



MONASH University

Australia

Department of Econometrics and Business Statistics

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

**Probabilistic Forecasts of Volatility
and its Risk Premia**

**Worapree Maneesoonthorn, Gael M. Martin,
Catherine S. Forbes and Simone Grose**

December 2010

Working Paper 22/10

Probabilistic Forecasts of Volatility and its Risk Premia

Worapree Maneesoonthorn, Gael M. Martin*,
Catherine S. Forbes and Simone Grose

Department of Econometrics and Business Statistics,
Monash University

December 20, 2010

Abstract

The object of this paper is to produce distributional forecasts of physical volatility and its associated risk premia using a non-Gaussian, non-linear state space approach. Option and spot market information on the unobserved variance process is captured by using dual ‘model-free’ variance measures to define a bivariate observation equation in the state space model. The premium for diffusive variance risk is defined as linear in the latent variance (in the usual fashion) whilst the premium for jump variance risk is specified as a conditionally deterministic dynamic process, driven by a function of past measurements. The inferential approach adopted is Bayesian, implemented via a Markov chain Monte Carlo algorithm that caters for the multiple sources of non-linearity in the model and the bivariate measure. The method is applied to empirical spot and option price data for the S&P500 index over the 1999 to 2008 period, with conclusions drawn about investors’ required compensation for variance risk during the recent financial turmoil. The accuracy of the probabilistic forecasts of the observable variance measures is demonstrated, and compared with that of forecasts yielded by more standard time series models. To illustrate the benefits of the approach, the posterior distribution is augmented by information on daily returns to produce Value at Risk predictions, as well as being used to yield forecasts of the prices of derivatives on volatility itself. Linking the variance risk premia to the risk aversion parameter in a representative agent model, probabilistic forecasts of relative risk aversion are also produced.

KEYWORDS: Volatility Forecasting; Non-linear State Space Models; Non-parametric Variance Measures; Bayesian Markov Chain Monte Carlo; VIX Futures; Risk Aversion.

JEL CODES: C11, C53, C58.

*Corresponding author. Email: gael.martin@monash.edu. This research has been supported by Australian Research Council (ARC) Discovery Grant DP0985234 and ARC Future Fellowship FT0991045.

1 Introduction

Volatility estimates play a central role in financial applications, with accurate forecasts of future volatility being critical for asset pricing, portfolio management and Value at Risk (VaR) calculations. Along with the information on volatility embedded in historical returns on a financial asset, the prices of options written on the asset also shed light on the option market's assessment of the volatility that is expected to prevail over the remaining life of the options. As such, many forecasting exercises have used both sources of market data to extract information on future volatility, with the relative accuracy of the options- and returns-based forecasts being gauged via a variety of means (e.g. Blair, Poon and Taylor, 2001, Martens and Zein, 2004, Pong, Shackleton, Taylor and Xu, 2004, Jiang and Tian, 2005, Koopman, Jungbacker and Hol, 2005, and Martin, Reidy and Wright, 2009).

Crucially, as option pricing occurs under the risk-neutralized measure for the underlying asset price process, any systematic disparity between returns- and option-implied volatility forecasts can be viewed as evidence of the option market having factored in non-zero prices for various risk factors, including variance risk. A recent literature has evolved in which this disparity has been used - in one way or another - to extract information on variance risk premia (e.g. Guo, 1998, Chernov and Ghysels, 2000, Pan, 2002, Jones, 2003, Eraker, 2004, Forbes, Martin and Wright, 2007, Eraker, 2008, Bollerslev, Tauchen and Zhou, 2009, Bollerslev, Sizova and Tauchen, 2009, Bollerslev and Todorov, 2009, Carr and Wu, 2009, Bollerslev, Gibson and Zhou, 2011, Duan and Yeh, 2010). However, in none of this work has the primary focus been the extraction of the risk premia for the purpose of improving the accuracy with which objective volatility can be forecast from the dual data source.

The primary aim of this paper is to combine option and spot price information with a view to producing accurate forecasts of the objective volatility process of the underlying. A non-Gaussian, non-linear state space framework is used to model volatility and its associated (time-varying) risk premia as latent state variables. Rather than link market price information to the state variables via complex theoretical option price formulae, we use direct non-parametric measures of volatility (see Britten-Jones and Neuberger, 2000, Barndorff-Nielsen and Shephard, 2002, Andersen, Bollerslev, Diebold and Labys, 2003, Jiang and Tian, 2005) to define a bivariate observation equation.

A secondary aim is to forecast the variance risk premia factored into the options-based measure. Motivated by empirical evidence, the state space representation is based on a model in which random jumps can occur contemporaneously in the asset price and the variance. By making appropriate adjustments to the observable volatility measures, variation in the price jumps is modelled non-parametrically. An explicit parametric model is then adopted for the

latent variance, including its two associated risk premia: one that compensates for small and regular movements in the variance (the diffusive variance risk premium), and the other compensating for rare jumps (the jump variance risk premium). The diffusive risk premium is parameterized in the conventional way, as proportional to the latent variance itself. The jump variance risk premium is also allowed to be time varying. Specifically, a conditionally deterministic process, driven in part by the past ‘observed’ risk premium, is used to capture the dynamic behaviour of this component of the model. This aspect of our approach is somewhat similar in spirit to the analysis of Todorov (2010), in which past realized jumps are allowed to affect the compensation for future jump risk demanded by investors.

Probabilistic forecasts of the latent variance, the variance risk premia and the observable variance measures, are produced using Bayesian methods. This focus on probabilistic forecasting, whilst inherent to the Bayesian inferential paradigm, is also consistent with more general developments in the recent forecasting literature, in which *distributional* forecasts *per se* are viewed as the primary object of interest. (See, for example, Corradi and Swanson, 2006, Gneiting, Balabdaoui and Raftery, 2007, Gneiting, 2008, Geweke and Amisano, 2010, McCabe, Martin and Harris, 2010). The Bayesian predictive distributions are produced via a Markov chain Monte Carlo (MCMC) algorithm that caters for the non-linearities in the model and that allows for multi-move sampling of the latent variances. The conditionally deterministic specification for the jump risk premium is computationally convenient, with the posterior distribution of the risk premium at any time point - including future time points - able to be estimated from the MCMC draws of the parameters to which the premium is functionally related.

The method is applied to empirical spot and option price data for the S&P500 index over the 1999 to 2008 period. Distributional forecasts are produced for all latent and observable quantities of interest, for an evaluation period including both the lead up to the recent global financial crisis and the peak of the crisis at the end of 2008. Most notably, the extraction of forecasts for the variance risk premia enables a picture to be constructed of the extent to which investors’ expectations of future risks - and, correspondingly, their demand for compensation - is affected by extreme daily movements in the market. The accuracy of the probabilistic forecasts of the (observable) measures of variance is compared with that of forecasts produced by standard time series models for these quantities, using predictive log scores.

Illustrations of how the predictions can be used in financial applications are provided. First, the model is augmented by observations on daily returns to produce probabilistic forecasts of returns themselves, from which VaR predictions are extracted. Secondly, we illustrate how the predictions of the latent variance may be used to estimate the prices of

futures written on volatility itself. Finally, coupled with a particular form of representative agent model, forecasts of the variance risk premia are transformed into probabilistic forecasts of the relative risk aversion of the representative investor.

The remainder of the paper is organized as follows. Section 2 describes the model assumed to underlie both the spot and option price data, including the variance risk premia that form part of that model. Section 3 outlines the state space approach that we use to analyze the model, including the dynamic risk premia that we embed within it. A description of the MCMC algorithm used to estimate the latent variables and static parameters and to produce the forecasts is provided in Section 4, with all details of the component of the algorithm that relates to the sampling of the latent variances provided in Appendix A. The results of an extensive empirical investigation of intraday spot and option price data for the S&P500 index from July 1999 to December 2008 are reported in Section 5. Some conclusions are given in Section 6.

2 Objective and Risk Neutral Distributions

The spot price P_t and objective stochastic variance V_t are assumed to evolve according to the following bivariate process,

$$d \ln P_t = \left(\mu - \frac{V_t}{2}\right)dt + \sqrt{V_t}dB_t^p + dJ_t^p \quad (1)$$

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

where $dJ_t^i = Z_t^i dN_t$, $i = \{p, v\}$, $Z_t^p|Z_t^v \sim N(\mu_p + \rho_J Z_t^v, \sigma_p^2)$, $Z_t^v \sim \text{Exp}(\mu_v)$, $P(dN_t = 1) = \delta_J dt$ and $P(dN_t = 0) = (1 - \delta_J) dt$. Under this specification, random jumps in the price and variance occur contemporaneously at rate δ_J , but with magnitudes determined respectively by a normal and an exponential distribution. The magnitudes of jumps in the price and variance processes are assumed to be correlated, governed by ρ_J . The two Brownian increments dB_t^p and dB_t^v are correlated with a coefficient ρ ; however dB_t^i and dJ_t^i are assumed to be independent, for $i = \{p, v\}$. This model is often referred to in the literature as the stochastic volatility with contemporaneous jumps (SVCJ) model (e.g. Duffie, Pan and Singleton, 2000, Eraker, Johannes and Polson, 2003, Eraker, 2004, Broadie, Chernov and Johannes, 2007).¹

Based on this particular dynamic model, equilibrium arguments can be used to produce

¹Bates (2000) and Pan (2002) have proposed extensions of this model in which the jump frequency (δ_J) depends on the level of variance.

the following risk-neutral distribution,

$$d \ln P_t = \left(r - \frac{V_t}{2}\right)dt + \sqrt{V_t}dB_t^{*p} + dJ_t^{*p} - \mu_p^* \delta_J dt \quad (3)$$

$$dV_t = \kappa^*[\theta^* - V_t]dt + \sigma_v \sqrt{V_t}dB_t^{*v} + dJ_t^{*v}, \quad (4)$$

under which options on the underlying asset are priced, where r denotes the risk-free interest rate (assumed constant), $dJ_t^{*i} = Z_t^{*i}dN_t^*$, $i = \{p, v\}$, $Z_t^{*p} \sim N(\mu_p^*, \sigma_p^2)$, $Z_t^{*v} \sim \text{Exp}(\mu_v^*)$ and we impose $\delta_J^* = \delta_J$. The term $-\mu_p^* \delta_J dt$ on the right hand side of (3) compensates the jump process.

Implicit in the move from (2) to (4) is the transformation

$$\kappa^*[\theta^* - V_t] = \kappa[\theta - V_t] - \lambda_D V_t,$$

where

$$\kappa^* = \kappa + \lambda_D; \quad \theta^* = \frac{\kappa\theta}{\kappa^*} \quad (5)$$

and λ_D is a scalar parameter. The term $\lambda_D V_t$ represents the premium associated with diffusive variance risk, with the value of λ_D determining the magnitude (and sign) of the premium factored into option prices for the risk associated with small and regular changes in the non-traded state variable, V_t . We interpret

$$\mu_p^* - \mu_p \text{ and} \quad (6)$$

$$\mu_v^* - \mu_v \quad (7)$$

as the premia for price and variance jump (size) risk respectively. As outlined below, our approach to estimating the model is such that $\mu_p^* - \mu_p$ in (6) is not identified, with our focus being solely on the identification and estimation of the variance risk premium parameters λ_D and $\mu_v^* - \mu_v$, with the latter denoted by λ_J hereafter. The specification in (7) for λ_J amounts to the assumption that investors factor into their option pricing a premium that equals the difference between the risk neutral expected mean jump *size* (for the variance) and the corresponding objective mean.²

Empirical estimates of λ_D and λ_J reported in the literature are (not surprisingly) model dependent. In pure stochastic volatility models (in which no jumps or jump premia are parameterized), estimates of λ_D are typically negative (see, for example, Guo, 1998, Forbes *et al.*, 2007, Bollerslev *et al.*, 2011), which implies slower reversion ($\kappa^* < \kappa$) to a higher

²Eraker (2008) considers a parameterization of the risk neutral distribution in which the *intensity* parameter δ_J differs between the objective and risk neutral processes and, hence, incorporates the premium. Duan and Yeh (2010) adopt a parameterization that allows the jump risk premium to reflect a premium for both jump *size* and jump *timing*.

mean level ($\theta^* > \theta$) under the risk neutral distribution. However, as outlined in some detail by Broadie *et al.* (2007), empirical conclusions regarding the significance and sign of λ_D are less clear cut once jumps (and associated premia) are included in the model specification. Overall, current empirical evidence points to λ_J being significantly greater than zero, with the significance of λ_D tending to be reduced accordingly. A significantly positive value for λ_J leads to the same *qualitative* result as a significantly negative value of λ_D . That is, either numerical outcome leads to a higher long run mean under the risk-neutral measure than under the objective measure, given that the two long-run means are given respectively by

$$E^*(V_t) = \frac{\kappa\theta + (\mu_v + \lambda_J)\delta_J}{\kappa + \lambda_D}$$

and

$$E(V_t) = \frac{\kappa\theta + \mu_v\delta_J}{\kappa}.$$

This implies, in turn, that (call) options are priced higher under the risk-neutral process, on average, than if they had been priced under the objective measure. That is, these signs for the risk premium parameters ($\lambda_D < 0$ and $\lambda_J > 0$ respectively) imply that investors are willing to pay a premium for options, as a hedge against movements in the spot price that result from either the diffusive or jump components of the random variance (or both).

As is clear from (5) and (7), observed option prices, assumed to be priced according to (3) and (4), can be used to identify the parameters of the *objective* process, and the risk premia, λ_D and λ_J , only if additional information on the objective parameters and/or λ_D and λ_J , is introduced. Previous analyses (based on a variety of versions of the model presented here) have solved this identification problem: by jointly estimating the objective and risk-neutral processes using option and spot price data (e.g. Chernov and Ghysels, 2000, Pan, 2002, Polson and Stroud, 2003, Eraker, 2004, Forbes *et al.*, 2007, Johannes, Polson and Stroud, 2009); by using option price data only to estimate (3) and (4), and extracting estimates of the risk premium parameters via separate return-based estimates of the objective parameters (e.g. Guo, 1998, Broadie *et al.*, 2007); or by imposing theoretical restrictions on the risk premia (Bates, 2000). Most importantly, in all of these studies, the link between observed market option prices and the underlying model in (3) and (4) occurs *indirectly*, via a parametric theoretical option price formula derived as the expected value of the discounted payoff of the option under the risk-neutral measure. In contrast, we link the observed option price data to the model in (3) and (4) *directly*, by using a non-parametric estimate of expected integrated variance over the life of the option, evaluated according to the risk-neutral process in (4). (See also Eraker, 2008, and Duan and Yeh, 2010, for recent applications of option-implied volatility measures in state space settings). Analogously, the

observed spot data is linked to the objective process in (1) and (2) by using high frequency returns to estimate the integrated variance associated with the objective process.

