Alternative Parametric Models for Spot Volatility in High Frequency: A Bayesian Approach *

Yaohan Chen [a], Jia Li [b], and Jun Yu [c]

[a] School of Big Data and Statistics, Anhui University
[b] School of Economics, Singapore Management University
[c] Department of Finance and Business Economics, University of Macau

Abstract

This paper proposes several alternative parametric models for spot volatility in high frequency, depending on whether or not jumps, seasonality, and announcement effects are included. Together with these alternative parametric models, nonlinear non-Gaussian state-space models are introduced based on the fixed-k estimator of spot volatility of Bollerslev, Li, and Liao (2021). According to Bollerslev, Li, and Liao (2021), the log fixed-k estimator of spot volatility equals the true log spot volatility plus a non-Gaussian random variable. Bayesian methods are introduced to estimate and compare these alternative models and to extract volatility from the estimated models. Simulation studies suggest that the Bayesian methods can accurately estimate the parameters, select the true model, and extract volatility. Empirical studies using high-frequency market index and individual stock prices reveal several important results. As an application of extracting volatility, we quantify the strategic value of information.

Keywords: Spot Volatility; High-frequency econometrics; Nonlinear non-Gaussian state-space model; MCMC; Strategic Trading
JEL Classification: C11, C32, C58, G12, G14

* Yaohan Chen, School of Big Data and Statistics, Anhui University, yaohan.chen@ahu.edu.cn. Jia Li, School of Economics, Singapore Management University, jiali@smu.edu.sg. Jun Yu, Department of Finance and Business Economics, University of Macau, junyu@um.edu.mo.

1 Introduction

Financial market volatility as the measure of risk plays a vital role both in finance theory and applications of asset pricing theory in practice (Engle, 2004). Acknowledging the fact that daily volatilities are time-varying, a strand of literature focuses on modeling daily volatility parametrically based on daily returns. Examples include the ARCH model of Engle (1982), the GARCH model of Bollerslev (1986), and the stochastic volatility model of Taylor (1982). As a by-product of volatility modeling, an estimate of daily volatility can be obtained after the model is estimated.

In a more recent strand of literature, daily realized volatilities (RVs) are used to estimate daily integrated volatility (IV). Daily RV is a nonparametrical method that is based on intraday returns, usually 5-minute returns; see Andersen and Bollerslev (1997) and Andersen, Bollerslev, Christoffersen, and Diebold (2013). By exploiting intraday information, 5-minute returns can estimate daily volatility more accurately than daily returns. Subsequently, considerable efforts have been made to find a reasonable model for daily RV, which is then used to forecast future daily RV: see Andersen, Bollerslev, Diebold, and Ebens (2001); Andersen, Bollerslev, Diebold, and Labys (2001, 2003); Gatheral, Jaisson, and Rosenbaum (2018); Wang, Xiao, and Yu (2022). Other than providing a more accurate estimate to IV, RV has been found a wide range of applications. For example, in an interesting paper, Bollerslev and Zhou (2002) use RVs, obtained from 5-minute returns, to construct GMM estimators for parameters in several parametric diffusion models.

However, most parametric models for daily volatilities (either RV or spot volatility) are not suitable for modeling spot volatilities in high frequencies. This is not surprising as spot volatilities in high frequencies have more complicated behavior than a standard parametric diffusion model can generate.

Based on 5-minute returns on S&P500 index futures from March 11, 2007, through March 9, 2012, Stroud and Johannes (2014) propose a high-frequency model where the total volatility has a multiplicative specification, including traditional autoregressive stochastic volatility components, seasonal components, and announcement components. They introduce a Bayesian method to estimate parameters in the model and find that all three components are important in the model.

The attempt to build a high-frequency model is important to enhance our understanding of the intraday behavior in volatility. It has potential important implications for asset pricing, volatility forecasting, trading, and risk management. However, the criticism to the use of daily returns as opposed to daily RV also applies here. That is, the use of 5-minute returns is less efficient than that based on returns in a higher frequency.

Apart from the literature where the quantity of interest is the daily RV, there is another strand of growing literature that tries to estimate spot volatility from ultra-high frequency data. In comparison to using RV as the proxy of volatility over a specific time interval, estimation of spot volatility plays the role of measuring return variations over a certain time point, expressed per unit of time (Andersen, Bollerslev, and Diebold, 2010). In the relatively earlier literature, Foster and Nelson (1996) firstly considered nonparametric estimation of spot volatility in the diffusion models.

Kristensen (2010) proposes a method using a kernel-weighted IV estimator as the estimation of spot volatility, provided that the bandwidth of the kernel vanishes. Later in Zu and Peter Boswijk (2014), the authors propose a spot volatility estimator based on the Two Scaled Realized Variance (TSRV) estimator in (see Zhang, Mykland, and Aït-Sahalia, 2005; Mykland and Zhang, 2008) and establish the corresponding asymptotic theory. In a recent study, Bollerslev, Li, and Liao (2021) establish a new theory for the conduct of nonparametric inference about the latent spot volatility. Unlike the theories that assume the number of observations in local estimation blocks go to infinite, the new theory treats the estimation block size k as fixed. As a result, the estimation error in the spot volatility estimator can be characterized by a scaled chi-square random variable. Bollerslev, Li, and Liao (2021) carry out an empirical application based on the intraday S&P 500 equity index. One of the important empirical results suggests that there exist jumps at FOMC news announcement times. However, this study makes no attempt to model the dynamics of spot volatility in high frequencies.

In this paper we propose several high-frequency models for the spot volatility based on the theory of Bollerslev, Li, and Liao (2021). All alternative specifications can be expressed as a nonlinear non-Gaussian state-space model. In particular, in all the alternative models, the observation equation, where the fixed-k estimator of spot volatility and the true spot volatility are linked, comes directly from the theory of Bollerslev, Li, and Liao (2021). The difference of the alternative models lies in how the dynamics in the latent spot volatility is specified.

We then conduct the Bayesian analysis of all the alternative models using Markov chain Monte Carlo (MCMC), including obtaining the posterior distributions for each parameter and each latent spot volatility. In particular, the posterior mean of latent spot volatility is the smoothed estimate of spot volatility. In addition, we make a Bayesian model comparison of alternative specifications via the Deviance Information Criterion (DIC).

The rest of the paper is organized as follows. In Section 2, relevant preliminary mathematical concepts are introduced. In Section 3, the fixed-k theory of Bollerslev, Li, and Liao (2021) is reviewed and used to motivate our modelling strategy. Section 4 introduces alternative volatility models. We also discuss the Bayesian methods for parameter estimation method, volatility extraction, and model comparison method. In Section 5, Monte Carlo experiments are designed to demonstrate that our proposed Bayesian methods in general work well. Section 6 contains empirical studies. Finally, Section 7 concludes this paper and briefly discusses the agenda for future work. More comprehensive discussions of technical details about MCMC and other additional results are contained in the appendix.

2 Mathematical Foundation

Before we introduce our high-frequency volatility models, we first clarify some relevant mathematical notations and related concepts. All random variables are defined on a fixed (complete) probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t\geq 0}$. Following Andersen, Bollerslev, Diebold, and Labys (2001), we adopt the assumption that logarithmic asset prices follow a univariate diffusion. In particular,

for the asset indexed by i, the logarithmic return is modeled as

$$p_i(t) - p_i(t-1) \equiv r_i(t) = \int_{t-1}^{t} \mu_i(s)ds + \int_{t-1}^{t} \sigma_i(s)dW(s), \tag{1}$$

where $W(s)$ stands for the standard Wiener process and hence, the corresponding volatility measure is based on the quadratic variation process, denoted by $\text{Qvar}_i(t)$, which yields

$$\text{Qvar}_i(t) = [p_i, p_i]_t - [p_i, p_i]_{t-1} = \int_{t-1}^{t} \sigma_i^2(s)ds. \tag{2}$$

This is commonly referred to the integrated volatility in the literature. It is broadly known in literature (see Andersen, Bollerslev, Diebold, and Labys, 2001; Barndorff-Nielsen and Shephard, 2002) that the integrated volatility over non-trivial time interval, such as a day, is an important quantity of interest in finance. Many nonparametric estimators for daily IV have been proposed. Arguably, the most widely used estimator is the daily RV based on 5-minute returns.

With the increasing availability of data sampled at ultra high frequencies, how to estimate the spot volatility, that is $\sigma_i^2(t)$, has drawn a growing interest in the literature. Based on the mathematical foundation just laid out, we focus on the following model for (log) price process as in the literature,

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t. \tag{3}$$

This is a continuous-time Itô semimartingale process with drift, diffusion and jump. Besides, for two random sequences a_n and b_n, we write $a_n \asymp b_n$, if $a_n/C \le b_n \le Ca_n$ for some finite constant $C \ge 1$.

3 Fixed-k Estimator of Spot Volatility

When the logarithmic price of an asset is characterized by the continuous-time Itô semimartingale process, Jacod, Li, and Liao (2021) suggest a way to estimate "spot covariance" in the general multivariate setting, $c_k(t) = \sigma_k(t)\sigma_k(t)^\top$, nonparametrically and uniformly as follows

$$\hat{c}_{n,j} \equiv \frac{1}{k_{n,j}\Delta_n} \sum_{i \in \mathcal{I}_{n,j}} \Delta_i^n X \Delta_i^n X^\top \mathbf{1}_{\{\|\Delta_i^n X\| \le u_n\}} \tag{4}$$

where

$$\begin{aligned}
\Delta_n &: T/n \\
\Delta_i^n X &: X_{i\Delta_n} - X_{(i-1)\Delta_n} \\
u_n &: \text{truncation threshold satisfying } u_n \asymp \Delta_n^{\varpi},\ \varpi \in (0, 1/2). \\
\mathcal{I}_{n,j} &: \text{set collecting indices of consecutive increments in } j\text{-th block,} \\
&\quad \text{such that } \{1, \ldots, n\} = \bigcup_{j=1}^{m_n} \mathcal{I}_{n,j} \text{ and } |\mathcal{I}_{n,j}| = k_{n,j}. \\
\mathcal{T}_{n,j} &: \text{Correspondingly, } [0,T] \text{ can be dissected as } [0,T] = \bigcup_{j=1}^{m_n} \mathcal{T}_{n,j} \\
&\quad t(n,j) \equiv (\min \mathcal{I}_{n,j} - 1)\,\Delta_n,
\end{aligned}$$

and

$$\mathcal{T}_{n,j} \equiv \begin{cases} [t(n,j), t(n,j+1)) & \text{if } 1 \le j < m_n \\ [t(n,m_n), T] & \text{if } \quad j = m_n. \end{cases}$$

$k_{n,j}$ denotes the block size. The issue of whether $k_{n,j}$ should be fixed or not is discussed in Bollerslev, Li, and Liao (2021). While it is commonly assumed $k_{n,j} \to \infty$, Bollerslev, Li, and Liao (2021) advocate a way of making inference for spot volatility with $k_{n,j} = k$ fixed. To see the link between the setting of Bollerslev, Li, and Liao (2021, henceforth Bollerslev, Li, and Liao (2021)) and ours, note that BLL2021QE set

$$\begin{aligned} \mathcal{I}_{n,j} &\equiv \{(j-1)k+1, \ldots, jk\} \\ \mathcal{T}_{n,j} &\equiv [(j-1)k\Delta_n, jk\Delta_n)\,. \end{aligned}$$

This setting is a special case of ours with $k_{n,j} = k$.

Essentially $(\hat{c}_{n,j})_{1\le j\le m_n}$ serves as the functional estimator of $(c_t)_{t\in[0,T]}$. More specifically, $(\hat{c}_{n,j})_{1\le j\le m_n}$ is identified with t-indexed functional estimator $(\hat{c}_{n,t})_{t\in[0,T]}$ such that

$$\hat{c}_{n,t} \equiv \hat{c}_{n,j} \quad \text{for} \quad t \in \mathcal{T}_{n,j} \quad \textit{and} \quad j \in \{1, \ldots, m_n\}. \tag{5}$$

The following theorem, which comes from Jacod, Li, and Liao (2021), develops the properties of the estimator when $k_{n,j} \to \infty$.

Theorem 3.1 (Jacod, Li, and Liao (2021)) *Under the Assumption 1 and Assumption 2 imposed in Jacod, Li, and Liao (2021) and $k_{n,j} \asymp \Delta_n^{-\rho}$ uniformly for all $j \in \{1, \ldots, m_n\}$ and $u_n \asymp \Delta_n^{\varpi}$ with $r \in (0, 1/2)$, $\rho \in (r, 1/2)$, and $\varpi \in ((1-\rho/2)/(2-r), 1/2)$. The following statements hold for some constant $\epsilon > 0$.*

(a) With

$$U_{n,j} \equiv k_{n,j}^{-1/2} \sum_{i\in\mathcal{I}_{n,j}} \left(\Delta_i^n W \Delta_i^n W^\top / \Delta_n - \mathbf{I}_d\right)$$

for each $1 \le j \le m_n$, we have

$$\max_{1\le j\le m_n} \sup_{t\in\mathcal{T}_{n,j}} \left\| k_{n,j}^{1/2} \left(\hat{c}_{n,t} - c_t\right) - \sigma_{t(n,j)} U_{n,j} \sigma_{t(n,j)}^\top \right\| = o_p\left(\Delta_n^\epsilon\right) \tag{6}$$

(b) If Assumption 2 holds, the following approximation result holds uniformly

$$\max_{1\le j\le m_n} \sup_{t\in\mathcal{T}_{n,j}} \left| k_{n,j}^{1/2} \left(f(\hat{c}_{n,j}) - f(c_t)\right) - \mathrm{tr}[\partial f(c_{t(n,j)}) \sigma_{t(n,j)} U_{n,j} \sigma_{t(n,j)}^\top] \right| = o_p\left(\Delta_n^\epsilon\right). \tag{7}$$

According to Theorem 9.3.2 of Jacod and Protter (2012), $k_{n,j} \to \infty$ and $k_{n,j}\Delta_n \to 0$ are needed to ensure the consistency of $\hat{c}_{n,t}$. The required conditions for the consistency is intuitive as they require

the local estimation block contain an increasing number of observations (i.e. $k_{n,j} \to \infty$), while at the same time the size of local estimation block shrinks to zero asymptotically (i.e. $k_{n,j}\Delta_n \to 0$).

Although this double asymptotic scheme theoretically justifies the consistency of the nonparametric estimation of the spot volatility, it requires a carefully chosen tuning sequence $k_{n,j}$, as manifest in the above theorem. The fixed-k theory established in BLL2021QE alleviates the concern about mimicking the double-asymptotic scheme. We are now in a position to review the fixed-k theory of BLL2021QE.

3.1 Fixed-k inference for volatility

BLL2021QE suggests a way to nonparametrically infer latent spot volatility of asset prices characterized by continuous-time Itô semimartingale process. The main contribution of BLL2021QE lies in that the estimation block size k is fixed.

By setting the estimation block size k fixed, the resulting spot volatility estimator of BLL2021QE is not consistent, but easy-to-calculate pointwise confidence intervals are available at any given point in time. In the univariate case, $c_t = \sigma_t^2$ is estimated by $\hat{c}_{n,j}$ where, for $t \in \mathcal{T}_{n,j}$ and $j \in \{1, \ldots, m_n\}$,

$$\hat{c}_{n,t} \equiv \hat{c}_{n,j}. \tag{8}$$

The only difference between (8) and (5) is that in (8) a fixed block size $k_{n,j} = k$ is used. Thus,

$$\hat{c}_{n,t} \equiv \hat{c}_{n,j} = \frac{1}{k\Delta_n} \sum_{i \in \mathcal{I}_{n,j}} (\Delta_i^n X)^2 \mathbf{1}_{\{|\Delta_i^n X| \le u_n\}}. \tag{9}$$

The following main theorem comes from BLL2021QE.

Theorem 3.2 (Bollerslev, Li, and Liao (2021)) *Suppose that the* Assumption 1 *imposed in BLL2021QE holds, then for any finite subset $\mathcal{M} \subseteq \{1, \ldots, m_n\}$, there exists a collection of independent random variables $\left(\bar{S}_j\right)_{j \in \mathcal{M}}$ such that for any $j \in \mathcal{M}$ and $t \in \mathcal{T}_{n,j}$,*

$$\frac{\hat{c}_{n,t}}{c_t} - \bar{S}_j = O_p\left(\Delta_n^{(2-r)\varpi \wedge (1/2)}\right) = o_p(1), \tag{10}$$

where

$$\bar{S}_j = (k\Delta_n)^{-1} \sum_{i \in \mathcal{I}_{n,j}} \left(W_{i\Delta_n} - W_{(i-1)\Delta_n}\right)^2,$$

is a $\bar{\chi}_k^2$-distributed random variable. The $\bar{\chi}_k^2$ refers to the scaled chi-square distribution such that

$$\bar{\chi}_k^2 \equiv Z_k/k, \text{ with } Z_k \sim \chi_k^2. \tag{11}$$

In companion with this definition, we have the scaled inverse-chi-square distribution

$$\bar{\chi}_k^{-2} \equiv k/Z_k, \text{ with } Z_k \sim \chi_k^2. \tag{12}$$

With k fixed, (10) suggests that $\frac{\hat{c}_{n,t}}{c_t}$ can be approximated by a scaled chi-square distributed random variable, that is,

$$\frac{\hat{c}_{n,t}}{c_t} \xrightarrow{d} \bar{\chi}^2_k \tag{13}$$

as $n \to \infty$ (hence $\Delta_n \to 0$). Taking the log transformation, we have

$$\ln \hat{c}_{n,t} - \ln c_t = \ln \bar{\chi}^2_k = \ln Z_k - \ln k. \tag{14}$$

Definition 1 (Log chi-square distribution) *Let $\mathcal{Z}_k$ denotes the log chi-square distribution, that is*

$$\mathcal{Z}_k = \ln Z_k = \ln \chi^2_k.$$

By Lee (2012, page 379), we have the following results for the log chi-square distribution.[1]

- For the density function, we have [2]

$$p(\mathcal{Z}) = \frac{1}{2^{k/2}\Gamma(k/2)} \exp\left\{\frac{1}{2}k\mathcal{Z} - \frac{1}{2}\exp(\mathcal{Z})\right\} \qquad (-\infty < \mathcal{Z} < \infty).$$

- For the moment generating function, we have

$$\mathrm{MGF(t)} = 2^t \Gamma(t + (k/2)) / \Gamma(k/2).$$

- For the mean and variance, we have

$$\begin{aligned} \mathbb{E}[\mathcal{Z}] &= \ln 2 + \psi(k/2), \\ \mathbb{V}[\mathcal{Z}] &= \psi'(k/2), \end{aligned}$$

where

$$\psi(z) = \frac{\mathrm{d}}{\mathrm{d}z}\log\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}.$$

Remark 3.1 *It is well-known that $\Gamma(z+1) = z\Gamma(z)$. If we differentiate both sides of the equation, we have $\Gamma'(z+1) = \Gamma(z) + z\Gamma'(z)$. If we then divide both sides of the equation by z and substitute $\Gamma(z)$ by $\Gamma(z+1)/z$, we have*

$$\psi(z+1) = \frac{1}{z} + \psi(z).$$

---

[1] For notational simplicity, we suppress the degrees-of-freedom parameter k in the expressions.

[2] It is easy to show that

$$\log p(\mathcal{Z}) = -\frac{k}{2}\log 2 - \log\Gamma(k/2) + \frac{1}{2}k\mathcal{Z} - \frac{1}{2}\exp(\mathcal{Z}).$$

This formula plays an important role in our acceptance-rejection sampling algorithm in the context of the algorithm of Kim, Shephard, and Chib (1998).

This formula suggests that the variance of log chi-square distribution decreases as k increases. In particular, it can be shown that

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\log\Gamma(z) = \frac{\mathrm{d}}{z}\psi(z) = \psi'(z) = \sum_{j=0}^{\infty}\frac{1}{(z+j)^2}.$$

This result has the implication for the choice of the fixed local estimation block size (i.e. k). The larger the local estimation block size is, the less the variance of the estimated spot volatility. This is consistent with the usual "bias-variance" trade-off (see Jacod, Li, and Liao, 2021; Bollerslev, Li, and Liao, 2021). According to the trade-off, a small k results in a more noisy but a less biased nonparametric estimate of the spot volatility. To make this "bias-variance" trade-off more formally, Bollerslev, Li, and Liao (2021) show that

$$\left\|\hat{c}_{n,j} - c_{(j-1)k\Delta_n}\right\|_2 \leq K\left(k^{-1/2} + (k\Delta_n)^{\kappa}\right), \tag{15}$$

where $\hat{c}_{n,j}$ is the nonparametric estimation of spot volatility over the jth block of length $k\Delta_n$, $\|\cdot\|_2$ is the standard L^2-norm, and κ is the "smoothness" parameter of the volatility process. (15) directly implies the "bias-variance" trade-off. In the following discussion, we mostly focus on the case where the local estimation window size is fixed at $k = 5$ and the price data is sampled at the frequency of every one minute within one day. Thus, $\Delta_n = 1/390 \approx 0.002$.

Although unobserved spot volatility is indexed continuously in our model, to facilitate nonparametric estimation of spot volatility, following (8), we assume there exists a surjective function that maps $t \in [0,T]$ to $j \in \{1,\ldots,m_n\}$.[3] To ensure our notations to be consistent wit those in the literature (such as Chernov, Ronald Gallant, Ghysels, and Tauchen (2003)), we split each day into M disjoint blocks and the size of each block is fixed as k. If T represents the total number of trading days, then $n = k(MT)$ and $\Delta_n = T/n = 1/(kM)$. In this case, the total number of blocks for the T trading days is $m_n = MT$. Alternatively, we may have the following representation. For any $t \in [0,T]$ and $r \in (0,1]$ or the corresponding discretized counterpart $t^\circ \in [0,T]$ and $r^\circ \in \{1/M, 2/M, \ldots, M/M = 1\}$, we express

$$t = \lfloor t- \rfloor + r,$$

and

$$t^\circ = \lfloor t- \rfloor + r^\circ,$$

where $\lfloor x- \rfloor$ denotes the greatest integer less than x.

[3] As implied by the fixed-k inference theory, instead of estimating $\ln c_{n,t}$ for $t \in [0,T]$, we estimate $\ln c_{n,j}$ sampled at discrete time points with $j \in \{1,\ldots,m_n\}$. In the state-space framework, therefore, we assume t is uniquely mapped to $j \in \{1,\ldots,m_n\}$ where m_n is the number of local estimation blocks. More specifically, for $c_{n,t}$ on the local estimation block $((j-1)k\Delta_n, jk\Delta_n]$, we treat $c_{n,t} \equiv c_{n,j}$.

Based on the fixed k-inference theory, we set up the following class of state-space model

$$\begin{cases} \ln\left(\hat{c}_{n,t^\circ}\right) &= \ln\left(c_{n,t^\circ}\right) + \epsilon_{t^\circ}, \quad \epsilon_{t^\circ} \sim \ln \bar{\chi}^2_k, & (16) \\ \ln\left(c_{n,t^\circ}\right) &= \text{alternative models}. & (17) \end{cases}$$

Clearly, the observation equation comes from the fixed-k theory. Since ϵ_{t° is not a Gaussian variable, a model in this class is a nonlinear non-Gausian state-space model. Alternative model specifications will be introduced in the next section.

4 Alternative Model Specifications

We now specify several alternative models for the latent log spot volatility, $\ln\left(c_{n,t^\circ}\right)$. Following Stroud and Johannes (2014), in our most general specification, we assume that $\ln\left(c_{n,t^\circ}\right)$ can be decomposed into several components:

$$\ln\left(c_{n,t^\circ}\right) = \mu + h_{t^\circ} + s_{t^\circ} + a_{t^\circ}, \tag{18}$$

where h_{t° is stochastic volatility process, s_{t° the seasonal component, a_{t° the announcement component. By shutting down different components in $\ln\left(c_{n,t^\circ}\right)$ or having different specification for h_{t°, we end up with alternative models.

4.1 Alternative models

4.1.1 Model 1

If we shut down a_{t° and s_{t° in (18) and impose AR(1) structure for h_{t° with associated intercept μ, we have our first model – the benchmark model. That is,

$$\begin{aligned} \ln\left(c_{n,t^\circ}\right) &= \mu + h_{t^\circ} \\ h_{t^\circ} &= \phi h_{t^\circ - 1/M} + e_{t^\circ}, \quad e_{t^\circ} \sim \mathcal{N}\left(0, \sigma_e^2\right). \end{aligned}$$

In this model, data is contained in $\{\ln \hat{c}_{n,t^\circ}\}$ with $\{h_{t^\circ}\}$ being latent variables. The parameters of the model are $\left\{\phi, \mu, \sigma_e^2\right\}$. For simplicity, we call this benchmark model – **Model 1**. To ensure the $\{h_{t^\circ}\}$ process to be stationary, we assume $\phi \in (0,1)$ and the distribution for the initial state is

$$\ln(c_{n,0}) \sim \mathcal{N}\left(\mu, \frac{\sigma_e^2}{1-\phi^2}\right). \tag{19}$$

The above model is different from the log square transformation of the lognormal stochastic volatility model widely studied in the literature; see, for example Harvey, Ruiz, and Shephard (1994).

The lognormal stochastic volatility model is given by

$$r_t = \sigma exp(h_t/2)\epsilon_t,\ \epsilon_t \sim \mathcal{N}(0,1), \tag{20}$$

$$h_t = \phi h_{t-1} + \sigma_h \eta_t,\ \eta_t \sim \mathcal{N}(0,1). \tag{21}$$

When applying the log square transformation to (20), we have

$$\ln r_t^2 = \mu + h_t + \ln \epsilon_t^2. \tag{22}$$

Clearly, $\ln \epsilon_t^2$ is a log chi-square random variable. This is contrast with the scaled chi-square random variable used in **Model 1**.

For the prior specification, we first introduce the auxiliary parameters ϕ^* and σ_e^{*2} respectively as in Kim, Shephard, and Chib (1998),

$$\begin{aligned} \phi &= 2\phi^* - 1 \\ \sigma_e &= \exp\left(\frac{1}{2}\ln \sigma_e^{*2}\right), \end{aligned}$$

and then use the following priors:

$$\begin{aligned} \phi^* &\sim \text{Beta}\left(\alpha_{\phi^*}, \beta_{\phi^*}\right) \\ \sigma_e^{*2} &\sim \text{I.G.}\left(\alpha_{\sigma_e^*}, \beta_{\sigma_e^*}\right) \\ \mu &\sim \mathcal{N}(0, 100) \end{aligned}$$

where $\alpha_{\phi^*}, \beta_{\phi^*}, \alpha_{\sigma_e^*}, \beta_{\sigma_e^*}$ are hyperparameters.

MCMC is applied to obtain the correlated random draws from the posterior distributions of μ, ϕ^* and $\ln\left(\sigma_e^{*2}\right)$. These draws can be regarded as correlated random draws from the original parameters. Based on the MCMC draws, we may obtain the posterior mean, quantiles, variance for each parameter.

Remark 4.1 *Spot volatility forecasting within the state-space model is straightforward. Specifically, using the extracted log-volatility (demeaned) at the last period in sample, i.e.* h_{t° *and* $t^\circ = T$ *and the corresponding parameters evaluated at the posterior mean as the estimation, denoted by* $\hat{\phi}$, $\hat{\sigma}_e^2$, $\hat{\mu}$ *respectively, any future k-th spot log-volatility can accordingly be calculated as* $c_{t^\circ+k/M} = \exp\left(\hat{\phi}^k h_{t^\circ} + \hat{\mu}\right)$. *Besides, the corresponding k-th forecasting variance can be calculated as* $\sum_{j=0}^{k-1} \hat{\phi}^j \hat{\sigma}_e^2$, *therefore the volatility prediction confidence interval can also be established.*

4.1.2 Model 2

Model 2 extends the benchmark model by combining the AR(1) structure and a discrete jump component in h_{t°. Specifically, **Model 2** is given by

$$\begin{aligned}\ln\left(c_{n,t^\circ}\right) &= \mu + h_{t^\circ} \\ h_{t^\circ} &= \phi h_{t^\circ - 1/M} + e_{t^\circ} + J_{t^\circ}\eta_{t^\circ}, \quad e_{t^\circ} \sim \mathcal{N}\left(0, \sigma_e^2\right), \quad \eta_{t^\circ} \sim \mathcal{N}\left(\mu_\eta, \sigma_\eta^2\right),\end{aligned}$$

where J_{t° is a jump indicator, defined by

$$J_{t^\circ} = \begin{cases} 1 & \text{with probability} \quad \kappa \\ 0 & \text{with probability} \quad 1-\kappa, \end{cases}$$

with κ being the jump probability, and η_{t° determines the jump size.

In this model, data is contained in $\{\ln \hat{c}_{n,t^\circ}\}$ with $\{h_{t^\circ}\}, \{J_{t^\circ}\}$ being latent variables. The parameters of the model are $\left\{\phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2\right\}$.

Following Chib, Nardari, and Shephard (2002), we assume the following conjugate priors for parameters in the jump component

$$\begin{aligned}\kappa &\sim \text{Beta}\left(\alpha_\kappa, \beta_\kappa\right), \\ \mu_\eta &\sim \mathcal{N}\left(\tilde{\mu}_\eta, \tilde{\sigma}_\eta^2\right), \\ \sigma_\eta^2 &\sim \text{I.G.}\left(\alpha_{\sigma_\eta}, \beta_{\sigma_\eta}\right),\end{aligned}$$

where I.G. denotes the Inverse-Gamma distribution.

