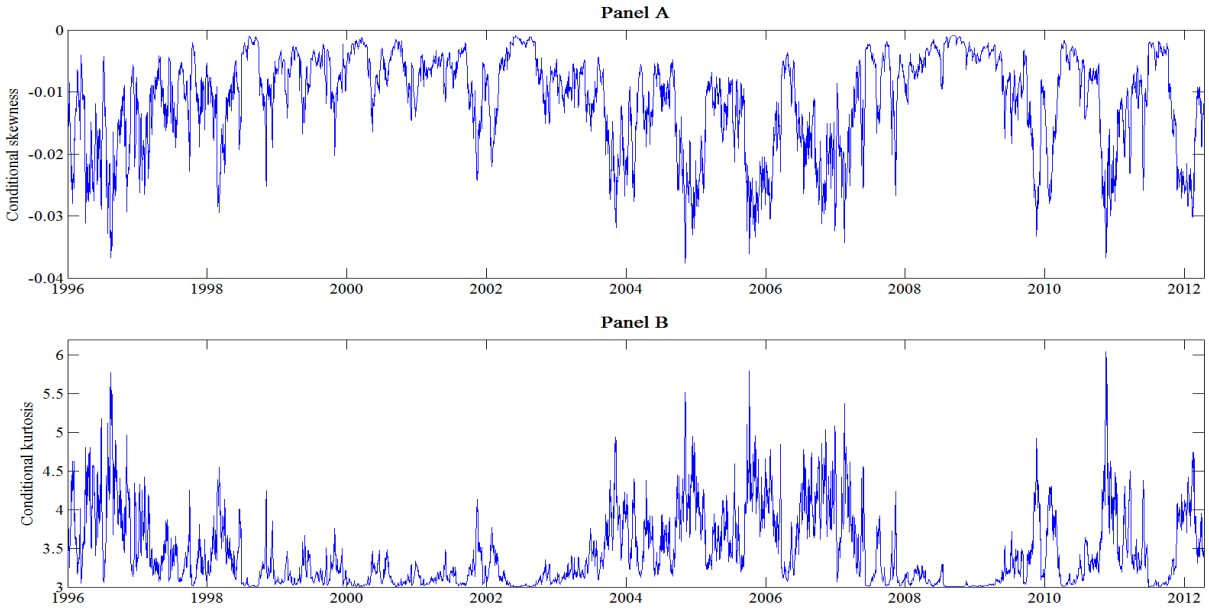

Figure 7: Marginal posterior means of the model implied conditional skewness (Panel A) and conditional kurtosis (Panel B) of each distribution $r_t|V_t$ associated with the sample period January 3, 1996 to May 15, 2012, inclusive. The marginal posterior mean is used as the relevant conditioning value of V_t for each t.

of the skewness and kurtosis of the conditional distribution $r_t|V_t$ implied by the assumed model, and as produced as an average of the relevant functions of the MCMC draws, are depicted in Panels A and B of Figure 7, respectively. (See Panel A of Table 1 for the analytical expressions). Given the conditioning on the MPM of V_t, at each t, the dynamics in these conditional moments are driven only by the dynamics in the price jump process, with periods of price jump clustering, as indicated by increases in δ_t^p (Panel A of Figure 5), being typically (but not always) associated with rises in the magnitude of both moments in Figure 7. In a period of medium volatility, between 2000 and 2004, and a period of high volatility, between 2008 and 2010, the conditional skewness and kurtosis revert towards Gaussian quantities of zero and three, respectively. This is consistent with the discussion in Section 2.3 where it was shown that the impact of price jumps is relatively small during periods of medium and high volatility. In contrast, price jumps were shown to exert their greatest influence on higher-order conditional moments during periods of low volatility, something that is now confirmed numerically by the appearance of consistently high (in magnitude) values of skewness and kurtosis during the (relatively) low volatility periods near the beginning of the sample, between 2004 and 2007 (approximately), and over the 2010-2011 period. Note that

a similar exercise was conducted to illustrate the role of the variance jump component on the conditional skewness and kurtosis of $r_t|V_{t-1}, \Delta J_t^p$. The resulting relationships between these moments and the intensity of variance jumps (not shown here), are also in line with the analytical results provided in Section 2.3.

5.4 Model selection and Bayes factors

Evaluation of hypotheses H_1 to H_7, described in (36) to (42) respectively in Section 4.2, is conducted by constructing Bayes factors for each pair of competing models. Table 3 reports the log of the Bayes factor for $\mathcal{M}_F$ versus $\mathcal{M}_i$, for $i = 1, 2, ..., 7$ (with each restricted model $\mathcal{M}_i$ corresponding to hypotheses H_i) with models indicated on the vertical axis used as the model in the numerator of the Bayes factor. Positive values for the log Bayes factor indicate that the corresponding ratio is greater than one, and hence the model of the numerator favoured, given a symmetric loss function. The relevant posterior model probabilities are also reported in the last row of Table 3.