In the following section we outline the state space model based on the observed volatility measures, with both risk premium parameters, λ_D and λ_J , assumed to be constant. In Section 3.2 we extend the model to allow for a dynamic model for λ_J . In common parlance, the model we adopt for λ_J is *observation-driven*, with the value of λ_J at time t , λ_{Jt} , given by a deterministic function of λ_{Jt-1} and the ‘observed’ value of λ_{Jt-1} , denoted by l_{Jt-1} . In specifying l_{Jt-1} , we exploit recent theoretical developments in Carr and Wu (2009) (amongst others) to link the difference between the two observed measures of variance over the option maturity period to λ_{Jt} .

3 A State Space Model Based on Spot Price and Option-Implied Volatility Measures

3.1 Constant risk premium parameters

Given the objective variance process in (2), we define quadratic variation over the horizon $t - 1$ to t (call this day t) as

$$Q\mathcal{V}_{t-1,t} = \int_{t-1}^t V_s ds + \sum_{t-1 < s \leq t}^{N_t} (Z_s^p)^2.$$

That is, $Q\mathcal{V}_{t-1,t}$ is equal to the sum of the integrated variance of the continuous sample path component of P_t ,

$$\mathcal{V}_{t-1,t} = \int_{t-1}^t V_s ds,$$

and the sum of the $N_t - N_{t-1}$ squared jumps that occur on day t . Denoting $r_{t_i} = \ln P_{t_i} - \ln P_{t_{i-1}}$ as the i th transaction (logarithmic) return, it is now standard knowledge (Barndorff-Nielsen and Shephard, 2002, and Anderson *et al.*, 2003) that

$$RV_t = \sum_{t_i \in [t-1,t]}^M r_{t_i}^2 \xrightarrow{p} Q\mathcal{V}_{t-1,t}, \quad (8)$$

where RV_t is referred to as realized variance and M is equal to the number of intraday returns on day t . On the other hand, bipower variation,

$$BV_t = \frac{\pi}{2} \sum_{t_i \in [t-1,t]}^M |r_{t_i}| |r_{t_{i-1}}|, \quad (9)$$

is a consistent estimator of $\mathcal{V}_{t-1,t}$ in the presence of price jumps (Barndorff-Nielsen & Shephard, 2004).³

With a view to avoiding the need to explicitly model the price jump process - and its associated risk premium - in the state space framework, we adopt BV_t (rather than RV_t) as a spot-price-based measure of the latent variance. With the value of M used in the calculation of BV_t assuming a finite value in practice, we view BV_t as a noisy measure of the unobserved integrated variance by specifying

$$BV_t = \mathcal{V}_{t-1,t} + u_{BV_t}; \quad u_{BV_t} \sim N(0, \sigma_{BV}^2 V_t^2), \quad (10)$$

where the latent variance underlying $\mathcal{V}_{t-1,t}$ evolves according to (2). The adoption of a Gaussian measurement error (conditional on the volatility path) is motivated, in part, by computational convenience, whilst the use of a state-dependent measurement error variance is motivated by the clear need for such a specification in the case of the empirical data analyzed in Section 5. However, the form of measurement error in (10) also has some additional justification, by being broadly consistent with existing asymptotic theory regarding BV_t , in a more limited setting. Specifically, Barndorff-Nielsen and Shephard (2006) demonstrate that under the assumption of *no* jumps, as $M \rightarrow \infty$, BV_t converges to a normal distribution with a variance that is an increasing function of integrated quarticity, $\text{IQ}_{t-1,t} = \int_{t-1}^t V_s^2 ds$.

As well as having spot-price based observations on the latent variance, we have option-based measurements via the following logic. Bollerslev and Zhou (2002) and Garcia, Lewis, Pastorello and Renault (2011), amongst others, derive a set of conditional moments for the integrated variance of Heston's (1993) stochastic volatility model, which corresponds to (2) above, with dJ_t^v set to zero. Defining $\mathcal{F}_t = \sigma\{V_s; s \leq t\}$ as the sigma-algebra generated by the point-in-time variance process, and using $E(\cdot|\mathcal{F}_t)$ to denote a conditional expectation with respect to the objective measure, the conditional mean for integrated variance under the physical measure, over the period from t to $t + \tau$, in this case can be expressed as a linear function of the point-in-time variance,

$$E(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) = E\left(\int_t^{t+\tau} V_s ds \middle| \mathcal{F}_t\right) = a_\tau V_t + b_\tau, \quad (11)$$

where

$$a_\tau = \frac{1}{\kappa} (1 - e^{-\tau\kappa}) \quad \text{and} \quad b_\tau = \tau\theta - \frac{\theta}{\kappa} (1 - e^{-\tau\kappa}). \quad (12)$$

Extending these existing results to cater for the distribution in (4), and using $E^*(\cdot|\mathcal{F}_t)$ to denote a conditional expectation with respect to the risk neutral measure, the risk-neutral

³Implicit in the results in (8) and (9) is the assumption that microstructure noise effects are absent. The formal incorporation of microstructure noise in the assumed process for intraday returns has led to modifications of RV_t and BV_t that are consistent estimators of $Q\mathcal{V}_{t-1,t}$ and $\mathcal{V}_{t-1,t}$ respectively, in the presence of such noise; see Martin *et al.* (2009) for a recent summary.

expectation of integrated variance over the horizon t to $t + \tau$ is given by

$$E^*(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) = a_\tau^* V_t + b_\tau^* + E^* \left(\int_t^{t+\tau} Z_s^{*v} dN_s^* \middle| \mathcal{F}_t \right), \quad (13)$$

where

$$a_\tau^* = \frac{1}{\kappa^*} (1 - e^{-\tau\kappa^*}) \quad \text{and} \quad b_\tau^* = \tau\theta^* - \frac{\theta^*}{\kappa^*} (1 - e^{-\tau\kappa^*}),$$

and where θ^* and κ^* are defined in (5). With the assumed independence of dJ_t^{*v} over time, plus the contemporaneous independence between Z_s^{*v} and dN_t^* , then

$$\begin{aligned} E^* \left(\int_t^{t+\tau} Z_s^{*v} dN_s^* \middle| \mathcal{F}_t \right) &= E^* \left(\int_t^{t+\tau} Z_s^{*v} dN_s^* \right) \\ &= \int_t^{t+\tau} E^* [Z_s^{*v}] E^* [dN_t^*] \\ &= \int_t^{t+\tau} \mu_v^* \delta_J ds \\ &= \tau [\mu_v + \lambda_J] \delta_J, \end{aligned}$$

given $\mu_v^* = \mu_v + \lambda_J$ from (7).

As shown by Jiang and Tian (2005) and Carr and Wu (2009) (and as based on the earlier work of Britten-Jones and Neuberger, 2000), the risk-neutral expectation of quadratic variation is implied by a continuum (over strike K) of option prices with maturity $\tau > 0$, $C(t + \tau, K)$, as

$$\begin{aligned} E^*(Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) &= E^*(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) + E^* \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \right] \\ &= 2 \int_0^\infty \frac{C(t + \tau, K) - C(t, K)}{K^2} dK. \end{aligned} \quad (14)$$

Hence, $E^*(Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t)$ is referred to as ‘model free’ implied variance. Importantly, this measure eschews the dependence of the ubiquitous Black-Scholes option-implied variance on the empirically invalid assumption of geometric Brownian motion for the underlying asset price.

Given an estimate of $E^*(Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t)$ in (14), based on a finite set of observed option prices on day t and denoted by MF_t^{QV} , we define the following option-based measurement equation,

$$\begin{aligned} MF_t &= MF_t^{QV} - \hat{E}^* \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \right] \\ &= E^*(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) + u_{MF_t}; \quad u_{MF_t} \sim N(0, \sigma_{MF}^2 V_t^2), \end{aligned} \quad (15)$$

where u_{MF_t} captures both the error associated with the discretization and truncation of the integral in (14) and the measurement error in $\hat{E}^* \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \right]$. As with (10), a (conditionally) Gaussian measurement error is adopted for convenience, with a state-dependent measurement error variance motivated again by empirical considerations. We follow Bollerslev *et al.* (2011) in estimating $E^* \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \right]$ as some constant proportion (c) of the estimated objective expectation, $\hat{E} \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \right]$, measured as follows.⁴ Following Tauchen and Zhou (2011), we identify the presence of a significant price jump on any day t by the realization of the statistic,

$$Z_{Jt} = \frac{RJ_t}{\sqrt{(2.61 - 2) M^{-1} \max \left(1, \frac{TP_t}{BV_t^2} \right)}}, \quad (16)$$

where $RJ_t = \frac{RV_t - BV_t}{RV_t}$, and where tri-power quarticity, TP_t , serves as a consistent estimator of the integrated quarticity, $\mathcal{I}Q_{t-1,t}$. Realized price jump variation on day t is then extracted by means of

$$JV_t = I(Z_{Jt} > Z_\alpha) \times \max(0, RV_t - BV_t), \quad (17)$$

with a time series of values for JV_t produced accordingly.⁵ As is consistent with the bulk of the empirical literature and with the theoretical assumption of independent jumps over time, JV_t (as based on the empirical data analyzed in Section 5) exhibits little autocorrelation.⁶ Hence, a short-memory autoregressive model of order one is used to produce point predictions, conditional on data up to time t , of $JV_{t,t+1}$, $JV_{t,t+2}$, ..., $JV_{t,t+\tau}$, with $\hat{E} \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \right]$ given as the aggregate of these τ predictions. The (modified) option-implied measure used in (15) is then

$$MF_t = MF_t^{QV} - c \hat{E} \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \right]. \quad (18)$$

⁴See Bollerslev and Todorov (2009) for an alternative method of estimating the risk-neutral expectation of jump variation, using short-maturity out-of-the-money options.

⁵Tri-power quarticity is computed as $TP_t = M \mu_{4/3}^{-3} \sum_{t_i \in [t-1, t]}^M |r_{t_i}|^{4/3} |r_{t_{i-1}}|^{4/3} |r_{t_{i-2}}|^{4/3}$, where $\mu_{4/3} = 2^{2/3} \Gamma(\frac{7}{6}) \Gamma(\frac{1}{2})^{-1}$ and $TP_t \rightarrow \mathcal{I}Q_{t-1,t}$ as $M \rightarrow \infty$. Under the null hypothesis of no price jumps, Z_{Jt} is asymptotically $N(0, 1)$ as $M \rightarrow \infty$. Thus, in testing whether a price jump is present on a particular day, the α -level critical value from a standard normal distribution (Z_α) applies.

⁶This assumption of independent jumps has been questioned of late, given the tendency for jumps to ‘cluster’ during certain time periods. See, for example, Aït-Sahalia, Cacho-Diaz and Laeven (2010).

Using an Euler discretization of (2), the full model we estimate is

$$BV_t = \mathcal{V}_{t-1,t} + \sigma_{BV} V_t \xi_{1t} \quad (19)$$

$$\begin{aligned} MF_t &= E^*(\mathcal{V}_{t,t+\tau} | \mathcal{F}_t) + \sigma_{MF} V_t \xi_{2t} \\ &= a_\tau^* V_t + b_\tau^* + \tau [\mu_v + \lambda_J] \delta_J + \sigma_{MF} V_t \xi_{2t} \end{aligned} \quad (20)$$

$$V_t = \kappa \theta \Delta t + (1 - \kappa \Delta t) V_{t-\Delta t} + \sigma_v \sqrt{\Delta t} \sqrt{V_{t-\Delta t}} \xi_{3t} + Z_t^v \Delta N_t \quad (21)$$

with

$$\xi_t = (\xi_{1t}, \xi_{2t}, \xi_{3t})' \stackrel{iid}{\sim} N(0_3, I_3) \text{ for all } t = 1, 2, \dots, T. \quad (22)$$

$$Z_t^v \sim \text{Exp}(\mu_v) \quad (23)$$

$$\Delta N_t \sim \text{Bernoulli}(\delta_J \Delta t). \quad (24)$$

Setting $\Delta t = 1$ in (21), the state equation for V_t describes the evolution of the point-in-time (annualized) variance from one day to the next. It is this variance quantity at time t that enters the function $E^*(\mathcal{V}_{t,t+\tau} | \mathcal{F}_t)$, and contributes to the measurement error variance in (19) and (20). We also assume that the (conditional) mean of BV_t in (19) is equal to V_t , by adopting a rectangular approximation to $\mathcal{V}_{t-1,t} = \int_{t-1}^t V_s ds$. The parameter θ is estimated as an annualized quantity, matching the annualized magnitude of the point-in-time variance, V_t . The parameter κ is treated as a daily quantity, measuring the rate of mean reversion in the annualized V_t per day. In accordance with this treatment of κ , $\tau = 22$ days and MF_t is modelled as an aggregated annualized variance over the trading month.

Finally, we conclude this section by acknowledging that the latent process adopted here for V_t is short-memory, as is typical in this literature. As is now a stylized fact, time series of *measured* volatility (both spot-price based and option-implied) exhibit, in contrast, long-memory characteristics; see Andersen *et al.* (2003) for an early illustration. Given that the primary focus of this paper is on producing accurate short-term (specifically, one-day-ahead) forecasts, we do not view this as a problem. Indeed, anticipating the empirical results reported in Section 5, the model-based one-day-ahead forecasts of the observable BV_t and MF_t out-perform reduced-form models fitted directly to the observable quantities that explicitly cater for the long-memory features in the latter.

3.2 A dynamic model for the jump risk premium

Carr and Wu (2009) propose a method of quantifying the variance risk premium using variance swaps. A variance swap is an over-the-counter contract with a payoff equal to the difference between quadratic variation, defined over the life of the swap contract, and the so-called variance swap rate, which is determined at time t .

Setting $t + \tau$ as the period at which the contract expires and denoting the price of the variance swap as p_t and its payoff as $x_{t+\tau}$, the no arbitrage conditions that underlie standard asset pricing theory imply that $p_t = kE^*(x_{t+\tau}|\mathcal{F}_t)$, where k is a constant (risk-neutral) discount factor. Defining the variance swap rate as $SW_{t,t+\tau}$ we have $x_{t+\tau} = Q\mathcal{V}_{t,t+\tau} - SW_{t,t+\tau}$. Given that the variance swap has zero market value at time t , it follows that $p_t = 0$ and

$$SW_{t,t+\tau} = E^*(Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) \quad (25)$$

as a consequence. As is consistent with the result in (14), Carr and Wu (2009) show that $SW_{t,t+\tau}$ can indeed be synthesized by a linear combination of τ -maturity option prices observed on day t .