4.1.3 Model 3

In **Model 3** we add the seasonal component to the benchmark model. The seasonal component for intraday volatility is used to capture the diurnal U-shaped patterns in high frequency financial data. There are several methods for modeling diurnal patterns, namely, the Fourier representation and a deterministic spline. Based on using 5-minute square return, when using a deterministic spline, Aït-Sahalia and Jacod (2014) and Christensen, Hounyo, and Podolskij (2018) document evidence of larger fluctuations near the opening and closing of the exchange than around lunch time. The model is given by

$$\begin{aligned}\ln\left(c_{n,t^\circ}\right) &= \mu + h_{t^\circ} + s_{t^\circ}, \\ h_{t^\circ} &= \phi h_{t^\circ - 1/M} + e_{t^\circ}, \quad e_{t^\circ} \sim \mathcal{N}\left(0, \sigma_e^2\right), \\ s_{t^\circ} \equiv \tilde{s}_{r^\circ} &= 12\left(1-b\right)\left(r^\circ - \frac{1}{2}\right)^2 + b, \quad r^\circ = t^\circ - \lfloor t - \rfloor, \quad t \in [0, T].\end{aligned}$$

The quadratic function $\tilde{s}_r = 12(1-b)\left(r-\frac{1}{2}\right)^2 + b$, is the only function within the class $f(r) = c(r-a)^2 + b$ that satisfies (i) $\int_0^1 \left(c(r-a)^2+b\right) dr = 1$; (ii) $\text{argmin}_r\, c(r-a)^2 + b = \frac{1}{2}$. The first condition is imposed for identification. The second condition assumes that the diurnal pattern reaches the minimum in the middle of a trading day, an empirical regularity that has been found in the literature.[4] There is a restriction in using the quadratic function. That is, it implies a symmetric diurnal pattern. In a recent study, Christensen, Hounyo, and Podolskij (2018) propose a nonparametric method to estimate the diurnal patten and find an asymmetric diurnal pattern. However, our approach can be easily extended to cover more complicated deterministic functions for diurnal pattern.

In this model, data is contained in $\{\ln \hat{c}_{n,t^\circ}\}$ with $\{h_{t^\circ}\}$ being latent variables. The parameters of the model are $\left\{\phi, \mu, \sigma_e^2, b\right\}$.

Since **Model 3** and **Model 1** share the same AR(1) specification for $\{h_{t^\circ}\}$, we use the same priors on $\left\{\phi, \mu, \sigma_e^2\right\}$ as before. For parameter b, we assume a flat prior on $[0,1]$, that is, $b \sim \mathcal{U}(0,1)$.

4.1.4 Model 4

Different from **Model 3** that includes the component to capture the diurnal pattern, **Model 4** include the component to capture macroeconomic news announcement effects. The motivation for incorporating announcement effects comes from recent empirical finance sdudies, for instance Lucca and Moench (2015) and Bernile, Hu, and Tang (2016). The specification of **Model 4** is given by

$$\begin{aligned}
\ln\left(c_{n,t^\circ}\right) &= \mu + h_{t^\circ} + a_{t^\circ}, \\
h_{t^\circ} &= \phi h_{t^\circ - 1/M} + e_{t^\circ}, \quad e_{t^\circ} \sim \mathcal{N}\left(0, \sigma_e^2\right), \\
a_{t^\circ} &= \sum_{q=1}^{Q} \sum_{l=0}^{L} \mathbf{1}_{t^\circ ql} \alpha_{ql},
\end{aligned}$$

where $\mathbf{1}_{t^\circ ql}$ is an indicator for news type q at time t° with $l = 0, 1, \ldots, L$ (i.e., $\mathbf{1}_{t^\circ ql} = 1$ if it is within l periods after type q announcement made at time $t^\circ - \frac{l}{M}$ and 0 otherwise), α_{ql} is the announcement effect for news type q at l periods after the announcement.

Again, since the specification for $\ln \hat{c}_{n,t^\circ}$ in **Model 4** is the same as that in **Model 1**, we use the same priors for parameters $\left\{\phi, \mu, \sigma_e^2\right\}$. Parameters $\left\{\alpha_{ql}\right\}_{q=1,\ldots,Q;l=1,\ldots,L}$ characterize announcement effects. The dimension of these parameters is $Q \times L$, and hence, it increases as Q and/or L increases. For instance, if $L = 5$ and $Q = 3$, we will have $L \times Q = 15$ parameters to determine the announcement effects. This would impose a great deal of computational challenges to the Bayesian analysis.

To alleviate the computational burden, we assume the announcement effects decay over time according to the following pattern, $\alpha_{ql} = \tilde{\alpha}_q \exp\left\{-\tilde{\beta}_q l\right\}$. This is relatively a parsimonious specification

[4] To satisfy the condition that $\text{argmin}_r\, c(r-a)^2 + b = \frac{1}{2}$, we have $c > 0$ and $a = \frac{1}{2}$. Substituting $a = \frac{1}{2}$ into $\int_0^1 \left(c(r-a)^2+b\right) dr = 1$ yields $\frac{1}{12}c + b = 1 \Rightarrow c = 12(1-b)$. Thus, the quadratic function is uniquely determined by single parameter b. The larger the value of b is, the less pronounced the quadratic volatility pattern.

that significantly reduces the dimension of the parameter space associated with announcement effects. The imposed decaying structure for the announcement effects is consistent with the intuition and the empirical evidence in the literature (see Stroud and Johannes, 2014; Lucca and Moench, 2015; Bernile, Hu, and Tang, 2016). Under this specification, the parameters are collected in $\left\{\tilde{\alpha}_q, \tilde{\beta}_q\right\}_{q=1}^{Q}$. When $L = 5$ and $Q = 3$, the number of parameters in connection to announcement effects reduce from 15 to $2 \times Q(=3) = 6$. The following priors are used for the new parameters,

$$\begin{aligned}\tilde{\alpha}_q &\sim \mathcal{N}\left(0, \tilde{\sigma}_q^2\right) \quad \text{for } q = 1, \cdots, Q, \\ \tilde{\beta}_q &\sim \mathcal{E}\left(-\tilde{\lambda}_q\right) \quad \text{for } q = 1, \cdots, Q,\end{aligned}$$

where $\mathcal{N}(\cdot,\cdot)$ and $\mathcal{E}(\cdot)$ denote normal distribution and exponential distribution respectively. Accordingly, in this model, data is contained in $\{\ln \hat{c}_{n,t^\circ}\}$ with $\{h_{t^\circ}\}$ being latent variables. The parameters of the model are $\left\{\phi, \mu, \sigma_e^2, \tilde{\alpha}_q, \tilde{\beta}_q\right\}$.

4.1.5 Model 5

In **Model 5**, we consider the model specification by combining all the specifications of **Models 1-3**. Thus, we have both jumps and diurnal patterns included in the model specifications. In other words, **Model 5** nests **Models 1-3**. We summarize the model structure of **Model 5** as follows with detailed explanations for notations kept in the description of **Models 1-3**.

$$\begin{aligned}\ln\left(c_{n,t^\circ}\right) &= \mu + h_{t^\circ} + s_{t^\circ}, \\ h_{t^\circ} &= \phi h_{t^\circ - 1/M} + e_{t^\circ} + J_{t^\circ}\eta_{t^\circ}, \quad e_{t^\circ} \sim \mathcal{N}\left(0, \sigma_e^2\right), \quad \eta_{t^\circ} \sim \mathcal{N}\left(\mu_\eta, \sigma_\eta^2\right), \\ s_{t^\circ} \equiv \tilde{s}_{r^\circ} &= 12\left(1-b\right)\left(r^\circ - \frac{1}{2}\right)^2 + b, \quad r^\circ = t^\circ - \lfloor t - \rfloor, \quad t \in [0, T].\end{aligned}$$

The parameters of nested **Model 5** are $\left\{\phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2, b\right\}$.

4.1.6 Model 6

Recall (18) for our general specified functional form of latent volatility process. By shutting down (or opening up) different components, we can obtain different model specifications. **Models 1-4** discussed in the previous subsections are about adding different components (i.e. jumps, diurnal components, and announcement effect components) respectively onto the single factor volatility model, **Model 1**. Alternatively speaking, built upon the benchmark model specification in **Model 1**, by combining different specifications in **Models 2-4**, we can at most obtain $C_3^0 + C_3^1 + C_3^2 + C_3^3 = 1 + 3 + 3 + 1 = 8$ different models. **Models 1-4** are about the subset of combinations (i.e. $C_3^0 + C_3^1$). We refer to the most comprehensive model specification that includes all the components as **Model 6**. Thus **Model 6 nests** all the specifications in **Models 1-4**. We summarize the model structure of **Model 6** as follows with detailed explanations for notations kept the same as in the description

of **Models 1-4**.

$$
\begin{aligned}
\ln\left(c_{n,t^\circ}\right) &= \mu + h_{t^\circ} + s_{t^\circ} + a_{t^\circ} \\
h_{t^\circ} &= \phi h_{t^\circ - 1/M} + e_{t^\circ} + J_{t^\circ}\eta_{t^\circ}, \quad e_{t^\circ} \sim \mathcal{N}\left(0, \sigma_e^2\right), \quad \eta_{t^\circ} \sim \mathcal{N}\left(\mu_\eta, \sigma_\eta^2\right) \\
s_{t^\circ} \equiv \tilde{s}_{r^\circ} &= 12\left(1-b\right)\left(r^\circ - \frac{1}{2}\right)^2 + b, \quad r^\circ = t^\circ - \lfloor t - \rfloor, \quad t \in [0, T]. \\
a_{t^\circ} &= \sum_{q=1}^{Q}\sum_{l=0}^{L} \mathbf{1}_{t^\circ ql}\alpha_{ql},
\end{aligned}
$$

The parameters of nested **Model 6** are $\left\{\phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2, b, \tilde{\alpha}_q, \tilde{\beta}_q\right\}$.

4.2 Bayesian analysis

4.2.1 MCMC method to sample from posteriors

With the state-space model summarized in (16) and (17), we design the corresponding Markov Chain Monte Carlo (MCMC) algorithms to estimate all the parameters involved (especially the latent variables of our major interests in this setting, the log spot volatility). MCMC as the leading modern Bayesian technique is quite suitable for the state-space model with a latent variable structure. The general idea of the MCMC method can be understood as a combination of various sampling algorithms such as the Metropolis-Hastings (M-H) algorithm, acceptance-rejection algorithm, Gibbs sampler, and the substitution sampler (see data augmentation algorithm in Tanner and Wong, 1987) by making draws from conditional distributions associated with the target posterior. It can be theoretically justified that as long as we can let the Markov chain run long enough, those draws taken from blocks of conditional distributions constitute the target posterior distribution. Accordingly, we can use these draws to summarize the posterior mean (or mode) as the estimation of target parameters and latent variables. Gilks, Richardson, and Spiegelhalter (1995), Bolstad (2009), and Gelman, Carlin, Stern, Dunson, Vehtari, and Rubin (2013) are all good references for technical details of MCMC methods. We exploit $\boldsymbol{h}$ to denote the sequence of h_{t°, $\boldsymbol{J}$ to denote all the jump indicators, and $\boldsymbol{\eta}$ to denote all the jump sizes. For the sake of description simplicity, we let $\tilde{\boldsymbol{h}} = \boldsymbol{h} + \mu$. Meanwhile, we use $\mathbf{y}$ to denote the sequence of log transformation of nonparametric estimator of volatility, that is the sequence of $\ln\left(\hat{c}_{n,t^\circ}\right)$. We summarize the main MCMC sampling steps for each model in this section while leaving the technical details in the appendix A.

- **Model 1**

  (1) Initialize $\left\{\boldsymbol{h}, \phi, \mu, \sigma_e^2\right\}$.

  (2) Sample $\tilde{\boldsymbol{h}} \mid \phi, \mu, \sigma_e^2, \mathbf{y}$.

  (3) Sample $\left\{\phi, \mu, \sigma_e^2\right\} \mid \tilde{\boldsymbol{h}}$.

  (4) Go to (2).

- **Model 2**

  (1) Initialize $\left\{\boldsymbol{h}, \boldsymbol{J}, \boldsymbol{\eta}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2\right\}$.

  (2) Sample $\tilde{\boldsymbol{h}} \mid \boldsymbol{J}, \boldsymbol{\eta}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2, \mathbf{y}$.

  (3) Sample $\left\{\phi, \mu, \sigma_e^2\right\} \mid \tilde{\boldsymbol{h}}, \boldsymbol{J}, \boldsymbol{\eta}$.

  (4) Sample $\boldsymbol{J} \mid \boldsymbol{h}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2$.

  (5) Sample $\boldsymbol{\eta} \mid \boldsymbol{h}, \boldsymbol{J}, \phi, \mu, \sigma_e^2, \mu_\eta, \sigma_\eta^2$.

  (6) Sample $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\} \mid \boldsymbol{J}, \boldsymbol{\eta}$.

  (7) Go to (2).

  In step (2) above, we need $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$ to obtain the initial condition of $\boldsymbol{h}$. Detailed sampling steps for sampling latent components such as $\boldsymbol{h}$, $\boldsymbol{J}$, $\boldsymbol{\eta}$ are covered in the appendix.

- **Model 3**

  (1) Initialize $\left\{\boldsymbol{h}, \phi, \mu, \sigma_e^2, b\right\}$.

  (2) Sample $\tilde{\boldsymbol{h}} \mid \phi, \mu, \sigma_e^2, b, \mathbf{y}$.

  (3) Sample $\left\{\phi, \mu, \sigma_e^2\right\} \mid \tilde{\boldsymbol{h}}$.

  (4) Sample $b \mid \boldsymbol{h}, \mu, \mathbf{y}$.

  (5) Go to (2).

  In step (4) above, we insert one M-H algorithm for sampling b.

- **Model 4**

  (1) Initialize $\left\{\boldsymbol{h}, \phi, \mu, \sigma_e^2, \tilde{\alpha}_q, \tilde{\beta}_q\right\}$.

  (2) Sample $\tilde{\boldsymbol{h}} \mid \phi, \mu, \sigma_e^2, \tilde{\alpha}_q, \tilde{\beta}_q, \mathbf{y}$.

  (3) Sample $\left\{\phi, \mu, \sigma_e^2\right\} \mid \tilde{\boldsymbol{h}}$.

  (4) Sample $\left\{\tilde{\alpha}_q, \tilde{\beta}_q\right\} \mid \boldsymbol{h}, \mu, \mathbf{y}$.

  (5) Go to (2).

  In step (4) above, we insert one M-H algorithm for sampling $\left\{\tilde{\alpha}_q, \tilde{\beta}_q\right\}$.

- **Model 5**

  This is a model nesting all the specifications from **Models 1-3**. The corresponding MCMC loop is summarized as follows

  (1) Initialize $\left\{\boldsymbol{h}, \boldsymbol{J}, \boldsymbol{\eta}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2, b\right\}$.

  (2) Sample $\tilde{\boldsymbol{h}} \mid \boldsymbol{J}, \boldsymbol{\eta}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2, b, \mathbf{y}$.

  (3) Sample $\left\{\phi, \mu, \sigma_e^2\right\} \mid \tilde{\boldsymbol{h}}, \boldsymbol{J}, \boldsymbol{\eta}$.

(4) Sample $\boldsymbol{J} \mid \boldsymbol{h}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2$.

(5) Sample $\boldsymbol{\eta} \mid \boldsymbol{h}, \boldsymbol{J}, \phi, \mu, \sigma_e^2, \mu_\eta, \sigma_\eta^2$.

(6) Sample $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\} \mid \boldsymbol{J}, \boldsymbol{\eta}$.

(7) Sample $b \mid \boldsymbol{h}, \mu, \mathbf{y}$.

(8) Go to (2).

- **Model 6**

  This is the largest model nesting all the components that are expected to obtain from MCMC. The MCMC loop is summarized as follows,

  (1) Initialize $\left\{\boldsymbol{h}, \boldsymbol{J}, \boldsymbol{\eta}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2, b, \tilde{\alpha}_q, \tilde{\beta}_q\right\}$.

  (2) Sample $\tilde{\boldsymbol{h}} \mid \boldsymbol{J}, \boldsymbol{\eta}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2, b, \tilde{\alpha}_q, \tilde{\beta}_q, \mathbf{y}$.

  (3) Sample $\left\{\phi, \mu, \sigma_e^2\right\} \mid \tilde{\boldsymbol{h}}, \boldsymbol{J}, \boldsymbol{\eta}$.

  (4) Sample $\boldsymbol{J} \mid \boldsymbol{h}, \phi, \mu, \sigma_e^2, \kappa, \mu_\eta, \sigma_\eta^2$.

  (5) Sample $\boldsymbol{\eta} \mid \boldsymbol{h}, \boldsymbol{J}, \phi, \mu, \sigma_e^2, \mu_\eta, \sigma_\eta^2$.

  (6) Sample $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\} \mid \boldsymbol{J}, \boldsymbol{\eta}$.

  (7) Sample $\left\{b, \tilde{\alpha}_q, \tilde{\beta}_q\right\} \mid \boldsymbol{h}, \mu, \mathbf{y}$.

  (8) Go to (2).

Remark 4.2 *To demonstrate the ideas of how we sample the latent volatility process, we use t as the discrete timing index. Besides, given that $\hat{c}_{n,t}$ as the nonparametric estimation of volatility is always a positive number, we emphasize it using square by rewriting the log transformation $\ln(\hat{c}_{n,t}) + \ln k$ as $\ln y_t^2 + \ln k$ and for the ease of notation as well. We use $\tilde{h}_t$ to denote the generally unobserved components at t temporarily within this remark. Alternatively speaking, $\tilde{h}_t$ refers to the latent volatility process in the following discussion. As we have discussed in section 3, this fixed-k inference theory suggests that there is a random variable $\mathcal{Z}$ following log chi-square distribution that characterizes the gap between $\ln y_t^2 + \ln k$ and $\tilde{h}_t$. Consequently, we have*

$$\log p(\mathcal{Z}) = \underbrace{-\frac{k}{2}\log 2 - \log\left(\Gamma\left(\frac{k}{2}\right)\right) + \frac{k}{2}(\ln y_t^2 + \ln k)}_{const} \underbrace{-\frac{k}{2}\tilde{h}_t - \frac{k y_t^2}{2}\exp\left\{-\tilde{h}_t\right\}}_{\log f^*(y_t \mid \tilde{h}_t)}. \tag{23}$$

We use "const" to emphasize that the associated term is constant from posterior perspective. As in Kim, Shephard, and Chib (1998), we note that $\exp\left(-\tilde{h}_t\right)$ is a convex function that can be bounded by a function linear in h_t. We use this fact to derive following inequality for constructing proposal

distribution used in M-H algorithm.

$$\begin{aligned}
\log f^*\left(y_t \mid \tilde{h}_t\right) &= -\frac{k}{2}\tilde{h}_t - \frac{k y_t^2}{2}\left\{\exp(-\tilde{h}_t)\right\} \\
&\leqslant -\frac{k}{2}\tilde{h}_t - \frac{k y_t^2}{2}\left\{\exp(-\tilde{h}_t^*)(1+\tilde{h}_t^*) - \tilde{h}_t \exp(-\tilde{h}_t^*)\right\} \\
&= \log g^*\left(y_t, \tilde{h}_t, \tilde{h}_t^*\right)
\end{aligned}$$

Hence,

$$f\left(\tilde{h}_t \mid y_t, \tilde{h}_{\setminus t}\right) \propto f\left(\tilde{h}_t \mid \tilde{h}_{\setminus t}\right) f^*\left(y_t \mid \tilde{h}_t\right) \leqslant f_N\left(\tilde{h}_t \mid \tilde{h}_t^*, v^2\right) g^*\left(y_t, \tilde{h}_t, \tilde{h}_t^*\right). \tag{24}$$

Note that terms collected in the exponential component (24) are of standard quadratic form in $\tilde{h}_t$, thus we can show that $f_N\left(\tilde{h}_t \mid \tilde{h}_t^*, v^2\right) g^*\left(y_t, \tilde{h}_t, \tilde{h}_t^*\right)$ is proportional to Gaussian density functional form $f_N\left(\tilde{h}_t \mid \mu_t, v^2\right)$. Specifically, for R.H.S. of (24), we have

$$\begin{aligned}
f_N\left(\tilde{h}_t \mid \tilde{h}_t^*, v^2\right) g^*\left(y_t, \tilde{h}_t, \tilde{h}_t^*\right) &= \frac{1}{\sqrt{2\pi}v}\exp\left\{-\frac{1}{2v^2}\left(\tilde{h}_t - \tilde{h}_t^*\right)^2\right\} \\
&\quad \times \exp\left\{\left(\frac{k y_t^2}{2}\exp\left(-\tilde{h}_t^*\right) - \frac{k}{2}\right)\tilde{h}_t - \frac{k y_t^2}{2}\exp(-\tilde{h}_t^*)(1+\tilde{h}_t^*)\right\}
\end{aligned}$$

If we rewrite and focus on the terms that correspond to $\tilde{h}_t$ in the exponential components of the expression above, we can derive following normal proposal distribution,

$$\begin{aligned}
-\frac{1}{2v^2}\left(\tilde{h}_t - \tilde{h}_t^*\right)^2 + \left(\frac{k y_t^2}{2}\exp\left(-\tilde{h}_t^*\right) - \frac{k}{2}\right)\tilde{h}_t &= -\frac{1}{2v^2}\tilde{h}_t^2 + \left[\left(\frac{k y_t^2}{2}\exp\left(-\tilde{h}_t^*\right) - \frac{k}{2}\right) + 2\cdot\frac{1}{2v^2}\tilde{h}_t^*\right]\tilde{h}_t - \frac{1}{2v^2}\tilde{h}_t^{*2} \\
&= -\frac{1}{2v^2}\left\{\tilde{h}_t^2 - \left[2v^2\left(\frac{k y_t^2}{2}\exp\left(-\tilde{h}_t^*\right) - \frac{k}{2}\right) + 2\tilde{h}_t^*\right]\tilde{h}_t\right\} - \frac{1}{2v^2}\tilde{h}_t^{*2} \\
&= -\frac{1}{2v^2}\left\{\tilde{h}_t^2 - 2\cdot\left[\frac{v^2}{2}\left(k y_t^2\exp\left(-\tilde{h}_t^*\right) - k\right) + \tilde{h}_t^*\right]\tilde{h}_t\right\} - \frac{1}{2v^2}\tilde{h}_t^{*2} \\
&= -\frac{1}{2v^2}\left\{\tilde{h}_t - \underbrace{\left[\frac{v^2}{2}\left(k y_t^2\exp\left(-\tilde{h}_t^*\right) - k\right) + \tilde{h}_t^*\right]}_{u_t}\right\}^2 \\
&\quad \underbrace{-\left[\frac{v^2}{2}\left(k y_t^2\exp\left(-\tilde{h}_t^*\right) - k\right) + \tilde{h}_t^*\right]^2 - \frac{1}{2v^2}\tilde{h}_t^{*2}}_{const}
\end{aligned}$$

Hence we use $f_N\left(\tilde{h}_t \mid u_t, v^2\right)$ as the proposal distribution in Metropolis-Hastings sampling procedure to sample from $f\left(\tilde{h}_t \mid y_t^2, \tilde{h}_{\setminus t}\right)$. Specifically, Metropolis-Hastings is applied in general for the j-th

step of Gibbs Sampling as follows,

(1) Generate $x \sim \mathcal{N}(u_t, v^2)$ *and* $u \sim \mathcal{U}(0,1)$

(2) Let [5]

$$\alpha = \min\left\{1, \frac{f_N(x;\tilde{h}_t^*, v^2) f_{\ln\chi_k^2}\left((\ln y_t^2 + \ln k); x\right)}{f_N\left(\tilde{h}_t^{(j-1)};\tilde{h}_t^*, v^2\right) f_{\ln\chi_k^2}\left((\ln y_t^2 + \ln k); \tilde{h}_t^{(j-1)}\right)} \times \frac{f_N\left(\tilde{h}_t^{(j-1)}; u_t, v^2\right)}{f_N\left(x; u_t, v^2\right)}\right\}$$

(3) If $u < \alpha$, $\tilde{h}_t^{(j)} = x$, *else,* $\tilde{h}_t^{(j)} = \tilde{h}_t^{(j-1)}$.

4.2.2 Deviance information criterion for model comparison

In this section, we focus on discussing how to compare alternative models when MCMC output is ready from each candidate model in the context of **Models 1-4**.

Deviance Information Criterion (DIC), proposed and well-discussed in Spiegelhalter, Best, Carlin, and van der Linde (2002, 2014), is a popular method for model selection when MCMC output is ready. There are a few nice features about DIC. First, DIC applies to a wide range of statistical models. Second, it does not suffer from Jeffreys-Lindley-Barlett's paradox. Third, it can be obtained even under improper priors. Finally, Li, Yu, and Zeng (2021) justify DIC by showing that DIC is an asymptotically unbiased estimator of the Kullback-Leibler divergence between the data generating process and the plug-in predictive distribution. DIC in general is given as follows

$$\text{DIC} = D\left(\bar{\boldsymbol{\vartheta}}\right) + 2P_D \tag{25}$$

where

$$\begin{aligned} D(\boldsymbol{\vartheta}) &= -2\ln p\left(\mathbf{y} \mid \boldsymbol{\vartheta}\right) \\ P_D &= -2\int [\ln p(\mathbf{y} \mid \boldsymbol{\vartheta}) - \ln p(\mathbf{y} \mid \bar{\boldsymbol{\vartheta}})] p(\boldsymbol{\vartheta} \mid \mathbf{y}) \mathrm{d}\boldsymbol{\vartheta}. \end{aligned}$$

$\bar{\boldsymbol{\vartheta}}$ refers to the posterior mean of parameter $\boldsymbol{\vartheta}$ and $\mathbf{y}$ generically denotes observable data. Specifically, for the DIC definition in (25), $D(\bar{\boldsymbol{\vartheta}})$ as the product of a negative number (-2) and log-observed-data likelihood evaluated at the posterior mean of parameters could be interpreted as the measure of model fit and hence it is expected to be minimized for fitting data purpose. While P_D as the second term in (25) could be interpreted as the product of a negative number (-2) and the posterior expected deviance between log-observed-data likelihood evaluated at different parameter values over parameter space and log-observed-data likelihood evaluated at the posterior mean. The

[5] By construction, we are using proposal distribution taking exactly the same shape as target distribution.

corresponding expectation is taken with respect to posterior distribution $p(\boldsymbol{\vartheta} \mid \mathbf{y})$. Since P_D is increasing with $\ln p\left(\mathbf{y} \mid \bar{\boldsymbol{\vartheta}}\right)$, there exists trade-off between these two terms $D\left(\bar{\boldsymbol{\vartheta}}\right)$ and P_D, therefore the objective to minimize DIC is reconciled with the goal of achieving balance between "model fit" and "model complexity". This also corresponds to the opinion that DIC can be understood as the Bayesian version of AIC. Computing DIC using conditional likelihood is straightforward based on posterior sampling from MCMC: (i) for $D\left(\bar{\boldsymbol{\vartheta}}\right)$, we just make $D(\boldsymbol{\vartheta})$ evaluated at posterior mean $\bar{\boldsymbol{\vartheta}}$; (ii) and p_D is calculated using posterior sample mean from MCMC. We also make a discussion using particle filter for approximating marginal likelihood (or the observed-data likelihood) in appendix B.

5 Monte Carlo Experiments

In this section, we conduct several Monte Carlo experiments to check the performance of the proposed Bayesian method in estimating parameters, extracting volatility estimates, and in comparing alternative models. Several data generation processes, which match with different alternative model specifications in the previous section, have been used to simulate data.

5.1 Experiment 1

In the first experiment, we simulate data according to,

$$\begin{aligned} dX_t &= \exp(h_t/2)\, dW_t \\ dh_t &= \kappa_h(\mu_h - h_t)\, dt + \sigma_h dB_t \\ \mathbb{E}[W_t B_t] &= \rho = 0 \end{aligned}$$

Specifically we simulate discrete data by setting $dt := \Delta_n = 1/390$, which implies that price data is sampled every one minute, leading to $6.5 \times 60 = 390$ returns as we assume there are 6.5 trading hours. Parameter specification for this experiment is as follows: $\kappa_h = 0.2$, $\mu_h = -5$, $\rho = 0$, $\sigma_h = 0.4$. For notational simplicity, we write $\ln(c_{n,t})$ as h_t. When extracting latent volatility, we focus our attention on the one-month interval (i.e., $T = 22$, assuming that there are 22 trading days within one month). In total we have $n = 22 \times 390 = 8580$ returns.[6]

First, to check the performance of the Bayesian method in estimating model parameters, we report the posterior means, posterior standard errors for all parameters. The corresponding results are summarized in the following figures with the vertical red dashed lines indicating the location of posterior means.

[6] We simulate price-level data using the described DGP each for a specific experiment. It works fine for different replicate experiments. Besides, we simulate price-level data by only accounting for diffusion process. This DGP scheme does not distort our general target since our approach relies on the distribution of the difference between the log fixed-k estimator of spot volatility and the true unobserved latent volatility. The derived distribution in BLL2021QE generally applies to continuous Itô semimartingale and therefore any fixed-k estimator of volatility associated price-level data generated from continuous Itô semimartingale does not affect the nonlinear non-Gaussian state-space model we establish. Similar DGP scheme has also been used in literature such as Xiu (2010).