A key message is that models in which both types of jumps are present are strongly supported by the data, and that models in which some kind of jump dynamic is allowed are even more favoured, with negative log Bayes factors appearing in the rows corresponding both to $\mathcal{M}_7$ (the model that excludes price and volatility jumps) and to $\mathcal{M}_6$ (the model that imposes constant price and variance jump intensities). Models $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$, which impose restrictions on the impact of price jumps on variance jump intensity, produce higher posterior probabilities than the full dynamic model, $\mathcal{M}_F$. Indeed, the model $\mathcal{M}_3$, in which feedback from a price jump (positive or negative) to the subsequent variance jump intensity is omitted, is assigned a posterior probability of 1.00 (to the level of precision recorded) in the set of eight models. The model $\mathcal{M}_4$, corresponding to the hypothesis that imposes a constant price jump intensity but still allows for a dynamic variance jump intensity, is favoured when compared with model $\mathcal{M}_5$, where the variance jump intensity is constant but the price jump intensity is dynamic. However, despite this apparent support for variance jump dynamics over price jump dynamics, it is worth remembering (as noted above) that the marginal posterior probabilities for all models other than $\mathcal{M}_3$ - in which both forms of dynamics co-exist - are negligible. The support for $\mathcal{M}_3$ is also consistent with the extremely small magnitudes of the point and interval estimates of β_{vp} and $\beta_{vp}^{(-)}$ from the full dynamic model, $\mathcal{M}_F$, reported in Table 2.

Table 3: Log Bayes factor in favour of model i (row) over model j (column), computed using the data from January 3, 1996 to May 15, 2012, inclusive.

	Model j							
Model i	$\mathcal{M}_F$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$	$\mathcal{M}_5$	$\mathcal{M}_6$	$\mathcal{M}_7$
$\mathcal{M}_1$ See H_1 in (36)	88	0	–	–	–	–	–	–
$\mathcal{M}_2$ See H_2 in (37)	14	-74	0	–	–	–	–	–
$\mathcal{M}_3$ See H_3 in (38)	143	55	130	0	–	–	–	–
$\mathcal{M}_4$ See H_4 in (39)	-167	-255	-181	-310	0	–	–	–
$\mathcal{M}_5$ See H_5 in (40)	-410	-497	-423	-553	-242	0	–	–
$\mathcal{M}_6$ See H_6 in (41)	-663	-751	-677	-806	-496	-253	0	–
$\mathcal{M}_7$ See H_7 in (42)	-6577	-6664	-6590	-6720	-6409	-6167	-5914	0
Posterior Model Probability	$6.8e^{-63}$	$9.3e^{-25}$	$5.6e^{-57}$	1.00	$1.6e^{-135}$	$8.8e^{-241}$	0.00	0.00

5.5 Evaluation of predictive densities

The model comparison exercise conducted in the previous section focusses on in-sample performance. It is also of interest, of course, to evaluate the forecasting value of the dynamic structures proposed in the paper. For this purpose we choose to assess the accuracy of the predictive returns distribution associated the most favoured model, $\mathcal{M}_3$, compared with that of model $\mathcal{M}_6$, in which constant jump intensities are imposed. Using the posterior draws of the static parameters based on the data from January 3, 1996 to May 15, 2012, inclusive, 242 one-step-ahead predictive densities are produced over the out-of-sample period comprised of May 16, 2012 to May 3, 2013, inclusive. Forecasts of the latent processes, $V_{T+1}, Z_T^v, \Delta N_T^v, \delta_{T+1}^p$ and δ_{T+1}^v, are produced using a fixed window of $n = 4070$ daily observations, with the forecasts updated recursively over the forecast evaluation period. The market conditions over this period are stable, with relatively low volatility levels observed, and the effect of jumps on the tail behaviour of the conditional return distribution is expected to be more marked as a consequence. Thus, this period thus serves well as an evaluation period for assessing the value of modelling jumps per se, as well as the added value of imposing a dynamic structure on the intensities.