In addition to the equality in (25), asset pricing theory allows the zero price of the variance swap to be linked to its payoff via the objective measure as

$$0 = E(m_{t,t+\tau}x_{t+\tau}|\mathcal{F}_t), \quad (26)$$

where $m_{t,t+\tau} = M_{t,t+\tau}/E(M_{t,t+\tau}|\mathcal{F}_t)$ is the normalized stochastic discount factor (or pricing kernel), with $E(M_{t,t+\tau}|\mathcal{F}_t) = e^{-r\tau}$ under the assumption of a constant risk-free interest rate. Given that $SW_{t,t+\tau}$ is known at time t , and using $E(m_{t,t+\tau}|\mathcal{F}_t) = 1$, (26) can be re-written as

$$\begin{aligned} SW_{t,t+\tau} &= E(m_{t,t+\tau}Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) \\ &= E(m_{t,t+\tau}|\mathcal{F}_t)E(Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) + cov(m_{t,t+\tau}, Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) \\ &= E(Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) + cov(m_{t,t+\tau}, Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t). \end{aligned}$$

Dividing through by $SW_{t,t+\tau}$, we produce an expression for the expected excess return on the variance swap investment

$$E\left(\frac{Q\mathcal{V}_{t,t+\tau}}{SW_{t,t+\tau}}|\mathcal{F}_t\right) - 1 = -cov\left(m_{t,t+\tau}, \frac{Q\mathcal{V}_{t,t+\tau}}{SW_{t,t+\tau}}|\mathcal{F}_t\right).$$

Alternatively, we can define the premium in variance payoff units as

$$E(Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) - SW_{t,t+\tau} = -cov(m_{t,t+\tau}, Q\mathcal{V}_{t,t+\tau}|\mathcal{F}_t),$$

and estimate the premium via an average of $Q\mathcal{V}_{t,t+\tau} - SW_{t,t+\tau}$.

Importantly, the above analysis highlights the fact that under the model outlined in Section 3.1, and with a focus on integrated variance only, the *conditional* variance risk premium defined over the maturity period τ is given by the following linear function of the

point in time latent variance V_t ,

$$\begin{aligned} E(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) - \left(SW_{t,t+\tau} - E^* \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \middle| \mathcal{F}_t \right] \right) &= E(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) - E^*(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) \\ &= a_\tau V_t + b_\tau - (a_\tau^* V_t + b_\tau^* + \tau \lambda_J \delta_J), \end{aligned} \quad (27)$$

with the terms on the R.H.S. of (27) following from (11) and (13).⁷ The conditional risk premium is also seen to be a linear function of λ_J , the premium for jump variance risk, and a non-linear function of λ_D , the parameter influencing the risk premium for diffusive risk, $\lambda_D V_t$. Bollerslev *et al.* (2011) - adopting a model that does not include variance jumps - propose a dynamic model for λ_D driven by observable low-frequency macro-finance variables. With the link made, via a particular equilibrium model, to the risk aversion parameter of a representative investor (referenced also in Section 5), they produce (indirectly) a model for time-varying risk aversion. Crucially, however, the analysis of Bollerslev *et al.* is based on monthly observations on option-based and realized variance measures, with conclusions drawn about the time variation of λ_D also linked to the monthly frequency. In addition, their model does not, by construction, include any premium for variance jump risk, so that the dynamic behaviour discerned in λ_D may well be confounded by variance jumps.

In contrast, our focus is on high-frequency (daily) variance data, with λ_J an explicit component of the specification. Based on the assumption that significant *daily* variation in λ_J , most notably in response to recent *observed* jumps in volatility, is more plausible, from a behavioural point of view, than daily variation in λ_D , we hypothesize that λ_J follows

⁷Note that the *unconditional* risk premium is given by

$$\begin{aligned} &E \left[E(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) - \left(SW_{t,t+\tau} - E^* \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \middle| \mathcal{F}_t \right] \right) \right] \\ &= E [E(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) - E^*(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t)] \\ &= E [(a_\tau V_t + b_\tau + \tau \mu_v \delta \tau) - (a_\tau^* V_t + b_\tau^* + \tau (\mu_v + \lambda_J) \delta_J)] \\ &= (a_\tau - a_\tau^*) E[V_t] + (b_\tau - b_\tau^*) - \tau \delta_J \lambda_J \\ &= (a_\tau - a_\tau^*) \left(\theta + \frac{\delta_J \mu_v}{\kappa} \right) + (b_\tau - b_\tau^*) - \tau \delta_J \lambda_J \\ &= (a_\tau - a_\tau^*) \theta + (b_\tau - b_\tau^*) + (a_\tau - a_\tau^*) \left(\frac{\delta_J \mu_v}{\kappa} \right) - \tau \delta_J \lambda_J \\ &= \underbrace{\frac{\theta \lambda_D}{\kappa^*} \left[\tau - \frac{1}{\kappa^*} (1 - e^{-\tau \kappa^*}) \right]}_{\text{Negative if } \lambda_D < 0} - \underbrace{\frac{\delta_J}{\kappa} [\tau \kappa \lambda_J + (a_\tau^* - a_\tau) \mu_v]}_{\text{Positive if } \lambda_J > 0 \text{ and } \lambda_D < 0 (\Rightarrow a_\tau^* > a_\tau)} \quad . \end{aligned}$$

Hence, the unconditional mean of the aggregate risk premium (over τ) defined in this way is negative if the risk premia have the anticipated signs. This result corresponds correctly to the spot price-based measure,

$\mathcal{V}_{t,t+\tau}$, being less, on average, than the option-price-based measure, $SW_{t,t+\tau} - E^* \left[\sum_{t < s \leq t+\tau}^{N_{t+\tau}} (Z_s^p)^2 \middle| \mathcal{F}_t \right]$.

a dynamic process, whilst λ_D is held constant. (See also Todorov, 2010). Note, however, that even with a constant value for λ_D , the diffusive risk premium over dt , $\lambda_D V_t dt$, is still a dynamic process via the assumed linear relationship with V_t . On the assumption that short memory dynamics drive λ_J , we specify a conditionally deterministic specification that mimics a generalized autoregressive heteroscedastic structure for volatility (Bollerslev, 1986), namely,

$$\lambda_{Jt} = \lambda_{J0} (1 - \alpha_1 - \alpha_2) + \alpha_1 \lambda_{Jt-1} + \alpha_2 l_{Jt-1}, \quad (28)$$

where l_{Jt-1} denotes the ‘observed’ value of λ_{Jt} at time $t - 1$. It is this value of λ_{Jt} that feeds into the risk-neutral expected integrated variance over the time to maturity, $E^*(\mathcal{V}_{t,t+\tau}|F_t)$, in (20).⁸

Motivated by (27), we set

$$\widehat{E}(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) - MF_t = a_\tau V_t + b_\tau - [a_\tau^* V_t + b_\tau^* + \tau \lambda_{Jt} \delta_J]$$

and solve for the observed value of λ_{Jt} as

$$l_{Jt} = \frac{[(a_\tau V_t + b_\tau) - (a_\tau^* V_t + b_\tau^*)] - [\widehat{E}(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) - MF_t]}{\tau \delta_J}, \quad (29)$$

at each point t , within the estimation algorithm. As a model-based estimate of the objective conditional expectation, $E(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t)$, needed to evaluate the right-hand-side of (29) at each t , we use the following linear function of BV_t ,

$$\widehat{E}(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t) = \frac{a_\tau}{a_1} BV_t + \left((b_\tau + \tau \mu_v \delta_J) - \frac{a_\tau (b_1 + \mu_v \delta_J)}{a_1} \right), \quad (30)$$

where a_1 and b_1 are simply those functions a_τ and b_τ in (12), evaluated at $\tau = 1$, respectively. The estimate in (30) is unbiased for $E(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t)$, as

$$\begin{aligned} & E\left(\frac{a_\tau}{a_1} BV_t + \left((b_\tau + \tau \mu_v \delta_J) - \frac{a_\tau (b_1 + \mu_v \delta_J)}{a_1} \right) \middle| \mathcal{F}_t\right) \\ &= \frac{a_\tau}{a_1} (a_1 V_t + b_1 + \mu_v \delta_J) + \left((b_\tau + \tau \mu_v \delta_J) - \frac{a_\tau (b_1 + \mu_v \delta_J)}{a_1} \right) \\ &= a_\tau V_t + b_\tau + \tau \mu_v \delta_J \\ &= E(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t). \end{aligned}$$

From (28), (29) and (30) it is seen that λ_{Jt} is a function only of a small number of static parameters, plus lagged values of the latent V_t and the observed BV_t and MF_t . This dynamic

⁸This specification for λ_{Jt} also mimics the structure of the autoregressive conditional duration model for durations (Engle and Russell, 1998) and the observation-driven model for count data analysed in Jung, Kukuk and Leisenfeld (2006) and Feigen, Gould, Martin and Snyder (2008).

specification for the jump risk premium is thus advantageous from an inferential point of view. It also has the advantage of avoiding the (potential) need to price an additional random risk factor in the model.

4 MCMC Algorithm and Priors

Given the complexity of the state space model represented by (19) to (24) (with λ_J now replaced by λ_{Jt} on day t , as modelled by (28)), the joint posterior distribution for all unknowns is analytically intractable. Hence, an MCMC algorithm is applied to produce draws from the joint posterior and those draws then used to estimate inferential quantities of interest, including predictive densities, in the usual way. To reduce notation, we define the vectors $V_{1:t} = (V_1, V_2, \dots, V_t)'$, $Z_{1:t}^v = (Z_1^v, Z_2^v, \dots, Z_t^v)'$, $\Delta N_{1:t} = (\Delta N_1, \Delta N_2, \dots, \Delta N_t)'$, $BV_{1:t} = (BV_1, BV_2, \dots, BV_t)'$ and $MF_{1:t} = (MF_1, MF_2, \dots, MF_t)'$, for $t = 1, 2, \dots, T$, and with $MF_{1:0}$ and $BV_{1:0}$ empty. Using this notation, the joint posterior density for all unknowns is given by $p(V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}, \phi | MF_{1:T}, BV_{1:T})$, where the vector of static parameters is given by $\phi = (\kappa, \theta, \sigma_{BV}, \sigma_{MF}, \sigma_v, \lambda_D, \mu_v, \delta_J, \lambda_{J0}, \alpha_1, \alpha_2)$. The joint posterior density satisfies

$$\begin{aligned} & p(V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}, \phi | MF_{1:T}, BV_{1:T}) \\ & \propto \left[\prod_{t=1}^T p(MF_t | MF_{1:t-1}, BV_{1:t-1}, V_t, \kappa, \theta, \sigma_{MF}, \lambda_D, \mu_v, \delta_J, \lambda_{J0}, \alpha_1, \alpha_2) \right. \\ & \quad \times p(BV_t | V_t, \sigma_{BV}) \times p(V_t | V_{t-1}, Z_t^v, \Delta N_t, \kappa, \theta, \sigma_v) \\ & \quad \left. \times p(Z_t^v | \Delta N_t, \mu_v) \times p(\Delta N_t | \delta_J) \right] \times p(\phi), \end{aligned}$$

where it is assumed that $V_0 = \theta + \frac{\mu_v \delta_J}{\kappa}$. The conditioning of MF_t on lagged values of $MF_{1:t-1}$ and $BV_{1:t-1}$ derives from the assumed structure for λ_{Jt} in (28).

The Gibbs-based MCMC algorithm is implemented in four main steps:

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

$$\begin{aligned} & p(V_{1:T} | Z_{1:T}^v, \Delta N_{1:T}, \phi, MF_{1:T}, BV_{1:T}) \\ & \propto \prod_{t=1}^T p(MF_t | MF_{1:t-1}, BV_{1:t-1}, V_t, \kappa, \theta, \sigma_{MF}, \lambda_D, \mu_v, \delta_J, \lambda_{J0}, \alpha_1, \alpha_2) \\ & \quad \times p(BV_t | V_t, \sigma_{BV}) \times p(V_t | V_{t-1}, Z_t^v, \Delta N_t, \kappa, \theta, \sigma_v); \end{aligned}$$

2. Generating $Z_{1:T}^v$ from

$$p(Z_{1:T}^v | V_{1:T}, \Delta N_{1:T}, \phi, MF_{1:T}, BV_{1:T}) \propto \prod_{t=1}^T p(V_t | V_{t-1}, Z_t^v, \Delta N_t, \kappa, \theta, \sigma_v) \times p(Z_t^v | \Delta N_t, \mu_v),$$

(T truncated normal random variables);

3. Generating $\Delta N_{1:T}$ from

$$p(\Delta N_{1:T} | V_{1:T}, Z_{1:T}^v, \phi, MF_{1:T}, BV_{1:T}) \propto \prod_{t=1}^T p(V_t | V_{t-1}, Z_t^v, \Delta N_t, \kappa, \theta, \sigma_v) \times p(\Delta N_t | \delta_J),$$

(T Bernoulli random variables);

4. Generating ϕ (with elements of ϕ blocked conveniently) from

$$\begin{aligned} & p(\phi | V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}, MF_{1:T}, BV_{1:T}) \\ & \propto \left[\prod_{t=1}^T p(MF_t^{IV} | MF_{1:t-1}^{IV}, BV_{1:t-1}, V_t, \kappa, \theta, \sigma_{MF}, \lambda_D, \mu_v, \delta_J, \lambda_{J0}, \alpha_1, \alpha_2) \right. \\ & \quad \left. \times p(BV_t | V_t, \sigma_{BV}) \times p(V_t | V_{t-1}, Z_t^v, \Delta N_t, \kappa, \theta, \sigma_v) \right] \times p(\phi). \end{aligned}$$

Obtaining draws of $V_{1:T}$ is the most challenging aspect of the simulation scheme, due to the presence of the state dependent errors in (19), (20) and (21). Whilst the algorithm of Stroud, Muller and Polson (2003) is used as the basis of our approach, we are not aware of this algorithm having been applied to a model with multiple sources of state dependence and a bi-variate measure. In brief, $V_{1:T}$ is drawn (in blocks) via a Metropolis Hastings (MH) sub-step involving augmentation by a pair of mixture indicator vectors, one associated with the state equation and the other associated with the bivariate measurement equation. First, the augmentation variables are drawn from their respective full conditional distributions, given a previously sampled value of $V_{1:T}$. Then, a candidate vector $\tilde{V}_{1:T}$ is obtained from an approximating linear Gaussian model having error variances dependent upon the mixture indicator vectors, using the forward filter backward sampling (FFBS) method of Carter and Kohn (1994) and Frühwirth-Schnatter (1994). Further details of the mixture-based algorithm used to draw $V_{1:T}$ are provided in Appendix A.

The variance jumps, which occur with probability δ_J on any day t , shift the intercept term in (21) by an amount Z_t^v , where (by assumption) a variance jump on day t occurs contemporaneously with a price jump. Given the absence of raw price information in the model, we introduce information on the probability of price (and variance) jumps via the prior distribution, as follows. For a given sample period (details of which are given in Section 5) we identify the presence of significant price jumps on any day t by the realizations of the statistic described in (16) above. The prior for δ_J is then specified as a beta distribution with mean equal to the proportion of days throughout the sample on which significant jumps are found to occur. The variance of this distribution is used as a tuning parameter in the algorithm.