[Place Figure 1 about here]

Remark 5.1 (From DGP Dynamic to Block Dynamic) *Recall that in general data is generated from the following dynamic system, if we treat $dh_t = h_{t+dt} - h_t$ of the second equation above, we can rewrite the second equation characterizing latent spot volatility dynamics as follows*

$$\begin{aligned} h_{t+dt} - h_t &= \kappa_h (\mu_h - h_t)\, dt + \sigma_h dB_t \\ &\Leftrightarrow \\ (h_{t+dt} - \mu_h) - (h_t - \mu_h) &= -\kappa_h dt (h_t - \mu_h) + \sigma_h dB_t \\ &\Rightarrow \\ h_{t+dt} - \mu_h &= \underbrace{(1 - \kappa_h dt)}_{\phi_h}(h_t - \mu_h) + \underbrace{\sigma_h dB_t}_{\varepsilon_{t+dt}} \end{aligned}$$

Given the property of Brownian motion, $\sigma_h dB_t \sim \mathcal{N}\left(0, \sigma_h^2 dt\right)$. Data is generated from this continuous setting. But to apply nonparametric estimation of spot volatility in our established framework, we need to select k consecutive time intervals ($dt \equiv \Delta_i^n$) to construct "local estimation window". This implies that we need to move from "observation" dynamics to "block" dynamics. If we simulate data from system given above, we are modeling following dynamics (let $\tilde{\tilde{h}}_t = h_t - \mu_h$ denote the demeaned latent spot volatility)

$$\begin{aligned} \tilde{\tilde{h}}_{t+kdt} &= \phi_h \tilde{\tilde{h}}_{t+(k-1)dt} + \varepsilon_{t+kdt} \\ &= \phi_h \left(\phi_h \tilde{\tilde{h}}_{t+(k-2)dt} + \varepsilon_{t+(k-1)dt}\right) + \varepsilon_{t+kdt} \\ &= \phi_h(\phi_h(\phi_h \tilde{\tilde{h}}_{t+(k-3)dt} + \varepsilon_{t+(k-2)dt}) + \varepsilon_{t+(k-1)dt}) + \varepsilon_{t+kdt} \\ &\cdots \\ &= \phi_h^k \tilde{\tilde{h}}_t + \left(\phi_h^k \varepsilon_t + \cdots + \varepsilon_{t+kdt}\right) \end{aligned}$$

Thus the "block" dynamics should be given as follows

$$\tilde{\tilde{h}}_{j+1} = \phi_h^k \tilde{\tilde{h}}_j + e_{j+1}. \tag{26}$$

This could be interpreted as that $\tilde{\tilde{h}}_j$ is regarded as constant spot volatility within the j-th block constructed by k intervals. Variance of $e_{j+1} = \phi_h^k \varepsilon_t + \cdots + \varepsilon_{t+kdt}$ is given as follows

$$\sigma_e^2 = \frac{\sigma_h^2 dt (1 - \phi_h^{2(k+1)})}{1 - \phi_h^2}.$$

For instance, if $dt := \Delta_i^n = 1/390$, $\sigma_h = 0.4$, $k = 5$, and $\kappa_h = 0.2$, then $\phi_h = 1 - \kappa_h dt \approx 0.9995$,

$\phi_h^k \approx 0.9974$ *and*

$$\sqrt{\frac{\sigma_h^2 dt(1-\phi_h^{2(k+1)})}{1-\phi_h^2}} \approx 0.0496$$

These quantities should be reasonably compared with posterior means summarized from MCMC outputs. We extend discussions here in the appendix using another example.

Second, to check the performance of the smoothing and filtering methods in extracting latent volatility, we plot the simulated (true) volatility, smoothed volatility and nonparametric estimation of volatility in Figures:

[Place Figure 2 about here]

We also provide an auxiliary Figure 3 for demonstration purpose in companion with Figure 2. Specifically, we jointly plot the estimated volatility from fixed-k inference (blue solid line) and true simulated volatility (red solid line) based on DGP of **Experiment 1** in panel (a). Corresponding specifications are exactly the same as in the description of Figure 2. This noisy "gap" between estimated volatility and simulated volatility is approximately and visually characterized in distribution using the suggested scaled chi-square distribution in panel (b). This comparison implies that scaled chi-square distribution as one theoretically-suggested distribution in characterizing this noisy "gap" is empirically supported well for this experiment. Besides, given the reasonably good approximation of the "gap" using scaled chi-square distribution in the setting where data is generated by specifying sampling interval as $\Delta_n = 1/390$, this auxiliary figure can also serve as the supporting evidence for using one-minute return in practice and sampling interval specified as $\Delta_n = 1/390$ is reasonably small enough.

[Place Figure 3 about here]

For the results demonstrated above, the local estimation window size is fixed at $k = 5$. In this regard, we treat the unobserved latent volatility as constant every 5 minutes, and accordingly for this fixe-k scheme we have $M = 390/5 = 78$ local estimation blocks each day and $22 \times 78 = 1716$ local estimation blocks for 22 days.

5.2 Experiment 2

In this experiment, we simulate data according to,

$$\begin{aligned}
dX_t &= \exp\left(\tilde{h}_t/2\right) dW_t \\
\tilde{h}_t &= \mu_h + h_t \\
dh_t &= -\kappa_h h_t dt + \sigma_h dB_t + \mathbf{1}_{\{t=t^\circ\}} J_t \eta_t \\
\mathbb{E}\left[W_t B_t\right] &= \rho = 0 \\
J_t &\sim \text{Bernoulli}(\kappa) \\
\eta_t &\sim \mathcal{N}\left(\mu_\eta, \sigma_\eta^2\right)
\end{aligned}$$

Specifically we simulate discrete data by setting $dt := \Delta_i^n = 1/390$, which implies that price data is sampled every one minute, leading to $6.5 \times 60 = 390$ returns as we assume there are 6.5 trading hours. Parameter specification for this experiment is as follows: $\kappa_h = 0.2$, $\mu_h = -5$, $\rho = 0$, $\sigma_h = 0.4$, $\kappa = 0.0025$, $\mu_\eta = 0.8$, $\sigma_\eta = 1.2$. For notational simplicity, we write $\ln(c_{n,t})$ as h_t. $\kappa = 0.0025 \approx 1/(78 \times 5)$ suggests that we assume approximately there is one jump each week. When extracting latent volatility, we focus our attention on the one-month interval (i.e., $T = 22$, assuming that there are 22 trading days within one month). In toal we have $n = 22 \times 390 = 8580$ returns. Besides, the jump component is incorporated in the transition dynamics of the volatility process using $\mathbf{1}_{\{t=t^\circ\}}$ to ensure that latent spot volatility transition dynamics in DGP are reconciled with corresponding specifications in **Model 2**. In other words, we only consider jumps that happen at the end of each local estimation block.

First, to check the performance of the Bayesian method in estimating model parameters, we report the posterior means, posterior standard errors for all parameters. The corresponding results are summarized in the following figures with the vertical red dashed lines indicating the location of posterior means.

[Place Figure 4 about here]

Second, to check the performance of the smoothing and filtering methods in extracting latent volatility, we plot the simulated (true) volatility, smoothed volatility and nonparametric estimation of volatility in Figures:

[Place Figure 5 about here]

5.3 Experiment 3

In this experiment, we simulate data according to,

$$\begin{aligned} dX_t &= \exp\left(\tilde{h}_t/2\right) dW_t \\ \tilde{h}_t &= \mu_h + h_t + s_t \\ dh_t &= -\kappa_h h_t dt + \sigma_h dB_t \\ \mathbb{E}\left[W_t B_t\right] &= \rho = 0 \\ s_t &= 12(1-b)\left(t - \lfloor t- \rfloor - \frac{1}{2}\right)^2 + b \end{aligned}$$

where s_t takes the quadratic functional form as described for **Model 3**, thus

$$s_t = 12(1-b)\left(t - \lfloor t- \rfloor - \frac{1}{2}\right)^2 + b.$$

All the parameters except for the parameter that specifies intraday volatility pattern inherit from **Experiment 1**. We specify b for different cases: one for $b = 0.4$, which represents the case when daily volatility exhibits relatively strong diurnal U-shaped patterns; while the other for $b = 0.8$, which represents the case when daily volatility exhibits relatively weak diurnal U-shaped patterns.

First, to check the performance of the Bayesian method in estimating model parameters, we report the posterior means, posterior standard errors for all parameters. The corresponding results are summarized in the following figures with the vertical red dashed lines indicating the location of posterior means.

[Place Figure 6 about here]

[Place Figure 7 about here]

Second, to check the performance of the smoothing and filtering methods in extracting latent volatility, we plot the simulated (true) volatility, smoothed volatility and nonparametric estimation of volatility in Figures:

[Place Figure 8 about here]

5.4 Experiment 4

In this experiment, we simulate data according to,

$$\begin{aligned} dX_t &= \exp\left(\tilde{h}_t/2\right) dW_t \\ \tilde{h}_t &= \mu_h + h_t + a_{t^\circ} \\ dh_t &= -\kappa_h h_t dt + \sigma_h dB_t \\ \mathbb{E}\left[W_t B_t\right] &= \rho = 0 \\ a_{t^\circ} &= \sum_{q=1}^{Q} \sum_{l=0}^{L} \mathbf{1}_{t^\circ ql} \alpha_{ql} \\ \alpha_{ql} &= \tilde{\alpha}_q \exp\left\{-\tilde{\beta}_q l\right\} \end{aligned}$$

where a_{t° inherits the corresponding specification directly from **Model 4** such that

$$a_{t^\circ} = \sum_{q=1}^{Q} \sum_{l=0}^{L} \mathbf{1}_{t^\circ ql} \alpha_{ql}$$

where $\mathbf{1}_{t^\circ ql}$ is indicator for news type q at t° with $l = 0, 1, \ldots, L$ (i.e., $\mathbf{1}_{t^\circ ql} = 1$ if it is l periods after type q announcement made at time $t^\circ - \frac{l}{M}$ and 0 otherwise). We currently focus on the single-type announcement effect, thus $Q = 1$ in this experiment. Besides, L is the parameter capturing the longest length of periods for which the announcement effect survives once it is made. L is set equal to 5 in this experiment. To simulate announcement indicators, we assume that the announcement happens at a rate approximately equal to $0.004 \approx 1/(78 \times 3)$. We specify the announcement effects by setting $\tilde{\alpha}_q = 0.8$, $\tilde{\beta}_q = 0.1$.

First, to check the performance of the Bayesian method in estimating model parameters, we report the posterior means, posterior standard errors for all parameters. The corresponding results are summarized in the following figures with the vertical red dashed lines indicating the location of posterior means.

[Place Figure 9 about here]

Second, to check the performance of the smoothing and filtering methods in extracting latent volatility, we plot the simulated (true) volatility, smoothed volatility and nonparametric estimation volatility in Figures:

[Place Figure 10 about here]

Remark 5.2 (Comparison of jumps and announcement impulse) *For the data generating process of **Experiment 3** described above associated with **Model 3** and the data generating process*

*of **Experiment 4** associated with **Experiment 4**, we exploit notation $\mathbf{1}_{\{t=t^\circ\}}J_t\eta_t$ and a_{t° to capture announcement shocks and jump shocks in this continuous-time setting respectively. By writing a_{t° in equation $h_t = \mu + h_t + s_t + a_{t^\circ}$, in which for any $t \in (\lfloor t- \rfloor, t^\circ]$, $a_t = a_{t^\circ}$, we make data generated in this continuous-time scheme reconciled with model specified as in (18). By contrast, the introduced indicator function $\mathbf{1}_{\{t=t^\circ\}}$ selects jump shocks that happen between two consecutive local estimation blocks so that to make the simulated jump shocks reconciled with model specification as in $h_{t^\circ} = \phi h_{t^\circ - 1/M} + e_{t^\circ} + J_{t^\circ}\eta_{t^\circ}$, since jump shocks are implicitly assumed to be measured at the same frequency with the local estimation block based on this equation that establishes a model incorporating jump shocks.*

5.5 Experiment 5

In this experiment, we simulate data by accommodating all the components in **Experiments 1-3**.

$$\begin{aligned}
dX_t &= \exp\left(\tilde{h}_t/2\right) dW_t \\
\tilde{h}_t &= \mu_h + h_t + s_t \\
dh_t &= -\kappa_h h_t dt + \sigma_h dB_t + \mathbf{1}_{\{t=t^\circ\}} J_t \eta_t \\
\mathbb{E}[W_t B_t] &= \rho = 0 \\
J_t &\sim \text{Bernoulli}(\kappa) \\
\eta_t &\sim \mathcal{N}\left(\mu_\eta, \sigma_\eta^2\right) \\
s_t &= 12(1-b)\left(t - \lfloor t- \rfloor - \frac{1}{2}\right)^2 + b
\end{aligned}$$

All the corresponding notations inherit directly those in **Experiments 1-3** and the values assigned to these parameters as well.

First, to check the performance of the Bayesian method in estimating model parameters, we report the posterior means, posterior standard errors for all parameters. The corresponding results are summarized in the following figures with the vertical red dashed lines indicating the location of posterior means.

[Place Figure 11 about here]

Second, to check the performance of the smoothing and filtering methods in extracting latent volatility, we plot the simulated (true) volatility, smoothed volatility and nonparametric estimation of volatility in Figures:

[Place Figure 12 about here]

5.6 Experiment 6

In this experiment, we simulate data by accommodating all the components in **Experiments 1-4** via following data generating process

$$\begin{aligned} dX_t &= \exp\left(\tilde{h}_t/2\right) dW_t \\ \tilde{h}_t &= \mu_h + h_t + s_t + a_{t^\circ} \\ dh_t &= -\kappa_h h_t dt + \sigma_h dB_t + \mathbf{1}_{\{t=t^\circ\}} J_t \eta_t \\ a_{t^\circ} &= \sum_{q=1}^{Q} \sum_{l=0}^{L} \mathbf{1}_{t^\circ ql} \alpha_{ql} \\ \alpha_{ql} &= \tilde{\alpha}_q \exp\left\{-\tilde{\beta}_q l\right\} \end{aligned}$$

All the corresponding notations inherit directly those in **Experiments 1-4** and the values assigned to these parameters as well.

First, to check the performance of the Bayesian method in estimating model parameters, we report the posterior means, posterior standard errors for all parameters. The corresponding results are summarized in the following figures with the vertical red dashed lines indicating the location of posterior means.

[Place Figure 13 about here]

Second, to check the performance of the smoothing and filtering methods in extracting latent volatility, we plot the simulated (true) volatility, smoothed volatility and nonparametric estimation of volatility in Figures:

[Place Figure 14 about here]

5.7 Results and discussions on model comparison

For all the experiments described in the previous part, we summarize corresponding results in the following table, where **Model i=1,2,3,4** and **DGP i=1,2,3,4** refer to model specifications and data generating processes discussed in the previous subsections respectively. For each data generating process **DGP i=1,2,3,4**, DIC based on conditional likelihood is reported along with the decomposed components in (25). In general, the model with a relatively smaller DIC should be preferred to the model with a relatively larger DIC (highlighted in bold). It is possible to see from this Monte Carlo experiment that standard deviance information criterion can in general select the alternative model associated with the true data generating process. Since for **DGP i=1,2,3**, it is by design that there is **no observed announcement indicators**, hence we do not compare **Model 4** with **Model i=1,2,3** for **DGP i=1,2,3**.

[Place Table 1 about here]

Remark 5.3 *The reason why* ***Model 1*** *and* ***Model 3*** *have much larger* DIC *when the true data is generated from stochastic volatility process with jumps (****DGP 2****) is that if there indeed exists "jump" components in the dynamic transition scheme (****DGP 2****) at relatively high frequency (for instance,* κ *specified in this experiment as* $\kappa \approx 0.01$ *suggests that approximately on average we have one volatility jump each day), model specifications without accommodating these potential "jumps" usually feature "poor" model fitting. We leave more discussions on this issue in the appendix.*

6 Empirical Study

This section discusses the empirical applications of our proposed estimation method of spot volatility. We see from discussions in extant literature that one prominent feature associated with spot volatility estimation by fixing local estimation window size is the "noise" introduced, which is also one motivating incentive for our proposed methodology using MCMC techniques to smooth out those "noises" of nonparametric estimation of spot volatility with local estimation-window size fixed. We first demonstrate some applications of the proposed methodology in tracking volatilities associated with individual assets. Then we discuss a broader application for quantifying private information closely connected with return volatility in high frequency in the studies of financial microstructure.

Data used for our empirical study is collected mainly from two sources. For the data corresponding to the U.S. equity market, we mainly collect it from the NYSE TAQ database (Trade and Quote database).[7] We follow the procedure suggested in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009), which has been encompassed in `highfrequency` package maintained at `CRAN` (Boudt, Kleen, and Sjørup, 2021). This procedure aims to eliminate non-zero trades and filter for valid sales conditions.[8] Aside from that, we merge trade entries that have the same timestamp into a single one. Thus, if there are multiple observations available for a specific timestamp, we take the median of these multiple observations as the corresponding observation associated with that timestamp. For the Chinese stock market, we mainly use one-minute price-level data of the CSI 300 index futures.

[7] This database contains intraday transactions data (trades and quotes). Generally there are 3 kinds of data products: **Trade & Quote Daily Product** (09/10/2003-present), **Trade & Quote Monthly Product** (01/01/1993-12/31/2014) and **NYSE Reg Sho Data** (01/01/2005-07/31/2007). In general, The TAQ Daily and Monthly data products are nearly identical whereas the key difference arises from that **Monthly Data Product** is delivered a whole month at a time, typically 60-90 days after the last trading day of the month. **Daily Data Product** is delivered one day at a time, hours after trading stops, and is available on WRDS the next day. We retain our focus on using **Trade & Quote Daily Product** as it is actively maintained and is of relatively higher quality in the sense that sampling intervals are more refined for the **Trade & Quote Daily Product** such that timestamps are provided at milliseconds (10^{-3} `secs`) granularity through March 2015, and in microseconds (10^{-6} `secs`) starting in April 2015.

[8] For more about sales conditions, readers may refer to NYSE online documentation about daily TAQ trade files at https://www.nyse.com/publicdocs/nyse/data/Daily_TAQ_Client_Spec_v3.3.pdf. By implementing this procedure, we essentially retain our focus on stocks exchanged in a single exchange market (for instance, `T/Q` refers to the NASDAQ exchange market).

6.1 Extracting spot volatility for individual asset

For the sake of mitigating the effect of microstructure "noise", it is a standard thumb rule in literature to have price data sampled at one-minute sampling frequency, (see Zhang, Mykland, and Aït-Sahalia, 2005). For this one-minute sampling scheme, we calculate the corresponding one-minute return within trading hours each day from 9:30 to 16:00 by taking the close-open return (i.e. close-open log price difference) as the return for the interval from 9:30 to 9:31 and close-close returns (i.e. close-close log price difference) as the returns for the following sampling intervals up till to 16:00, which is usually the ending trading hours each day.[9] More specifically, we should expect 391 observed data at price level and accordingly $6.5 \times 60 = 390$ calculated one-minute returns each day. We use two market indices (S&P 500 index ETF representing the U.S. market and CSI 300 index futures representing the Chinese market respectively) and one individual stock (Apple Inc.) as the data source. We demonstrate both the extracted volatility (red dashed line) and the corresponding nonparametric estimation of volatility (blue solid line) with local estimation block size fixed (i.e. fixed $k = 5$, every 5-minutes) as follows,

S&P 500 index ETF from TAQ in November 2015

[**Place Figure 15 about here**]

Apple Inc. Stock Price in August 2017

[**Place Figure 16 about here**]

CSI 300 Index futures in January 2020

[**Place Figure 17 about here**]

CSI 300 Index futures in August 2020

[**Place Figure 18 about here**]

For each index, the extracted volatility is based on **Model 5**, the nested model including all the specifications of **Models 1-3**. We also make a model comparison across **Models 1-3** based on DIC to check whether we need to extend the benchmark model specification (**Model 1**) to incorporate either jumps in the volatility dynamics or the diurnal pattern of the latent volatility process. Specifically for each year (S&P 500 index ETF in 2015, Apple Inc. Stock Price in 2017, and CSI 300 Index futures in 2020) and each month we apply our proposed modeling framework and methodology to extract volatility based on model specifications corresponding to **Model 1**, **Model 2**, and **Model 3**. Results are summarized in Table 2, Table 3, and Table 4 respectively.

[9] Timing scheme for trading hours corresponding to CSI 300 index futures. From 2010 to 2015, we have pirce-level data sampled at one-minute frequency from 9:15 to 15:15 each day; while from 2016 to 2021, we have data sampled at one-minute frequency from 9:30 to 15:00 each day.

[Place Table 2 about here]

[Place Table 3 about here]

[Place Table 4 about here]

These results suggest that for most cases, we need to allow for the jump specifications in modeling the latent volatility process and that is why we apply **Model 5** that nests both the jump specifications and intraday diurnal pattern specification for extracting volatility. We also summarize the posterior mean and standard deviations of parameter based on MCMC draws as follows. The associated diagnostic check using MCMC trace plot is summarized in the appendix C.3.

[Place Table 5 about here]

[Place Table 6 about here]

[Place Table 7 about here]

By comparing the posterior mean of b summarized in the last column of Table 5, Table 6, and Table 7, we find that: the volatility of individual stock return exhibits a relatively stronger (lower b) diurnal pattern while the volatility associated with market indices exhibits a relatively weaker (higher b) diurnal pattern in the corresponding periods.

6.2 Spot volatility, liquidity, and strategic value of information

In this section, we specifically focus on one application of our proposed methodology for extracting volatility to study the financial market microstructure, thus the value of information in strategic trading.

6.2.1 Value of information in continuous-time setting

Volatility plays a vital role in modern financial market microstructure studies since essentially speaking, all the studies about the financial market structure are about recovering insiders' private trading information in comparison to relatively uninformed liquidity traders. In other words, the central questions of paramount importance are what the value of asset-specific information is to a strategic trader and how to practically quantify such a kind of value, which is also the amount investors would pay for information. One motive intuition for quantifying the value of information is to use two components: (i) the extent to which specific information can offer speculator the reduction in uncertainty; (ii) liquidity associated with assets for which the acquired information can be used to trade quickly without generating adverse effects on the assets' prices. Inspired by

Grossman and Stiglitz (1980) and the subsequent studies in Kyle (1985) and Back (1992), this intuition has been justified using the ratio of uncertainty about the asset's fundamentals and the asset's illiquidity measure. Given the recent finding in empirical literature (Collin-Dufresne and Fos, 2015; Kacperczyk and Pagnotta, 2019; Akey, Grégoire, and Martineau, 2022) that private information is hardly reflected by equity prices, Kadan and Manela (2024) extend the modeling framework of Kyle (1985) and Back (1992) (comprehensively summarized in Back (2017)) and proposes that in equilibrium the ex-ante dollar expected profits of informed trader over a specific interval indexed from 0 to 1 can serve as the measure of value of information to strategic trader, which is given by [10]

$$\Omega = \frac{\sigma_v^2}{\lambda} P_0, \tag{27}$$

where σ_v^2 characterizes volatility associated with private information of informed traders and P_0 refers to the initial price of the specific asset over this timing interval. Besides, we follow the convention in literature as in Back (2017) assuming that private information is denoted by $\tilde{v}$ and follows log-normal distribution such that $\ln \tilde{v} \sim \mathcal{N}\left(\mu, \sigma_v^2\right)$. λ in (27) is widely known as Kyle's Lambda, initially proposed in Kyle (1985), as the measure of sensitivity of assets' return to share order flow (Lee and Ready, 1991; Ellis, Michaely, and O'Hara, 2000; Holden and Jacobsen, 2014). Kyle's lambda serves as one alternative measure reflecting financial market turbulence, which is usually high during periods in which the whole financial market is exposed to a systematic crisis such as the 2008 financial depression and more recent years Covid-19 global pandemic crisis. As we can see from (27) that σ_v^2 as the major component characterizing the value of information to strategic trader also measures the magnitude of reduction in uncertainty that speculator would have if had acquired corresponding information. More importantly, as we will see in the following discussion that although σ_v^2 is originally the measure of uncertainty associated with information (for instance, at price level), it can be directly used as the measure of volatility associated with the logarithmic return of asset.

In standard literature corresponding to financial market microstructure, it is usually assumed that observed cumulative share orders in the continuous-time modeling framework, denoted by Y_t, can be decomposed into cumulative share orders of informed trader, denoted by X_t; and cumulative share orders of uninformed trader, denoted by Z_t, thus

$$Y_t = X_t + Z_t. \tag{28}$$

Dynamics specified for is directly characterized via Brownian motion as $dZ_t = \sigma_z dB_t$. Then results contained in theorem 3 and example 2 in Back (1992) and Back (2017) suggest that in equilibrium

$$\frac{dP_t}{P_t} = \lambda dY_t \quad \text{and} \quad dX_t = \frac{\frac{\ln \bar{v} - \mu}{\lambda} - Y_t}{1 - t} dt, \quad \bar{v} = \mathbb{E}\left[\tilde{v}\right], \tag{29}$$

[10] Alternatively, it is possible to interpret 0 as the starting point of specific interval while 1 as the ending point of specific interval.

where P_t refers to price process associated with target asset in equilibrium. Then it is straightforward to see that

$$\begin{aligned}
\frac{dP_t}{P_t} &= \lambda dY_t \\
&= \lambda dX_t + \lambda dZ_t \\
&= \lambda dX_t + \frac{\sigma_v}{\sigma_z}\sigma_z dB_t \\
&= \lambda \frac{\frac{\ln \tilde{v}-\mu}{\lambda} - Y_t}{1-t} dt + \sigma_v dB_t,
\end{aligned}$$

which yields one important implication for the empirical strategy that σ_v^2 as the measure of volatility associated with private information can be used as the proxy for volatility associated with logarithmic return. Specifically, for two consecutive asset prices P_{τ_i} and $P_{\tau_{i-1}}$, by applying log-linearization we have

$$r_i := \ln\left(\frac{P_{\tau_i}}{P_{\tau_{i-1}}}\right) \approx \frac{P_{\tau_i} - P_{\tau_{i-1}}}{P_{\tau_{i-1}}},$$

where the R.H.S. of the equation above can be regarded as the discretized approximation of $\frac{dP_t}{P_t}$.

6.2.2 Estimating value of information using spot volatility

In this section we discuss empirical strategies using spot volatility to estimate value of information. The general idea is simple: suppose we have a specific way to estimate Kyles lambda, denoted by $\hat{\lambda}$, then we can construct a spot information value associated with informed traders by using our proposed spot volatility estimation as the proxy of asset return volatility over each tiny interval using Bayesian techniques based upon the fixed-k inference theory. Thus

$$\hat{\Omega} = \frac{\hat{\sigma}_v^2}{\hat{\lambda}} P_0. \tag{30}$$

This will be a more desirable method in comparison to the widely adopted alternative using integrated volatility (i.e. annualized realized volatility) instead since the model developed in Back (1992) is naturally a continuous-time extension of the discrete-time model of Kyle (1985) and accordingly spot volatility should be preferred rather than integrated volatility. In practice, one feasible method to estimate λ is by regressing the asset's return on share order flow over that interval. The theoretical justification for using regression to back out λ originated from modeling insider trading in the continuous-time setting initialized by Back (1992). Recently, this idea is discussed more comprehensively in Back (2017) and Kadan and Manela (2024).

$$r_i = \lambda y_i + \varepsilon_i, \tag{31}$$

where $r_i = p_{\tau_i} - p_{\tau_{i-1}} = \ln P_{\tau_i} - \ln P_{\tau_{i-1}}$ (log-return of assets over specific interval indexed by i) and $y_i = Y_{\tau_i} - Y_{\tau_{i-1}}$ (share order flow over specific interval indexed by i).

Construction of proxy for share order flow is still an active area corresponding to financial market microstructure, which is mainly based on designing trade classification algorithms to identify trading direction. We follow the convention in extant literature and borrow mainly the ideas from Holden and Jacobsen (2014) using following "order imbalance" as the proxy of share order flow,[11]

$$\text{Order Imbalance} = \frac{Buys - Sells}{Buys + Sells}, \tag{32}$$

where the trading classification scheme to identify the trading direction, Buys (+1) and Sells (−1), is inherited from Chakrabarty, Li, Nguyen, and Van Ness (2007) and Holden and Jacobsen (2014). Although there exists a fact that Kyle's lambda, λ, essentially of specific format as the ratio of volatilities (volatility associated with private information relative to volatility associated with uniformed trading orders) should be a positive number in theory (Kyle, 1985; Back, 1992), the estimated Kyle's lambda from regression, $\hat{\lambda}$, does not necessarily meet this requirement in practice. We find empirically that this regression implementation works fine as a way to back out this information measure since there are just a few exceptions for which the obtained $\hat{\lambda}$ is negative throughout **Trade & Quote Daily Product** database (09/10/2003-present) as demonstrated in the following figure

[Place Figure 19 about here]

Besides, we use following figure to demonstrate that for each trading day within a specific month (November 2018 in Figure 20) as one way to justify the empirically documented positive relationship between logarithmic returns of target assets and corresponding share order flows

[Place Figure 20 about here]

and finally the estimated private information value associated insider trading is demonstrated as follows based on **Model 5**.