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

$$\begin{aligned}
&p\left(r_{T+1}|r_{T-n+1:T}, \ln BV_{T-n+1:T}, \Delta N_{T-n+1:T}^p, Z_{T-n+1:T}^p, \mathcal{M}_i\right) \\
&= \int p\left(r_{T+1}|V_{T+1}, Z_{T+1}^p, \Delta N_{T+1}^p\right) p\left(V_{T+1}|V_T, Z_T^v, \Delta N_T^v, \phi_i\right) \\
&p\left(V_T, Z_T^v, \Delta N_T^v|r_{T-n+1:T}, \ln BV_{T-n+1:T}, \Delta N_{T-n+1:T}^p, Z_{T-n+1:T}^p, \phi_i\right) \\
&p\left(\Delta N_{T+1}^p|\Delta N_{T-n+1:T}^p, \phi_i\right) p\left(Z_{T+1}^p|V_{T+1}, \phi_i\right) d\left(V_{T+1}, V_T, Z_T^v, \Delta N_T^v, \phi_i\right). \qquad (45)
\end{aligned}$$

We evaluate (45) using two different criteria. First, the relative predictive performance of the two competing models is evaluated based on the cumulative difference in log score (CLS), computed over the evaluation period as

$$CLS\left(k\right) = \sum_{t=T+1}^{T+k} \ln\left[\frac{p\left(r_t^o|r_{k:t-1}, \ln BV_{k:t-1}, \Delta N_{k:t-1}^p, Z_{k:t-1}^p, \mathcal{M}_3\right)}{p\left(r_t^o|r_{k:t-1}, \ln BV_{k:t-1}, \Delta N_{k:t-1}^p, Z_{k:t-1}^p, \mathcal{M}_6\right)}\right], \qquad (46)$$

for $k = 1, ..., 242$ (see also Geweke and Amisano 2010, for discussion of the use of CLS in predictive performance evaluation). The notation $p\left(r_t^o|r_{k:t-1}, \ln BV_{k:t-1}, \Delta N_{k:t-1}^p, Z_{k:t-1}^p, \mathcal{M}_3\right)$ denotes the (estimated) predictive density - conditional on model $\mathcal{M}_3$ - evaluated at the observed value of r_t^o, with the conditioning dataset of time subscript $k : t-1$ always retaining

the sample size $n = 4070$. The notation $p\left(r_t^o | r_{k:t-1}, \ln BV_{k:t-1}, \Delta N_{k:t-1}^p, Z_{k:t-1}^p, \mathcal{M}_6\right)$ is correspondingly defined. A positive value of CLS indicates that the preferred model, $\mathcal{M}_3$, with the dynamic jump intensity structures, outperforms the simpler model, $\mathcal{M}_6$, that imposes constant jump intensities.