In a similar fashion, the prior for μ_v is specified as inverse gamma, with a mean value proportionate to the average magnitude of

$$JV_t \cdot I(r_t < 0) \tag{31}$$

over the sample of days on which significant large price *falls* are in evidence (indicated by $I(r_t < 0)$, where r_t is the daily logarithmic return), with JV_t as defined in (17). That is, *a priori*, we assume that the magnitude of the average jump in variance is some proportion of the average magnitude of the square of *negative* price jumps, reflecting the prior belief of a negative value for the parameter (ρ_J) relating the price and variance jumps (Eraker *et al.*, 2003, and Eraker, 2004). Once again, both the (a-priori) proportionate relationship between μ_v and the average of the sample quantity in (31), plus the variance of the prior distribution for μ_v , are viewed as tuning parameters in the algorithm.

With non-informative priors being invoked for the standard deviation parameters, σ_{BV} and σ_{MF} , simulation of these parameters is standard, via inverted gamma distributions. The volatility of volatility parameter, σ_v , is produced analogously, with the restriction $\sigma_v^2 < 2\kappa\theta$ ensuring the positivity of the latent variance processes. The joint prior for κ , θ and λ_D is uniform, subject to $\kappa\theta > \sigma_v^2/2$ and $\lambda_D < 0$, with the associated univariate conditional posteriors being non-standard due to the fact that the conditional mean function in (20) is non-linear in all three parameters. We use the structure of the model to define Gaussian kernels and produce candidate draws for each of κ , θ and λ_D with separate MH sub-steps. The parameters of the jump premium process are also uniform *a priori*, subject to $\lambda_{J0} > 0$, $\alpha_1 > 0$, $\alpha_2 > 0$ and $\alpha_1 + \alpha_2 < 1$, resulting in a joint normal posterior truncated by the inequality constraints. Draws of the vector $\lambda_{J1:T} = (\lambda_{J1}, \lambda_{J2}, \dots, \lambda_{JT})'$ are produced automatically from (the degenerate) $p(\lambda_{J1:T} | V_{1:T}, Z_{1:T}^v, \Delta N_{1:T}, \phi, MF_{1:T}, BV_{1:T})$ via the conditionally deterministic relationship in (28), with $l_0 = \lambda_{J0}$. Given the form of independent, informative priors invoked for μ_v and δ_J (as described above), inverse gamma and beta candidates are adopted for the respective MH sub-steps for these two parameters.

5 Empirical Application

5.1 Data description

In this section we report results of the application of the algorithm to daily variance measures constructed from intraday spot and option price data for the S&P500 index, for the period July 26, 1999 to December 31, 2008.⁹ The first 1961 observations (corresponding to daily variance measures for the period July 26, 1999 to June 21, 2007) are used as the *initial* dataset for estimation, with the remaining 386 observations being reserved for the evaluation of the one-step-ahead probabilistic forecasts, based on an expanding sample window. The evaluation period covers: the period immediately preceding the global financial crisis (June

⁹The numerical results reported in this section have been produced using the JAVA programming language.

22, 2007 to July 27, 2007); the early period of the crisis - during which defaults on sub-prime mortgages began to impact on the viability of financial institutions and the availability of credit (late July 2007 to August 2008); and the period of historically unprecedented high levels of stock market volatility (towards the end of 2008). All index data has been supplied by the Securities Industries Research Centre of Asia Pacific (SIRCA) on behalf of Reuters, with the raw index data having been cleaned using the methods of Brownlees and Gallo (2006).¹⁰ The raw option-implied measure (corresponding to MF_t^{QV}) is based on the publicly available VIX_t^2 , associated with day t , constructed by the Chicago Board Option Exchange (CBOE) using the model-free methodology. See the CBOE website (www.cboe.com) for more details.

The realized bipower variation measure in (9) is based on fixed five minute sampling, with a ‘nearest price’ method used to construct artificial returns five minutes apart. Note that the various forms of microstructure noise-adjusted measures that have appeared in the literature have their prime motivation in the case of data on traded assets, rather than observations on a constructed index. However, one could argue that the presence of stale prices in the index at the point of any recorded up-date, plus the inherent discreteness in the underlying prices, induce a form of noise. With this view we use a subsampled (or averaged) version of the five-minute based measure as an additional form of noise adjustment (i.e. in addition to sampling the observations at fixed five minute intervals).

In Figure 1 we reproduce plots of BV_t and MF_t for $c = 0, 1, 1.2$ and 1.5 (Panels A to D respectively) for the full sample period. The plot of MF_t in Panel A (with $c = 0$) corresponds to the raw CBOE measure $MF_t = MF_t^{QV} = VIX_t^2$. The area highlighted in grey at the end of each plot represents the period reserved for the evaluation of the probabilistic forecasts. The empirical regularity of the option-implied variance exceeding (in the main) the realized variance is in evidence in all panels. Despite the option-implied measure being much less noisy than the (daily) returns-based measure, both measures exhibit broadly similar fluctuations, with there being only a slight tendency for the peaks in MF_t to lag those in BV_t . The magnitude and variability of the realized variance and the option-implied measures increased significantly during the global financial crisis, demonstrated here with the sharp increase in both series towards the end of the sample. The plots in Panels B to D show a reduction in the extreme movements in the option-implied measure, relative to the raw measure, MF_t^{QV} , as befits an adjustment that removes from MF_t^{QV} (an estimate of) the risk neutral expectation of price jump variation. However, there is little discernible difference in the three MF_t series, across the different values of c employed in adjusting the raw MF_t^{QV} for price jump variation.

¹⁰The authors would like to acknowledge the excellent research assistance of Chris Tse in producing the realized variance and bipower series.

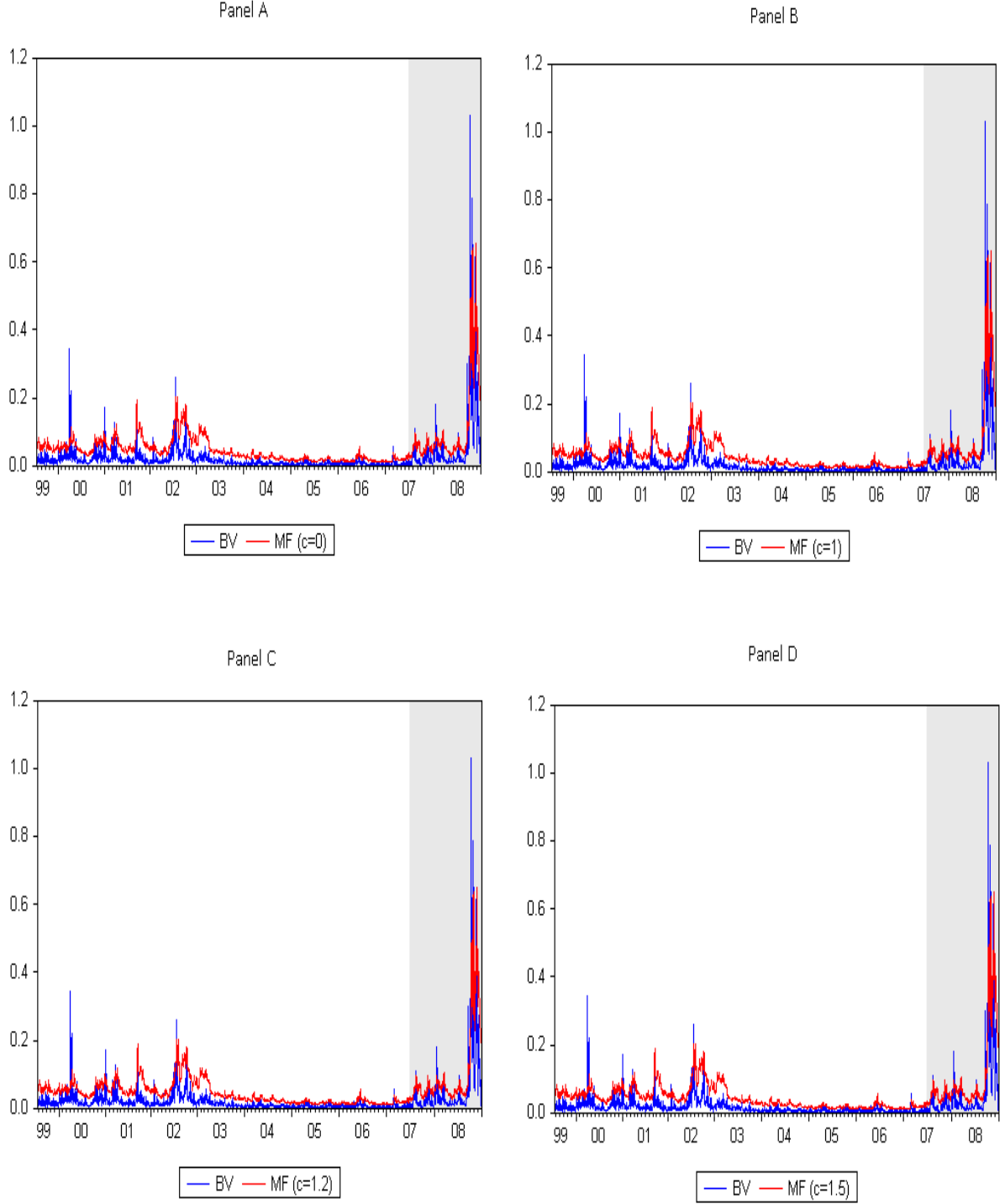


Figure 1: Plots of BV_t and MF_t^{QV} (Panel A) and BV_t and MF_t for $c = 1.0, 1.2$ and 1.5 (Panel B, C and D, respectively).

5.2 Empirical Estimates

In Table 1 we report the results based on estimation of the model in (19) to (24) and (28), for the initial sample period, July 26, 1999 to June 21, 2007. Marginal posterior means (MPM) and 95% interval estimates are calculated from 50,000 MCMC draws, following a 50,000 draw burn-in period. We report the results based on BV_t and MF_t for three different values of c , $c = 1.0, 1.2, 1.5$. The first value of c implies that the risk neutral expectation of jump variation is equivalent to the objective expectation, while the second and third values (respectively) imply that the risk neutral expectation is 20% and 50% larger than the objective expectation.¹¹ Convergence of the algorithm is assessed using the visual inspection of cumulative sum statistics; see, for example, Bauwens and Lubrano (1998).

Both point and interval estimates of the parameters are very robust overall to the value of c . Most notably, for all values of c , the results indicate a high level of persistence in both the latent variance process (small value of κ) and the process for λ_{Jt} (high values of α_1 and α_2). The latter result supports the decision to model the variance jump premium as a dynamic process and to produce distributional forecasts for this quantity. As anticipated, the (time series) average of λ_{Jt} , $\bar{\lambda}_J$, (as estimated here by the average of the T MCMC draws of each λ_{Jt}) is positive, whilst the (constant) value of λ_D is negative. The magnitude of λ_D is small, consistent with values reported by Eraker (2004), indicating a slight increase in the persistence in volatility under the risk neutral measure. However, the magnitude of $\bar{\lambda}_J$ is large, in particular in comparison with the estimated magnitude of the mean of the actual variance jumps, μ_v , over the period. This result indicates the extreme sensitivity of the market to this aspect of the latent variance, with the implied risk neutral expectation of (variance) jump size (μ_v^*) being many orders of magnitude larger than the *actual* mean jump size. The point (and interval) estimates of δ_J indicate that the probability of a variance jump occurring on any one day ranges from approximately 2% to 5%. These magnitudes slightly exceed those reported in previous studies (based on different model specifications and for earlier sample periods), as summarized by Broadie *et al.* (2007), where point estimates of the probability of variance jumps vary from 0.3% to 2%. The estimates of the mean of the variance jumps reported in Broadie *et al.*, μ_v , range from 0.018 to 0.037 (in annualized decimal form), and are thus slightly higher than the magnitudes for this parameter recorded in Table 1, but are broadly consistent nevertheless.

¹¹Note that the *a priori* assumption being adopted here is that price jump premium is non-negative, i.e. that $c \geq 1$.

Table 1: Empirical results for the S&P500 stock index for July 26, 1999 to June 21, 2007, based on the measures, BV_t and MF_t . The value of c recorded is that used in the production of MF_t in (18).

	$c = 1.0$	$c = 1.2$	$c = 1.5$
Parameter	MPM (95% HPD)	MPM (95% HPD)	MPM (95% HPD)
κ	0.01706 (0.01309, 0.02091)	0.018246 (0.01386, 0.02259)	0.01707 (0.01300, 0.02110)
θ	0.00881 (0.00529, 0.01215)	0.00870 (0.00542, 0.01183)	0.00885 (0.00507, 0.01249)
σ_v	0.00815 (0.00761, 0.00867)	0.00809 (0.00746, 0.00885)	0.00891 (0.00828, 0.00951)
σ_{BV}	0.60822 (0.58624, 0.63073)	0.59575 (0.57292, 0.61979)	0.59418 (0.57166, 0.61844)
σ_{MF}	0.26501 (0.19478, 0.46334)	0.24634 (0.18140, 0.43857)	0.24424 (0.17803, 0.43837)
μ_v	0.00756 (0.00597, 0.00977)	0.00868 (0.00683, 0.01110)	0.00800 (0.00633, 0.01037)
δ_J	0.03713 (0.02495, 0.05202)	0.03684 (0.02421, 0.05097)	0.03778 (0.02482, 0.05308)
λ_D	-0.00710 (-0.01727, -0.00027)	-0.00804 (-0.01876, -0.00028)	-0.00759 (-0.01765, -0.00026)
λ_{J0}	0.07668 (0.00382, 0.15897)	0.08782 (0.00446, 0.18314)	0.08129 (0.00415, 0.16924)
α_1	0.38226 (0.01803, 0.87587)	0.38545 (0.01799, 0.87734)	0.38403 (0.01778, 0.87614)
α_2	0.28115 (0.01148, 0.67120)	0.26495 (0.01026, 0.64011)	0.26627 (0.00997, 0.61428)
$\bar{\lambda}_J$	0.36964	0.36047	0.34810

5.3 Probabilistic forecasts of latent variables

Bayesian predictions of the latent variance, V_{T+1} , and the dynamic jump risk premium, λ_{JT+1} , are produced by estimating the (marginal) predictive densities, $p(V_{T+1}|MF_{1:T}, BV_{1:T})$ and $p(\lambda_{JT+1}|MF_{1:T}, BV_{1:T})$, respectively. In the usual fashion, the marginal predictives are estimated by taking draws from the respective conditional predictives (for V_{T+1} and λ_{JT+1}), given MCMC draws of the relevant conditioning quantities. The predictive density of the diffusive risk premium, $p(\lambda_D V_{T+1}|MF_{1:T}, BV_{1:T})$, is estimated via draws of the product, $\lambda_D V_{T+1}$, composed from the posterior draws of λ_D and V_{T+1} .