[Place Figure 21 about here]

7 Conclusion

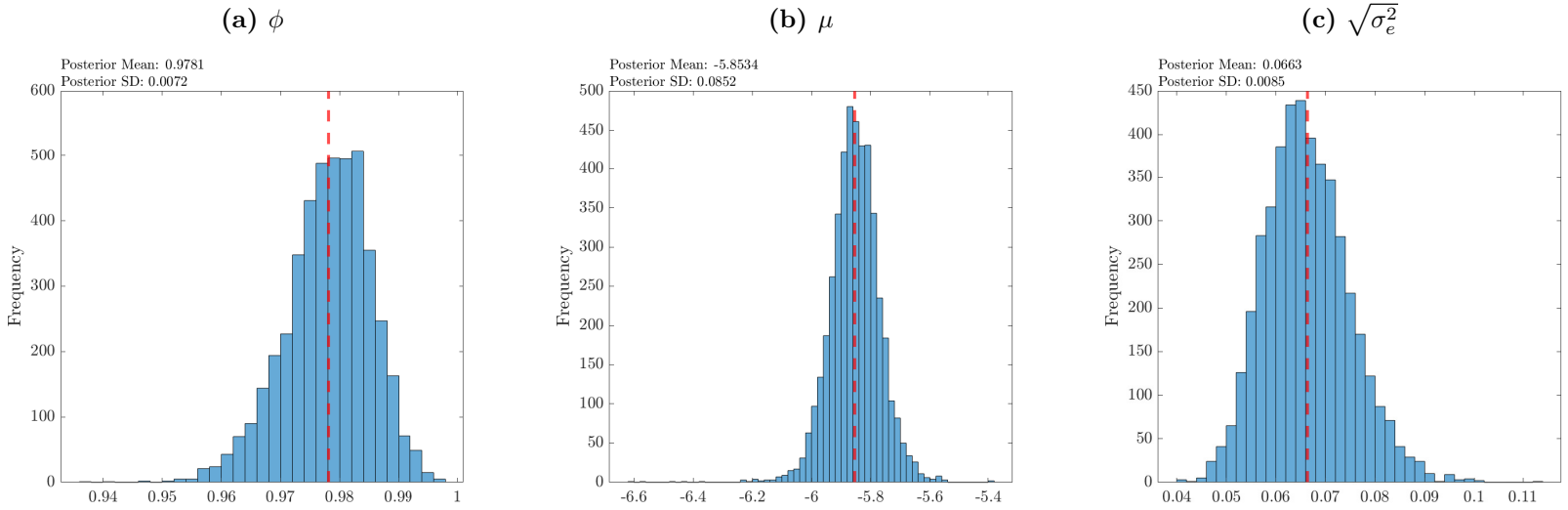
One of the main contributions from BLL2021QE is to establish the fixed-k inference theory for time-varying spot volatility. Although the fixed-k inference for spot volatility is applicable as we do not require the local estimation window size increases as the sample size increases in the standard manner, once the local estimation window size is fixed, the corresponding nonparametric estimation

[11] We are grateful to Professor Craig W. Holden for kindly sharing their `SAS` codes.

of spot volatility is inconsistent and rather noisy. To handle this issue, in this paper we build several parametric models for spot volatility in high frequency based on the fixed-k theory. All our models can be cast into a nonlinear non-Gaussian state-space form. We then develop Bayesian methods to estimate alternative model specifications, extract spot volatility, and compare alternative models. The main advantage of the method we propose in this paper is that we can eliminate much noise induced from the fixed-k inference for spot volatility in the high-frequency setting. Simulation studies show that the proposed Bayesian methods work well and the underlying spot volatility can be accurately extracted. We empirically demonstrate how our proposed method works in practice by applying it to S&P 500 index ETF, Apple Inc. stock, and CSI 300 Index futures respectively to extract spot volatility. Then we discuss how the extracted volatility relates to the strategic value of information for the informed trader (i.e. insider) in the financial market in detail, which shows the practical usefulness of the proposed method for the studies of financial market microstructure.

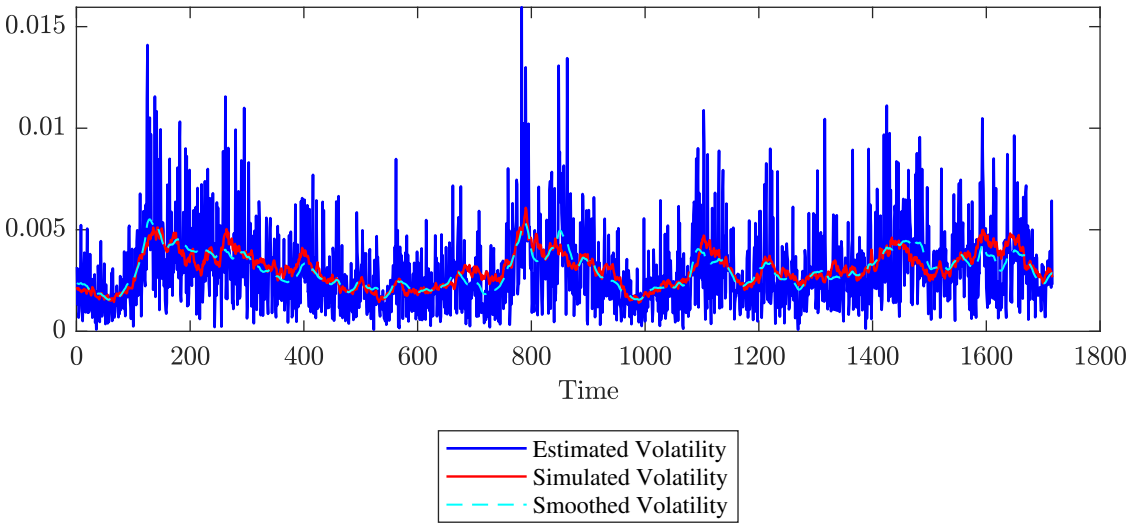
Figures and Tables

Figure 1. Posterior Summary of Parameters for **Model 1**



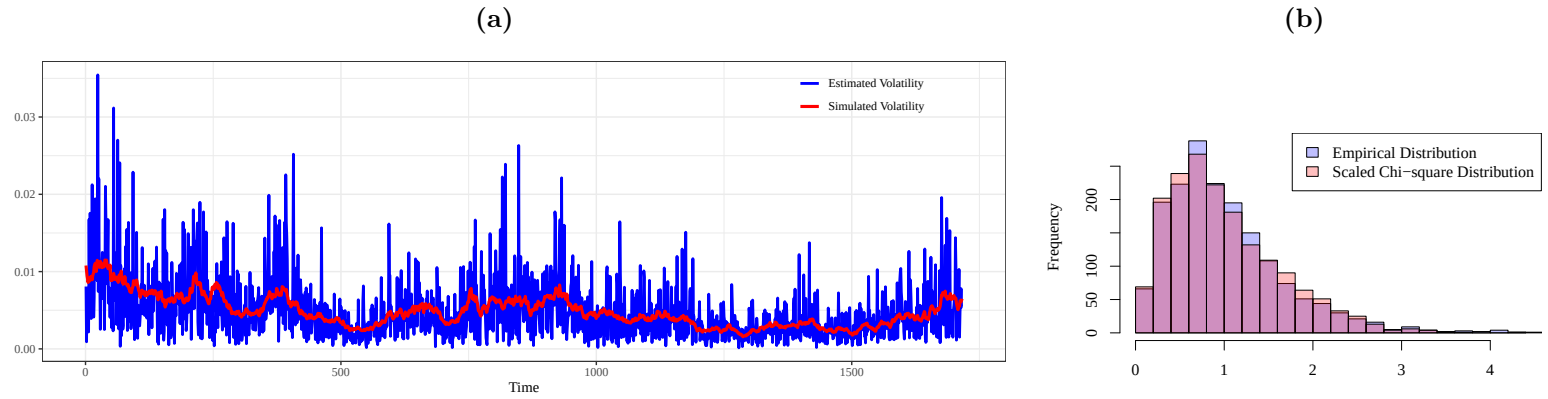
Note: Vertical red dashed lines indicate the location of the posterior mean of the corresponding parameters.

Figure 2



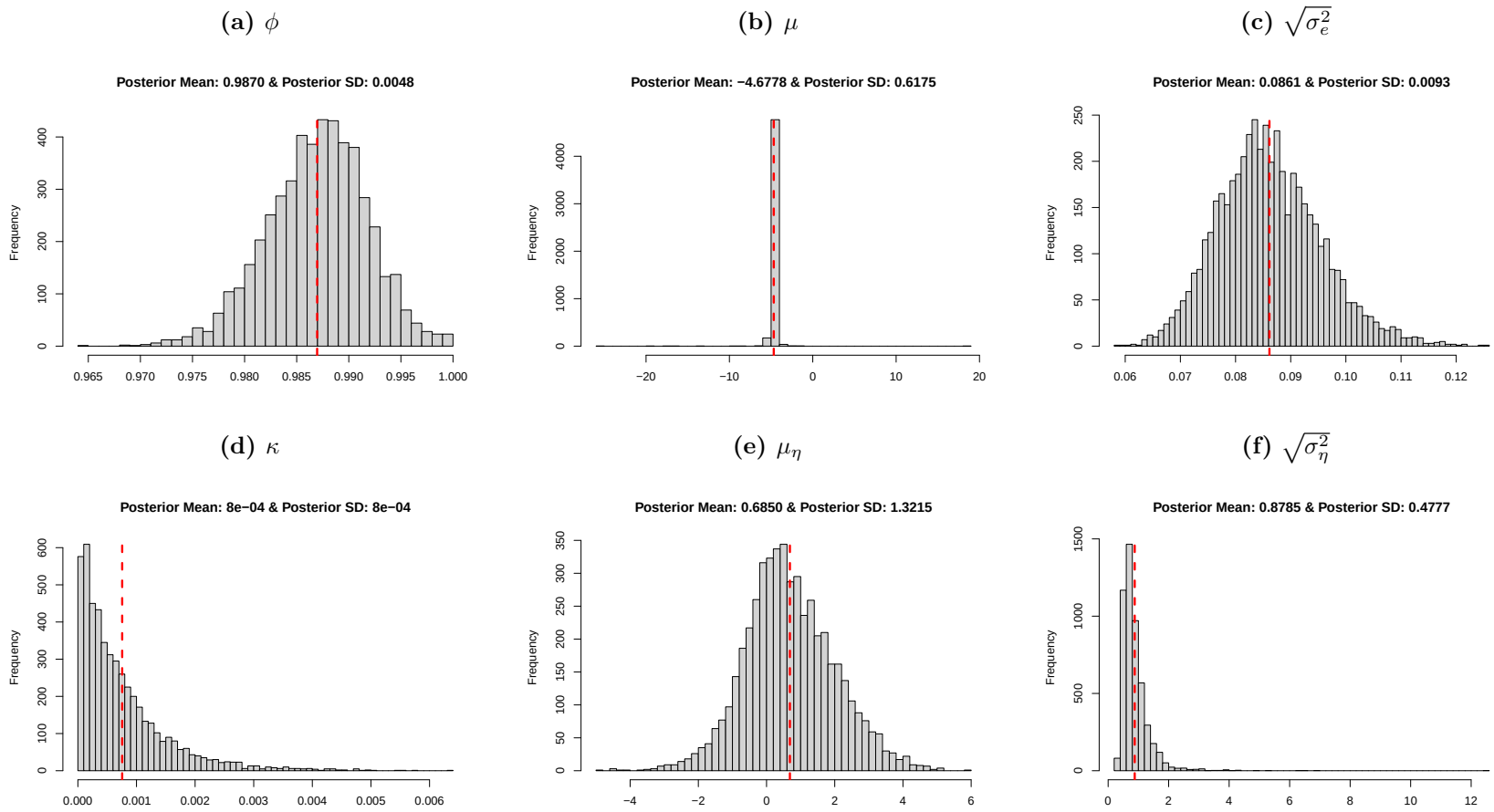
Note: In the figure above we jointly plot estimated volatility from fixed-k inference (blue solid line), smoothed volatility from MCMC (cyan dashed line) and true simulated volatility (red solid line) based on the DGP of **Experiment 1** described in section 5. As what we have claimed in the main context that sampling interval specified for this Monte Carlo experiment is $\Delta_n = 1/390$ and number of observations contained in each local estimation block is fixed at $k = 5$, thus for every 5 minutes we obtain the corresponding locally estimated volatility. Besides, for this sampling scheme we have $390/5 = 78$ local estimation blocks for each day and totally $22 \times 78 = 1716$ local estimation blocks as observations for 22 trading days within one month. For this experiment, we run totally 1000000 MCMC loops with the initial 100000 loops as burn-in samplings to be discarded. Every 20 samplings are saved for posterior analysis.

Figure 3



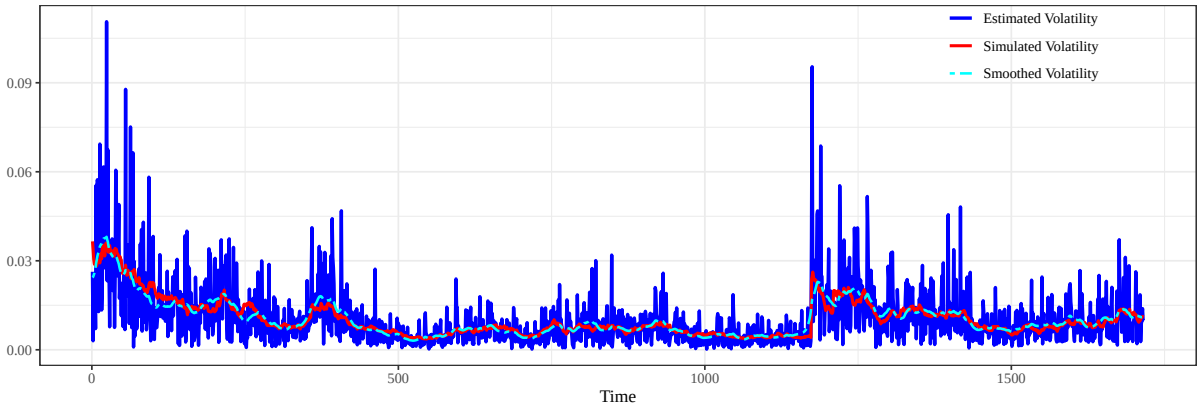
Note: This is an auxiliary figure for the demonstration purpose in companion with Figure 2. We jointly plot the estimated volatility from fixed-k inference (blue solid line) and the true simulated volatility (red solid line) based on DGP of **Experiment 1** in panel (a). Corresponding specifications are exactly the same as in the description of Figure 2. Panel (b) summarizes the histogram of gaps between the estimated volatility from fixed-k inference and the underlying true volatility process along with the histogram of random variables simulated from the scaled chi-square distribution.

Figure 4. Posterior Summary of Parameters for **Model 2**



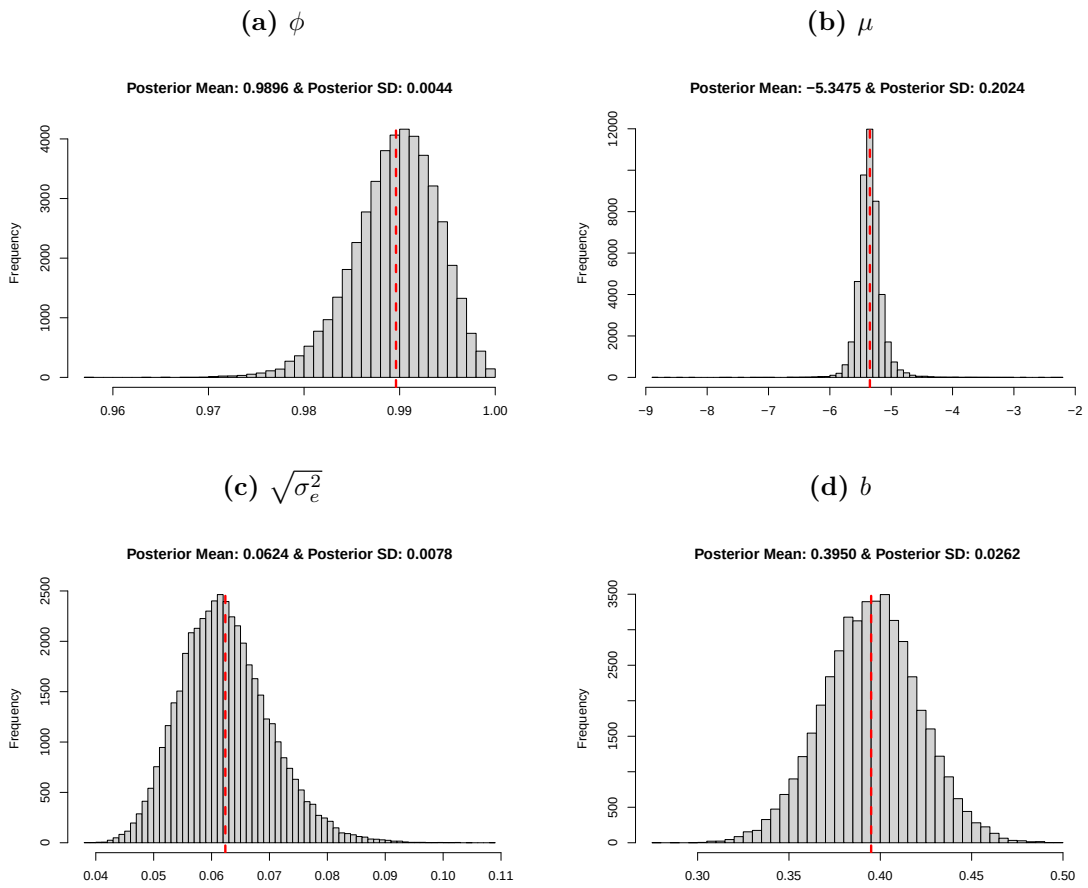
Note: Vertical red dashed lines indicate the location of the posterior mean of the corresponding parameters.

Figure 5



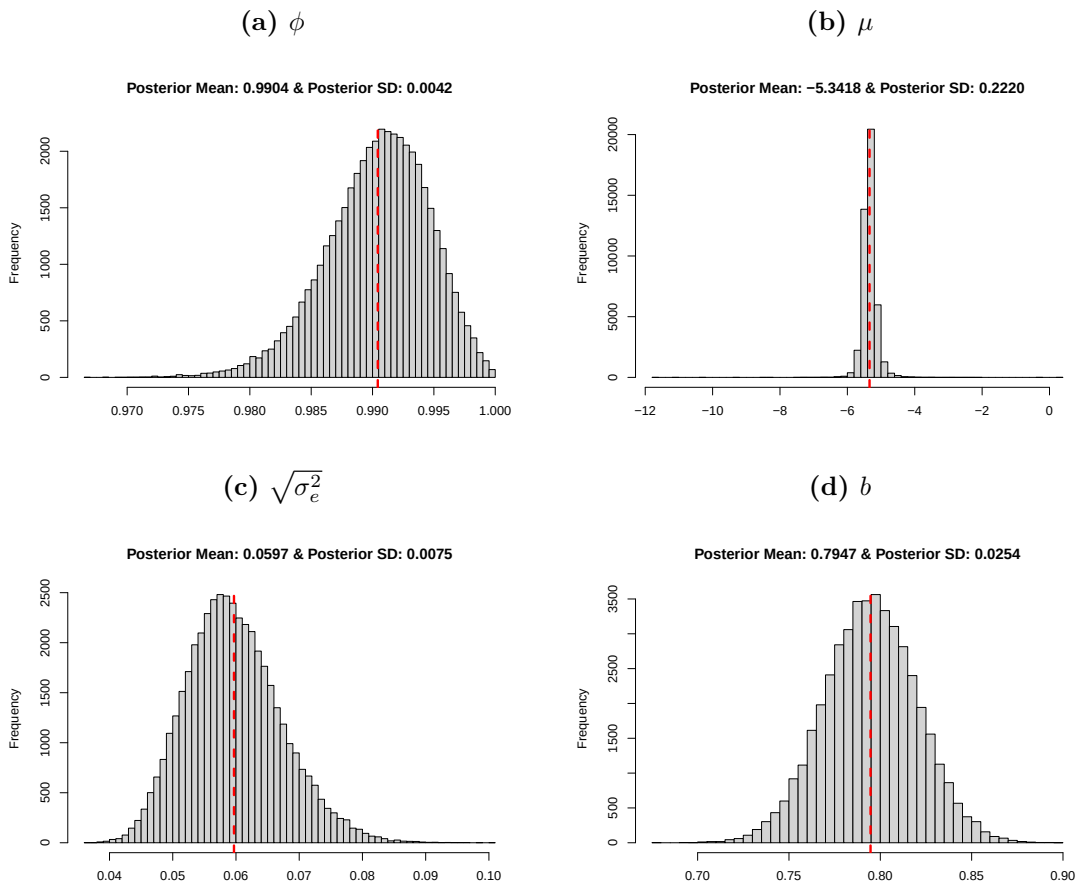
Note: In the figure above we jointly plot estimated volatility from fixed-k inference (blue solid line), smoothed volatility from MCMC (cyan dashed line) and true simulated volatility (red solid line) based on the DGP of **Experiment 2** described in section 5. Sampling interval for this experiment is still specified the same as in **Experiment 1**, $\Delta_n = 1/390$. We use fixed $k = 5$ observations contained in each local estimation block to obtain locally nonparametric estimation of volatility. But for this experiment, we accommodate 2 stochastic volatility factors and allow jumps in the transition scheme of one stochastic volatility factor. For this experiment, we run totally 1100000 MCMC loops with the initial 100000 loops as burn-in samplings to be discarded. Every 100 samplings are saved for posterior analysis.

Figure 6. Posterior Summary of Parameters for **Model 3** (strong diurnal pattern)



Note: Vertical red dashed lines indicate the location of the posterior mean of the corresponding parameters.

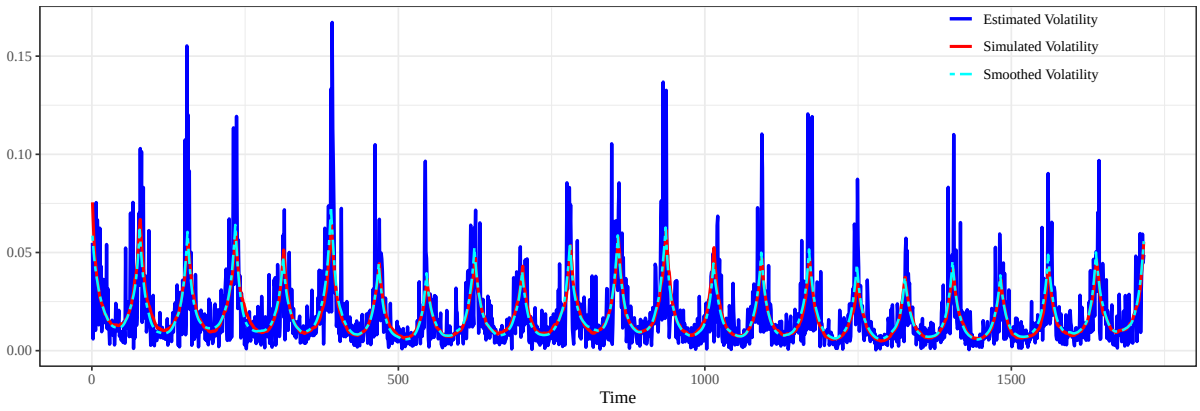
Figure 7. Posterior Summary of Parameters for **Model 3** (weak diurnal pattern)



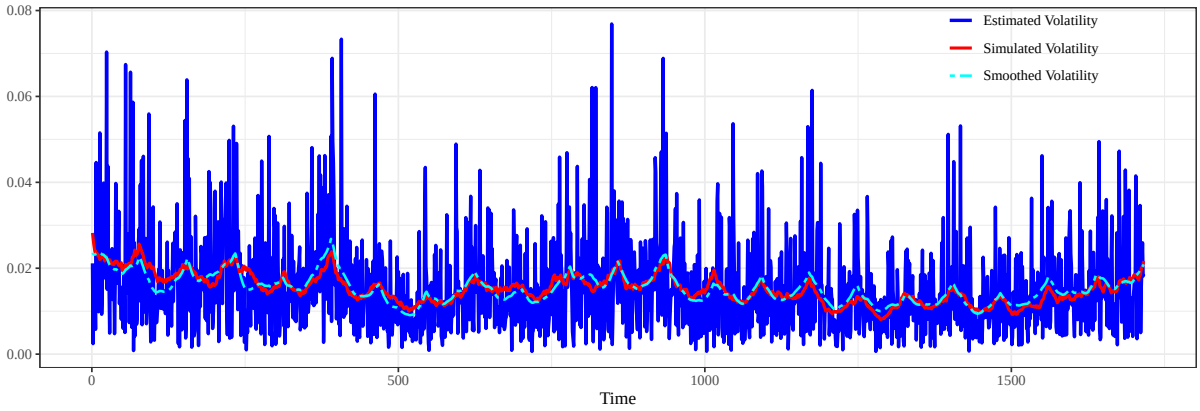
Note: Vertical red dashed lines indicate the location of the posterior mean of the corresponding parameters.

Figure 8

(a) Intraday-pattern parameter $b := 0.4$

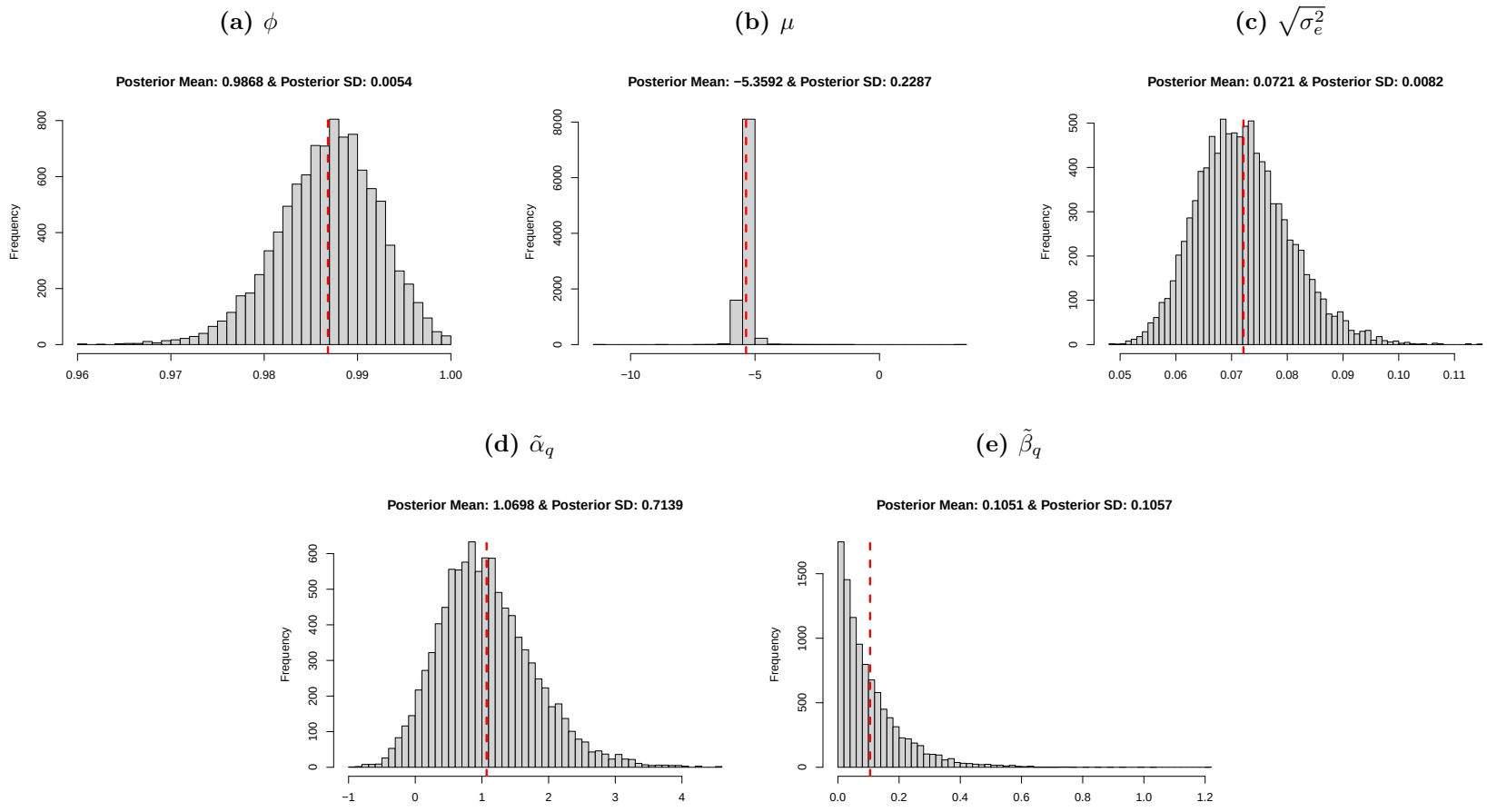


(b) Intraday-pattern parameter $b := 0.8$



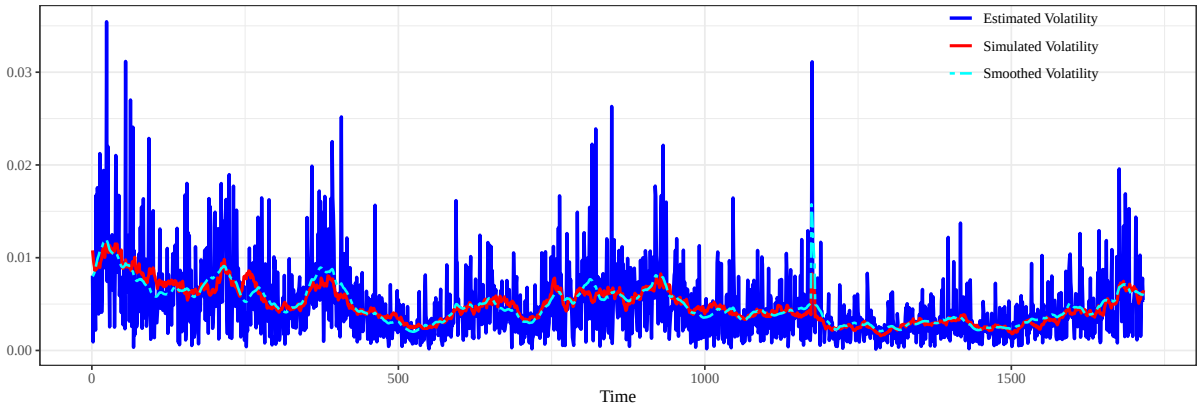
Note: In the figure above we jointly plot estimated volatility from fixed-k inference (blue solid line), smoothed volatility from MCMC (cyan dashed line) and true simulated volatility (red solid line) based on the DGP of **Experiment 3** described in section 5. Sampling interval for this experiment is still specified the same as in **Experiment 1**, $\Delta_n = 1/390$. We use fixed $k = 5$ observations contained in each local estimation block to obtain locally nonparametric estimation of volatility. But for this experiment, we accommodate 2 stochastic volatility factors and allow jumps in the transition scheme of one stochastic volatility factor. Besides, a quadratic functional form is exploited to capture the intraday volatility pattern as in discussing **Model 3** as one alternative model specification.

