Ca' Foscari
University
of Venice

Master Degree
in Quantitative Finance and Risk Management

Final Thesis

Bayesian Inference for Discrete Time Stochastic Volatility

Supervisor
Ch. Prof. Stefano Federico Tonellato

Graduand
Thao Nguyen Tran
Matriculation Number 906116

Academic Year
2025 / 2026

Contents

Appendices

List of Tables

List of Figures

Introduction

Modeling and forecasting financial volatility is a central problem in financial econometrics, with important implications for risk management, asset pricing, portfolio construction, and financial stability assessment. Volatility plays a fundamental role in measuring market uncertainty and directly affects the assessment of financial risk and the construction of predictive distributions for future returns.

Empirical financial returns exhibit several well-documented stylized facts that complicate volatility modeling, including volatility clustering, strong persistence in conditional variance, heavy-tailed return distributions, and asymmetric responses of volatility to return shocks (Cont, 2001). These features indicate that volatility evolves dynamically over time and cannot generally be treated as constant.

To account for such behaviour, Engle (1982) introduced the autoregressive conditional heteroskedasticity (ARCH) model, which was later generalized by Bollerslev (1986) through the generalized ARCH (GARCH) framework. In these models, conditional variance is specified as a deterministic function of past returns and past conditional variances. Owing to their tractability and forecasting performance, ARCH and GARCH models have become standard tools in empirical finance. Despite their success, the deterministic variance structure of ARCH-type models may limit their ability to capture smoothly evolving latent volatility processes and persistent uncertainty observed in financial markets (Taylor, 1986; Jacquier et al., 1994).

Stochastic volatility (SV) models provide a more flexible framework by treating volatility as an unobserved stochastic process evolving over time. Early contributions to SV modeling can be traced back to Taylor (1986), who introduced the idea of modeling financial volatility as a latent process. Since then, SV models have been extensively studied in the financial econometrics literature. A major contribution was provided by Jacquier et al. (1994), who developed one of the first Bayesian estimation procedures for SV models using a single-move Metropolis–Hastings algorithm for the latent volatility states. Although computationally demanding, this approach established the feasibility of Bayesian inference for SV models and stimulated subsequent research on more efficient sampling methods. A major breakthrough was provided by Kim et al. (1998), who introduced an offset mixture of normals to approximate the log-chi-square distribution, facilitating efficient Bayesian estimation of SV models. Subsequent research focused on

improving the efficiency of latent-state simulation through simulation smoothing (Durbin and Koopman, 2001) and block-sampling methods for Gaussian state-space models and Gaussian Markov random fields (McCausland et al., 2011; Rue, 2001). Most recently, the Ancillarity-Sufficiency Interweaving Strategy (ASIS) proposed by Yu and Meng (2011) and adapted for SV by Kastner and Frühwirth-Schnatter (2017) has established a robustly efficient sampling scheme that maintains high efficiency across all parameter ranges, overcoming the breakdown of standard samplers when the volatility of volatility is small.

In the standard SV specification, observed returns are driven by observation disturbances (or return shocks), while volatility evolves according to a separate latent stochastic process driven by latent innovations (volatility shocks) (see chapter 2). More precisely, the return shocks represent short-run noises affecting observed returns directly, whereas the volatility shocks govern the evolution of latent uncertainty over time. This separation between observation disturbances and latent innovations distinguishes SV models from deterministic volatility specifications such as GARCH models, where volatility is updated mechanically as a function of past returns.

Although the standard Gaussian SV model captures persistent volatility dynamics, it remains insufficient in several important respects. First, financial returns often exhibit extreme observations that occur more frequently than predicted by the Gaussian distribution. Ignoring this feature may lead to systematic underestimation of tail risk and poor density forecasting performance during periods of market stress. A common extension is therefore to replace Gaussian return shocks with Student-t disturbances, allowing the model to accommodate excess kurtosis behavior (Harvey et al., 1994b; Jacquier et al., 2004; Fridman and Harris, 1998; Liesenfeld and Jung, 2000; Lee and Koopman, 2004).