Recursive forecasts of V_{T+1} , λ_{JT+1} and $\lambda_D V_{T+1}$ (based on expanding windows) are produced over the evaluation period: June 22, 2007 to December 31, 2008. Given the robustness of the results to the value of c (in (18)), we summarize the predictive results for $c = 1.5$ only. The three predictive densities are estimated using every 10^{th} of the 50,000 draws of all unknowns (after a 50,000 draw burn-in). The draws of ϕ are only updated every 60 days (approximately) throughout the evaluation period¹². In Figure 2, plots of both the marginal predictive means and 95% predictive intervals for V_{T+1} , λ_{JT+1} and $\lambda_D V_{T+1}$ are given respectively in Panels A, B and C. As is clear, as the crisis deepens, both the level of the latent variance itself and the premium demanded by investors for variance risk (of both the diffusive and jump type), increases. The degree of uncertainty associated with all three latent variables, as measured by the magnitude of the prediction intervals, is also *markedly* larger in the extreme crisis period, in late 2008, than in the earlier (and pre-) crisis period.

5.4 Probabilistic forecasts of observables

Using the hierarchical structure of the state space model, the Bayesian predictives for the observable variance quantities, namely $p(BV_{T+1}|MF_{1:T}, BV_{1:T})$ and $p(MF_{T+1}|MF_{1:T}, BV_{1:T})$, are estimated using the MCMC draws of all unknowns, including V_{T+1} and λ_{JT+1} . These predictive distributions are summarized in Figure 3, with the observed values also displayed. Over the full evaluation period, the empirical coverage of the 95% prediction intervals for BV_{T+1} and MF_{T+1} are 94.82% and 92.28%, respectively, with these figures highlighting the accuracy with which our approach predicts the observable variance measures (most notably the objective measure), even during the height of the financial crisis.

Given the current practice of producing forecasts of these measures via univariate, observation-driven time series models, it is of interest to also compare the accuracy of our method with such simpler alternatives. With an extensive comparative evaluation of the accuracy of al-

¹²The (static) parameter estimates are very robust across sub-periods, justifying this attempt to ease the computational burden associated with the production of the forecasts.

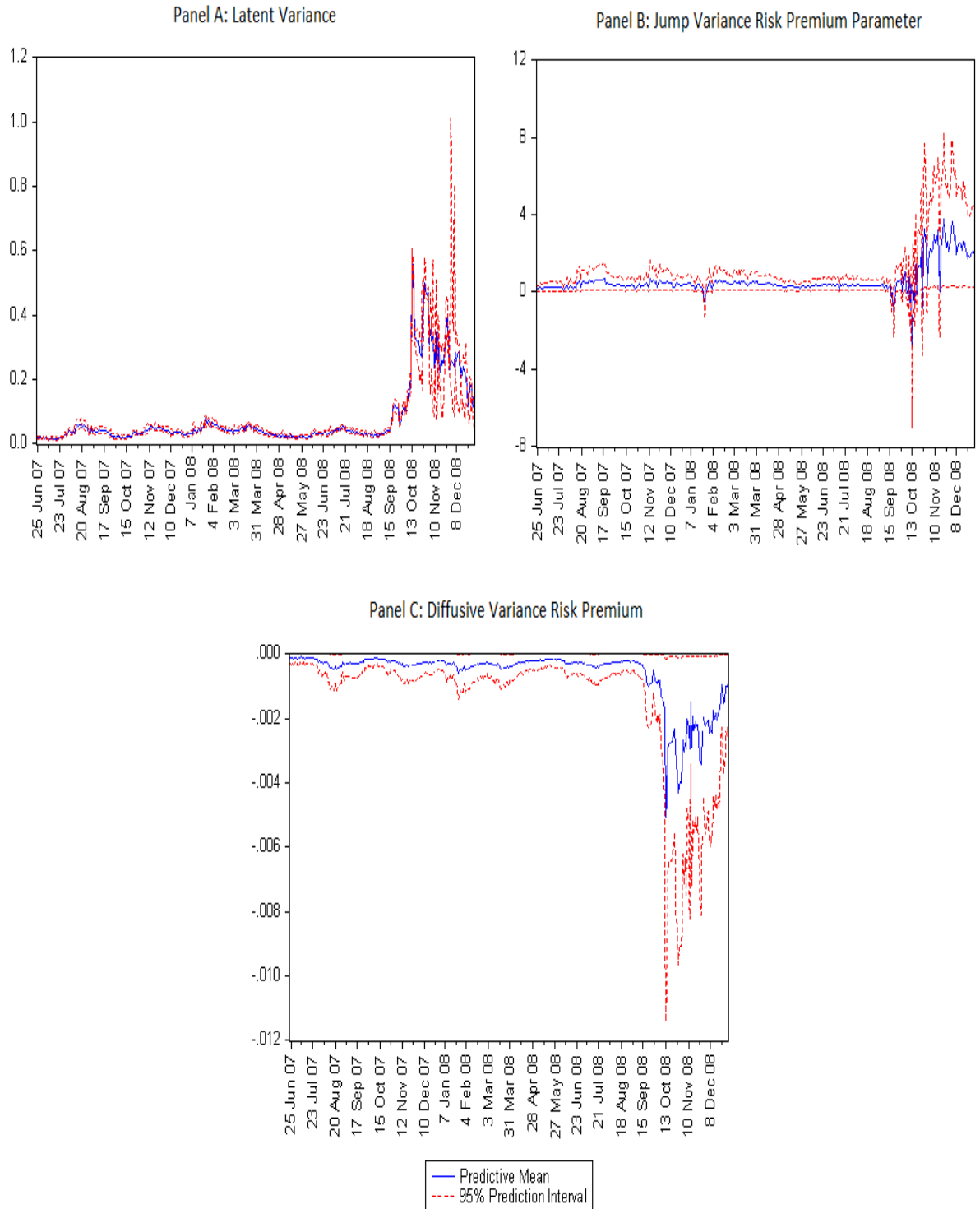


Figure 2: One-step-ahead predictive means and 95% predictive intervals of V_{T+1} (Panel A), λ_{JT+1} (Panel B) and $\lambda_D V_{T+1}$ (Panel C) for June 22, 2007 to December 31, 2008.

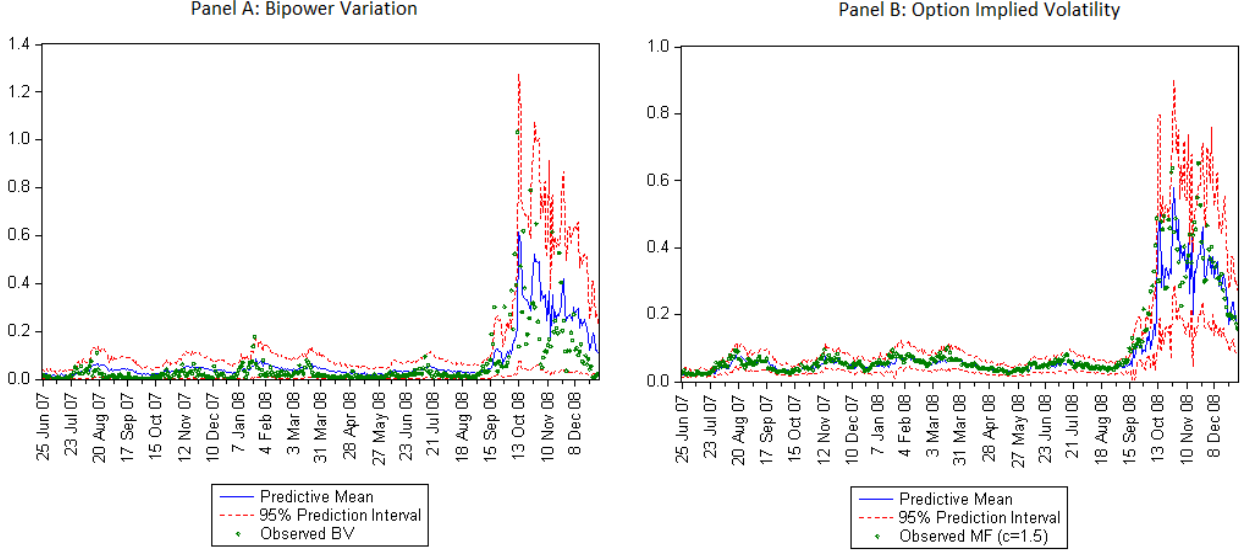


Figure 3: One-step-ahead predictive means, 95% predictive intervals and observed values of BV_t (Panel A) and MF_t (Panel B) from June 22, 2007 to December 31, 2008.

ternative volatility forecasts being beyond the scope of this paper, we compare our state space-based forecasts with forecasts from a single model only for each of BV_t and MF_t . Specifically, for BV_t we consider the heterogeneous autoregressive (HAR) specification introduced by Corsi (2004) and used by Andersen, Bollerslev and Diebold (2007) and Bollerslev, Kretschmer, Pigorsch and Tauchen (2009), amongst many others,

$$BV_t = \beta_0 + \beta_D BV_{t-1} + \beta_W BV_{t-5,t} + \beta_M BV_{t-22,t} + \varepsilon_t, \quad (32)$$

where $BV_{t-h,t} = h^{-1} \sum_{i=1}^h BV_{t-i}$. In order to best replicate typical statistical practice in the area, we estimate this model using quasi-maximum likelihood estimation (QMLE), based on Gaussian errors. For MF_t we adopt a second-order autoregressive model for the first differenced series,

$$\Delta MF_t = \beta_0 + \beta_1 \Delta MF_{t-1} + \beta_2 \Delta MF_{t-2} + \eta_t, \quad (33)$$

where $\Delta MF_t = MF_t - MF_{t-1}$, and a Gaussian distributional assumption is again adopted and QMLE applied. (See Ahoniemi, 2006, for discussion of time series models for option implied volatility).

Following Geweke and Amisano (2010), we compare the accuracy of the Bayesian marginal predictive densities for BV_t and MF_t , $\hat{p}(BV_t^o | MF_{1:t-1}, BV_{1:t-1})$ and $\hat{p}(MF_t^o | MF_{1:t-1}, BV_{1:t-1})$, with that of the predictive densities produced by the competing univariate models, $\tilde{p}(BV_t^o | MF_{1:t-1}, BV_{1:t-1}, \hat{\theta}_{QMLE,t-1})$ and $\tilde{p}(MF_t^o | MF_{1:t-1}, BV_{1:t-1}, \hat{\theta}_{QMLE,t-1})$, respectively.

Here the competing predictive densities are constructed using the QMLE of the unknown parameter vector of the model being assessed, $\hat{\theta}_{QMLE,t-1}$, as based on data up to and including period $t-1$. To mimic the approach adopted for ϕ in the state space model, $\hat{\theta}_{QMLE,t-1}$ is up-dated only every 60 days (approximately). The height of each predictive density is evaluated at the observed realization, BV_t^o or MF_t^o , over the evaluation period $t = 1, 2, \dots, 386$, with the cumulative difference in log scores (CLS),

$$\begin{aligned} CLS_{BV} &= \sum_{t=T+1}^{T+386} \ln \left[\frac{\hat{p}(BV_t^o | MF_{1:t-1}, BV_{1:t-1})}{\tilde{p}(BV_t^o | MF_{1:t-1}, BV_{1:t-1}, \hat{\theta}_{QMLE,t-1})} \right] \text{ and} \\ CLS_{MF} &= \sum_{t=T+1}^{T+386} \ln \left[\frac{\hat{p}(MF_t^o | MF_{1:t-1}, BV_{1:t-1})}{\tilde{p}(MF_t^o | MF_{1:t-1}, BV_{1:t-1}, \hat{\theta}_{QMLE,t-1})} \right], \end{aligned}$$

used to measure the relative performance of the two methods. A positive CLS value indicates that the Bayesian marginal predictive outperforms the competitor.

Over the full evaluation period, $CLS_{BV} = 1381.05$ and $CLS_{MF} = 445.34$, indicating that the state space model provides superior predictions of *both* BV_t and MF_t . Plots of the cumulative sums of both CLS_{BV} and CLS_{MF} (from $t = T + 1$ to $t = T + k$, for $k = 2, 3, \dots, 386$) are produced in Figure 4 to gauge relative predictive performance over the crisis period as each new observation contributes to the cumulative difference in log scores. From the plot in Panel A, it is clear that the state space model for BV_t *consistently* dominates the HAR specification over the full sample period. In the case of the predictions of MF_t , despite the early inferior performance of the state space model, as the extreme period of the global financial crisis starts to unfold towards the end of 2008, the state space model completely dominates the simple univariate alternative, yielding the large positive value for CLS_{MF} for the full period.

5.5 Applications of the volatility and risk premia forecasts

Whilst the (overall) accuracy of the state space-based forecasts of the observed variance measures themselves is certainly worthy of note, a more convincing testimony to the (predictive) worth of any model comes from its ability to produce accurate forecasts of financial quantities into which volatility is an *input*. With this in mind, we conduct two exercises. First, in Section 5.5.1, we use our bi-variate state space model, *augmented* by an additional measurement equation based on daily returns (r_t), to produce recursive predictive distributions for r_{T+1} , from which the daily 5% and 1% VaRs are extracted. Conventional statistics for both unconditional coverage and independence of exceedances are reported for the evaluation period, and compared with corresponding statistics based on a univariate HAR model for realized variance. Secondly, in Section 5.5.2 we document the accuracy of estimates of

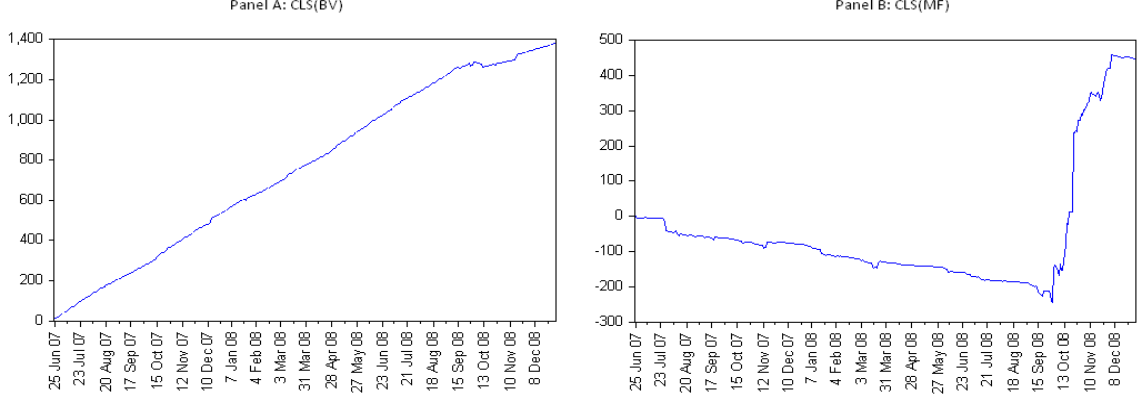


Figure 4: Cumulative sum of the CLS_{BV} (Panel A) and the CLS_{MF} (Panel B) for June 22, 2007 to December 31, 2008.

the settlement prices of VIX futures, as based on our state space-based forecasts of the latent variance. In this case the comparator is a series of forecasts produced by fitting the model in (33) to the observed VIX_t^2 ($= MF_t^{QV}$) itself. Accuracy is gauged by comparing the empirical coverage of the relevant probability intervals with the nominal level, as well as comparing the mean squared errors associated with the relevant point predictions.