Figure 9. Posterior Summary of Parameters for **Model 4**



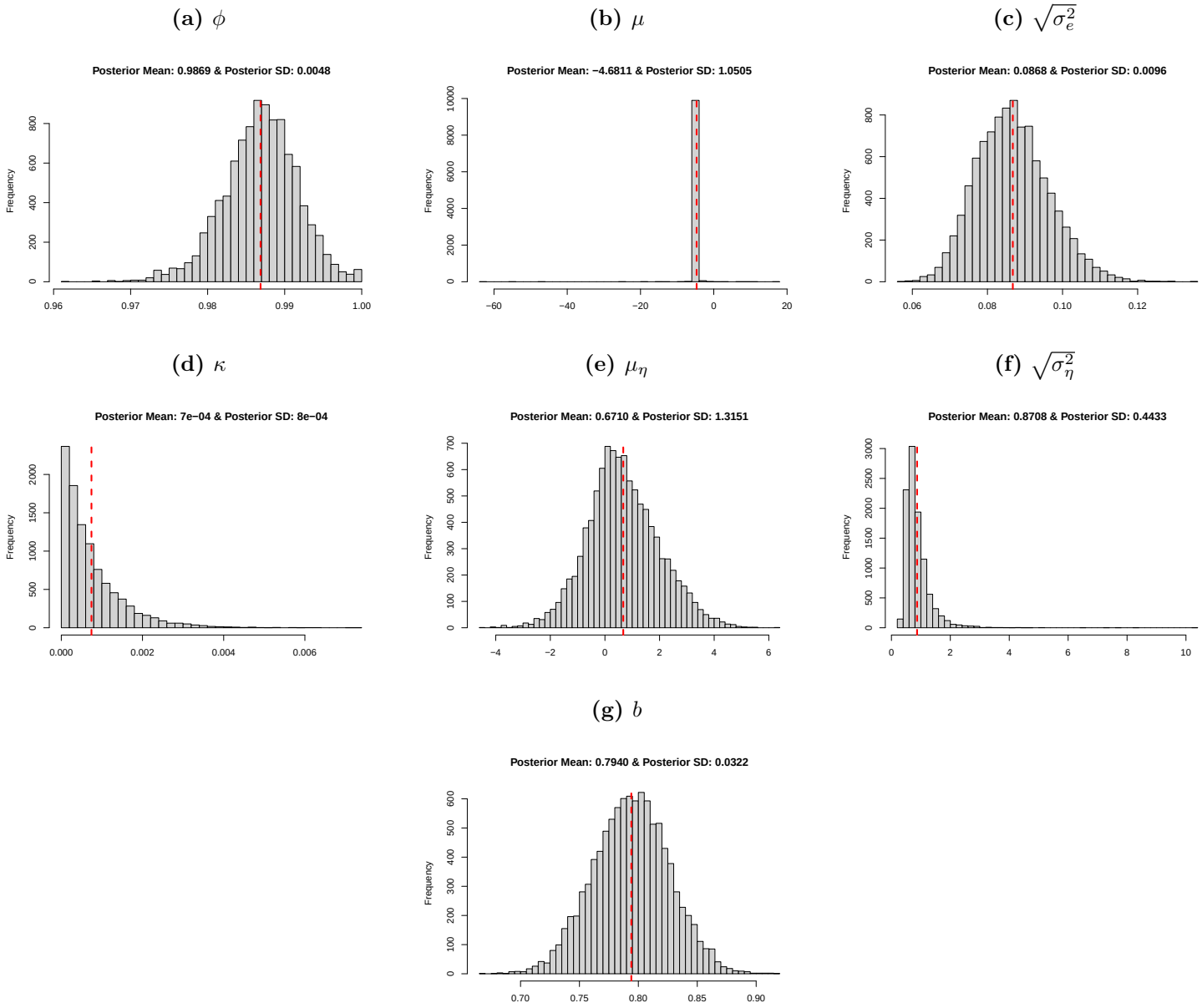
Note: Vertical red dashed lines indicate the location of the posterior mean of the corresponding parameters.

Figure 10



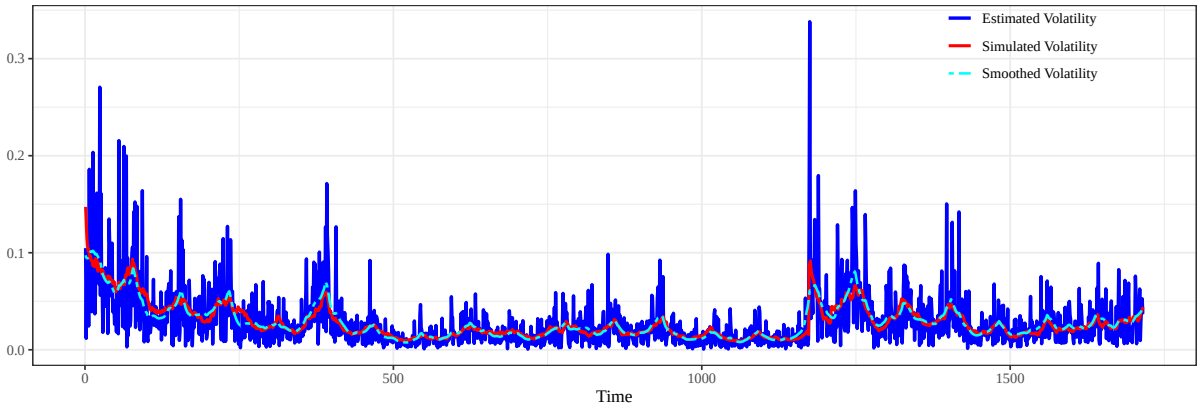
Note: In the figure above we jointly plot estimated volatility from fixed-k inference (blue solid line), smoothed volatility from MCMC (cyan dashed line) and true simulated volatility (red solid line) based on the DGP of **Experiment 4** described in section 5. Sampling interval for this experiment is still specified the same as in **Experiment 1**, $\Delta_n = 1/390$. We use fixed $k = 5$ observations contained in each local estimation block to obtain locally nonparametric estimation of volatility. Currently for this experiment, single announcement is introduced but it is allowed for these announcements to take effects for several periods. Specifically, we assume that once an announcement is made at specific timing point, it can take effect up to $L = 5$ periods.

Figure 11. Posterior Summary of Parameters for **Model 5**



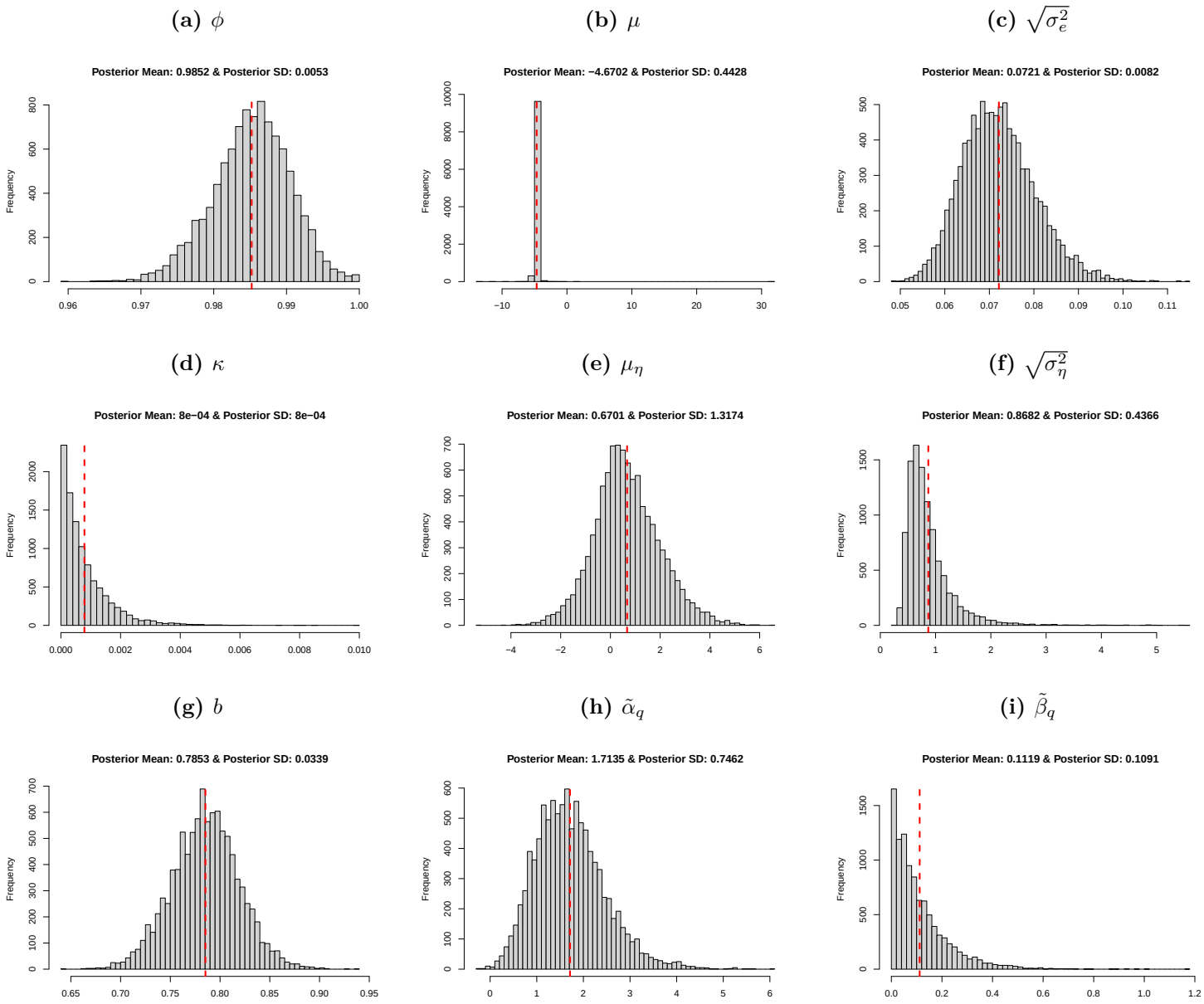
Note: Vertical red dashed lines indicate the location of the posterior mean of the corresponding parameters.

Figure 12



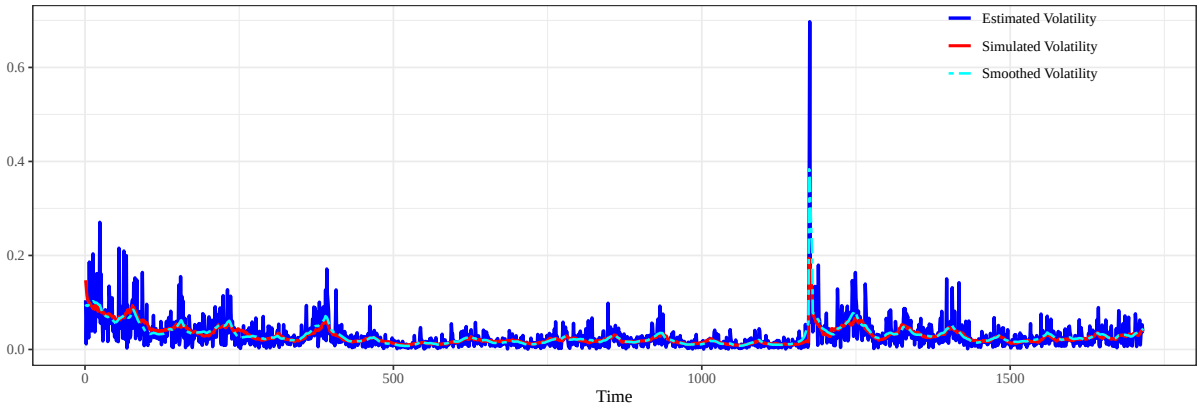
Note: In the figure above we jointly plot estimated volatility from fixed-k inference (blue solid line), smoothed volatility from MCMC (cyan dashed line) and true simulated volatility (red solid line) based on the DGP of **Experiment 5** described in section 5. Sampling interval for this experiment is still specified the same as in **Experiment 1**, $\Delta_n = 1/390$. We use fixed $k = 5$ observations contained in each local estimation block to obtain locally nonparametric estimation of volatility. Currently for this experiment, single announcement is introduced but it is allowed for these announcements to take effects for several periods. Specifically, we assume that once an announcement is made at specific timing point, it can take effect up to $L = 5$ periods.

Figure 13. Posterior Summary of Parameters for **Model 6**



Note: Vertical red dashed lines indicate the location of the posterior mean of the corresponding parameters.

Figure 14

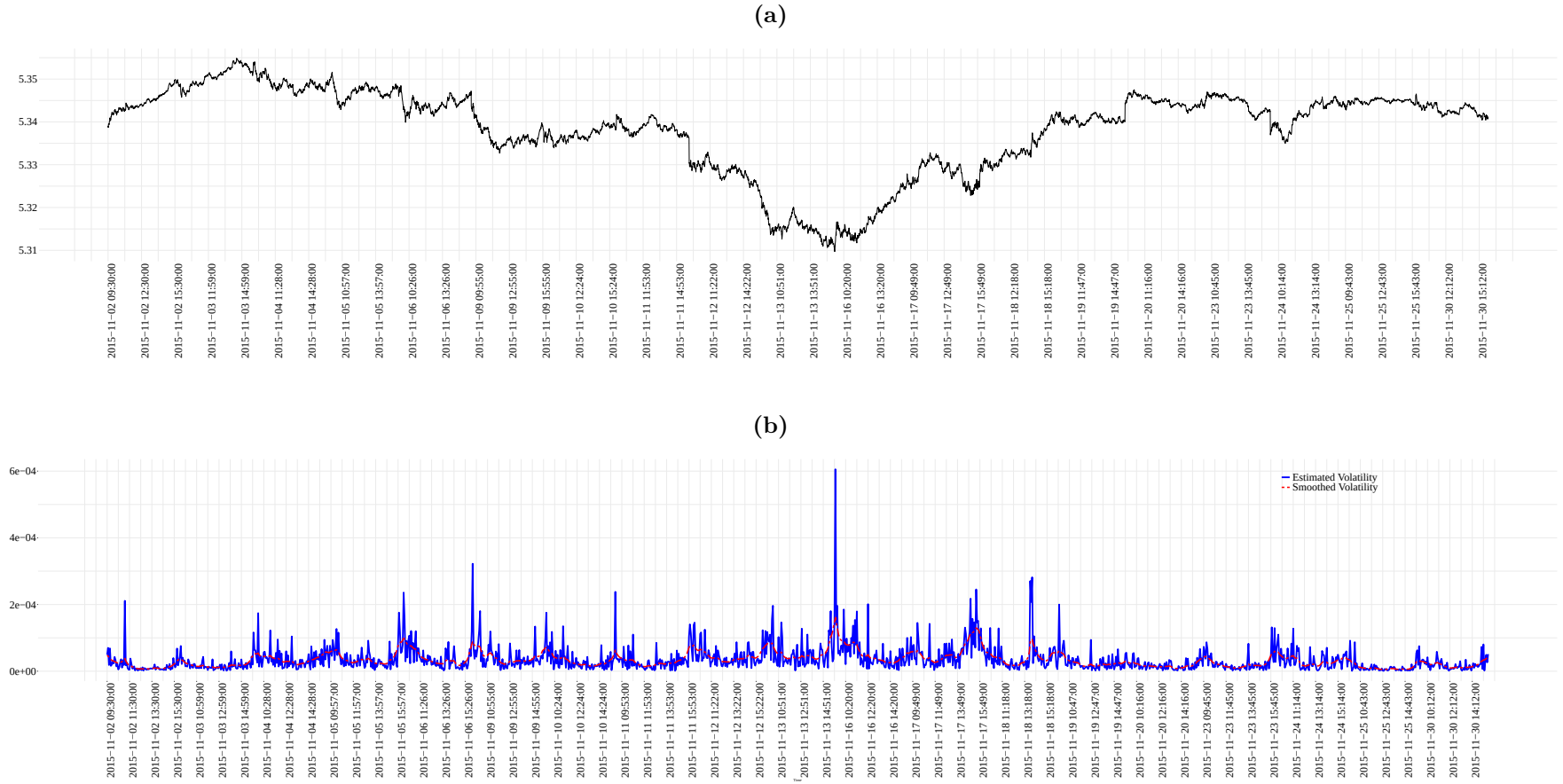


Note: In the figure above we jointly plot estimated volatility from fixed-k inference (blue solid line), smoothed volatility from MCMC (cyan dashed line) and true simulated volatility (red solid line) based on the DGP of **Experiment 6** described in section 5. Sampling interval for this experiment is still specified the same as in **Experiment 1**, $\Delta_n = 1/390$. We use fixed $k = 5$ observations contained in each local estimation block to obtain locally nonparametric estimation of volatility. Currently for this experiment, single announcement is introduced but it is allowed for these announcements to take effects for several periods. Specifically, we assume that once an announcement is made at specific timing point, it can take effect up to $L = 5$ periods.

Table 1

	DGP 1			**DGP 2**			**DGP 3**			**DGP 4**		
	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D
Model 1	**-3310.18**	-3470.29	80.05	1222718.45	1209944.77	6386.84	-3243.91	-3745.30	250.69	-3243.12	-3444.81	100.85
Model 2	-3308.32	-3498.27	94.98	**-3210.47**	-3568.32	178.92	-3241.08	-3749.67	254.30	-3243.14	-3469.41	113.13
Model 3	-3297.71	-3473.93	88.11	56578.51	54348.87	1114.82	**-3329.11**	-3470.51	70.70	-3242.52	-3446.71	102.10
Model 4	–	–	–	–	–	–	–	–	–	**-3272.01**	-3496.99	112.49

Figure 15

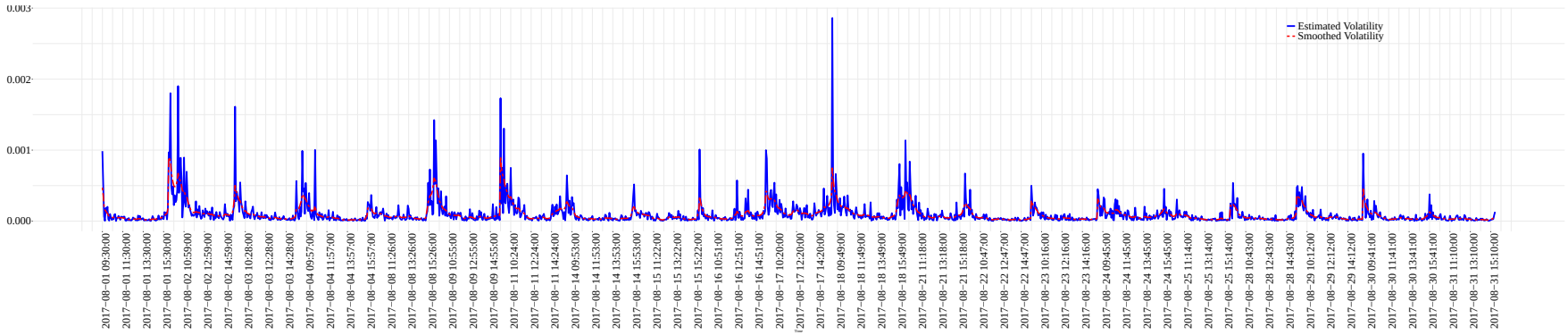


Note: In the figure above we report empirical results using price of S&P 500 index ETF sampled at the one-minute frequency. Panel ((a)) plots the log-price level of S&P 500 index. Panel ((b)) plots the nonparametric estimates of spot volatility (blue solid line) and the smoothed volatility estimate (red dashed line) using our Bayesian techniques. Results demonstrated in this figure correspond to S&P 500 index ETF within November 2015.

Figure 16

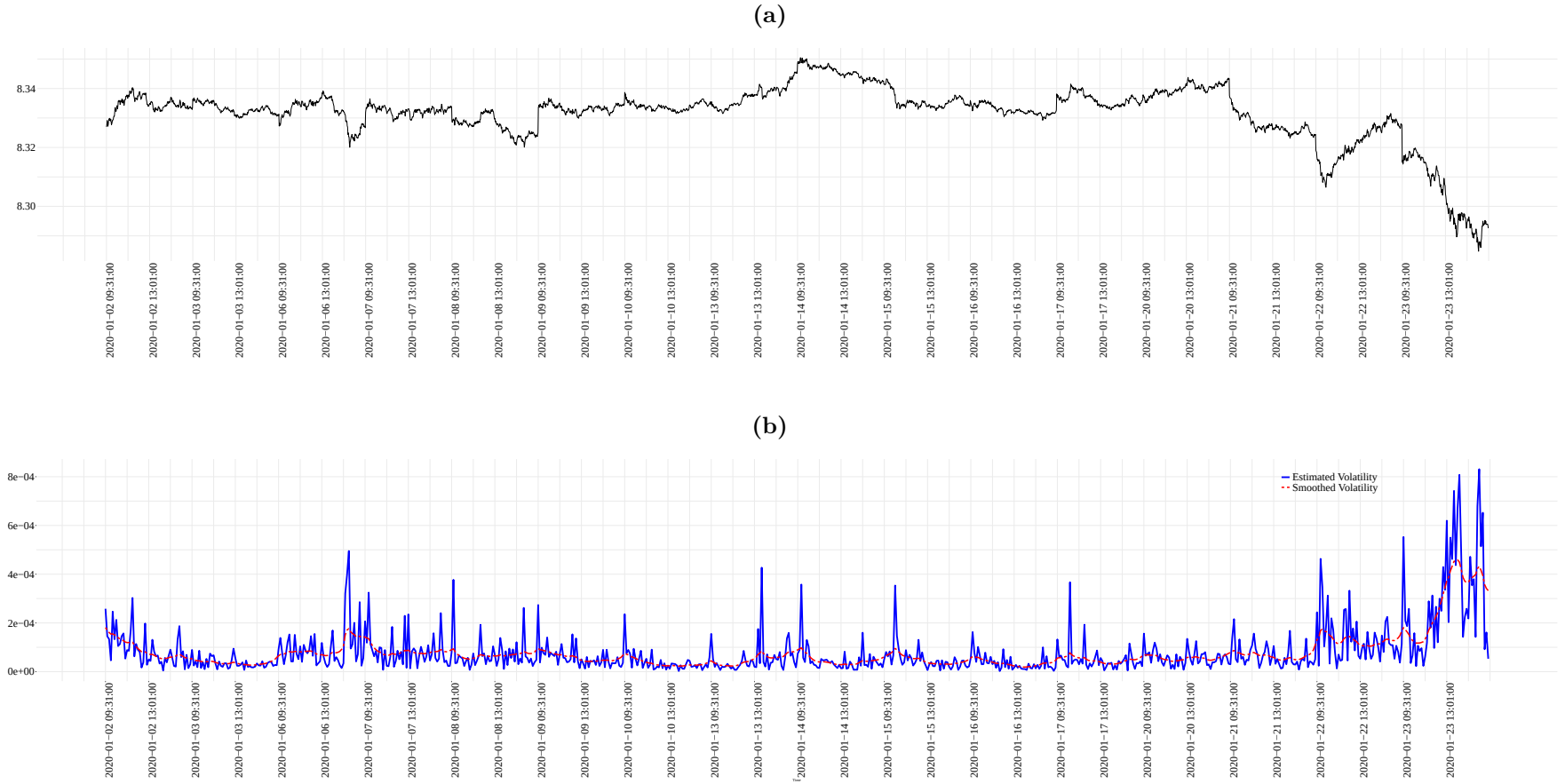
(a)

(b)



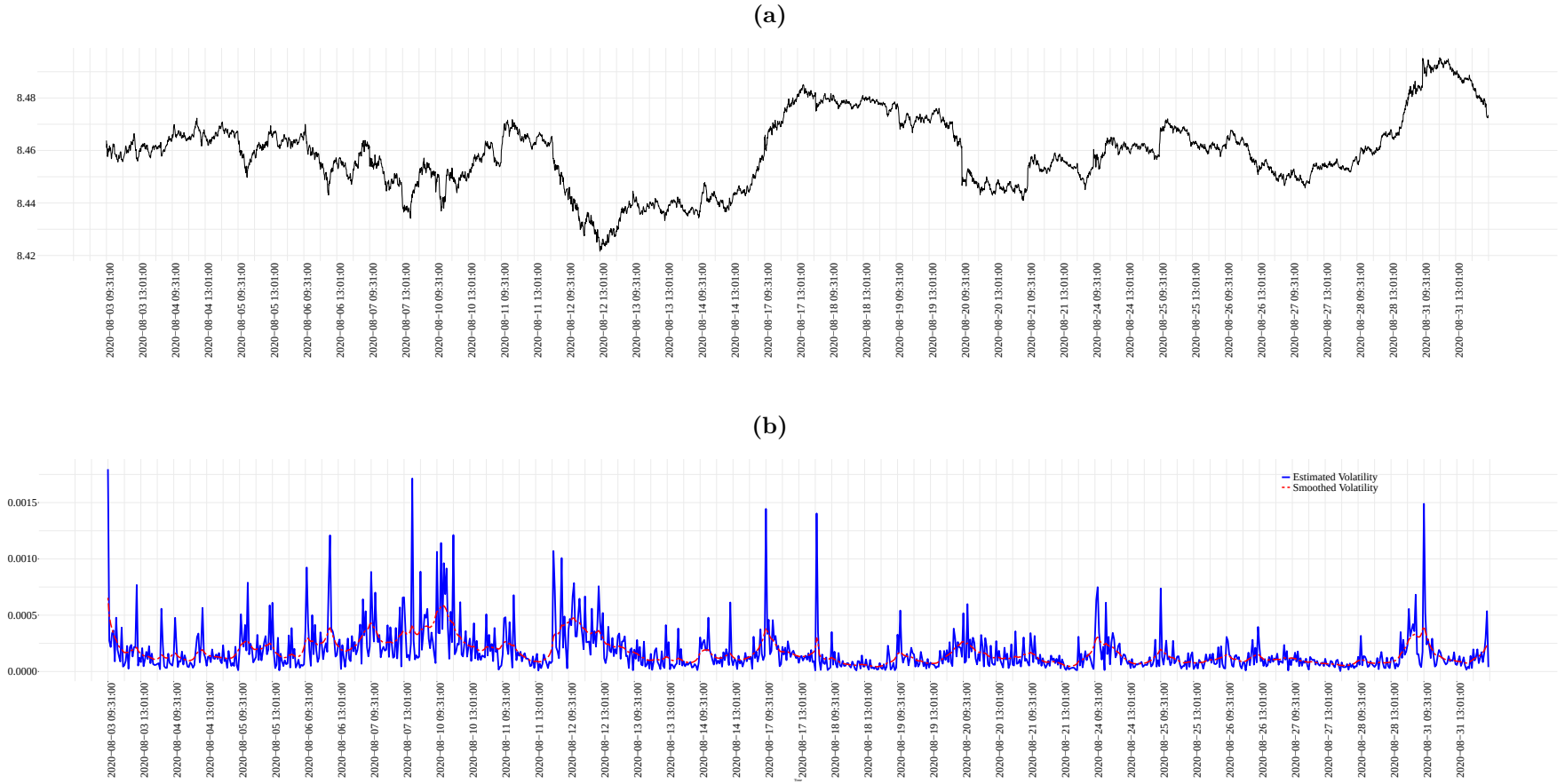
Note: In the figure above we report the empirical results using the one-minute price data of Apple Inc. Panel ((a)) plots the log-price of Apple Inc. Panel ((b)) plots the nonparametric estimates of spot volatility (blue solid line) and the smoothed volatility estimate (red dashed line) using our Bayesian techniques. Results demonstrated in this figure correspond to Apple stock within August 2017.