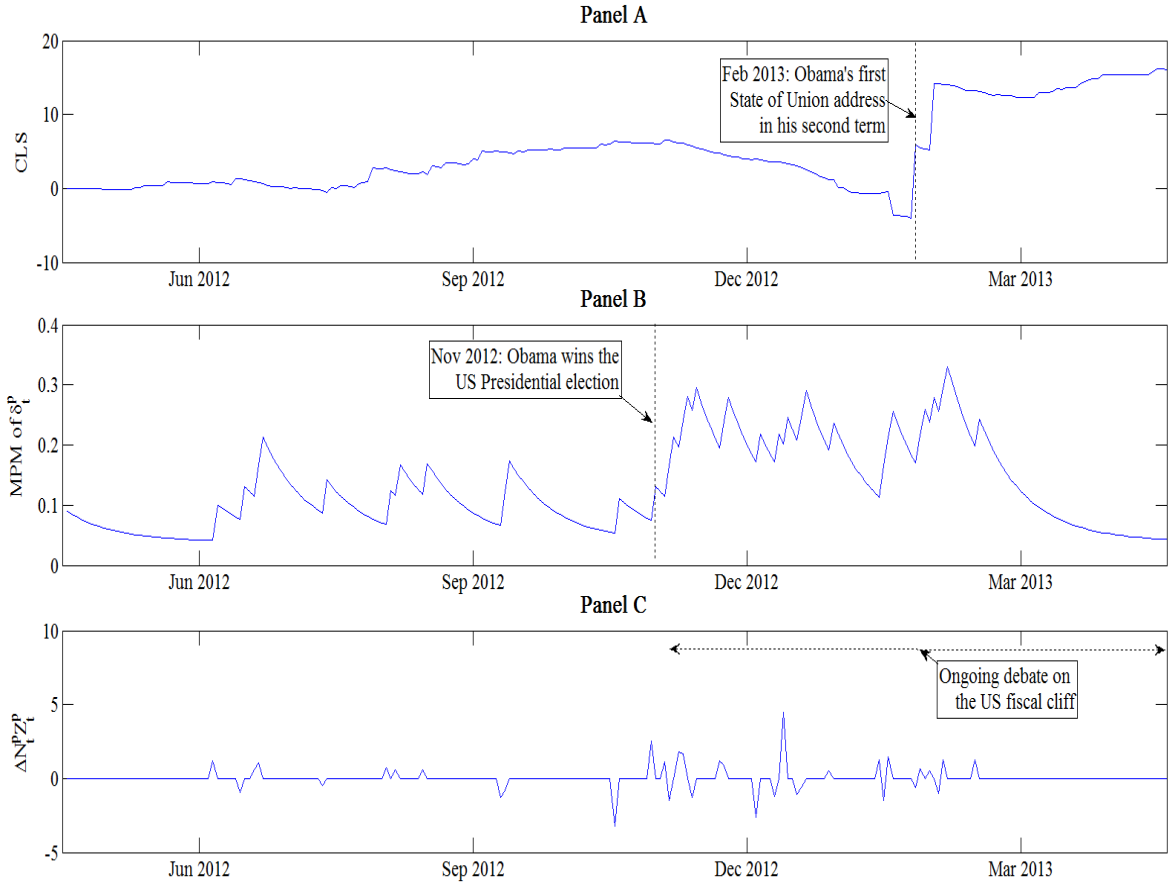

Figure 8: Cumulative log score for the most preferred model ($\mathcal{M}_3$), relative to model $\mathcal{M}_6$ (with static jumps) (Panel A); the marginal posterior mean of the predicted price jump intensity (Panel B); and observed price jumps (Panel C). The out-of-sample period is May 16, 2012 to May 3, 2013 inclusive.

In Figure 8, the CLS of the predictive return distribution over the out-of-sample period is depicted (Panel A), alongside the marginal posterior predictive mean of the δ_t^p (Panel B) and the observed price jump (Panel C). As can be observed, the CLS at the final time point is approximately 16, providing evidence that the dynamic model $\mathcal{M}_3$ outperforms the model with constant jump intensities, $\mathcal{M}_6$, with the greatest gain in predictive performance occurring around February 2013. As evident from the plot of observed price jumps in Panel

C, this sub-period has the greatest concentration of price jumps during the full evaluation period, and overlaps with President Obama's first State of Union address after his re-election.

Table 4: Unconditional coverage and independence tests; p-values for the one-step-ahead VaR predictions between May 16, 2012 and May 3, 2013, inclusive.

		5% VaR	1% VaR
$\mathcal{M}_3$	Unconditional coverage	0.74	0.78
	Independence	0.30	0.85
	Empirical coverage	4.55%	0.83%
$\mathcal{M}_6$	Unconditional coverage	0.41	0.78
	Independence	0.16	0.85
	Empirical Coverage	6.20%	0.83%
$\mathcal{M}_7$	Unconditional coverage	0.06	0.02
	Independence	0.07	0.52
	Empirical coverage	7.85%	2.89%

Secondly, we assess the performance of the models in predicting the one-step-ahead 5% and 1% VaR for the market portfolio associated with the S&P500 market index. Given the relative predictive accuracy of the model with dynamic jumps, plus the implications - as highlighted in Section 2.3 - of the presence of jumps per se for the tails of the return distribution, the ability of the three models, $\mathcal{M}_3$ (dynamic jumps), $\mathcal{M}_6$ (constant jumps), and $\mathcal{M}_7$ (no jumps), are assessed. In Table 4, we report the p-values of the Christoffersen (1998) tests of correct unconditional coverage and independence of exceedances of the (predicted) VaR. Models that produce forecasts that fail to reject both of these tests are deemed adequate in predicting VaR. Both models that include price and volatility jumps, $\mathcal{M}_3$ and $\mathcal{M}_6$, provide satisfactory VaR predictions at both the 5% and 1% levels, with the relevant null hypotheses not rejected in all cases. The results also indicate that both models capture the very extreme movements in returns equally well (with equivalent empirical coverage at the 1% level). However, the model with dynamic jump structures, $\mathcal{M}_3$, does better in capturing the returns movements induced by jumps (in either price or volatility) that are (relatively) smaller in magnitude, with a coverage at the 5% level that is closer to the nominal value than is the empirical coverage for $\mathcal{M}_6$. The dynamic model $\mathcal{M}_3$ could also be deemed superior in the sense that a slight *overestimation* of risk (at the 5% level) implied by an empirical coverage of 4.55% is preferable to an underestimation of risk implied by the empirical coverage of 6.20% for the restricted model $\mathcal{M}_6$. Model $\mathcal{M}_7$, on the other hand, performs relatively poorly in predicting extreme movements in returns, with the null hypotheses of the unconditional

coverage test (for both the 5% and 1% VaR) and the independence test (for the 1% VaR) being rejected at the 10% significance level. The empirical coverages of 7.85% for the 5% VaR and 2.89% for the 1% VaR also indicate that $\mathcal{M}_7$ produces a severe underestimation of risk. These results further highlight the importance of price and volatility jumps in this empirical setting and, moreover, confirm the added value of employing dynamic structures for both types of jumps.