Finally, in the spirit of that Bollerslev *et al.* (2011), in Section 5.5.3 predictions of the latent variance and the variance risk *premia* are used to extract a sequence of one-step-ahead predictions of the relative risk aversion of the representative agent. Whilst it is not possible to conduct an assessment of accuracy in this case, due to the absence of ‘observations’ on risk aversion, it is still of interest to document this outcome of our model and to calibrate the results with comparable results recorded in the literature.

5.5.1 Value at risk (VaR) prediction

Predicting the one-day-ahead 5% and 1% VaR for the market portfolio associated with the S&P500 index is equivalent to calculating the 5% and 1% quantile, respectively, for the predictive distribution for the portfolio return. Although our state space model does not explicitly model the return, we provide a method for augmenting the inferences drawn from the model based on the variance measures to produce, in turn, forecast distributions for the future return. Specifically, draws from the posterior distribution of the volatility model, conditional upon the spot- and option-based variance measures, are *resampled* to reflect additional conditioning on observed end-of-day returns. The model for the logarithmic price return, $r_t = \ln(P_t) - \ln(P_{t-1})$, conditional on $(V_t, V_{t-1}, Z_t^v, \Delta N_t)'$, is based on an initial Euler

approximation of (1),

$$r_t = \left(\mu - \frac{V_t}{2} \right) - \frac{\rho}{\sigma_v} (V_t - \kappa\theta - (1 - \kappa) V_{t-1} - Z_t^v \Delta N_t) + (\mu_p + \rho_J Z_t^v) \Delta N_t + \sqrt{V_t (1 - \rho^2) + \sigma_p^2 \Delta N_t} \xi_{4t}, \quad (34)$$

with $\{\xi_{4t}\} \stackrel{iid}{\sim} N(0, 1)$, assumed to be independent of $\{\xi_{3t}\}$, $\{\xi_{2t}\}$ and $\{\xi_{1t}\}$ in (19), (20) and (21) respectively. In order to render the resampling method computationally efficient, we make the simplifying assumption that $\sigma_p^2 = 0$, and introduce an unknown scale parameter, σ_r , absorbing all constant factors in the error variance, so that the error term in (34) collapses to $\sigma_r \sqrt{V_t} \xi_{4t}$. In addition, we introduce an additional regression parameter, β_v , resulting in a final model for returns given by

$$r_t = \mu + \beta_v V_t - \rho \left(\frac{V_t - \kappa\theta - (1 - \kappa) V_{t-1} - Z_t^v \Delta N_t}{\sigma_v} \right) + \mu_p \Delta N_t + \rho_J Z_t^v \Delta N_t + \sigma_r \sqrt{V_t} \xi_{4t}. \quad (35)$$

Draws of $V_{1:T}$, $Z_{1:T}^v$, $\Delta N_{1:T}$, θ , κ and σ_v , produced via the application of the MCMC algorithm described in Section 4, are resampled (as per the description in Appendix B), with draws of the return-specific parameters in (34), μ , β_v , ρ , μ_p , ρ_J and σ_r , then produced by exploiting the regression structure in (35) in the usual way. With standard informative priors used for the return-specific parameters, the resampling method exploits the (closed-form) solution for the marginal likelihood for the vector of returns $r_{1:T} = (r_1, r_2, \dots, r_T)'$, conditional on $V_{1:T}$, $Z_{1:T}^v$, $\Delta N_{1:T}$, θ , κ and σ_v .

In Table 3 we report the empirical coverage statistics for the 5% and 1% VaR predictions produced by our (augmented) state space approach, along with the corresponding statistics associated with an HAR model fitted to realized variance (HAR-RV). As with the HAR model adopted above for the bipower measure, we assume a Gaussian distribution for the innovations and estimate the model using QMLE. We also report, for both approaches, the p -values associated with the tests of correct unconditional coverage and independence of exceedances of the VaR (Christoffersen, 1998). The former test assesses whether the empirical coverage differs significantly from the nominal level, while the latter tests for independence in the sequence of returns that exceed the VaR. An acceptable series of VaR predictions should fail to reject both of these hypotheses. Results are reported for an evaluation period that excludes the final four months of 2008, and for the full evaluation period. Overall, the state space approach provides more accurate coverage than the HAR-RV model, with empirical coverages that are a good deal closer to the nominal levels than those of the HAR-RV model. Its empirical coverages over the June 22, 2007 to August 31, 2008 period, for both the 5% and 1% VaR, are also not significantly different from their respective nominal levels (at the 5% significance level). In addition, the state space approach accurately captures

the dynamics in returns - over both periods considered and for both VaR levels - as the null hypothesis of independence in the exceedances is not rejected in all four cases. The HAR-RV model, on the other hand, fails the independence test in both periods, and for both the 5% and 1% VaR cases.

Table 3: Empirical coverage and p -values for the unconditional coverage (UC) and independence (IND) tests for 5% and 1% VaR predictions

		Excluding the extreme period June 22, 07 to Aug 31, 08		Overall period June 22, 07 to Dec 31, 08	
		State Space	HAR-RV	State Space	HAR-RV
5% VaR	Empirical Coverage	7.31%	13.29%	10.10%	14.51%
	UC Test	0.0845	0.0000	0.0000	0.0000
	IND Test	0.0619	0.0006	0.5816	0.0046
1% VaR	Empirical Coverage	0.33%	7.97%	2.59%	7.77%
	UC Test	0.1762	0.0000	0.0088	0.0000
	IND Test	0.9348	0.0409	0.4652	0.0243

5.5.2 Pricing VIX futures

The Chicago Board of Exchange (CBOE) introduced trading on futures contracts based on the VIX in March 2004. For a given futures contract with maturity date T_M , the settlement price quoted on day t , used for marking-to-market purposes, reflects the market's expectation of the VIX value at the time of maturity. The square of the VIX ($= MF^{QV}$ in our notation) at time T_M is, in turn, a representation of the risk neutral expectation of $QV_{T_M, T_M+\tau}$, namely, quadratic variation over the period τ dated from T_M . It is thus of interest to ascertain the degree to which a model-based prediction of $\sqrt{E^*(QV_{T_M, T_M+\tau} | \mathcal{F}_{T_M})}$ accords with observed settlement prices on the VIX .

Under our assumed model,

$$\begin{aligned}
E[E^*(Q\mathcal{V}_{T_M, T_M+\tau}|\mathcal{F}_{T_M})|\mathcal{F}_t] &= E[E^*(\mathcal{V}_{T_M, T_M+\tau}|\mathcal{F}_{T_M})|\mathcal{F}_t] \\
&\quad + E\left[E^*\left(\sum_{T_M < s \leq T_M+\tau}^{N_{T_M+\tau}} (Z_s^p)^2 \middle| \mathcal{F}_{T_M}\right) \middle| \mathcal{F}_t\right] \\
&= a_\tau^* E(V_{T_M}|\mathcal{F}_t) + b_\tau^* + \tau \delta_J [\mu_v + E(\lambda_{JT_M}|\mathcal{F}_t)] \\
&\quad + E\left[E^*\left(\sum_{T_M < s \leq T_M+\tau}^{N_{T_M+\tau}} (Z_s^p)^2 \middle| \mathcal{F}_{T_M}\right) \middle| \mathcal{F}_t\right],
\end{aligned}$$

where $E(V_{T_M}|\mathcal{F}_t)$ and $E(\lambda_{JT_M}|\mathcal{F}_t)$ are $(T_M - t)$ -steps-ahead point predictions of the stochastic variance and the jump variance risk premium, respectively. These two quantities can be obtained via our assumed models for V_t and λ_{Jt} , (21) and (28) respectively, whilst

$$E\left[E^*\left(\sum_{T_M < s \leq T_M+\tau}^{N_{T_M+\tau}} (Z_s^p)^2 \middle| \mathcal{F}_{T_M}\right) \middle| \mathcal{F}_t\right] = E\left[cE\left(\sum_{T_M < s \leq T_M+\tau}^{N_{T_M+\tau}} (Z_s^p)^2 \middle| \mathcal{F}_{T_M}\right) \middle| \mathcal{F}_t\right]$$

is produced from the univariate time series modelling of price jump variation as described in Section 3.1), with $c = 1.5$ imposed.

Using posterior draws of all unknowns, draws of $\sqrt{E[E^*(Q\mathcal{V}_{T_M, T_M+\tau}|\mathcal{F}_{T_M})|\mathcal{F}_t]}$ can be produced for $t = 1, 2, \dots, 386$, over the evaluation period: June 22, 2007 to December 31, 2008, for futures contracts with the closest maturity (i.e. the smallest value of T_M at any given point in time t). A point estimate of $\sqrt{E[E^*(Q\mathcal{V}_{T_M, T_M+\tau}|\mathcal{F}_{T_M})|\mathcal{F}_t]}$ is given by the sample mean of the draws and the 95% probability interval produced from the draws in the usual way. As a comparator, the model in (33) is fitted to the observed VIX_t^2 (with a Gaussian distributional assumption adopted for the error term) and used to produce predictions of VIX . As reported in Table 4, the 95% interval constructed for the state space-based prediction covers approximately 82% of the observed daily settlement values, whilst the univariate time series model produces a coverage of only 66%. Further, the accuracy of the (estimated) posterior mean of $\sqrt{E[E^*(Q\mathcal{V}_{T_M, T_M+\tau}|\mathcal{F}_{T_M})|\mathcal{F}_t]}$ as a point predictor of the settlement price - as measured by the mean squared error - is also superior to that of the univariate comparator.

5.5.3 Prediction of Risk Aversion

In modelling both variance risk premia (diffusive and jump) as dynamic processes, we are effectively modelling time variation in the risk aversion of the representative investor. In the spirit of Bollerslev *et al.* (2011), we link the risk aversion parameter, γ , associated with the power utility function for a representative investor, to both $\lambda_D V_t$ and λ_{Jt} . As per the analytical demonstration in Appendix C, this leads to the following relationship between risk

Table 4: Prediction of *VIX* future prices: empirical coverage of 95% prediction intervals and mean squared error of point predictions

	State Space	Time Series
95% Interval Coverage	82.12%	65.80%
Mean Squared Error	18.66	27.20

aversion in period $T + 1$, γ_{T+1} , the two risk premia, $\lambda_D V_{T+1}$ and $\lambda_J V_{T+1}$, and certain other unknowns in the model,

$$\gamma_{T+1} = \frac{\lambda_D V_{T+1} - \delta_J \lambda_{JT+1}}{\rho \sigma_v V_{T+1} + \rho_J \delta_J \mu_v^2}. \quad (36)$$

Thus, draws of all unknowns on the right hand side of (36) (with draws of the *full* set of these unknowns, including ρ and ρ_J , obtained as part of the resampling exercise in Section 5.5.1) can be used to produce draws of γ_{T+1} and those draws used, in turn, to estimate the predictive density of γ_{T+1} . Similar to the plots presented in Figure 2, a plot summarizing the predictive distribution of γ_{T+1} is produced in Figure 5. In the earlier period, in particular, the mean values of γ_{T+1} fluctuate around values that are broadly consistent with the range of estimates - produced via very different means - that have been reported for this parameter in the literature (see Cochrane, 2005 and Bollerslev *et al.*, 2011 for some recent discussion and documentation of these values). However, as is to be expected, given the relationship between γ_{T+1} and the variance risk premia, the implied risk aversion of the representative investor increases *dramatically* during the extreme crisis period towards the end of 2008, together with a widening of the predictive bounds. In particular, the widening bounds can be interpreted as an increased dispersion in the behaviour of investors during the crisis. That is, the confusion that all investors suffer during such a period can be seen as translating here into uncertain estimates of the risk aversion level associated with a rational representative investor. (See Bollerslev and Todorov, 2009, for related discussion).

6 Conclusions

This paper is the first to combine non-Gaussian, non-linear state space techniques with the Bayesian inferential methodology for the purpose of producing probabilistic predictions of objective volatility and its associated risk premia, using both option and returns-based volatility measures. In the usual fashion, the premium for diffusive variance risk is linear

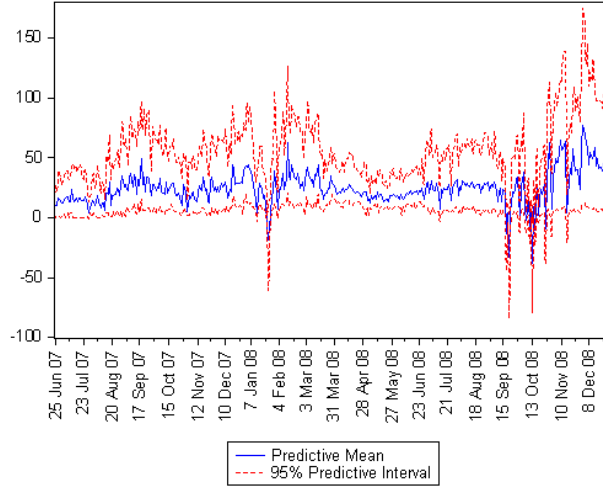


Figure 5: One-step-ahead predictive mean and 95% predictive interval of γ_{T+1} (the relative risk aversion parameter for a representative investor with power utility) for June 22, 2007 to December 31, 2008.

in the latent variance, while a conditionally deterministic process driven by a past function of the measurements is adopted for the premium for jump variance risk. An empirical investigation using the S&P 500 market portfolio sheds light on the changes in the predictive distributions of the latent quantities of interest, including the latent objective variance, over the period of the recent global financial crisis. Our state space approach is shown to provide highly accurate predictive coverage over the full out-of-sample period, most notably for the observable objective measure. Our method also yields superior predictions of the observable measure of objective variance to those of a widely-used univariate time-series model. More accurate predictive performance (overall) is demonstrated for the case of the option-implied variance measure, with the predictive performance of the state space model, relative to that of a univariate time series specification, becoming completely dominant as the market itself becomes more volatile. The accuracy with which latent volatility is predicted translates into relatively accurate predictions of related quantities - namely the VaR on the market portfolio and the market prices of futures written on the *VIX* index. Via a particular form of representative agent model, we link the dynamic risk premia to the risk aversion parameter, enabling probabilistic forecasts of the risk aversion of a representative investor to be produced. Our model quantifies the changes in investor risk aversion over the financial crisis period, where, as might be anticipated, risk aversion is seen to increase dramatically in both level and variability as the crisis deepens. Most notably, the results point to the fact that *extreme* values for this behavioural parameter would have been predicted, with

non-negligible probability, during the height of the stock market turmoil.