Figure 17



Note: In the figure above we report the empirical results using the one-minute price data of CSI 300 index futures in January 2020. Panel ((a)) plots the log-price level of CSI 300 index futures. Panel ((b)) plots the nonparametric estimates of spot volatility (blue solid line) and the smoothed volatility estimate (red dashed line) using our Bayesian techniques. Results demonstrated in this figure correspond to CSI 300 index futures within January 2020.

Figure 18



Note: In the figure above we report the empirical results using the one-minute price data of CSI 300 index futures in August 2020. Panel ((a)) plots the log-price level of CSI 300 index futures. Panel ((b)) plots the nonparametric estimates of spot volatility (blue solid line) and the smoothed volatility estimate (red dashed line) using our Bayesian techniques. Results demonstrated in this figure correspond to CSI 300 index futures within August 2020.

Table 2. Model Selected for S&P 500 Index ETF in 2015

Date	Model	DIC	$D(\boldsymbol{\vartheta})$	p_D
January 2015	1	-2727.40	-3282.86	277.73
February 2015	2	-2066.28	-2907.14	420.43
March 2015	1	-2498.42	-3200.40	350.99
April 2015	2	-2133.17	-2876.71	371.77
May 2015	2	-2250.85	-3043.06	396.10
June 2015	2	-2283.65	-3422.16	569.25
July 2015	1	-2257.16	-2885.81	314.33
August 2015	1	-2830.98	-3445.31	307.16
September 2015	3	-2982.72	-3332.93	175.10
October 2015	1	-3080.86	-3598.36	258.75
November 2015	1	-2606.25	-3096.89	245.32
December 2015	1	-2849.65	-3422.49	286.42

Table 3. Model Selected for Apple Inc. Stock Price in 2017

Date	Model	DIC	$D(\boldsymbol{\vartheta})$	p_D
January 2017	2	-1714.08	-2551.75	418.83
February 2017	2	-1233.57	-1793.62	280.03
March 2017	2	-1821.20	-2678.98	428.89
April 2017	2	-1631.22	-2231.49	300.14
May 2017	2	-2631.12	-3648.79	508.83
June 2017	2	-2731.98	-3771.19	519.61
July 2017	2	-2327.21	-3295.71	484.25
August 2017	2	-2709.67	-3644.26	467.30
September 2017	2	-2318.04	-3247.47	464.72
October 2017	2	-2704.97	-3748.87	521.95
November 2017	2	-2342.86	-3292.15	474.65
December 2017	1	-2377.74	-3633.71	627.98

Table 4. Model Selected for CSI 300 Index in 2020

Date	Model	DIC	$D(\vartheta)$	p_D
January 2020	1	-1367.30	-1595.93	114.32
February 2020	1	-1747.86	-2071.24	161.69
March 2020	1	-1809.46	-2192.64	191.59
April 2020	1	-1711.00	-2061.63	175.32
May 2020	2	-1568.94	-1856.89	143.98
June 2020	2	-1713.23	-2274.82	280.79
July 2020	1	-2126.11	-2472.58	173.24
August 2020	2	-1785.07	-2146.15	180.54
September 2020	1	-1874.93	-2230.75	177.91
October 2020	2	-1298.32	-1720.46	211.07
November 2020	2	-1749.23	-2343.43	297.10
December 2020	2	-1797.38	-2286.43	244.53

Table 5. Posterior Summary of **Model 5** Parameters for S&P 500 ETF

	ϕ	μ	$\sqrt{\sigma_e^2}$	κ	μ_η	$\sqrt{\sigma_\eta^2}$	b
January 2015	0.9314	-10.6814	0.2327	0.0023	0.4756	0.7925	0.6505
	0.0145	0.0920	0.0213	0.0031	1.0793	0.3836	0.0592
February 2015	0.8630	-11.8068	0.4635	0.0127	0.9162	0.7700	0.6232
	0.0250	0.1109	0.0367	0.0133	0.7617	0.2799	0.0843
March 2015	0.9205	-11.7109	0.2670	0.0029	1.2659	0.8791	0.6716
	0.0163	0.0907	0.0268	0.0031	0.8765	0.3713	0.0629
April 2015	0.8234	-11.8096	0.4126	0.0085	0.3941	0.7304	0.5651
	0.0310	0.0742	0.0348	0.0115	0.8503	0.3124	0.0658
May 2015	0.8987	-12.1515	0.3722	0.0054	1.2657	0.9002	0.5700
	0.0189	0.1060	0.0310	0.0061	0.9199	0.3801	0.0790
June 2015	0.7930	-11.9509	0.4837	0.0155	1.4376	0.8143	0.6031
	0.0321	0.0748	0.0395	0.0109	0.5847	0.2466	0.0621
July 2015	0.9416	-11.9417	0.2564	0.0025	1.5408	0.9219	0.6165
	0.0137	0.1257	0.0259	0.0024	0.9558	0.3958	0.0708
August 2015	0.9846	-10.9692	0.2260	0.0025	1.0233	0.8391	0.3343
	0.0052	0.7358	0.0179	0.0025	0.8693	0.3647	0.0694
September 2015	0.9659	-10.4311	0.1523	0.0012	0.5647	0.8383	0.5911
	0.0083	0.1180	0.0153	0.0014	1.1843	0.4297	0.0473
October 2015	0.9555	-11.1649	0.1888	0.0022	0.7109	0.7978	0.8046
	0.0099	0.1095	0.0180	0.0021	0.8601	0.3648	0.0532
November 2015	0.9414	-11.4897	0.2135	0.0018	0.4866	0.8158	0.6810
	0.0138	0.1020	0.0215	0.0024	1.1374	0.4057	0.0600
December 2015	0.9698	-11.0310	0.1868	0.0021	0.9683	0.8345	0.7201
	0.0080	0.2104	0.0187	0.0019	0.9039	0.3597	0.0570

Table 6. Posterior Summary of **Model 5** Parameters for Apple Inc. Stock

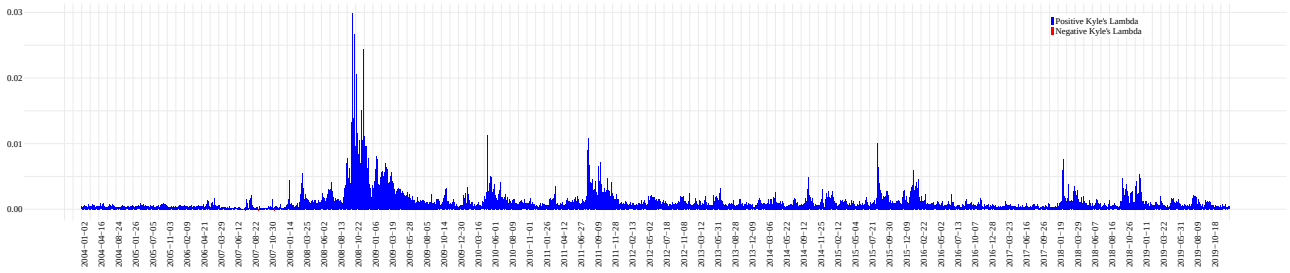
	ϕ	μ	$\sqrt{\sigma_e^2}$	κ	μ_η	$\sqrt{\sigma_\eta^2}$	b
January 2017	0.8336	-11.6393	0.4606	0.0175	1.6388	0.7839	0.4773
	0.0283	0.0999	0.0374	0.0099	0.5376	0.2390	0.0799
February 2017	0.8363	-11.5400	0.4528	0.0164	1.2551	0.8021	0.5053
	0.0311	0.1226	0.0438	0.0149	0.7643	0.2800	0.1041
March 2017	0.8005	-11.4342	0.4771	0.0153	1.1766	0.7884	0.5941
	0.0296	0.0846	0.0360	0.0136	0.6884	0.2618	0.0697
April 2017	0.8148	-11.2624	0.4239	0.0132	0.6342	0.7379	0.5785
	0.0377	0.0872	0.0408	0.0149	0.7696	0.2943	0.0730
May 2017	0.8907	-11.2462	0.4094	0.0119	0.9838	0.7432	0.3568
	0.0196	0.1115	0.0307	0.0103	0.6161	0.2369	0.0768
June 2017	0.9145	-10.8191	0.3869	0.0109	1.0729	0.7394	0.3435
	0.0152	0.1269	0.0300	0.0085	0.5542	0.2337	0.0785
July 2017	0.8456	-11.1228	0.4563	0.0148	0.6327	0.6979	0.4765
	0.0241	0.0975	0.0312	0.0153	0.6697	0.2468	0.0749
August 2017	0.8947	-10.6961	0.3686	0.0083	0.7524	0.7393	0.4341
	0.0185	0.1015	0.0273	0.0088	0.7337	0.2872	0.0711
September 2017	0.9123	-10.8906	0.4057	0.0110	1.3838	0.8018	0.4142
	0.0165	0.1400	0.0305	0.0077	0.5950	0.2513	0.0876
October 2017	0.8673	-11.3134	0.4316	0.0108	0.7922	0.7307	0.4636
	0.0177	0.0951	0.0261	0.0110	0.7128	0.2624	0.0742
November 2017	0.9202	-11.0970	0.4336	0.0141	1.2448	0.7748	0.2424
	0.0172	0.1767	0.0323	0.0100	0.5673	0.2356	0.1010
December 2017	0.8127	-11.1758	0.5669	0.0078	1.9642	3.1212	0.6861
	0.0207	0.0839	0.0286	0.0038	0.9528	0.8653	0.0767

Table 7. Posterior Summary of **Model 5** Parameters for CSI 300 Index Futures

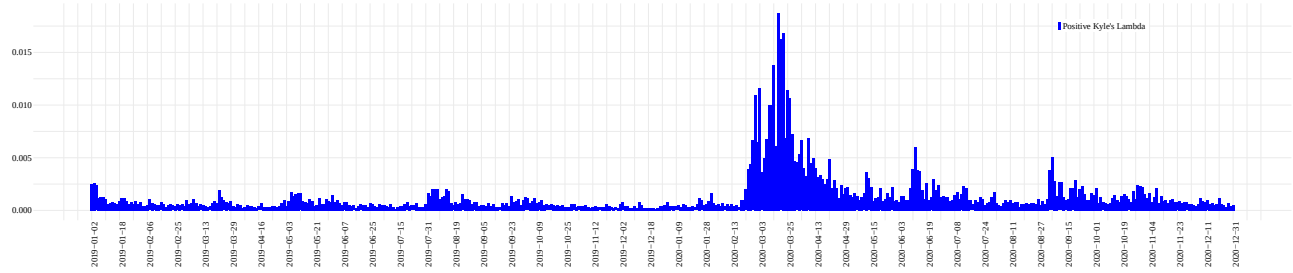
	ϕ	μ	$\sqrt{\sigma_e^2}$	κ	μ_η	$\sqrt{\sigma_\eta^2}$	b
January 2020	0.9461	-10.6461	0.2139	0.0031	0.5133	0.8219	0.8715
	0.0207	0.2011	0.0340	0.0040	1.1557	0.4347	0.0583
February 2020	0.8404	-9.9389	0.3371	0.0063	1.9077	2.9523	0.9678
	0.0318	0.0763	0.0260	0.0033	1.0855	0.9737	0.0276
March 2020	0.9532	-9.3718	0.2260	0.0047	1.1011	0.8282	0.9262
	0.0131	0.1693	0.0242	0.0040	0.7615	0.3338	0.0455
April 2020	0.8893	-10.7519	0.2662	0.0043	0.5160	0.7759	0.8996
	0.0315	0.0872	0.0370	0.0058	0.9981	0.3499	0.0505
May 2020	0.9174	-11.0502	0.2159	0.0029	0.4834	0.8201	0.8263
	0.0208	0.0978	0.0280	0.0036	1.1455	0.4057	0.0558
June 2020	0.7761	-11.0128	0.3640	0.0118	0.7256	0.7275	0.7924
	0.0339	0.0677	0.0300	0.0112	0.6980	0.2619	0.0534
July 2020	0.9430	-9.2816	0.2159	0.0023	0.5143	0.8143	0.9107
	0.0142	0.1233	0.0210	0.0030	1.1364	0.3920	0.0469
August 2020	0.9091	-9.8210	0.2396	0.0031	0.4617	0.8025	0.8289
	0.0291	0.0936	0.0350	0.0042	1.1146	0.3946	0.0547
September 2020	0.9013	-10.3231	0.2362	0.0031	0.5099	0.8049	0.8579
	0.0250	0.0814	0.0287	0.0040	1.0755	0.3928	0.0509
October 2020	0.7939	-10.6007	0.3586	0.0114	0.7909	0.7751	0.8349
	0.0485	0.0794	0.0439	0.0123	0.7867	0.3156	0.0620
November 2020	0.7933	-10.6782	0.3498	0.0116	1.1399	0.7755	0.8509
	0.0393	0.0655	0.0344	0.0087	0.5798	0.2553	0.0520
December 2020	0.8011	-10.7099	0.3386	0.0072	0.5606	0.7650	0.9509
	0.0467	0.0626	0.0405	0.0094	0.8860	0.3544	0.0352

Figure 19

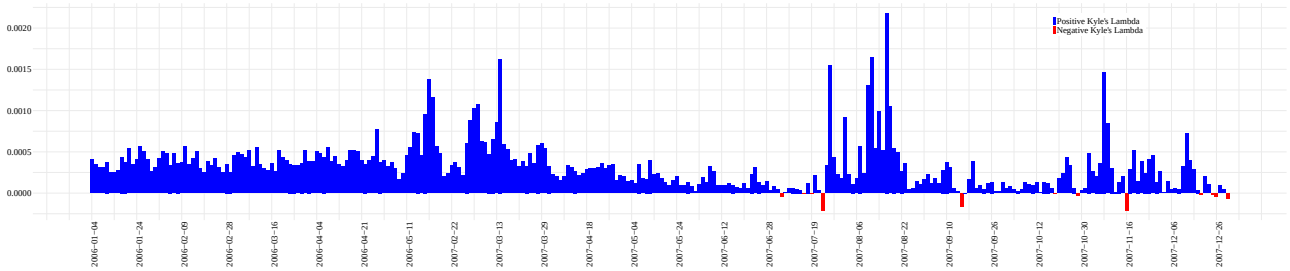
(a)



(b)

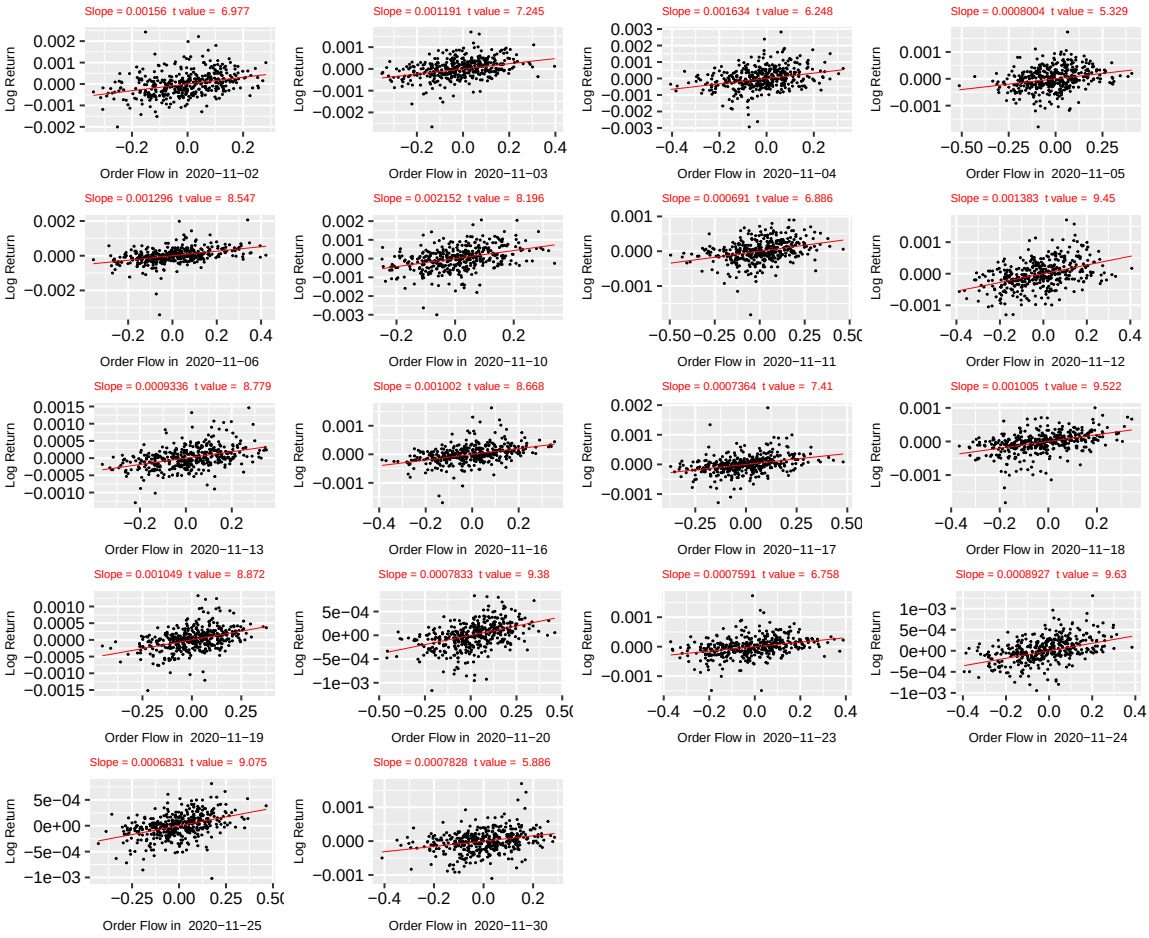


(c)



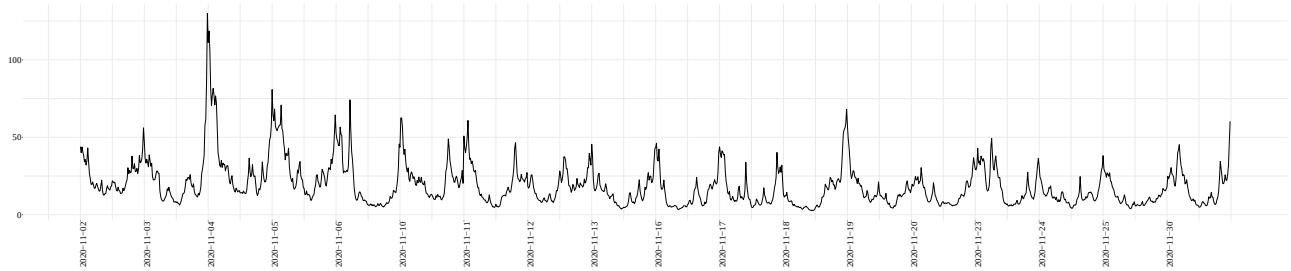
Note: In the figure above, we demonstrate daily estimated Kyle's lambda associated with S&P 500 ETF by regressing one-minute log-return of S&P 500 on one-minute share order flow. Panel (a) demonstrates daily estimated Kyle's lambda over the trading days from January 2004 to December 2019. For the specific focus on Covid-19 pandemic period, panel (b) specifically demonstrates corresponding estimation over trading days in 2020. Finally, we use panel (c) as the complementary to (a) and (b) by demonstrating that using daily high-frequency log-return (at the one-minute frequency) and share order flow to back out Kyle's lambda generally works fine (blue bar indicates desired positive estimation) with just a few exceptions (red bar indicates a negative estimation of Kyle's lambda that is not ideally consistent with the theoretical prediction).

Figure 20



Note: In the figure above, we demonstrate the basic regression structure corresponding to regressing one-minute log-return of S&P 500 ETF on corresponding share order flow over active trading days within November 2020 (18 active trading days in total within that month). The major implication of this figure is that the log-return-flow pattern (captured by Kyle's λ) associated with insider trading theory under the continuous-time setting is possible to be unveiled via this daily regression.

Figure 21



Note: In the figure above, we demonstrate our estimated private information value associated with insider trading based on estimated spot volatility using the proposed modeling framework and Bayesian sampling techniques. Specifically, for each day we use univariate regression by regressing one-minute log-return of S&P 500 ETF on corresponding share order flow over each active trading day; spot volatility is estimated by fixing the local estimation window size as $k = 5$ (i.e, every 5-minutes, so that would be $390/5 = 78$ locally estimated spot volatilities in each trading day and hence 78 locally quantified private information). Results demonstrated in this figure correspond to November 2020.

References

Aït-Sahalia, Y., and J. Jacod (2014): *High Frequency Econometrics.* Princeton University Press. [Cited on page 10.]

Akey, P., V. Grégoire, and C. Martineau (2022): "Price revelation from insider trading: Evidence from hacked earnings news," *Journal of Financial Economics*, 143(3), 1162–1184. [Cited on page 29.]

Andersen, T. G., and T. Bollerslev (1997): "Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *The Journal of Finance*, 52(3), 975–1005. [Cited on page 1.]

Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold (2013): "Chapter 17 - Financial Risk Measurement for Financial Risk Management," vol. 2 of *Handbook of the Economics of Finance*, pp. 1127–1220. Elsevier. [Cited on page 1.]

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2010): "CHAPTER 2 - Parametric and Nonparametric Volatility Measurement," in *Handbook of Financial Econometrics: Tools and Techniques*, ed. by Y. AIT-SAHALIA, and L. P. HANSEN, vol. 1 of *Handbooks in Finance*, pp. 67–137. North-Holland, San Diego. [Cited on page 1.]

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001): "The distribution of realized stock return volatility," *Journal of Financial Economics*, 61(1), 43–76. [Cited on page 1.]

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96(453), 42–55. [Cited on pages 1, 2, and 3.]

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71(2), 579–625. [Cited on page 1.]

Back, K. (1992): "Insider Trading in Continuous Time," *The Review of Financial Studies*, 5(3), 387–409. [Cited on pages 29, 30, and 31.]

Back, K. E. (2017): *Asset pricing and portfolio choice theory.* Oxford University Press. [Cited on pages 29 and 30.]

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009): "Realized kernels in practice: trades and quotes," *The Econometrics Journal*, 12(3), C1–C32. [Cited on page 26.]

Barndorff-Nielsen, O. E., and N. Shephard (2002): "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64(2), 253–280. [Cited on page 3.]

Bernile, G., J. Hu, and Y. Tang (2016): "Can information be locked up? Informed trading ahead of macro-news announcements," *Journal of Financial Economics*, 121(3), 496–520. [Cited on pages 11 and 12.]

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31(3), 307–327. [Cited on page 1.]

Bollerslev, T., J. Li, and Z. Liao (2021): "Fixed-k inference for volatility," *Quantitative Economics*, 12(4), 1053–1084. [Cited on pages 2, 4, 5, 7, 18, and 31.]

Bollerslev, T., and H. Zhou (2002): "Estimating stochastic volatility diffusion using conditional moments of integrated volatility," *Journal of Econometrics*, 109(1), 33–65. [Cited on page 1.]

Bolstad, W. M. (2009): *Understanding Computational Bayesian Statistics.* New Jersey: John Wiley & Sons, Inc. [Cited on page 13.]

Boudt, K., O. Kleen, and E. Sjørup (2021): "Analyzing intraday financial data in R: The highfrequency package," Working paper. [Cited on page 26.]

Chakrabarty, B., B. Li, V. Nguyen, and R. A. Van Ness (2007): "Trade classification algorithms for electronic communications network trades," *Journal of Banking & Finance*, 31(12), 3806–3821. [Cited on page 31.]

Chernov, M., A. Ronald Gallant, E. Ghysels, and G. Tauchen (2003): "Alternative models for stock price dynamics," *Journal of Econometrics*, 116(1), 225–257, Frontiers of financial econometrics and financial engineering. [Cited on page 7.]

Chib, S., F. Nardari, and N. Shephard (2002): "Markov chain Monte Carlo methods for stochastic volatility models," *Journal of Econometrics*, 108(2), 281–316. [Cited on page 10.]

Christensen, K., U. Hounyo, and M. Podolskij (2018): "Is the diurnal pattern sufficient to explain intraday variation in volatility? A nonparametric assessment," *Journal of Econometrics*, 205(2), 336–362. [Cited on pages 10 and 11.]

Collin-Dufresne, P., and V. Fos (2015): "Do Prices Reveal the Presence of Informed Trading?," *The Journal of Finance*, 70(4), 1555–1582. [Cited on page 29.]

Ellis, K., R. Michaely, and M. O'Hara (2000): "The Accuracy of Trade Classification Rules: Evidence from Nasdaq.," *Journal of Financial & Quantitative Analysis*, 35(4), 529–551. [Cited on page 29.]

Engle, R. (2004): "Risk and Volatility: Econometric Models and Financial Practice," *The American Economic Review*, 94(3), 405–420. [Cited on page 1.]

Engle, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50(4), 987–1007. [Cited on page 1.]

Foster, D. P., and D. B. Nelson (1996): "Continuous Record Asymptotics for Rolling Sample Variance Estimators," *Econometrica*, 64(1), 139–174. [Cited on page 1.]

Gatheral, J., T. Jaisson, and M. Rosenbaum (2018): "Volatility is rough," *Quantitative Finance*, 18(6), 933–949. [Cited on page 1.]

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013): *Bayesian data analysis*, Chapman & Hall/CRC Texts in Statistical Science Series. CRC, Boca Raton, Florida. [Cited on page 13.]

Gilks, W., S. Richardson, and D. Spiegelhalter (1995): *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC Interdisciplinary Statistics. Taylor & Francis. [Cited on page 13.]

Grossman, S. J., and J. E. Stiglitz (1980): "On the Impossibility of Informationally Efficient Markets," *The American Economic Review*, 70(3), 393–408. [Cited on page 29.]

Harvey, A., E. Ruiz, and N. Shephard (1994): "Multivariate Stochastic Variance Models," *The Review of Economic Studies*, 61(2), 247–264. [Cited on page 8.]

Holden, C., and S. Jacobsen (2014): "Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions," *The Journal of Finance*, 69(4), 1747–1785. [Cited on pages 29 and 31.]

Jacod, J., J. Li, and Z. Liao (2021): "Volatility coupling," *The Annals of Statistics*, 49(4), 1982 – 1998. [Cited on pages 3, 4, and 7.]

Jacod, J., and P. Protter (2012): *Discretization of Process.* Springer. [Cited on page 4.]

Kacperczyk, M., and E. S. Pagnotta (2019): "Chasing Private Information," *The Review of Financial Studies*, 32(12), 4997–5047. [Cited on page 29.]

Kadan, O., and A. Manela (2024): "Liquidity and the strategic value of information," *Review of Finance*, 29(1), 1–32. [Cited on pages 29 and 30.]

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, 65(3), 361–393. [Cited on pages 6, 9, and 15.]

Kristensen, D. (2010): "Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach," *Econometric Theory*, 26(1), 60–2093. [Cited on page 2.]

Kyle, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–1335. [Cited on pages 29, 30, and 31.]

Lee, C. M., and M. J. Ready (1991): "Inferring Trade Direction from Intraday Data.," *Journal of Finance*, 46(2), 733–746. [Cited on page 29.]

Lee, P. M. (2012): *Bayesian Statistics: An Introduction, 4th edition.* Wiley. [Cited on page 6.]

Li, Y., J. Yu, and T. Zeng (2021): "Deviance Information Criterion for Model Selection: A Theoretical Justification," Working paper, Remin University of China , Singapore Management University and Zhe Jiang University. [Cited on page 17.]

Lucca, D. O., and E. Moench (2015): "The Pre-FOMC Announcement Drift," *The Journal of Finance*, 70(1), 329–371. [Cited on pages 11 and 12.]

Malik, S., and M. K. Pitt (2011): "Particle filters for continuous likelihood evaluation and maximisation," *Journal of Econometrics*, 165(2), 190–209. [Cited on page A-12.]

Mykland, P. A., and L. Zhang (2008): "Inference for volatility-type objects and implications for hedging," *Statistics and Its Interface*, 1(2), 255–278. [Cited on page 2.]

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002): "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64(4), 583–639. [Cited on page 17.]

——— (2014): "The deviance information criterion: 12 years on," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(3), 485–493. [Cited on page 17.]

Stroud, J. R., and M. S. Johannes (2014): "Bayesian Modeling and Forecasting of 24-Hour High-Frequency Volatility," *Journal of the American Statistical Association*, 109(508), 1368–1384. [Cited on pages 1, 8, 12, and A-12.]

Tanner, M. A., and W. H. Wong (1987): "The Calculation of Posterior Distributions by Data Augmentation," *Journal of the American Statistical Association*, 82(398), 528–540. [Cited on page 13.]

Taylor, S. J. (1982): "Financial returns modelled by the product of two stochastic processes : a study of daily sugar prices, 1961-79," *Time series analysis : Theory and Practice*, 1, 203–226. [Cited on page 1.]

Wang, X., W. Xiao, and J. Yu (2022): "Modeling and forecasting realized volatility with the fractional Orstein-Uhlenbeck process," Working paper, forthcoming in *Journal of Econometrics*. [Cited on page 1.]

Xiu, D. (2010): "Quasi-maximum likelihood estimation of volatility with high frequency data," *Journal of Econometrics*, 159(1), 235–250. [Cited on page 18.]