6 Conclusions

In this paper, a stochastic volatility model incorporating dynamic behaviour in price and variance (and, hence, volatility) jumps is proposed. The model allows price and variance jumps to cluster over time, and for the occurrence of a price jump to impact on the likelihood of a subsequent variance jump, via a bivariate Hawkes process. The implications of the jump structures for the tail behaviour of the conditional return distributions are investigated, highlighting that the contribution of jumps to non-Gaussian tail behaviour is greatest in low volatility periods. A nonlinear state space model based on both daily returns and non-parametric measures of volatility and price jumps for the S&P500 market index is constructed, with a hybrid Gibbs-MH MCMC algorithm used to analyze the model. Perhaps not surprisingly, our investigation suggests that the price jump intensity possesses qualitatively different time series behaviour from that of the variance jump intensity. Clusters of inflated price jump intensities are relatively short-lived and scattered throughout the sample period, whilst clusters of high volatility 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 in evidence for the price jump intensity.

Several different hypotheses that impose various restrictions on the price and variance jump processes are explored using Bayes factors, with the overall conclusion being in favour of the hypothesis that specifies dynamics in both price and variance jump intensity, but that does not allow for feedback from past price jumps to the variance jump intensity. The dynamic structures imposed on the occurrences of price and variance jumps are also shown to add value to the predictions of the return distributions in an out-of-sample context. These findings have implications for the pricing of assets, including both the S&P500 index itself and any derivatives written on the index, such as S&P500 index futures and options, and potentially any derivatives written on the VIX (volatility) index.

Acknowledgements. The authors would like to thank Yacine Aït-Sahalia, John Maheu and Herman van Dijk for very constructive comments on an early draft of this paper, plus the participants at the Peter C.B. Phillips PhD Summer Camp, Singapore Management University, 2012; the New Zealand Econometrics Study Group, 2013; and the 7th Rimini Bayesian Econometrics Workshop, 2013. The research has been supported by Australian Research Council Future Fellowship FT0991045.

Appendix A: Adjustment terms for the central moments of $r_t|V_{t-1}, \Delta J_t^p$. Due to the positivity restriction on the (discretized) volatility process, the implied higher-order central moments of $r_t|V_{t-1}, \Delta J_t^p$ require adjustment terms that result from the truncation of the distribution $V_t|V_{t-1}, \Delta J_t^v, r_{t-1}, \Delta J_{t-1}^p$ at zero. This appendix documents the adjustment terms that are abbreviated in Panel B of Table 1 in Section 2.3. In the following we define

$$V_t^* = \frac{-\left(\overline{V}_t^d\right)}{\sigma_v \sqrt{\left(1-\rho^2\right) V_{t-1}}}.$$

The adjustment term for $E\left(r_t|V_{t-1}, \Delta J_t^p\right)$ is

$$g_1\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right) = \gamma\sigma_v\sqrt{\left(1-\rho^2\right) V_{t-1}}\phi\left(V_t^*\right)\left(1-\Phi\left(V_t^*\right)\right)^{-1}.$$

The adjustment term for $Var\left(r_t|V_{t-1}, \Delta J_t^p\right)$ is

$$\begin{aligned} g_2\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right) &= \left(\frac{1-\overline{V}_t^d}{\gamma}\right) g_1\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right) \\ &\quad - \left[g_1\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right)\right]^2. \end{aligned}$$

The adjustment term for $E\left(\left(r_t - E\left(r_t|V_{t-1}, \Delta J_t^p\right)\right)^3 |V_{t-1}, \Delta J_t^p\right)$ is

$$\begin{aligned} g_3\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right) &= 2\left[g_1\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right)\right]^3 + \left(\gamma\overline{V}_t^d - \frac{3}{\gamma}\right) g_1\left[\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right)\right]^2 \\ &\quad + 2\gamma\sigma_v^2\left(1-\rho^2\right) V_{t-1}\gamma^2\overline{V}_t^d\phi\left(V_t^*\right)\left(1-\Phi\left(V_t^*\right)\right)^{-2} \\ &\quad + g_1\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right)\left(\gamma^2\left(\left(\overline{V}_t^d\right)^2 + \sigma_v^2\left(1-\rho^2\right) V_{t-1}\right) - 3\overline{V}_t^d\right). \end{aligned}$$