References

- [1] Ahoniemi, K. 2006. Modeling and Forecasting Implied Volatility - an Econometric Analysis of the VIX Index, *Discussion paper* No. 129, *Helsinki Center of Economic Research, University of Helsinki, Finland*.
- [2] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R.J.A. 2010. Modeling Financial Contagion Using Mutually Exciting Jump Processes, *NBER Working Paper*, No. w15850.
- [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] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. 2003. Modelling and Forecasting Realized Volatility, *Econometrica*, 71: 579-625.
- [5] 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.
- [6] Barndorff-Nielsen, O.E. and Shephard, N. 2004. Power and Bipower Variation with Stochastic Volatility and Jumps, *Journal of Financial Econometrics*, 2: 1-37.
- [7] Barndorff-Nielsen, O.E. and Shephard, N. 2006. Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation, *Journal of Financial Econometrics*, 4: 1-30.
- [8] Bates, D.S. 2000. Post-87 Crash Fears in the S&P 500 Futures Option Market, *Journal of Econometrics*, 94: 181-238.
- [9] Bauwens, L and Lubrano, M. 1998. Bayesian Inference on GARCH Models using the Gibbs Sampler, *Econometrics Journal*, 1: C23-C46.
- [10] Blair, B.J, Poon, S-H. and Taylor, S.J. 2001. Forecasting S&P100 Volatility: the Incremental Information Content of Implied Volatilities and High Frequency Index Returns, *Journal of Econometrics*, 105: 5-26.
- [11] Bollerslev, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31: 307-327.

- [12] Bollerslev, T., Gibson, M. and Zhou, H. 2011. Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities, *Journal of Econometrics*, 160: 235–245.
- [13] Bollerslev, T., Kretschmer, U., Pigorsch, C. and Tauchen, G. 2009. A Discrete-Time Model for Daily S&P500 Returns and Realized Variations: Jumps and Leverage Effects, *Journal of Econometrics*, 150: 151-166.
- [14] Bollerslev, T., Sizova, N. and Tauchen, G. 2009. Volatility in Equilibrium: Asymmetries and Dynamic Dependencies, *CREATES Research Paper No. 2009-5, School of Economics and Management, University of Aarhus, Denmark*.
- [15] Bollerslev, T., Tauchen, G. and Zhou, H. 2009. Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies*, 22: 4463-4492.
- [16] Bollerslev, T. and Todorov, V. 2009. Tails, Fears and Risk Premia, *CREATES Research Paper No. 2009-26, School of Economics and Management, University of Aarhus, Denmark*.
- [17] Bollerslev, T. and Zhou, H. 2002. Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility, *Journal of Econometrics*, 109: 33-65.
- [18] Breeden, D.T. 1979. An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities, *Journal of Financial Economics*, 7: 265-296.
- [19] Britten-Jones, M. and Neuberger, A. 2000. Option Prices, Implied Price Processes and Stochastic Volatility, *The Journal of Finance*, LV: 839-866.
- [20] Broadie, M., Chernov, M. and Johannes, M. 2007. Model Specification and Risk Premia: Evidence from Futures Options, *The Journal of Finance*, LXII: 1453-1490.
- [21] 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.
- [22] Carr, P. and Wu, L. 2009. Variance Risk Premiums, *Review of Financial Studies*, 22: 1311-1341.
- [23] Carter, C.K. and Kohn, R. 1994. On Gibbs Sampling for State Space Models, *Biometrika*, 81: 541-553.

- [24] Chernov, M. and Ghysels, E. 2000. A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation, *Journal of Financial Economics*, 56: 407-458.
- [25] Christoffersen, P. F. 1998. Evaluating Interval Forecasts. *International Economic Review*, 39: 841-862.
- [26] Cochrane, J.H. 2005. *Asset Pricing*, Second Edition, Princeton University Press.
- [27] Corradi, V and Swanson, N. 2006. Predictive Density and Conditional Confidence Interval Accuracy Tests, *Journal of Econometrics*, 135: 187–228.
- [28] Corsi, F. 2004. A Simple Long Memory Model of Realized Volatility, *Working Paper, University of Southern Switzerland*.
- [29] Cox, J.C., J.E. Ingersoll and Ross, S.A. 1985. A Theory of the Term Structure of Interest Rates, *Econometrica*, 53: 385-408.
- [30] Duan, J.C. and Yeh, C.Y. 2010. Jump and Volatility Risk Premiums Implied by VIX, *Journal of Economic Dynamics and Control*, 34: 2232-2244.
- [31] Duffie, D., Pan J. and Singleton, K. 2000. Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica*, 68: 1343-1376.
- [32] Engle, R.F. and Russell, J.R. 1998. Autoregressive Conditional Duration: A New Approach for Irregularly Spaced Transaction Data, *Econometrica*, 66: 987-1007.
- [33] Eraker, B. 2004. Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. *The Journal of Finance*, LIX: 1367-1403.
- [34] Eraker, B. 2008. The Volatility Premium, *Working Paper*, available for download from <http://www.marginalq.com/eraker/>.
- [35] Eraker, B., Johannes, M. and Polson, N. 2003. The Impact of Jumps in Volatility and Returns, *The Journal of Finance*, LVIII: 1269-1300.
- [36] Feigen, P.D., Gould, P., Martin, G.M. and Snyder, R.D. 2008. Feasible Parameter Regions for Alternative Discrete State Space Models, *Statistics and Probability Letters*, 78: 2963-2970.
- [37] Forbes C.S., Martin, G.M. and Wright J. 2007. Inference for a Class of Stochastic Volatility Models Using Option and Spot Prices: Application of a Bivariate Kalman

- Filter, *Econometric Reviews, Special Issue on Bayesian Dynamic Econometrics*, 26: 387-418.
- [38] Frühwirth-Schnatter, S. 1994. Data Augmentation and Dynamic Linear Models, *Journal of Time Series Analysis*, 15: 183-202.
 - [39] Garcia, R., Lewis, M., Pastorello, S. and Renault, E. 2011. Estimation of Objective and Risk-Neutral Distributions Based on Moments of Integrated Volatility, *Journal of Econometrics*, 160: 22-32.
 - [40] Geweke, J. and Amisano, G. 2010. Comparing and Evaluating Bayesian Prediction Distributions of Asset Returns, *International Journal of Forecasting*, 26: 216-230.
 - [41] Gneiting, T. 2008. Editorial: Probabilistic forecasting, *Journal of the Royal Statistical Society (A)*, 171: 319-321.
 - [42] Gneiting, T, Balabdaoui, F and Raftery, A. 2007. Probabilistic Forecasts, Calibration and Sharpness, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69: 243–268.
 - [43] Guo, D. 1998. The Risk Premium of Volatility Implicit in Currency Options, *Journal of Business and Economic Statistics*, 16: 498-507.
 - [44] 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.
 - [45] Jiang, G.J. and Tian, Y.S. 2005. The Model-Free Implied Volatility and its Information Content, *The Review of Financial Studies*, 18: 1305-1342.
 - [46] Johannes, M., Polson, N.G. and Stroud, J.R. 2009. Optimal Filtering of Jump-Diffusions: Extracting Latent States from Asset Prices, *Review of Financial Studies*, 22: 2759-2799.
 - [47] Jones, C. 2003. The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, *Journal of Econometrics*, 116: 181-224.
 - [48] Jung, R.C., Kukuk, M. and Leisenfeld, R. (2006) Time Series of Count Data: Modeling, Estimation and Diagnostics, *Computational Statistics and Data Analysis*, 51: 2350-2364.

- [49] Koopman, S. J., Jungbacker, B. and Hol, E. 2005. Forecasting Daily Variability of the S&P100 Stock Index using Historical, Realized and Implied Volatility Measurements, *Journal of Empirical Finance*, 12: 445-475.
- [50] Martens, M. and Zein, J. 2004. Predicting Financial Volatility: High-Frequency Time Series Forecasts Vis-a-Vis Implied Volatility, *Journal of Futures Markets*, 24: 1005-1028.
- [51] Martin, G.M., Reidy, R. and Wright, J. 2009. Does the option market produce superior forecasts of noise-corrected volatility measures? *Journal of Applied Econometrics*, 24: 77-104.
- [52] McCabe, B.P.M., Martin, G.M. and Harris, D.G. 2010. Efficient Probabilistic Forecasts of Counts. *Journal of the Royal Statistical Society, Series B*, forthcoming.
- [53] Pan, J. 2002. The Jump-risk Premia Implicit in Options: Evidence from an Integrated Time-series Study, *Journal Of Financial Economics*, 63: 3-50.
- [54] Polson, N. G. and Stroud, J.R.. 2003. Bayesian Inference for Derivative Prices, *Bayesian Statistics*, 7: 641-650.
- [55] Pong, S., Shackleton, M.B., Taylor, S.J. and Xu, X. 2004. Forecasting Currency Volatility: a Comparison of Implied Volatilities and AR(FI)MA Models, *Journal of Banking and Finance*, 28: 2541-2563.
- [56] Smith. A.F.M. and Gelfand, A.E. 1992. Bayesian Statistics without Tears: a Sampling Resampling Perspective, *The American Statistician*, 46: 84-88.
- [57] Stroud, J.R., Muller, P. and Polson, N.G. 2003. Nonlinear State-space Models with State-Dependent Variances, *Journal of the American Statistical Association*, 98: 377-386.
- [58] Tauchen G. and Zhou, H. 2011. Realized Jumps on Financial Markets and Predicting Credit Spreads, *Journal of Econometrics*, 160: 102–118
- [59] Todorov, V. 2010. Variance Risk-Premium Dynamics: The Role of Jumps, *Review of Financial Studies*, 23: 345-383.
- [60] Zellner, A. 1986. On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior Distribution. In Goel, P.K. and Zellner, A. (Eds) *Bayesian Inference and Decision Techniques: essays in honour of Bruno de Finetti*, 233-243, Amsterdam, North Holland.

Appendix A: Generation of $V_{1:T}|\phi, MF_{1:T}, BV_{1:T}$. As the state variable, V_t , appears in the error terms of the state equation in (21) and the measurement equations in (19) and (20), a closed form representation of the conditional posterior distribution for the stochastic variance vector, $V_{1:T}$, is not available. In this paper we extend an approach suggested by Stroud *et al.* (2003) and augment the state space model with two mixture indicator vectors, $z_{1:T}^v = (z_1^v, z_2^v, \dots, z_T^v)'$ associated with the latent variance vector, and $z_{1:T}^o = (z_1^o, z_2^o, \dots, z_T^o)'$ associated with the observed bivariate variance measure. Note that, by assumption, the value of V_0 is fixed with $V_0 = \theta + \frac{\mu_v \delta_J}{\kappa}$ and, hence, a mixture indicator variable is not required for V_0 . Each component of a mixture indicator vector takes on the value of an integer from 1 to K , and defines a suitable linearization of the relevant state or observation equation. The mixture indicator vectors are then used to establish a candidate draw, within a MH subchain, given the previously sampled latent variance vector, $V_{1:T}$, conditional upon all other parameters ϕ , jump occasions $\Delta N_{1:T}$ and variance jump sizes $Z_{1:T}^v$. The algorithm essentially constructs an expanded state variable, $\{z_{1:T}^v, z_{1:T}^o, V_{1:T}\}$, and produces MCMC draws from $p(z_{1:T}^v, z_{1:T}^o, V_{1:T} | BV_{1:T}, MF_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \phi)$, so that after discarding the draws of $z_{1:T}^v$ and $z_{1:T}^o$ the latent volatilities are seen as draws from the desired full conditional posterior distribution, $p(V_{1:T} | BV_{1:T}, MF_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \phi)$.

Prior to sampling, a grid of values over the space of plausible stochastic variance values, $\{\bar{\mu}_1, \bar{\mu}_2, \dots, \bar{\mu}_K\}$, along with a set of associated bandwidth parameters, $\{\bar{\sigma}_1, \bar{\sigma}_2, \dots, \bar{\sigma}_K\}$, are established. These values provide prior probability weights for the indicator elements, with

$$p(z_t^v = k | V_{t-1}) = \frac{\bar{\sigma}_k^{-1} \phi\left(\frac{\bar{\mu}_k - V_{t-1}}{\bar{\sigma}_k}\right)}{C(V_{t-1})}, \quad k = 1, 2, \dots, K$$

and

$$p(z_t^o = k | V_t) = \frac{\bar{\sigma}_k^{-1} \phi\left(\frac{\bar{\mu}_k - V_t}{\bar{\sigma}_k}\right)}{C(V_t)}, \quad k = 1, 2, \dots, K,$$

where $\phi(\cdot)$ denotes the probability density function for the standard normal distribution and

$$C(V_t) = \sum_{k=1}^K \bar{\sigma}_k^{-1} \phi\left(\frac{\bar{\mu}_k - V_t}{\bar{\sigma}_k}\right), \quad \text{for all } t = 0, 1, 2, \dots, T.$$

It is possible to generalize to different locations and/or bandwidth grids for each of the state and observation indicators; however, to keep the algorithm as simple as possible we have chosen them to be the same here. In addition, to keep the expressions uncluttered, in what follows the requisite conditioning on $\{BV_{1:T}, MF_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \phi\}$ has been suppressed.

The MH algorithm for the expanded state variable $\{z_{1:T}^v, z_{1:T}^o, V_{1:T}\}$ proceeds as follows:

1. Sample $z_{1:T}^v$ given $\{z_{1:T}^o, V_{1:T}\}$ from T independent multinomial distributions, with

$$p^a(z_t^v = k | z_{1:T}^o, V_{1:T}) = \frac{(\sigma_v \bar{\sigma}_k \sqrt{\bar{\mu}_k})^{-1} \phi\left(\frac{V_t - E[V_t | V_{t-1}]}{\sigma_v \sqrt{\bar{\mu}_k}}\right) \phi\left(\frac{\bar{\mu}_k - V_{t-1}}{\bar{\sigma}_k}\right)}{\sum_{j=1}^K (\sigma_v \bar{\sigma}_j \sqrt{\bar{\mu}_j})^{-1} \phi\left(\frac{V_t - E[V_t | V_{t-1}]}{\sigma_v \sqrt{\bar{\mu}_j}}\right) \phi\left(\frac{\bar{\mu}_j - V_{t-1}}{\bar{\sigma}_j}\right)},$$

for $t = 1, 2, \dots, T$, where $E[V_t | V_{t-1}] = \kappa\theta + (1 - \kappa)V_{t-1} + Z_t^v \Delta N_t$.