Second, empirical evidence suggests that negative return shocks tend to increase future volatility more strongly than positive shocks of the same magnitude, a phenomenon commonly referred to as the leverage effect, originally discussed by Black (1976). Standard SV specifications do not capture this asymmetry because observation disturbances and latent innovations are assumed to be independent. Leverage SV specifications address this issue by allowing correlation between return shocks and volatility shocks, thereby introducing asymmetric volatility responses consistent with observed equity market behavior (Harvey and Shephard, 1996; Jacquier et al., 2004; Yu, 2005; Omori et al., 2007; Nakajima and Omori, 2012).

In addition, while these univariate SV extensions improve the modeling of individual return series, many practical financial applications require the joint analysis of multiple assets. Portfolio allocation, risk management, and systemic risk assessment depend not only on individual volatility dynamics but also on the time-varying dependence structure across financial returns. Consequently, multivariate volatility models are considered to capture both dynamic variances and evolving cross-asset correlations in a coherent

probabilistic framework (Harvey et al., 1994b; Jacquier et al., 1994; Pitt and Shephard, 1999).

Since volatility is latent rather than directly observable, SV models are formulated within a state-space framework. The latent-state formulation naturally accommodates volatility clustering and persistent changes in uncertainty while allowing volatility shocks to evolve independently from observation disturbances. However, the latent-variable structure of SV models makes likelihood-based inference analytically intractable because its estimation requires integration over the entire latent volatility trajectory (Jacquier et al., 1994; Kim et al., 1998).

Bayesian methods (see chapter 1) provide a coherent framework for handling this uncertainty by jointly estimating latent states and model parameters through simulation-based methods. In particular, Markov Chain Monte Carlo (MCMC) methods provide a practical framework for conducting posterior inference in complex high-dimensional latent-variable models (Jacquier et al., 1994). By sampling jointly from the posterior distribution of model parameters and latent states, these methods naturally account for parameter uncertainty in predictive distributions and density forecasts.

In multivariate financial applications, additional challenges arise because the number of covariance parameters grows rapidly with the dimension of the system. Modeling time-varying covariance matrices directly quickly becomes computationally infeasible in high-dimensional settings. Factor stochastic volatility (FSV) models, formally developed by Pitt and Shephard (1999), address this problem by introducing a small number, K, of latent common factors that explain co-movement across assets. This factor structure substantially reduces dimensionality while preserving flexible time-varying covariance dynamics. As a result, FSV models provide a parsimonious framework for modeling multivariate financial volatility and cross-sectional dependence (Chib et al., 2006).

A central objective of volatility modeling is not only in-sample estimation but also predictive performance. Within the Bayesian framework, prediction is based on the posterior predictive distribution, which integrates over parameter and latent-state uncertainty, and therefore provides a coherent probabilistic framework for density forecasting and model comparison. Following the framework of Gneiting and Raftery (2007), predictive densities can be evaluated using strictly proper scoring rules, with particular emphasis on the logarithmic score. Because the logarithmic score rewards models that assign high probability to realized observations, it provides a natural criterion for comparing competing SV specifications, especially informative when differences in tail behavior and volatility dynamics are of primary interest.

While the extensions of standard SV model have each received considerable attention in the literature, their relative contributions to model performance are not always clear. Heavy-tailed innovations, leverage effects, and latent factor structures address different aspects of financial return dynamics, and their benefits may vary depending on the

empirical setting and inferential objective. A systematic comparison within a unified Bayesian framework therefore provides useful insight into the extent to which these modeling components improve volatility estimation, predictive performance, and the representation of dependence structures. This thesis studies several SV specifications within a unified state-space framework. The analysis considers the baseline SV model together with its extensions incorporating heavy-tailed return innovations and leverage effects. In addition, a multivariate FSV specification is examined in order to capture joint volatility dynamics across multiple financial return series.

This thesis is motivated by the following research questions.

- To what extent do heavy-tailed innovations and leverage effects improve the ability of stochastic volatility models to capture the main features of financial returns?
- How do these extensions affect posterior inference and predictive performance relative to the baseline Gaussian specification?
- In a multivariate setting, can latent factor structures provide an effective and computationally tractable representation of common volatility dynamics across assets?