The adjustment term for $E\left(\left(r_t - E\left(r_t|V_{t-1}, \Delta J_t^p\right)\right)^4 |V_{t-1}, \Delta J_t^p\right)$ is

$$g_4\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right) = -6\left[g_1\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right)\right]^4 - \frac{4\gamma\overline{V}_t^d}{\phi\left(V_t^*\right)}\left[g_1\left(\phi\left(V_t^*\right), \Phi\left(V_t^*\right)\right)\right]^3$$

$$
\begin{aligned}
&+\left(\frac{12}{\gamma}+4\gamma\overline{V}_t^d\right)\left[g_1\left(\phi\left(V_t^*\right),\Phi\left(V_t^*\right)\right)\right]^3\\
&+\left(3\gamma^2\left(\overline{V}_t^d\right)^2-\frac{3}{\gamma^2}+2\gamma^2\sigma_v^2\left(1-\rho^2\right)V_{t-1}+\frac{6\overline{V}_t^d+2\gamma^2\sigma_v^2\left(1-\rho^2\right)V_{t-1}}{\phi\left(V_t^*\right)}\right)\\
&\times\left[g_1\left(\phi\left(V_t^*\right),\Phi\left(V_t^*\right)\right)\right]^2\\
&+\left(\gamma^3\left(\overline{V}_t^d\right)^3-3\gamma\left(\overline{V}_t^d\right)^2-\gamma^3\overline{V}_t^d\sigma_v^2\left(1-\rho^2\right)V_{t-1}-\left(V_t^*\right)\right)\\
&\times g_1\left(\phi\left(V_t^*\right),\Phi\left(V_t^*\right)\right)\phi\left(V_t^*\right)\\
&+\left(3\gamma\left(\overline{V}_t^d\right)^2-\frac{3\overline{V}_t^d}{\gamma}-\gamma\left(\sigma_v^2\left(1-\rho^2\right)V_{t-1}\right)^2+\left(2\gamma^3\overline{V}_t^d-5\gamma\right)\sigma_v^2\left(1-\rho^2\right)V_{t-1}\right)\\
&\times g_1\left(\phi\left(V_t^*\right),\Phi\left(V_t^*\right)\right).
\end{aligned}
$$

Note that when variance jumps are absent, $\delta_{t-1}^v=0$, and

$$
Var\left(r_t|V_{t-1},\Delta J_t^p\right)=\overline{V}_t^d+\gamma^2\sigma_v^2V_{t-1}\left(1-\rho^2\right)+g_2(\cdot,\cdot)
$$

and

$$
E\left(\left(r_t-E\left(r_t|V_{t-1},\Delta J_t^p\right)\right)^4|V_{t-1},\Delta J_t^p\right)=3\left(Var\left(r_t|V_{t-1},\Delta J_t^p\right)\right)^2+3\sigma_v^2V_{t-1}\left(1-\rho^2\right)+g_4(\cdot,\cdot).
$$

Hence, it follows that in this case,

$$
\begin{aligned}
ku\left(r_t|V_{t-1},\Delta J_t^p,\delta_{t-1}^v\right) &= \frac{E\left(\left(r_t-E\left(r_t|V_{t-1},\Delta J_t^p,\delta_{t-1}^v\right)\right)^4|V_{t-1},\Delta J_t^p,\delta_{t-1}^v\right)}{\left[Var\left(r_t|V_{t-1},\Delta J_t^p,\delta_{t-1}^v\right)\right]^2}\\
&= 3+\frac{3\sigma_v^2V_{t-1}\left(1-\rho^2\right)+g_4\left(.,.\right)}{\left(\overline{V}_t^d+\gamma^2\sigma_v^2V_{t-1}\left(1-\rho^2\right)+g_2\left(.,.\right)\right)^2}>3.
\end{aligned}
$$

Clearly, if $\sigma_v^2=0$, standardized kurtosis collapses to a value of 3.

Appendix B: MCMC algorithm. The MCMC algorithm for sampling from the joint posterior in (35) can be broken down into four main steps:

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

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

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

$$\begin{aligned}
&p\left(\Delta N^v_{1:T}|V_{1:T}, Z^v_{1:T}, r_{1:T}, \ln BV_{1:T}, \Delta N^p_{1:T}; \phi\right) \\
&\propto p\left(Z^v_1|\Delta N^v_1, \mu_v\right) \times p\left(\Delta N^v_1|\delta^v_1\right) \times \left[\prod_{t=2}^{T} p\left(\Delta N^p_t|\Delta N^p_{1:t-1}; \alpha_p, \beta_{pp}, \delta^p_0\right)\right. \\
&\times p\left(V_t|V_{t-1}, Z^v_{t-1}, \Delta N^v_{t-1}, r_{t-1}, \Delta N^p_{t-1}; \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right) \\
&\left.\times\ p\left(Z^v_t|\Delta N^v_t, \mu_v\right) \times p\left(\Delta N^v_t|\Delta N^v_{1:t-1}, \Delta N^p_{1:t-1}, Z^p_{1:t-1}; \alpha_v, \beta_{vv}, \beta_{vp}, \beta^{(-)}_{vp}, \delta^v_0\right)\right],
\end{aligned}$$