2. Sample $z_{1:T}^o$ given $\{z_{1:T}^v, V_{1:T}\}$ from T independent multinomial distributions, with

$$p^a(z_t^o = k | z_{1:T}^v, V_{1:T}) = \frac{(\sigma_{BV} \sigma_{MF} \bar{\sigma}_k \bar{\mu}_k^2)^{-1} \phi\left(\frac{BV_t - V_t}{\sigma_{BV} \bar{\mu}_k}\right) \phi\left(\frac{MF_t - E^*(\mathcal{V}_{t,t+\tau} | \mathcal{F}_t)}{\sigma_{MF} \bar{\mu}_k}\right) \phi\left(\frac{\bar{\mu}_k - V_t}{\bar{\sigma}_k}\right)}{\sum_{j=1}^K (\sigma_{BV} \sigma_{MF} \bar{\sigma}_j \bar{\mu}_j^2)^{-1} \phi\left(\frac{BV_t - V_t}{\sigma_{BV} \bar{\mu}_j}\right) \phi\left(\frac{MF_t - E^*(\mathcal{V}_{t,t+\tau} | \mathcal{F}_t)}{\sigma_{MF} \bar{\mu}_j}\right) \phi\left(\frac{\bar{\mu}_j - V_t}{\bar{\sigma}_j}\right)},$$

for $t = 1, 2, \dots, T$, where $E^*(\mathcal{V}_{t,t+\tau} | \mathcal{F}_t) = a_\tau^* V_t + b_\tau^* + \tau[\mu_v + \lambda_{Jt}] \delta_J$.

3. Sample $\tilde{V}_{1:T}$ given $\{z_{1:T}^v, z_{1:T}^o\}$ using an FFBS sampling algorithm and a candidate state space model designed to have a smoothed state distribution that closely approximates the smoothed state distribution of the augmented model, given $\{z_{1:T}^v, z_{1:T}^o\}$. The candidate state space model for day $t = 1, 2, \dots, T - 1$ is given by the four dimensional measurement equation

$$\begin{pmatrix} BV_t \\ MF_t \\ \bar{\mu}_{z_t^o} \\ \bar{\mu}_{z_{t+1}^v} \end{pmatrix} = \begin{bmatrix} 0 \\ b_\tau^* + \tau[\mu_v + \lambda_{Jt}] \delta_J \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ a_\tau^* \\ 1 \\ 1 \end{bmatrix} V_t + \epsilon_t,$$

$$\epsilon_t \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_{BV}^2 \bar{\mu}_{z_t^o}^2 & 0 & 0 & 0 \\ 0 & \sigma_{MF}^2 \bar{\mu}_{z_t^o}^2 & 0 & 0 \\ 0 & 0 & \bar{\sigma}_{z_t^o}^2 & 0 \\ 0 & 0 & 0 & \bar{\sigma}_{z_{t+1}^v}^2 \end{bmatrix} \right),$$

and, for day $t = T$, the trivariate measurement equation

$$\begin{pmatrix} BV_T \\ MF_T \\ \bar{\mu}_{z_T^o} \end{pmatrix} = \begin{bmatrix} 0 \\ b_\tau^* + \tau[\mu_v + \lambda_{JT}] \delta_J \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ a_\tau^* \\ 1 \end{bmatrix} V_T + \epsilon_T,$$

$$\epsilon_T \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_{BV}^2 \bar{\mu}_{z_T^o}^2 & 0 & 0 \\ 0 & \sigma_{MF}^2 \bar{\mu}_{z_T^o}^2 & 0 \\ 0 & 0 & \bar{\sigma}_{z_T^o}^2 \end{bmatrix} \right).$$

The candidate state space model uses the scalar state equation

$$V_t = (1 - \kappa) V_{t-1} + (\kappa\theta + Z_t^v \Delta N_t) + \omega_t, \quad \omega_t \sim N(0, \sigma_v^2 \bar{\mu}_{z_t^v}),$$

for all $t = 1, 2, \dots, T$, with $V_0 = \theta + \frac{\mu_v \delta_J}{\kappa}$.

4. Accept $\tilde{V}_{1:T}$ with probability determined by the MH acceptance rule

$$\min \left\{ 1, \prod_{t=1}^T \left[\frac{C(V_{t-1})p^a(V_t|V_{t-1})p(\tilde{V}_t|\tilde{V}_{t-1})C(V_t)p^a(BV_t, MF_t|V_t)p(BV_t, MF_t|\tilde{V}_t)}{C(\tilde{V}_{t-1})p^a(\tilde{V}_t|\tilde{V}_{t-1})p(V_t|V_{t-1})C(\tilde{V}_t)p^a(BV_t, MF_t|\tilde{V}_t)p(BV_t, MF_t|V_t)} \right] \right\}$$

where $V_{1:T}$ denotes the previous draw in the Markov chain,

$$p^a(V_t|V_{t-1}) = C(V_{t-1})^{-1} \sum_{k=1}^K \left[\left(\bar{\sigma}_k \sigma_v \sqrt{\bar{\mu}_k} \right)^{-1} \times \phi \left(\frac{V_t - (\kappa\theta + Z_t^v \Delta N_t) - (1 - \kappa) V_{t-1}}{\sigma_v \sqrt{\bar{\mu}_k}} \right) \phi \left(\frac{\bar{\mu}_k - V_{t-1}}{\bar{\sigma}_k} \right) \right],$$

and

$$p(V_t|V_{t-1}) = \left(\sigma_v \sqrt{V_{t-1}} \right)^{-1} \phi \left(\frac{V_t - (\kappa\theta + Z_t^v \Delta N_t) - (1 - \kappa) V_{t-1}}{\sigma_v \sqrt{V_{t-1}}} \right),$$

and similarly,

$$p^a(BV_t, MF_t|V_t) = C(V_t)^{-1} \sum_{j=1}^K \left(\sigma_{BV} \sigma_{MF} \bar{\sigma}_j \bar{\mu}_j^2 \right)^{-1} \phi \left(\frac{BV_t - V_t}{\sigma_{BV} \bar{\mu}_j} \right) \times \phi \left(\frac{MF_t - E^*(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t)}{\sigma_{MF} \bar{\mu}_j} \right) \phi \left(\frac{\bar{\mu}_j - V_t}{\bar{\sigma}_j} \right),$$

and

$$p(BV_t, MF_t|V_t) = \left(\sigma_{BV} \sigma_{MF} V_t^2 \right)^{-1} \phi \left(\frac{BV_t - V_t}{\sigma_{BV} V_t} \right) \phi \left(\frac{MF_t - E^*(\mathcal{V}_{t,t+\tau}|\mathcal{F}_t)}{\sigma_{MF} V_t} \right).$$

5. Discard $z_{1:T}^v$ and $z_{1:T}^o$.

Appendix B: Sampling/Importance Resampling from the Augmented State Space

Model Given the specification for returns in (35), the joint posterior density for *all* unknowns, conditional on the three vectors of observed data, $r_{1:T}$, $MF_{1:T}$, and $BV_{1:T}$, can be decomposed as follows

$$\begin{aligned} & p(\mu, \beta_v, \rho, \mu_p, \rho_J, \sigma_r, V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \phi | r_{1:T}, MF_{1:T}, BV_{1:T}) \\ \propto & p(\mu, \beta_v, \rho, \mu_p, \rho_J, \sigma_r, |r_{1:T}, V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v) \\ & \times p(r_{1:T} | V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v) \\ & \times p(V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \phi | MF_{1:T}, BV_{1:T}). \end{aligned} \quad (37)$$

In order to sample from (37), we first draw from $p(V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \phi | MF_{1:T}, BV_{1:T})$ using the MCMC algorithm for the volatility model as described in Section 4. Then, given the

vector of observed returns, $r_{1:T}$, and using the draws of $(V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v)$, a draw is produced from the *candidate* density,

$$q(\mu, \beta_v, \rho, \mu_p, \rho_J, \sigma_r, V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \phi) \propto p(V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \phi | MF_{1:T}, BV_{1:T}) \\ \times p(\mu, \beta_v, \rho, \mu_p, \rho_J, \sigma_r | r_{1:T}, V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v). \quad (38)$$

Under a standard Gaussian-inverse gamma g -prior framework (Zellner, 1986), the return-specific parameters, $\mu, \beta_v, \rho, \mu_p, \rho_J$, and σ_r , are sampled from the resulting Gaussian-inverse gamma posterior, conditional upon the draws of the latent variables, $V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v$, and the static parameters contained in ϕ , specifically θ, κ and σ_v . Given that, up to a constant of proportionality, the candidate density in (38) differs from the target density (37) only by the factor $p(r_{1:T} | V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v)$, a feasible sampling/importance resampling (SIR) algorithm (Smith and Gelfand, 1992) to correct the initial draws is implemented, so that a sample from the joint posterior distribution in (37) is obtained. Crucially, given the assumed structure of (35), including the assumption of normality for the innovations, and the use of a conjugate prior structure, $p(r_{1:T} | V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v)$ is known in closed form. Hence, given M available draws of $(V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v)$ from the MCMC algorithm, the draws are resampled from the discrete distribution defined by the weights

$$w^{(i)} = \frac{p(r_{1:T} | V_{1:T}^{(i)}, \Delta N_{1:T}^{(i)}, Z_{1:T}^{v(i)}, \theta^{(i)}, \kappa^{(i)}, \sigma_v^{(i)})}{\sum_{k=1}^M p(r_{1:T} | V_{1:T}^{(k)}, \Delta N_{1:T}^{(k)}, Z_{1:T}^{v(k)}, \theta^{(k)}, \kappa^{(k)}, \sigma_v^{(k)})}, \quad i = 1, 2, \dots, M.$$

The reweighted draws of $V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v$ are used to produce draws from $p(\mu, \beta_v, \rho, \mu_p, \rho_J, \sigma_r | r_{1:T}, V_{1:T}, \Delta N_{1:T}, Z_{1:T}^v, \theta, \kappa, \sigma_v)$. Each draw of the full set of unknowns is then used to produce the conditional (Gaussian) predictive distribution for r_{T+1} . The (marginal) predictive, $p(r_{T+1} | r_{1:T}, MF_{1:T}, BV_{1:T})$ is produced by averaging the conditional predictives in the usual way.

Appendix C: Transformation to Risk Aversion The equilibrium frameworks of Breen (1979) and Cox, Ingersoll and Ross (1985) lead to factor risk premiums that are equal to the negative of the covariance between changes in the factor and the rate of change in the marginal utility of wealth. For the SVCJ model adopted here, this implies that

$$(\lambda_D V_t - \delta_J \lambda_J) dt = -cov\left(\frac{dw_t}{w_t}, dV_t\right), \quad (39)$$

where $dV_t = dV_t^d + dJ_t^v$, with $dV_t^d = \kappa[\theta - V_t]dt + \sigma_v \sqrt{V_t} dB_t^v$ representing the diffusive component of the volatility process. Thus, the variance risk premium can be decomposed

into the diffusive variance risk premium

$$\lambda_D V_t dt = -\text{cov} \left(\frac{dw_t}{w_t}, dV_t^d \right) \quad (40)$$

and the jump variance risk premium,

$$-\delta_J \lambda_J dt = -\text{cov} \left(\frac{dw_t}{w_t}, dJ_t^v \right), \quad (41)$$

where w_t denotes the marginal utility of wealth for the representative investor. Following Bollerslev *et al.* (2011), we adopt the canonical power utility function,

$$U_t = e^{-\beta t} \frac{W_t^{1-\gamma}}{1-\gamma}, \quad \gamma > 0,$$

where β denotes the constant subjective discount rate and γ is the risk aversion parameter, from which it follows that,

$$w_t = e^{-\beta t} W_t^{-\gamma}.$$

Proxying wealth by the value of the market portfolio and recognizing that, in our empirical setting, the price process in (1) refers to a market stock index, Ito's lemma yields

$$\begin{aligned} dw_t &= \frac{\partial w_t}{\partial P_t} dP_t + \frac{\partial w_t}{\partial t} dt + \frac{1}{2} \frac{\partial^2 w_t}{\partial P_t^2} (dP_t)^2 \\ &= (-\gamma P_t^{-\gamma-1} e^{-\delta t}) dP_t + (\gamma(1+\gamma) P_t^{-\gamma-2} e^{-\delta t}) (dP_t)^2 \\ &= -\gamma P_t^{-1} w_t dP_t + \gamma(1+\gamma) P_t^{-2} w_t (dP_t)^2, \end{aligned}$$

that is,

$$\frac{dw_t}{w_t} = -\gamma P_t^{-1} dP_t + \gamma(1+\gamma) P_t^{-2} (dP_t)^2. \quad (42)$$

Substituting (42) into (40) gives the following relation

$$\begin{aligned} \lambda_D V_t dt &= -\text{cov} \left(-\gamma P_t^{-1} dP_t + \gamma(1+\gamma) P_t^{-2} (dP_t)^2, dV_t^d \right) \\ &= \gamma P_t^{-1} \text{cov} (dP_t, dV_t^d) - \gamma(1+\gamma) P_t^{-2} \text{cov} ((dP_t)^2, dV_t^d) \\ &= \gamma P_t^{-1} \text{cov} (dP_t, dV_t^d), \text{ since } \text{cov} ((dP_t)^2, dV_t^d) = 0, \end{aligned}$$

which implies that

$$\lambda_D = \gamma \rho \sigma_v. \quad (43)$$

Similarly, substituting (42) into (41) yields

$$\begin{aligned} \delta_J \lambda_J dt &= \text{cov} \left(-\gamma P_t^{-1} dP_t + \gamma(1+\gamma) P_t^{-2} (dP_t)^2, dJ_t^v \right) \\ &= -\gamma P_t^{-1} \text{cov} (dP_t, dJ_t^v), \text{ since } \text{cov} ((dP_t)^2, dJ_t^v) = 0, \end{aligned}$$

which implies that

$$\delta_J \lambda_J = -\gamma \delta_J \rho_J \mu_v^2. \quad (44)$$

Finally, substituting (43) and (44) into (39) and solving for γ , we obtain

$$\gamma = \frac{\lambda_D V_t - \delta_J \lambda_J}{\rho \sigma_v V_t + \rho_J \delta_J \mu_v^2}.$$

We are, of course, adopting a dynamic model for λ_J , which, combined with dynamics of the stochastic variance V_t , implies a dynamic model for γ .