The thesis aims to examine how heavy tails, leverage structures, and latent factor representations affect volatility dynamics, posterior inference, and multivariate dependence behavior in financial returns. Particular emphasis is placed on understanding the extent to which these model extensions improve the representation of persistent uncertainty, extreme return behavior, and common volatility movements across assets.

Beyond predictive density evaluation, the analysis also investigates posterior parameter estimation, latent volatility behavior, MCMC convergence diagnostics, and computational aspects of Bayesian inference in high-dimensional latent-variable models. Particular attention is paid to the trade-off between model flexibility and computational tractability, especially in multivariate FSV models where increasing model complexity may lead to more challenging posterior inference.

The remainder of the thesis is organized as follows. Chapter 1 introduces the Bayesian framework and computational methods used for posterior inference and predictive analysis. Chapter 2 presents the stochastic volatility models considered in the thesis, including univariate and multivariate factor specifications. Chapter 3 reports the results. Chapter 4 concludes and discusses limitations and directions for future research.

Chapter 1

Bayesian Framework and Computational Methods

This chapter introduces the Bayesian framework and computational methods employed throughout the thesis. It presents the key concepts of Bayesian inference, discusses Markov Chain Monte Carlo (MCMC) methods for posterior simulation, and outlines the construction and evaluation of predictive distributions.

1.1 Bayesian Framework

Bayesian inference provides a coherent framework for statistical modeling by combining prior information with the likelihood of the observed data. Rather than treating unknown parameters as fixed quantities, the Bayesian approach represents uncertainty through probability distributions and updates prior beliefs after observing the data (Robert, 2001).

Let $\mathbf{Y} = [Y_1, \ldots, Y_T]'$ a random sample, i.e. a T-variate random vector and $\mathbf{y} = \{y_1, \ldots, y_T\}$ the vector of observed data, i.e. the sample realization of $\mathbf{Y}$, with $\mathbf{y} \in \mathcal{Y}$, where $\mathcal{Y}$ represents the sample space. Let $\theta \in \Theta \subseteq \mathbb{R}^k$ be a vector of unknown parameters, where Θ represents the parameter space. The central object of Bayesian inference is the posterior distribution, which combines information from the likelihood and the prior distribution.

1.1.1 Likelihood, Prior, and Posterior

The likelihood function is defined as

$$L(\theta \mid y) = f(y \mid \theta), \tag{1.1}$$

where $f(y|\theta)$ denotes the probability density function of $\mathbf{Y}$ evaluated at the observed realization $\mathbf{y}$ conditional on the parameter θ. In general, the likelihood function summarizes the information about the unknown parameter vector θ contained in the observed data. In the Bayesian framework, inference is based on combining the information contained in the likelihood with prior beliefs about the parameters, represented by a prior distribution $\pi(\theta)$, through Bayes' theorem.

Combining the likelihood with the prior distribution yields the joint probability density function of the data and parameters

$$p(y, \theta) = f(y \mid \theta)\, \pi(\theta). \tag{1.2}$$

Conditioning on the observed data and applying Bayes' theorem yields the posterior distribution of θ,

$$\pi(\theta \mid y) = \frac{f(y \mid \theta)\, \pi(\theta)}{m(y)} \propto\ f(y \mid \theta)\, \pi(\theta) \tag{1.3}$$

where $m(y) = \int_\Theta f(y \mid \theta)\, \pi(\theta)\, d\theta$ is the marginal likelihood, which does not depend on θ. The proportionality highlights the fact that the posterior distribution can be expressed as the product of the likelihood times the prior density function of θ, up to a normalizing constant.

The posterior distribution $\pi(\theta|y)$ summarizes the uncertainty about the unknown parameters after observing the data and forms the basis for Bayesian estimation and prediction (Robert, 2001).