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005): "A Tale of Two Time Scales," *Journal of the American Statistical Association*, 100(472), 1394–1411. [Cited on pages 2 and 27.]

Zu, Y., and H. Peter Boswijk (2014): "Estimating spot volatility with high-frequency financial data," *Journal of Econometrics*, 181(2), 117–135. [Cited on page 2.]

Appendix

A MCMC Procedure for Stochastic Volatility Model with Jumps

Since we are doing Bayesian analysis using Gibbs Sampling, the most preliminary and prominent thing that is necessary to keep in mind is which should be regarded as **data** and which should be regarded as **parameters**. For discussions contained in this part, we suppress the subscript n that denotes sample size and temporarily use t as the discrete timing index (i.e. t° used in the main context) for the local estimation block. Besides, we use T to denote the total number of local estimation blocks. Recall we as econometricians have:

$$\{\ln \hat{c}_t\}_{t=1}^T := \left\{\ln y_t^2\right\}_{t=1}^T \ : \ \textbf{data}, \text{ for instance, nonparametric estimation of spot volatility}$$

$$\{\ln c_t\}_{t=1}^T := \{h_t\}_{t=1}^T \ : \ \textbf{parameters}, \text{ latent spot volatility}$$

$$\left\{\phi, \mu, \sigma_e^2\right\} \ : \ \textbf{parameters}, \text{ volatility parameters}$$

$$\{J_t\}_{t=1}^T \ : \ \textbf{parameters}, \text{ jump component indicator}$$

$$\{\eta_t\}_{t=1}^T \ : \ \textbf{parameters}, \text{ jump component magnitude}$$

$$\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\} \ : \ \textbf{parameters}, \text{ jump parameters}$$

The target is to obtain the posterior sampling of **parameters** conditional on data by switching between different parameter space blocks. As we can see from the description specifying **data** and **parameters**, our analysis using Gibbs Sampling could be in general divided into two parts that are contained in the following two subsections respectively. For other parameters involved in the alternative models in the main context such as parameter b that specifies the intraday pattern in **Model 3** and $\left\{\tilde{\alpha}_q, \tilde{\beta}_q\right\}$ that specifies the announcement effect in **Model 4**, we just need to insert additional acceptance-rejection sampling or Metropolis-Hastings sampling step within the Gibbs sampling loop. Thus, for a specific sampling step, with all other parameters fixed at the conditionally sampled value, we just need to subtract that from the measurement equation in the state-space model.

A.1 Sampling volatility parameters and latent spot volatilities

- Sampling $\left\{\tilde{h}_t\right\}_{t=1}^T$. This sampling step is conditional on data $\left\{y_t^2\right\}_{t=1}^T$ and updated (or initial in the first MCMC loop) $\left\{\phi, \mu, \sigma_e^2\right\}$, $\{J_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$, and $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$. For the sake of description

simplicity, recall transition dynamics of latent spot volatility as follows

$$\begin{aligned}\tilde{h}_{t+1} &= \mu + \phi\left(\tilde{h}_t - \mu\right) + J_{t+1}\eta_{t+1} + e_{t+1}, \quad e_{t+1} \sim \mathcal{N}(0, \sigma_e^2) \\ \tilde{h}_t &= \mu + \phi\left(\tilde{h}_{t-1} - \mu\right) + J_t\eta_t + e_t, \quad e_t \sim \mathcal{N}(0, \sigma_e^2),\end{aligned}$$

and this transition dynamics implicitly suggest the initial condition as follows

$$\tilde{h}_1 \sim \left(\underbrace{\mu + \frac{\kappa\mu_\eta}{1-\phi}}_{\mu_1}, \underbrace{\frac{\kappa\left[\sigma_\eta^2 + (1-\kappa)^2\mu_\eta^2\right] + (1-\kappa)(\kappa\mu_\eta)^2 + \sigma_e^2}{1-\phi^2}}_{\varsigma_1^2} \right).$$

Our target is to obtain the posterior distribution of $\left\{\tilde{h}_t\right\}_{t=1}^T$ conditional on **data** and all the rest parameters: $\left\{\phi, \mu, \sigma_e^2\right\}$, $\left\{J_t\right\}_{t=1}^T$, $\left\{\eta_t\right\}_{t=1}^T$, and $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$. Given the normality assumption for $\left\{e_t\right\}_{t=1}^T$, it is standard to show that

$$\begin{aligned}\tilde{h}_t \mid \tilde{h}_{\setminus t} &\propto \exp\left\{ -\frac{\left[\tilde{h}_{t+1} - \mu - \phi(\tilde{h}_t - \mu) - J_{t+1}\eta_{t+1}\right]^2}{2\sigma_e^2} - \frac{\left[\tilde{h}_t - \mu - \phi(\tilde{h}_{t-1} - \mu) - J_t\eta_t\right]^2}{2\sigma_e^2} \right\} \\ &\propto \exp\left\{ \underbrace{-\frac{1}{2}\frac{1+\phi^2}{\sigma_e^2}}_{\text{coef\_a}}\left(\tilde{h}_t - \mu\right)^2 + \underbrace{\phi\frac{\tilde{h}_{t+1} - \mu - J_{t+1}\eta_{t+1} + \tilde{h}_{t-1} - \mu - J_t\eta_t}{\sigma_e^2}}_{\text{coef\_b}}\left(\tilde{h}_t - \mu\right) \right\} \\ &\propto \exp\left\{ -\frac{1}{2 \times \left(1/\sqrt{\text{coef\_a}}\right)^2}\left[\left(\tilde{h}_t - \mu\right) - \frac{\text{coef\_b}}{\text{coef\_a}}\right]^2 \right\}\end{aligned}$$

which suggests that posterior distribution of $\tilde{h}_t \mid \tilde{h}_{\setminus t}$ is normal distribution with mean as

$$\tilde{h}_t^* = \mu + \frac{\phi}{1+\phi^2}\left(\tilde{h}_{t+1} - \mu - J_{t+1}\eta_{t+1} + \tilde{h}_{t-1} - \mu - J_t\eta_t\right)$$

and variance as

$$v^2 = \frac{\sigma_e^2}{1+\phi^2}$$

If $t = 1$, note that

$$\begin{aligned}\tilde{h}_1 \mid \tilde{h}_2 &\propto \exp\left\{ -\frac{\left(\tilde{h}_1 - \mu_1\right)^2}{2\varsigma_1^2} \right\} \exp\left\{ -\frac{\left(\tilde{h}_2 - \mu - \phi(\tilde{h}_1 - \mu) - J_2\eta_2\right)^2}{2\sigma_e^2} \right\} \\ &\propto \exp\left\{ -\frac{1}{2}\left(\frac{\phi^2}{\sigma_e^2} + \frac{1}{\varsigma_1^2}\right)\left(\tilde{h}_1 - \mu\right)^2 + \left(\phi\frac{\tilde{h}_2 - \mu - J_2\eta_2}{\sigma_e^2} + \frac{\mu_1 - \mu}{\varsigma_1^2}\right)\left(\tilde{h}_1 - \mu\right) \right\}\end{aligned}$$

which suggests that $\tilde{h}_1$ is sampled from normal distribution with mean equal to

$$\mu + \left(\frac{\phi^2}{\sigma_e^2} + \frac{1}{\varsigma_1^2}\right)^{-1} \left(\phi \frac{\tilde{h}_2 - \mu - J_2\eta_2}{\sigma_e^2} + \frac{\kappa\mu_\eta}{(1-\phi)\varsigma_1^2}\right)$$

and variance equal to

$$\left(\frac{\phi^2}{\sigma_e^2} + \frac{1}{\varsigma_1^2}\right)^{-1}$$

If $t = T$, $\tilde{h}_T \mid \tilde{h}_{T-1}$ is simply sampled from normal distribution $\mathcal{N}\left(\mu + \phi\left(\tilde{h}_{T-1} - \mu\right) + J_T\eta_T, \sigma_e^2\right)$. Recall that

$$\left(\ln y_t^2 + \ln k\right) \mid \tilde{h}_t \;\sim\; \ln \chi_k^2$$

and our previous discussion basically shows how $\tilde{h}_t \mid \tilde{h}_{\setminus t}$ is conditionally distributed. But to formally sample $\tilde{h}_t$ from the full conditional posterior distribution $f\left(\tilde{h}_t \mid y_t^2, \tilde{h}_{\setminus t}\right)$, we should note the difference between the scenario where there are jumps and the scenario where there are no jumps. The key difference lies in the mean and variance associated with proposal distribution that is used to construct Metropolis-Hastings sampling within the Gibbs sampling loop. To show the difference more specifically, we still exploit $\tilde{h}_t^*$ to denote the derived conditional mean of $f\left(\tilde{h}_t \mid \tilde{h}_{\setminus t}\right)$ and the mean and variance associated with $f\left(\tilde{h}_t \mid \tilde{h}_{\setminus t}\right)$ can be summarized as follows,

$$\tilde{h}_t^* = \begin{cases} \mu + \left(\dfrac{\phi^2}{\sigma_e^2} + \dfrac{1}{\varsigma_1^2}\right)^{-1} \left(\phi \dfrac{\tilde{h}_2 - \mu - J_2\eta_2}{\sigma_e^2} + \dfrac{\kappa\mu_\eta}{(1-\phi)\varsigma_1^2}\right) & \text{if} \quad t = 1 \\ \mu + \dfrac{\phi}{1+\phi^2}\left(\tilde{h}_{t+1} - \mu - J_{t+1}\eta_{t+1} + \tilde{h}_{t-1} - \mu - J_t\eta_t\right) & \text{if} \quad 2 \leqslant t \leqslant T-1 \\ \mu + \phi\left(\tilde{h}_{T-1} - \mu\right) + J_t\eta_T & \text{if} \quad t = T \end{cases}$$

and

$$v^2 = \begin{cases} \left(\dfrac{\phi^2}{\sigma_e^2} + \dfrac{1}{\varsigma_1^2}\right)^{-1} & \text{if} \quad t = 1 \\ \dfrac{\sigma_e^2}{1+\phi^2} & \text{if} \quad 2 \leqslant t \leqslant T-1 \\ \sigma_e^2 & \text{if} \quad t = T \end{cases}$$

Since all other techniques apply similarly here, for each $\tilde{h}_t$ we exploit $f_N\left(u_t, v^2\right)$ as the proposal distribution with u_t as follows

$$u_t = \left[\frac{v^2}{2}\left(k y_t^2 \exp\left(-\tilde{h}_t^*\right) - k\right) + \tilde{h}_t^*\right]$$

- Sampling μ. This sampling step is conditional on data $\{y_t^2\}_{t=1}^T$, updated $\{\phi, \sigma_e^2\}$, $\{\tilde{h}_t\}_{t=1}^T$, $\{J_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$ and $\{\kappa, \mu_\eta, \sigma_\eta^2\}$. Conditional on updated $\{J_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$, and $\{\tilde{h}_t\}_{t=1}^T$ updated within Gibbs loop, sampling μ is straightforward as follows since $e_t = h_t - \phi h_{t-1} - (1-\phi)\mu - J_t\eta_t$ is normally distributed with zero mean and variance equal to σ_e^2. For the sake of description simplicity, we define

$$\xi_1 = \tilde{h}_1 \sim \mathcal{N}\left(\mu, \frac{\sigma_e^2}{1-\phi^2}\right) \text{ without jumps}$$

$$\xi_t = \tilde{h}_t - \phi\tilde{h}_{t-1}, \text{ for } t \geqslant 2$$

It is straightforward to show that if there exist jumps components specified, then $\xi_1 = \tilde{h}_1$ is distributed with mean (μ_1) equal to $\mu + \frac{\kappa\mu_\eta}{1-\phi}$ and variance (ς_1^2) equal to $\frac{\kappa\left[\sigma_\eta^2 + (1-\kappa)^2\mu_\eta^2\right] + (1-\kappa)(\kappa\mu_\eta)^2 + \sigma_e^2}{1-\phi^2}$. Recall the functional form of joint normal density function as follows

$$f_N\left(\xi_1 - \frac{\kappa\mu_\eta}{1-\phi}; \mu, \varsigma_1^2\right) \cdot \prod_{t=2}^{T} f_N\left(\xi_t - J_t\eta_t; (1-\phi)\mu, \sigma_e^2\right)$$

$$= \frac{1}{\sqrt{2\pi}\varsigma_1} \exp\left\{-\frac{\left((\xi_1 - \frac{\kappa\mu_\eta}{1-\phi} - \mu\right)^2}{2\varsigma_1^2}\right\} \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left\{-\frac{\left[\xi_t - J_t\eta_t - (1-\phi)\mu\right]^2}{2\sigma_e^2}\right\}$$

If $\pi(\mu) \sim \mathcal{N}(a, b^2)$, then the joint likelihood is given as follows

$$f_N(\mu; a, b^2) \cdot f_N\left(\xi_1 - \frac{\kappa\mu_\eta}{1-\phi}; \mu, \varsigma_1^2\right) \cdot \prod_{t=2}^{T} f_N\left(\xi_t - J_t\eta_t; (1-\phi)\mu, \sigma_e^2\right)$$

$$= \frac{1}{\sqrt{2\pi}b} \frac{1}{\sqrt{2\pi}\varsigma_1} \left[\frac{1}{\sqrt{2\pi}\sigma_e}\right]^{T-1} \exp\left\{-\frac{1}{2b^2}(\mu - a)^2 - \frac{1}{2\varsigma_1^2}(\xi_1 - \frac{\kappa\mu_\eta}{1-\phi} - \mu)^2 - \frac{1}{2\sigma_e^2}\sum_{t=2}^{T}\left[\xi_t - J_t\eta_t - (1-\phi)\mu\right]^2\right\}$$

$$\propto \exp\left\{-\frac{1}{2b^2}(\mu - a)^2 - \frac{1}{2\varsigma_1^2}(\xi_1 - \frac{\kappa\mu_\eta}{1-\phi} - \mu)^2 - \frac{1}{2\sigma_e^2}\sum_{t=2}^{T}\left[\xi_t - J_t\eta_t - (1-\phi)\mu\right]^2\right\}$$

Again by applying the method of completing squares, we are able to show that conjugate posterior distribution of μ is Gaussian distribution with mean equal to

$$\left[\frac{1}{b^2} + \frac{1}{\varsigma_1^2} + \frac{(T-1)(1-\phi)^2}{\sigma_e^2}\right]^{-1} \left[\frac{a}{b^2} + \frac{1}{\varsigma_1^2}\left(\xi_1 - \frac{\kappa\mu_\eta}{1-\phi}\right) + \frac{1-\phi}{\sigma_e^2}\sum_{t=2}^{T}(\xi_t - J_t\eta_t)\right]$$

and variance equal to

$$\left[\frac{1}{b^2} + \frac{1}{\varsigma_1^2} + \frac{(T-1)(1-\phi)^2}{\sigma_e^2}\right]^{-1}$$

- Sampling σ_e^2. This sampling step is conditional on data $\left\{y_t^2\right\}_{t=1}^T$ and updated $\{\phi, \mu\}$, $\left\{\tilde{h}_t\right\}_{t=1}^T$, $\{J_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$ and $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$. This sampling step is based on the following joint likelihood function

$$\begin{aligned}
\sigma_e^2 \mid \text{rest} \propto & \left(\sigma_e^2\right)^{-\alpha-1} \exp\left\{-\frac{\beta}{\sigma_e^2}\right\} \times \left(\sigma_e^2\right)^{-(T-1)/2} \times \left(\varsigma_1^2\right)^{-1/2} \times \\
& \exp\left\{-\frac{1}{2\varsigma_1^2}\left(\xi_1-\mu_1\right)^2-\frac{1}{2\sigma_e^2}\sum_{t=2}^{T}\left[\xi_t-(1-\phi)\mu-J_t\eta_t\right]^2\right\} \\
\propto & \left(\sigma_e^2\right)^{-\alpha-\frac{T-1}{2}-1} \exp\left\{-\left(\beta+\frac{1}{2}\sum_{t=2}^{T}\left[\xi_t-(1-\phi)\mu-J_t\eta_t\right]^2\right)\Big/\sigma_e^2\right\} \times \\
& \left\{\frac{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2+\sigma_e^2}{1-\phi^2}\right\}^{-1/2} \exp\left\{-\frac{1}{2}\frac{\left(1-\phi^2\right)\left(\xi_1-\mu_1\right)^2}{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2+\sigma_e^2}\right\}
\end{aligned}$$

This is not standard functional form of inverse-gamma distribution in terms of σ_e^2, hence additional Metropolis-Hastings step is needed using I.G. as proposal distribution. Note that

$$\begin{aligned}
\sigma_e^2 \mid \text{rest} \approx & \left(\sigma_e^2\right)^{-\alpha-\frac{T-1}{2}-1} \exp\left\{-\left(\beta+\frac{1}{2}\sum_{t=2}^{T}\left[\xi_t-(1-\phi)\mu-J_t\eta_t\right]^2\right)\Big/\sigma_e^2\right\} \\
\sim & \text{ I. G.}\left(\alpha+\frac{T-1}{2}, \beta+\frac{\sum_{t=2}^{T}\left[\xi_t-(1-\phi)\mu-J_t\eta_t\right]^2}{2}\right)
\end{aligned}$$

and accordingly we may use this inverse-gamma distribution as proposal distribution for sampling σ_e^2. Specifically, within the loop of Gibbs Sampling, given the current value of $\left(\sigma_e^2\right)^{(i-1)}$ at the i-th iteration, sampling $\left(\sigma_e^2\right)'$ from $\text{I. G.}\left(\alpha+\frac{T}{2}, \beta+\frac{\sum_{t=2}^{T}\left[\xi_t-(1-\phi)\mu-J_t\eta_t\right]^2}{2}\right)$. For the sake of discussion simplicity, we exploit $f_{\text{I.G.}}\left(x;\alpha,\beta\right)$ to denote probability density function evaluated at x for inverse-gamma distribution with parameter α and β, then the full conditional density function in terms of σ_e^2 could be rewritten as follows

$$\begin{aligned}
& \left\{\frac{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2+\sigma_e^2}{1-\phi^2}\right\}^{-1/2} \exp\left\{-\frac{1}{2}\frac{\left(1-\phi^2\right)\left(\xi_1-\mu_1\right)^2}{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2+\sigma_e^2}\right\} \\
& \qquad \times f_{\text{I.G.}}\left(\sigma_e^2;\alpha+\frac{T}{2}, \beta+\frac{\sum_{t=2}^{T}\left[\xi_t-(1-\phi)\mu-J_t\eta_t\right]^2}{2}\right)
\end{aligned}$$

This suggests that the proposed $\left(\sigma_e^2\right)'$ is accepted with probability equal to

$$\min\left\{1, \frac{\left\{\frac{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2+\left(\sigma_e^2\right)'}{1-\phi^2}\right\}^{-1/2}\exp\left\{-\frac{1}{2}\frac{(1-\phi^2)(\xi_1-\mu_1)^2}{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2+(\sigma_e^2)'}\right\}}{\left\{\frac{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2+\left(\sigma_e^2\right)^{(i-1)}}{1-\phi^2}\right\}^{-1/2}\exp\left\{-\frac{1}{2}\frac{(1-\phi^2)(\xi_1-\mu_1)^2}{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2+\left(\sigma_e^2\right)^{(i-1)}}\right\}}\right\}$$

- Sampling ϕ. This sampling step is conditional on data $\left\{y_t^2\right\}_{t=1}^T$ and updated $\left\{\mu, \sigma_e^2\right\}$, $\left\{\tilde{h}_t\right\}_{t=1}^T$, $\{J_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$ and $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$. If prior imposed on ϕ is

$$\pi(\phi) \propto \left\{\frac{(1+\phi)}{2}\right\}^{\phi^{(1)}-1}\left\{\frac{(1-\phi)}{2}\right\}^{\phi^{(2)}-1}$$

then full conditional density of ϕ is proportional

$$\pi(\phi) \propto \underbrace{\frac{1}{\sqrt{2\pi}\varsigma_1}\exp\left\{-\frac{\left(\xi_1-\frac{\kappa\mu_\eta}{1-\phi}-\mu\right)^2}{2\varsigma_1^2}\right\}\prod_{t=2}^{T}\frac{1}{\sqrt{2\pi}\sigma_e}\exp\left\{-\frac{\left[\xi_t-J_t\eta_t-(1-\phi)\mu\right]^2}{2\sigma_e^2}\right\}}_{f\left(\left\{\tilde{h}_t\right\}_{t=1}^T \mid \{J_t\}_{t=1}^T,\{\eta_t\}_{t=1}^T,\{\phi,\mu,\sigma_e^2\},\{\kappa,\mu_\eta,\sigma_\eta^2\}\right)}$$

where

$$\begin{aligned}
&\log f\left(\left\{\tilde{h}_t\right\}_{t=1}^T \mid \{J_t\}_{t=1}^T, \{\eta_t\}_{t=1}^T, \left\{\phi, \mu, \sigma_e^2\right\}, \left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}\right) \\
\propto\ & -\frac{1}{2}\log\varsigma_1^2 - \frac{\left(\xi_1-\frac{\kappa\mu_\eta}{1-\phi}-\mu\right)^2}{2\varsigma_1^2} - \frac{\sum_{t=2}^T\left[\xi_t-J_t\eta_t-(1-\phi)\mu\right]^2}{2\sigma_e^2} \\
\propto\ & \frac{1}{2}\log\left(1-\phi^2\right) - \frac{\left(\xi_1-\frac{\kappa\mu}{1-\phi}-\mu\right)^2\left(1-\phi^2\right)}{2\left\{\kappa\left[\sigma_\eta^2+(1-\kappa)^2\mu_\eta^2\right]+(1-\kappa)(\kappa\mu_\eta)^2\right\}} - \frac{\sum_{t=2}^T\left[\tilde{h}_t-\mu-J_t\eta_t-\phi\left(\tilde{h}_{t-1}-\mu\right)\right]^2}{2\sigma_e^2}
\end{aligned}$$

Recall that

$$\begin{aligned}
&\exp\left\{-\frac{\sum_{t=2}^T\left[\tilde{h}_t-\mu-J_t\eta_t-\phi\left(\tilde{h}_{t-1}-\mu\right)\right]^2}{2\sigma_e^2}\right\} \\
\propto\ & \exp\left\{-\frac{\sum_{t=2}^T\left(\tilde{h}_{t-1}-\mu\right)^2}{2\sigma_e^2}\phi^2 + \frac{\sum_{t=2}^T\left(\tilde{h}_t-\mu-J_t\eta_t\right)\left(\tilde{h}_{t-1}-\mu\right)}{\sigma_e^2}\phi\right\} \\
\propto\ & \exp\left\{-\frac{1}{2V_\phi}\left(\phi-\hat{\phi}\right)^2\right\}
\end{aligned}$$

where

$$V_\phi = \frac{\sigma_e^2}{\sum_{t=2}^T\left(\tilde{h}_{t-1}-\mu\right)^2}$$

and

$$\hat{\phi} = \frac{\sum_{t=2}^{T}\left(\tilde{h}_t - \mu - J_t\eta_t\right)\left(\tilde{h}_{t-1} - \mu\right)}{\sum_{t=2}^{T}\left(\tilde{h}_{t-1} - \mu\right)^2}.$$

This suggests that we can use $\mathcal{N}\left(\hat{\phi}, V_\phi\right)$ as proposal distribution to construct Metropolis-Hastings algorithm for sampling ϕ. Specifically, within the loop of Gibbs Sampling, given the current value of $\phi^{(i-1)}$ at the i-th iteration, sampling ϕ' from $\mathcal{N}\left(\hat{\phi}, V_\phi\right)$. Since the associated acceptance-rejection ratio is constructed as

$$\min\left\{1, \underbrace{\frac{\pi\left(\phi'\right) f\left(\left\{\tilde{h}_t\right\}_{t=1}^{T} \mid \left\{J_t\right\}_{t=1}^{T}, \left\{\eta_t\right\}_{t=1}^{T}, \left\{\phi', \mu, \sigma_e^2\right\}, \left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}\right)}{\pi\left(\phi^{(i-1)}\right) f\left(\left\{\tilde{h}_t\right\}_{t=1}^{T} \mid \left\{J_t\right\}_{t=1}^{T}, \left\{\eta_t\right\}_{t=1}^{T}, \left\{\phi^{(i-1)}, \mu, \sigma_e^2\right\}, \left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}\right)}}_{\text{part I}} \times \underbrace{\frac{f_N\left(\phi^{(i-1)}; \hat{\phi}, V_\phi\right)}{f_N\left(\phi'; \hat{\phi}, V_\phi\right)}}_{\text{part II}}\right\}$$

then $f_N\left(\phi^{(i-1)}; \hat{\phi}, V_\phi\right)$ and $f_N\left(\phi'; \hat{\phi}, V_\phi\right)$ would be cancelled with $\exp\left\{-\frac{1}{2V_\phi}\left(\phi^{(i-1)} - \hat{\phi}\right)^2\right\}$ contained in the denominator of part I and $\exp\left\{-\frac{1}{2V_\phi}\left(\phi' - \hat{\phi}\right)^2\right\}$ contained in the numerator of part I respectively. This suggests that the proposed value ϕ' is accepted as $\phi^{(i)}$ with probability equal to $\min\left\{1, \exp\left\{g(\phi') - g(\phi^{(i-1)})\right\}\right\}$ where

$$g\left(\phi\right) = \log\pi\left(\phi\right) + \frac{1}{2}\log\left(1-\phi^2\right) - \frac{\left(\xi_1 - \frac{\kappa\mu}{1-\phi} - \mu\right)^2\left(1-\phi^2\right)}{2\left\{\kappa\left[\sigma_\eta^2 + (1-\kappa)^2\mu_\eta^2\right] + (1-\kappa)\left(\kappa\mu_\eta\right)^2\right\}}.$$

If the proposed value is rejected, set $\phi^{(i)}$ equal to $\phi^{(i-1)}$.

So far we have completely described sampling procedure related to "volatility" components including latent spot volatility, $\left\{\tilde{h}_t\right\}_{t=1}^{T}$ and volatility parameter, $\left\{\phi, \mu, \sigma_e^2\right\}$. All the sampling procedure described within this subsection is conditional on data and remained "jump" components. In the next subsection, we proceed to discuss sampling procedure related to "jump" components, $\left\{J_t\right\}_{t=1}^{T}$, $\left\{\eta_t\right\}_{t=1}^{T}$ and $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$ conditional on latent spot volatility $\left\{\tilde{h}_t\right\}_{t=1}^{T}$ and volatility parameter, $\left\{\phi, \mu, \sigma_e^2\right\}$.