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

$$\begin{aligned}
&p\left(Z^v_{1:T}|V_{1:T}, \Delta N^v_{1:T}, r_{1:T}, BV_{1:T}, \Delta N^p_{1:T}; \phi\right) \\
&\propto p\left(Z^v_1|\Delta N^v_1, \mu_v\right) \left[\prod_{t=2}^{T} p\left(Z^v_t|\Delta N^v_t, \mu_v\right)\right. \\
&\left.\times\ p\left(V_t|V_{t-1}, Z^v_{t-1}, \Delta N^v_{t-1}, r_{t-1}, \Delta N^p_{t-1}; \kappa, \theta, \sigma_v, \mu, \gamma, \rho\right)\right],
\end{aligned}$$

4. Generating elements of ϕ from

$$\begin{aligned}
&p\left(\phi|V_{1:T}, Z^v_{1:T}, \Delta N^v_{1:T}, r_{1:T}, \ln BV_{1:T}, \Delta N^p_{1:T}\right) \\
&\propto p\left(V_{1:T}, Z^v_{1:T}, \Delta N^v_{1:T}, \phi|r_{1:T}, \ln BV_{1:T}, \Delta N^p_{1:T}\right).
\end{aligned}$$

The most challenging part of the algorithm is the generation of the state vector $V_{1:T}$, due to the nonlinear functions of V_t that feature in the measurement equations (26) and (29), and in the state equation (30). 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 *et al.* (2003). In the current context that involves augmenting the state space model with mixture indicator vectors corresponding to the latent variance vector $V_{1:T}$ and the bivariate observation vectors $r_{1:T}$ and $\ln BV_{1:T}$. Conditionally, the mixture indicators define suitable linearizations of the relevant state or observation equation and are used to establish a linear Gaussian candidate model for use within an MH subchain. Candidate vectors of $V_{1:T}$ are sampled and evaluated in blocks. With due consideration taken of the different model structure and data types, Appendix A of Maneesoonthorn *et al.* provides sufficient information for the details of this component of the algorithm applied herein to be extracted.

The elements of ϕ are sampled using MH subchains wherever necessary. Given observations on $\Delta N^p_{1:T}$ and $Z^p_{1:T}$, plus draws of $V_{1:T}$ and of all other unknowns that appear in (26), the parameters μ and γ can be treated as regression coefficients, with exact draws produced in the standard manner from a (truncated) Gaussian joint conditional posterior distribution,

as a consequence of the previously specified priors. The sampling of σ_{BV} is also standard, given the inverse gamma form of its conditional posterior. As described in Section 4.1, the parameters ρ and σ_v are sampled indirectly via the conditionals of $\psi = \rho\sigma_v$ and $\omega = \sigma_v^2 - \psi^2$, which take the form of normal and inverse gamma distributions, respectively. Conditional upon the draws of $V_{1:T}$, $\Delta N_{1:T}^v$ and $Z_{1:T}^v$, the parameters κ, θ, ψ and ω are drawn in blocks, taking advantage of the (conditionally) linear regression structure with truncated Gaussian errors, and with the constraint $\sigma_v^2 \leq 2\kappa\theta$ imposed.

The parameters associated with the price and volatility jump processes are dealt with as follows. The mean of the volatility 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 (33) and (34) define stationary processes, and that (16) and (17) are defined on the $[0, 1]$ interval. The intensity parameters δ_∞^p and δ_∞^v are then computed using the explicit relationships in (16) and (17), and the vectors $\delta_{1:T}^v$ and $\delta_{1:T}^p$ updated deterministically based on (33) and (34).

The algorithms for all sub-models described in Section 4.2, $\mathcal{M}_i$, for $i = 1, ..., 7$, proceed in an analogous way.

Appendix C: Marginal likelihood computation. The basic idea underlying the evaluation of (44) is the recognition that it can be re-expressed as

$$p\left(r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p | \mathcal{M}_i\right) = \frac{p\left(r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p | \phi_i, \mathcal{M}_i\right) p\left(\phi_i | \mathcal{M}_i\right)}{p\left(\phi_i | r_{1:T}, \ln BV_{1:T}, \Delta N_{1:T}^p, Z_{1:T}^p, \mathcal{M}_i\right)}, \quad (47)$$

for any point ϕ_i in the posterior support of model $\mathcal{M}_i$. The first component of the numerator on the right-hand-side of (47) is the likelihood, conditional on $\mathcal{M}_i$, marginal of the latent variables. That is,

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

The denominator on the right-hand-side of (47) is simply the conditional posterior density of the (static) parameter vector, also marginalized over the latent variables,

$$\begin{aligned}
&p\left(\phi_i|r_{1:T},\ln BV_{1:T},\Delta N^p_{1:T},Z^p_{1:T},\mathcal{M}_i\right)\\
&=\int p\left(\phi_i|r_{1:T},\ln BV_{1:T},\Delta N^p_{1:T},Z^p_{1:T},V^{(i)}_{1:T},Z^{v(i)}_{1:T},\Delta N^{v(i)}_{1:T},\mathcal{M}_i\right)\\
&p\left(V^{(i)}_{1:T}|Z^{v(i)}_{1:T},\Delta N^{v(i)}_{1:T};\phi_i,\mathcal{M}_i\right)p\left(Z^v_{1:T}|\Delta N^{v(i)}_{1:T};\phi_i,\mathcal{M}_i\right)\\
&p\left(\Delta N^{v(i)}_{1:T}|\Delta N^p_{1:T};\phi_i,\mathcal{M}_i\right)d\left(V^{(i)}_{1:T},Z^{v(i)}_{1:T},\Delta N^{v(i)}_{1:T}\right). \qquad (49)
\end{aligned}$$

The evaluation of (48) at a high density posterior point ϕ_i^* (say, the vector of marginal posterior means for the elements of ϕ_i) is a straightforward use of the output of a full MCMC run for model $\mathcal{M}_i$; namely, the product of the closed form representations of $p\left(r_{1:T}|V^{(i)}_{1:T},Z^p_{1:T},\Delta N^p_{1:T},Z^{v(i)}_{1:T},\Delta N^{v(i)}_{1:T};\phi_i,\mathcal{M}_i\right)$, $p\left(\ln BV_{1:T}|V^{(i)}_{1:T};\phi_i,\mathcal{M}_i\right)$ and $p\left(\Delta N^p_{1:T}|\phi_i,\mathcal{M}_i\right)$ is averaged over the draws of the latent states, $V^{(i)}_{1:T}$, $Z^{v(i)}_{1:T}$ and $\Delta N^{v(i)}_{1:T}$, and computed at the given point ϕ_i^*. Evaluation of (49) is more difficult, in particular when a combination of Gibbs and MH algorithms needs to be employed in the production of draws of ϕ_i. Defining $y=(r_{1:T},\ln BV_{1:T},\Delta N^p_{1:T},Z^p_{1:T})'$, and exploiting the structure of posterior density, we decompose $p\left(\phi_i^*|y,\mathcal{M}_i\right)$ into six constituent densities as:

$$\begin{aligned}
p\left(\phi_i^*|y,\mathcal{M}_i\right) &= p\left(\phi_{1i}^*|y,\mathcal{M}_i\right)p\left(\phi_{2i}^*|\phi_{1i}^*,y,\mathcal{M}_i\right)p\left(\phi_{3i}^*|\phi_{1i}^*,\phi_{2i}^*,y,\mathcal{M}_i\right)\\
&\quad ...p\left(\phi_{6i}^*|\phi_{1i}^*,\phi_{2i}^*,...,\phi_{5i}^*,y,\mathcal{M}_i\right), \qquad (50)
\end{aligned}$$

where $\phi_{1i}=\left(\sigma_{BV},\mu_v,\delta_0^p,\delta_0^v,\rho,\sigma_v,\mu_p\right)$, $\phi_{2i}=\left(\alpha_p,\alpha_v,\kappa,\gamma_p\right)$, $\phi_{3i}=\left(\beta_{pp},\beta_{vv},\theta,\sigma_p\right)$, $\phi_{4i}=\left(\beta_{vp},\mu\right)$, $\phi_{5i}=\left(\beta_{vp}^{(-)},\gamma\right)$ and $\phi_{6i}=\sigma_r$. Following the methods outlined by Chib (1995) and Chib and Jeliazkov (2001), five additional MCMC chains, each of which involves a different level of conditioning and, hence, a reduced number of free parameters, are then run to estimate each of the last five components of (50), in turn evaluated at ϕ_{ji}^*. $j=2,...,6$. The first component on the right hand side of (50), involving no such conditioning, is estimated from the output of the full MCMC chain, in the usual way. More details of the relevant reduced algorithms are available from the authors on request.

References

[1] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R.J.A. (2013), "Modeling Financial Contagion Using Mutually Exciting Jump Processes," *Working Paper*. Available at www.princeton.edu/~yacine/research.htm.

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

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

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

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

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

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

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

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

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

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

[12] 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.

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

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

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

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

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

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

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

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

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

[22] Fulop, A., Li, J. and Yu, J. (2012), "Bayesian Learning of Impacts of Self-Exciting Jumps in Returns and Volatility," *Working Paper*, *Singapore Management University School of Economics*. Available at http://ink.library.smu.edu.sg/soe_research/1325.

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

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

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

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

[27] ——— (1971b), "Point Spectra of Some Mutually Exciting Point Processes," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 33, 438-443.

[28] 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.

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

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

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

[32] 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.

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

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

[35] 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.

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

[37] 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.

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

[39] 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.

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

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