A.2 Sampling jump components and jump parameters

- Sampling $\left\{J_t\right\}_{t=1}^{T}$. This sampling step is conditional on data $\left\{y_t^2\right\}_{t=1}^{T}$ and updated $\left\{\tilde{h}_t\right\}_{t=1}^{T}$, $\left\{\phi, \mu, \sigma_e^2\right\}$, $\left\{\eta_t\right\}_{t=1}^{T}$ and $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$. Recall our previous definition that

$$\begin{aligned}
\xi_1 &= \tilde{h}_1 \\
\xi_t &= \tilde{h}_t - \phi\tilde{h}_{t-1} = (1-\phi)\,\mu + J_t\eta_t + e_t, \text{ for } t \geqslant 2
\end{aligned}$$

which implies that for $t \geqslant 2$

$$\begin{aligned}\xi_{t \mid J_t=1} &= (1-\phi)\,\mu + \eta_t + e_t \\ \xi_{t \mid J_t=0} &= (1-\phi)\,\mu + e_t\end{aligned}$$

Since η_t and e_t are assumed to be normally distributed independently with each other such that $\eta_t \sim \mathcal{N}(\mu_\eta, \sigma_\eta^2)$ and $e_t \sim \mathcal{N}(0, \sigma_e^2)$, this further implies that

$$\begin{aligned}\xi_{t \mid J_t=1} &\sim \mathcal{N}\left((1-\phi)\,\mu + \mu_\eta, \sigma_\eta^2 + \sigma_e^2\right) \\ \xi_{t \mid J_t=0} &\sim \mathcal{N}\left((1-\phi)\,\mu, \sigma_e^2\right)\end{aligned}$$

With κ updated within Gibbs sampling loop, updating $\{J_t\}_{t=1}^T$ as follows

$$\mathbb{P}\{J_t = 1 \mid \text{rest}\} = \frac{\kappa f_N\left(\xi_t; (1-\phi)\mu + \mu_\eta, \sigma_\eta^2 + \sigma_e^2\right)}{(1-\kappa) f_N\left(\xi_t; (1-\phi)\mu, \sigma_e^2\right) + \kappa f_N\left(\xi_t; (1-\phi)\mu + \mu_\eta, \sigma_\eta^2 + \sigma_e^2\right)}.$$

- Sampling $\{\eta_t\}_{t=1}^T$. This sampling step is conditional on data $\left\{y_t^2\right\}_{t=1}^T$, updated $\left\{\tilde{h}_t\right\}_{t=1}^T$, $\left\{\phi, \mu, \sigma_e^2\right\}$, $\{J_t\}_{t=1}^T$, and $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$. Given the dynamic transition system stated above, if $J_t = 1$, then

$$\begin{aligned}(\ln y_t^2 + \ln k) \mid \tilde{h}_t &\sim \ln \chi_k^2 \\ \tilde{h}_t &= \mu + \phi(\tilde{h}_{t-1} - \mu) + \eta_t + e_t\end{aligned}$$

η_t is regarded as parameter with prior distribution such that $\eta_t \sim \mathcal{N}\left(\mu_\eta, \sigma_\eta^2\right)$, therefore posterior distribution of η_t conditional on all the rest is based on the following joint likelihood function

$$f_N\left(\eta_t; \mu_\eta, \sigma_\eta^2\right) \cdot f_N\left(\tilde{h}_t - \phi\tilde{h}_{t-1} - (1-\phi)\mu; \eta_t, \sigma_e^2\right) \cdot f_{\ln \chi_k^2}\left(y_t^2; \tilde{h}_t\right)$$

Thus

$$\frac{1}{\sqrt{2\pi}\sigma_\eta} \exp\left\{-\frac{(\eta_t - \mu_\eta)^2}{2\sigma_\eta^2}\right\} \cdot \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left\{-\frac{\left(\tilde{h}_t - \phi\tilde{h}_{t-1} - (1-\phi)\mu - \eta_t\right)^2}{2\sigma_e^2}\right\} \cdot f_{\ln \chi_k^2}\left(y_t^2; \tilde{h}_t\right)$$

which further implies that

$$\begin{aligned}
\eta_t \mid \text{rest} &\propto \exp\left\{-\frac{\eta_t^2 - 2\eta_t\mu_\eta + \mu_\eta^2}{2\sigma_\eta^2} - \frac{\left(\tilde{h}_t - \phi\tilde{h}_{t-1} - (1-\phi)\mu\right)^2 - 2\left(\tilde{h}_t - \phi\tilde{h}_{t-1} - (1-\phi)\mu\right)\eta_t + \eta_t^2}{2\sigma_e^2}\right\} \\
&\propto \exp\left\{-\frac{1}{2}\underbrace{\left(\frac{1}{\sigma_\eta^2} + \frac{1}{\sigma_e^2}\right)}_{\text{coef_a}}\eta_t^2 + \underbrace{\left(\frac{\mu_\eta}{\sigma_\eta^2} + \frac{\tilde{h}_t - \phi\tilde{h}_{t-1} - (1-\phi)\mu}{\sigma_e^2}\right)}_{\text{coef_b}}\eta_t\right\} \\
&\propto \exp\left\{-\frac{1}{2(\sqrt{1/\text{coef_a}})^2}\left(\eta_t - \frac{\text{coef_b}}{\text{coef_a}}\right)^2\right\}
\end{aligned}$$

Therefore posterior distribution of $\eta_t \mid \text{rest}$ is normally distributed with mean equal to

$$\left(\frac{1}{\sigma_\eta^2} + \frac{1}{\sigma_e^2}\right)^{-1}\left(\frac{\mu_\eta}{\sigma_\eta^2} + \frac{\tilde{h}_t - \phi\tilde{h}_{t-1} - (1-\phi)\mu}{\sigma_e^2}\right)$$

and variance equal to

$$\left(\frac{1}{\sigma_\eta^2} + \frac{1}{\sigma_e^2}\right)^{-1}$$

- Sampling κ. This sampling step is conditional on data $\{y_t\}_{t=1}^T$, updated $\left\{\tilde{h}_t\right\}_{t=1}^T$, $\left\{\phi, \mu, \sigma_e^2\right\}$, $\{J_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$ and $\left\{\kappa, \mu_\eta, \sigma_\eta^2\right\}$. Since $\{J_t\}_{t=1}^T$ follows Bernoulli distribution, Bernoulli(κ), and by specification we as econometricians have prior knowledge on κ that κ follows Beta distribution Beta(α, β), then κ is sampled from posterior distribution conditional on $\{J_t\}_{t=1}^T$. Actually $\{J_t\}_{t=1}^T$ as binary random variables, realizations of the sum of $\{J_t\}_{t=1}^T$ lie in between 0 and T and we denote it as k. Given that

$$L(k \mid \kappa) := \mathbb{P}\left(\sum_{t=1}^T J_t = k \;\middle|\; T, \kappa\right) = \binom{T}{k}\kappa^k(1-\kappa)^{T-k}$$

and the probability density function of Beta distribution

$$\pi(\kappa \mid \alpha, \beta) = \frac{\kappa^{\alpha-1}(1-\kappa)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}.$$

Joint likelihood of k and κ is given as follows

$$L(k \mid \kappa) \cdot \pi(\kappa \mid \alpha, \beta)$$

and by integrating out κ from this joint likelihood as follows yields

$$\begin{aligned} f(k \mid T, \alpha, \beta) &= \int_0^1 L(k \mid \kappa) \cdot \pi(\kappa \mid \alpha, \beta)\, d\kappa \\ &= \begin{pmatrix} T \\ k \end{pmatrix} \frac{1}{\mathrm{B}(\alpha, \beta)} \int_0^1 \kappa^{k+\alpha-1}(1-\kappa)^{T-k+\beta-1} d\kappa \\ &= \begin{pmatrix} T \\ k \end{pmatrix} \frac{\mathrm{B}(k+\alpha, T-k+\beta)}{\mathrm{B}(\alpha, \beta)} \\ &= \begin{pmatrix} T \\ k \end{pmatrix} \frac{\mathrm{B}\left(\sum_{t=1}^T J_t + \alpha, T - \sum_{t=1}^T J_t + \beta\right)}{\mathrm{B}(\alpha, \beta)}. \end{aligned}$$

Given the property of Beta function and Gamma function, it is easy to show that conditional on realizations of $\{J_t\}_{t=1}^T$, posterior distribution of κ follows Beta distribution as $\mathrm{Beta}\left(\sum_{t=1}^T J_t + \alpha, T - \sum_{t=1}^T J_t + \beta\right)$.

- Sampling μ_η. This sampling step is conditional on data $\{y_t\}_{t=1}^T$, updated $\left\{\tilde{h}_t\right\}_{t=1}^T$, $\left\{\phi, \mu, \sigma_e^2\right\}$, $\{J_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$ and $\left\{\kappa, \sigma_\eta^2\right\}$. If we assume prior distribution for μ_η with prior mean equal to a and prior variance equal to b^2, then the joint density function is given as follows,

$$\begin{aligned} &\left(\frac{1}{\sqrt{2\pi}b}\right) \exp\left\{-\frac{(\mu_\eta - a)^2}{2b^2}\right\} \times \left(\frac{1}{\sqrt{2\pi}\sigma_\eta}\right)^{T^{(1)}} \prod_{i=1}^{T^{(1)}} \exp\left\{-\frac{(\eta_i - \mu_\eta)^2}{2\sigma_\eta^2}\right\} \\ \propto\ & \exp\left\{-\frac{(\mu_\eta - a)^2}{2b^2} - \sum_{i=1}^{T^{(1)}} -\frac{(\eta_i - \mu_\eta)^2}{2\sigma_\eta^2}\right\} \\ \propto\ & \exp\left\{-\frac{1}{2}\left(\frac{1}{b^2} + \frac{T^{(1)}}{\sigma_\eta^2}\right)\mu_\eta^2 + \left(\frac{a}{b^2} + \sum_{i=1}^{T^{(1)}} \frac{\eta_i}{\sigma_\eta^2}\right)\mu_\eta\right\}, \end{aligned}$$

where $T^{(1)}$ refers to the total number of observations when there are jumps, i.e., $J_t = 1$. Again by applying the method of completing squares, we can show that μ_η is updated from posterior distribution following distribution with the mean equal to

$$\left(\frac{1}{b^2} + \frac{T^{(1)}}{\sigma_\eta^2}\right)^{-1} \left(\frac{a}{b^2} + \sum_{i=1}^{T^{(1)}} \frac{\eta_i}{\sigma_\eta^2}\right)$$

and variance equal to

$$\left(\frac{1}{b^2} + \frac{T^{(1)}}{\sigma_\eta^2}\right)^{-1}.$$

- Sampling σ_η^2. This sampling step is conditional on data $\{y_t\}_{t=1}^T$, updated $\left\{\tilde{h}_t\right\}_{t=1}^T$, $\left\{\phi, \mu, \sigma_e^2\right\}$, $\{J_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$ and $\{\kappa, \mu_\eta\}$. For $J_t = 1$, η_t is normally distributed with mean equal to μ_η and variance equal to σ_η^2. We denote $T^{(1)}$ as the number of elements in $\{J_t\}_{t=1}^T$ which $J_t = 1$. The full conditional density function in terms of σ_η^2 with inverse-gamma prior distribution imposed on σ_η^2 is summarized as follows

$$\sigma_\eta^2 \mid \text{rest} \propto \left(\sigma_\eta^2\right)^{-\alpha-1} \exp\left\{-\frac{\beta}{\sigma_\eta^2}\right\} \times \left(\frac{1}{\sqrt{2\pi}\sigma_\eta}\right)^{T^{(1)}} \times \prod_{i=1}^{T^{(1)}} \exp\left\{-\frac{(\eta_i - \mu_\eta)^2}{2\sigma_\eta^2}\right\}.$$

  If we impose inverse-gamma distribution as the prior distribution for σ_η^2, based on the previous remark, we can sample σ_η^2 from inverse-gamma distribution as the posterior distribution,

$$\sigma_\eta^2 \mid \text{rest} \sim \text{I.G.}\left(\alpha + \frac{T^{(1)}}{2}, \beta + \frac{1}{2}\sum_{i=1}^{T^{(1)}} (\eta_i - \mu_\eta)^2\right).$$

B Particle Filter for Approximating Likelihood Function

We follow Malik and Pitt (2011) and Stroud and Johannes (2014) by constructing a particle filter algorithm to approximate the marginal likelihood function that is used for calculating information criteria based on the marginal likelihood function. Our particle filter algorithm is summarized as follows. For the sake of notation simplicity, we use t for discrete timing index and replace the corresponding nonparametric estimation of volatility from 1 to t, $\{\ln(\hat{c}_{n,1:t})\}$, with $\ln y_{1:t}^2$. We collect all the latent state variables into Z_t such that

$$Z_t = \left\{\tilde{h}_t, J_t, \eta_t\right\}$$

where we use $\tilde{h}_t = h_t + \mu$. Besides, we implicitly assume that parameters that specify the models are fixed at specific values.

$$\begin{aligned} p\left(Z_t \mid \ln y_{1:t}^2\right) &= p\left(\tilde{h}_t \mid \ln y_{1:t}^2\right) p\left(J_t, \eta_t \mid \tilde{h}_t, \ln y_{1:t}^2\right) \\ &= p\left(\tilde{h}_t \mid \ln y_{1:t}^2\right) p\left(Z_t^* \mid \tilde{h}_t, \ln y_{1:t}^2\right) \end{aligned}$$

where for the sake of notation simplicity, we collect "jump" related latent variables jointly into $Z_t^* = \{J_t, \eta_t\}$. Our analysis begins from the observation that the first part of this factorization, $p\left(\tilde{h}_t \mid \ln y_{1:t}^2\right)$ is **unavailable** analytically while the second part of this factorization, $p\left(Z_t^* \mid \tilde{h}_t, \ln y_{1:t}^2\right)$ is **available** in closed form. This actually has been demonstrated in the description of MCMC sampling algorithm in the previous section. We then use the result in Malik and Pitt (2011) that the likelihood function for a fixed parameter value at $\boldsymbol{\theta}$ can be approximated using the output from APF as

$$\mathcal{L}\left(\left\{\ln y_t^2\right\}_{t=1}^T \mid \boldsymbol{\theta}\right) = \prod_{t=1}^T \left(\frac{1}{N}\sum_{i=1}^N \pi_t^{(i)}\right)\left(\frac{1}{N}\sum_{i=1}^N w_t^{(i)}\right) \tag{B.1}$$

where N denotes the number of particles used in APF. The generic idea for obtaining $\left\{\pi_t^{(i)}\right\}$ and $\left\{w_t^{(i)}\right\}$ is as follows:

1. Start with sample $Z_{t-1}^{(i)} = \left(\tilde{h}_{t-1}^{(i)}, Z_{t-1}^{*(i)}\right) \sim p\left(Z_{t-1} \mid \ln y_{1:t-1}^2\right)$. For the initial state $Z_0^{(i)}$, we sample $\tilde{h}_0^{(i)}$ from initial stationary distribution following mix-normal such that

$$\tilde{h}_0^{(i)} \sim \mathcal{N}\left(\mu + \frac{\kappa\mu_\eta}{1-\phi}, \frac{\kappa\left[\sigma_\eta^2 + (1-\kappa)^2\mu_\eta^2\right] + (1-\kappa)(\kappa\mu_\eta)^2 + \sigma_e^2}{1-\phi^2}\right)$$

   and $J_0^{(i)}$, $\eta_0^{(i)}$ from Bernoulli(κ) and $\mathcal{N}\left(\mu_\eta, \sigma_\eta^2\right)$ respectively.

2. Compute $\pi_t^{(i)} \propto p\left(\ln y_t^2 \mid \hat{\tilde{h}}_t^{(i)}\right)$, where

$$\hat{\tilde{h}}_t^{(i)} = \mathbb{E}\left(\tilde{h}_t \mid Z_{t-1}^{(i)}, \ln y_{1:t-1}^2\right).$$

Moreover, note that dynamics of latent variables are specified such that $\tilde{h}_t - \mu = \phi\left(\tilde{h}_{t-1} - \mu\right) + J_t\eta_t + e_t$ and $e_t \sim \mathcal{N}\left(0, \sigma_e^2\right)$, $J_t \sim \text{Bernoulli}\left(\kappa\right)$, $\eta_t \sim \mathcal{N}\left(\mu_\eta, \sigma_\eta^2\right)$, then conditional on $Z_{t-1}^{(i)} = \left\{\tilde{h}_{t-1}^{(i)}, J_{t-1}^{(i)}, \eta_{t-1}^{(i)}\right\}$, we have $\hat{\tilde{h}}_t^{(i)}$ calculated as follows

$$\begin{aligned}
\hat{\tilde{h}}_t^{(i)} &= \mathbb{E}\left(\tilde{h}_t \mid Z_{t-1}^{(i)}, \ln y_{1:t-1}^2\right) \\
&= \phi\tilde{h}_{t-1}^{(i)} + \left(1-\phi\right)\mu + \mathbb{E}\left[J_t\eta_t\right] + \mathbb{E}\left[e_t\right] \\
&= \phi\tilde{h}_{t-1}^{(i)} + \left(1-\phi\right)\mu + \kappa\mu_\eta
\end{aligned}$$

3. With the results calculated in step 2, we then generate

$$k^i \sim \mathcal{M}\left(\pi_t^{(1)}, \ldots, \pi_t^{(N)}\right)$$

where $\mathcal{M}(\cdot)$ refers to the generic multinomial distribution specified by $\left\{\pi_t^{(i)}\right\}$.

4. Generate $\tilde{\tilde{h}}_t^{(i)} \sim p\left(\tilde{h}_t \mid Z_{t-1}^{(k^i)}, \ln y_{1:t-1}^2\right)$. This is a mix-normal distribution such that if $J_t = 1$, $\mathcal{N}\left(\left(1-\phi\right)\mu + \mu_\eta, \sigma_\eta^2 + \sigma_e^2\right)$; if $J_t = 0$, $\mathcal{N}\left(\left(1-\phi\right)\mu, \sigma_e^2\right)$.

$$\tilde{\tilde{h}}_t^{(i)} \mid Z_{t-1}^{(k^i)}, \ln y_{1:t-1}^2 \sim \begin{cases} \mathcal{N}\left(\phi\tilde{h}_{t-1} + \left(1-\phi\right)\mu + \mu_\eta, \sigma_\eta^2 + \sigma_e^2\right) & \text{if } J_t = 1 \\ \mathcal{N}\left(\phi\tilde{h}_{t-1} + \left(1-\phi\right)\mu, \sigma_e^2\right) & \text{if } J_t = 0 \end{cases}$$

5. Compute $w_t^{(i)} \propto p\left(\ln y_t^2 \mid \tilde{\tilde{h}}_t^{(i)}\right) / \pi_t^{(k^i)}$.

6. Generate

$$j^i \sim \mathcal{M}\left(w_t^{(1)}, \ldots, w_t^{(N)}\right)$$

and set $\tilde{h}_t^{(i)} = \tilde{\tilde{h}}_t^{(j^i)}$.

7. Generate $Z_t^{*(i)} \sim p\left(Z_t^* \mid \tilde{h}_t^{(i)}, \ln y_{1:t}^2\right)$. For fixed $\boldsymbol{\theta}$, recall that $Z_t^{*(i)} = \left\{J_t^{(i)}, \eta_t^{(i)}\right\}$ and we have derived in subsection A.2 that

$$J_t^{(i)} \mid \tilde{h}_t^{(i)}, \ln y_{1:t}^2 \sim \text{Bernoulli}\left(\frac{\kappa f_N\left(\xi_t^{(i)}; (1-\phi)\mu + \mu_\eta, \sigma_\eta^2 + \sigma_e^2\right)}{\left(1-\kappa\right) f_N\left(\xi_t^{(i)}; (1-\phi)\mu, \sigma_e^2\right) + \kappa f_N\left(\xi_t^{(i)}; (1-\phi)\mu + \mu_\eta, \sigma_\eta^2 + \sigma_e^2\right)}\right)$$

and $\eta_t^{(i)} \mid \tilde{h}_t^{(i)}, \ln y_{1:t}^2$ is normally distributed with mean equal to

$$\left(\frac{1}{\sigma_\eta^2} + \frac{1}{\sigma_e^2}\right)^{-1}\left(\frac{\mu_\eta}{\sigma_\eta^2} + \frac{\xi_t^{(i)} - (1-\phi)\mu}{\sigma_e^2}\right)$$

and variance equal to

$$\left(\frac{1}{\sigma_\eta^2}+\frac{1}{\sigma_e^2}\right)^{-1}$$

where $\xi_t^{(i)} = \tilde{h}_t^{(i)} - \phi\tilde{h}_{t-1}^{(i)}$.

C Additional Discussions

C.1 DGP Dynamic and Block Dynamic

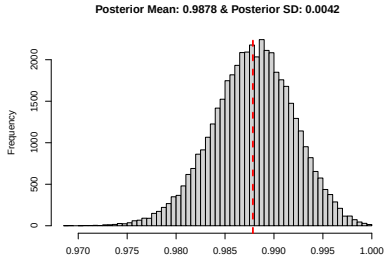
We summarize one another example demonstrating the insight in Remark 5.1. Specifically, we set $dt := \Delta_i^n = 1/390$, $\sigma_h = 1.6$, $k = 5$ and $\kappa_h = 0.5$, then $\phi_h = 1 - \kappa_h dt \approx 0.9987$, $\phi_h^k \approx 0.9936$ and

$$\sqrt{\sigma_e^2} = \sqrt{\frac{\sigma_h^2 dt(1-\phi_h^{2(k+1)})}{1-\phi_h^2}} \approx 0.1978$$

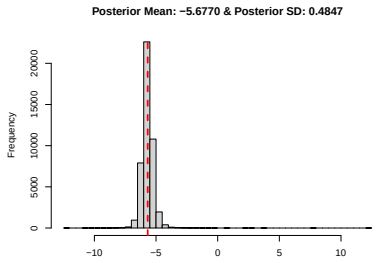
while the corresponding MCMC samplings are summarized as follows

Figure C.1

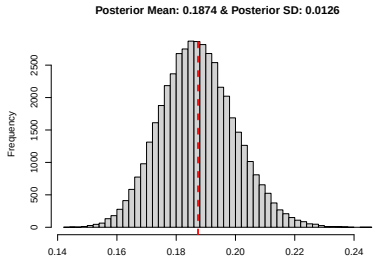
(a) ϕ



(b) μ



(c) $\sqrt{\sigma_e^2}$



C.2 Announcements and jumps

In this remark, we make more discussions about model comparisons when the data generating process is **DGP 4**, which corresponds to the concern about whether we need to specify a model accommodating announcement effects when there are observed announcements. For the results demonstrated above, we simulate announcement indicators at a relatively high-frequency such that approximately on average there is one announcement each day ($1/(78) \approx 0.013$). And if we decrease the rate at which announcement is made (for instance, $1/(78 \times 2) \approx 0.0064$, $1/(78 \times 3) \approx 0.004$, $1/(78 \times 10) \approx 0.001$). If seldom are there any announcements, for instance approximately there is only one announcement made every half-month ($1/(78 \times 10) \approx 0.001$), model specification with announcement effects would not be that distinguishable (using DIC) in comparison to other model specifications. Corresponding results are summarized in Table C.1. The basic point is that when data generated from a specific model is hard to be distinguished from data generated from other models, this will be reflected in using DIC for model comparison. Similar scheme applies for jump specifications, which is summarized in Table C.2, where we compare **Model 1**, **Model 2** and **Model 3** across different **DGP 3** with different specifications of rates at which jump happens and different jump magnitudes (size).

Table C.1

	Announcement rate														
	0.013			0.0064			0.004			0.001			0.0006		
	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D
Model 1	-3243.12	-3444.81	100.85	-3255.95	-3436.43	90.24	-3282.65	-3455.12	86.24	-3299.98	-3468.52	84.27	-3310.18	-3470.29	80.05
Model 2	-3243.14	-3469.41	113.13	-3255.66	-3463.36	103.85	-3282.28	-3484.30	101.01	-3298.39	-3495.35	98.48	-3308.29	-3498.62	95.16
Model 3	-3242.52	-3446.71	102.10	-3255.06	-3438.39	91.67	-3280.30	-3457.39	87.55	-3298.94	-3469.80	85.43	-3309.44	-3471.69	81.12
Model 4	**-3272.01**	-3496.99	112.49	**-3276.67**	-3489.01	106.17	**-3287.22**	-3492.75	102.76	-3298.69	-3501.63	101.47	-3308.29	-3497.44	94.58

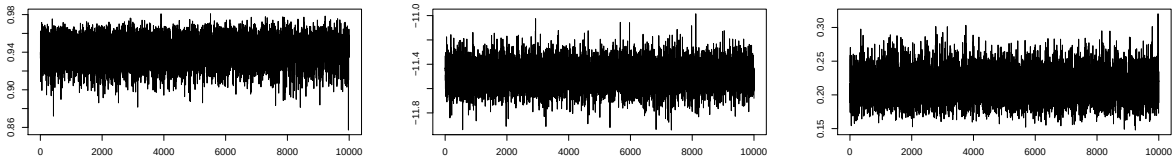
Table C.2

	Jump rate & jump magnitude											
	$\kappa = 0.013,\ \mu_\eta = 0.8$			$\kappa = 0.0064,\ \mu_\eta = 2$			$\kappa = 0.004,\ \mu_\eta = 2.6$			$\kappa = 0.001,\ \mu_\eta = 2.6$		
	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D	DIC	$D(\boldsymbol{\vartheta})$	p_D
Model 1	1222718.45	1209944.77	6386.84	850881.90	840710.30	5085.80	1176678	1174184	1247.13	385010.2	382755.9	1127.18
Model 2	**-3210.47**	-3568.32	178.92	**-3055.91**	-3492.26	218.17	**-3056.22**	-3624.52	284.15	**-2954.89**	-3367.02	206.06
Mdoel 3	56578.51	54348.87	1114.82	23271.85	19608.22	1831.82	59762.12	56986.58	1387.77	-2045.05	-2576.56	265.75

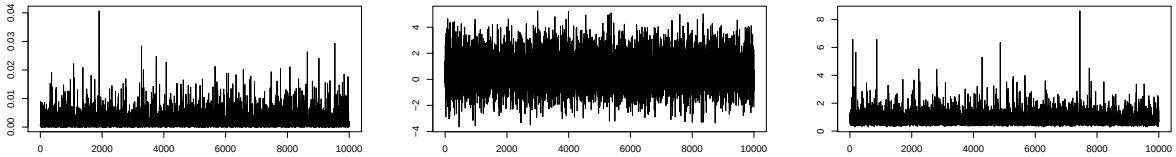
C.3 MCMC convergence trace plot on empirical part

C.3.1 MCMC trace plot of Model 5 on S&P 500 ETF

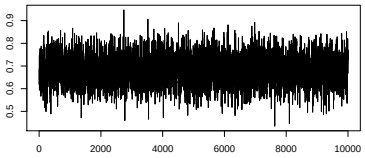
(a) ϕ **(b)** μ **(c)** $\sqrt{\sigma_e^2}$



(d) κ **(e)** μ_η **(f)** $\sqrt{\sigma_\eta^2}$



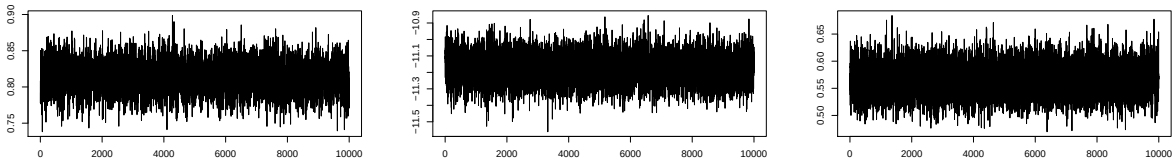
(g) b



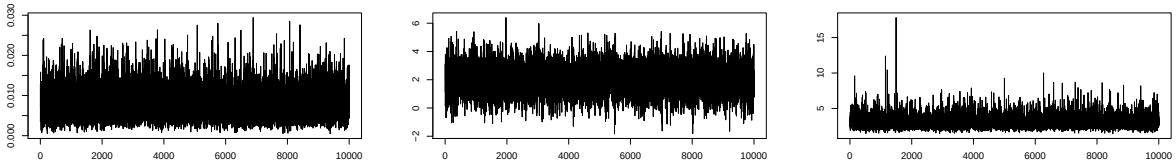
Note: In the figure above, we demonstrate the trace plot of MCMC samplings corresponding to the application of **Model 5** S&P 500 ETF in November 2015. Thus, the trace plot of posterior MCMC samplings of $\left\{\phi, \mu, \sqrt{\sigma_e^2}, \kappa, \mu_\eta, \sqrt{\sigma_\eta^2}, b\right\}$. This result is based on 1100000 MCMC samplings with the initial 100000 samplings as burn-in samplings. We save one draw every 100 samplings to construct the posterior sample for the remaining samplings.

C.4 MCMC trace plot of Model 5 on Apple Inc. stock price

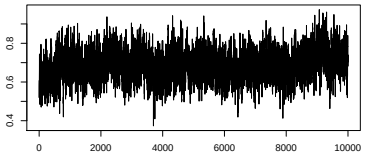
(a) ϕ **(b)** μ **(c)** $\sqrt{\sigma_e^2}$



(d) κ **(e)** μ_η **(f)** $\sqrt{\sigma_\eta^2}$



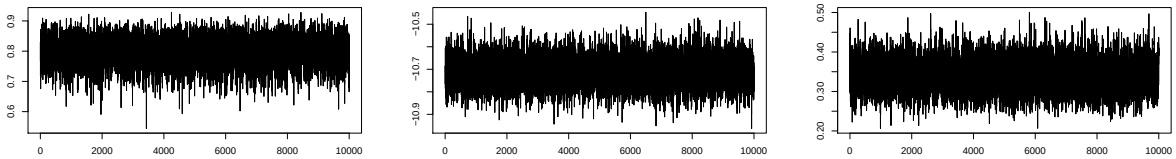
(g) b



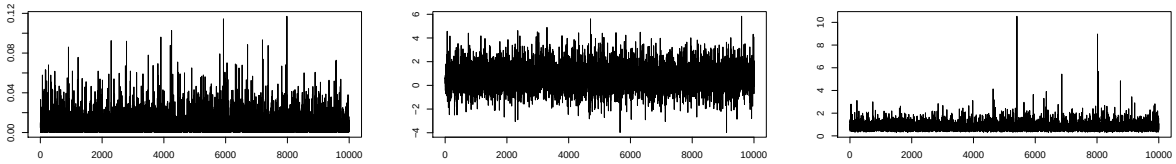
Note: In the figure above, we demonstrate the trace plot of MCMC samplings corresponding to the application of **Model 5** Apple Inc. stock price in August 2017. Thus, the trace plot of posterior MCMC samplings of $\left\{\phi, \mu, \sqrt{\sigma_e^2}, \kappa, \mu_\eta, \sqrt{\sigma_\eta^2}, b\right\}$. This result is based on 1100000 MCMC samplings with the initial 100000 samplings as burn-in samplings. We save one draw every 100 samplings to construct the posterior sample for the remaining samplings.

C.4.1 MCMC trace plot of Model 5 on CSI 300 index futures

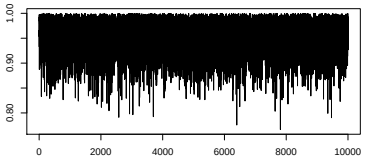
(a) ϕ **(b)** μ **(c)** $\sqrt{\sigma_e^2}$



(d) κ **(e)** μ_η **(f)** $\sqrt{\sigma_\eta^2}$



(g) b



Note: In the figure above, we demonstrate the trace plot of MCMC samplings corresponding to the application of **Model 5** CSI 300 index futures in August 2020. Thus, the trace plot of posterior MCMC samplings of $\left\{\phi, \mu, \sqrt{\sigma_e^2}, \kappa, \mu_\eta, \sqrt{\sigma_\eta^2}, b\right\}$. This result is based on 1100000 MCMC samplings with the initial 100000 samplings as burn-in samplings. We save one draw for every 100 samplings to construct the posterior sample for the remaining samplings.