1.1.2 Posterior Inference

Given the posterior distribution $\pi(\theta \mid y)$, inference is typically based on summaries of this distribution. For a generic function $g(\theta)$, the posterior expectation is defined as

$$\mathbb{E}_\pi[g(\theta) \mid y] = \int_\Theta g(\theta)\, \pi(\theta \mid y)\, d\theta. \tag{1.4}$$

In many applications, the marginal likelihood $m(y)$ cannot be evaluated analytically because it involves a high-dimensional integration over the unknown quantities in the model. Consequently, Bayesian inference generally relies on simulation-based methods such as MCMC (see section 1.2). Based on the resulting posterior draws $\{\theta^{(m)}\}_{m=1}^M$, posterior expectations can be approximated by:

$$\mathbb{E}_\pi[g(\theta) \mid y] \approx \frac{1}{M} \sum_{m=1}^{M} g(\theta^{(m)}) \tag{1.5}$$

The Monte Carlo approximation above can be used not only to approximate posterior

expectations, but also to approximate posterior quantiles and higher-order moments (Robert and Casella, 2004).

While posterior expectations are among the most commonly used summaries, Bayesian inference often requires alternative summaries, including posterior quantiles and point estimators derived under different loss functions. From a decision theoretic perspective (Robert, 2001), a loss function $\ell(\theta, q)$ quantifies the penalty incurred when the decision or estimator q is taken while the true parameter value is θ. The corresponding Bayes risk is the posterior expected loss

$$R(q) = \mathbb{E}_\pi\big[\ell(\theta, q) \mid y_{1:T}\big], \tag{1.6}$$

which serves as the objective function in Bayesian decision problems. A Bayes estimator is any decision rule q^* that minimizes this posterior risk:

$$q^* = \arg\min_q R(q). \tag{1.7}$$

Different choices of loss functions lead to different posterior summaries. For example, under quadratic loss, the Bayes estimator is the posterior mean, whereas under the absolute loss, the Bayes estimator is the posterior median. Among the various loss functions, the asymmetric check losses is of particular interest because its corresponding Bayes estimator is the posterior quantile (see section B.1),

$$\ell_\alpha(\theta, q) = \big(\alpha - \mathbf{1}\{\theta < q\}\big)\,(\theta - q). \tag{1.8}$$

Consequently, the posterior α-quantile can be characterized as the solution to the Bayesian decision problem

$$\arg\min_q \mathbb{E}_\pi\left[\ell_\alpha(\theta, q) \mid y_{1:T}\right]. \tag{1.9}$$

In general, the asymmetric check loss provides a useful illustration of how different decision criteria lead to different posterior summaries, highlighting the close connection between Bayesian inference and decision theory.

1.1.3 Latent Variable Models

The discussion above has focused on Bayesian inference for models characterized by a finite-dimensional parameter vector θ. In many econometric applications, however, the observed data are also driven by latent quantities that evolve over time and cannot be directly observed. In such settings, inference must account not only for uncertainty about the model parameters but also about the latent states themselves. This leads naturally to the class of latent variable and state-space models (Harvey, 1989; Durbin and Koopman, 2001).

A defining feature of state-space models is the presence of an unobserved latent process. Let $h_{1:T} = \{h_1, \ldots, h_T\}$ denote latent states and let θ represent the static model parameters such as μ, ϕ, and σ_η (see chapter 2). Conditionally on the latent states, observations are independent with density

$$f(y_{1:T} \mid h_{1:T}, \theta) = \prod_{t=1}^{T} f(y_t \mid h_t, \theta), \tag{1.10}$$

while the latent process evolves according to a Markov transition density, such that, conditionally on h_{t-1} and θ, h_t is independent of h_s for all $s < t - 1$ (see section A.1). Hence,

$$f(h_{1:T} \mid \theta) = f(h_1 \mid \theta) \prod_{t=2}^{T} f(h_t \mid h_{t-1}, \theta). \tag{1.11}$$

The likelihood is obtained by integrating out the entire latent trajectory

$$f(y_{1:T} \mid \theta) = \int f(y_{1:T} \mid h_{1:T}, \theta)\, f(h_{1:T} \mid \theta)\, dh_{1:T}. \tag{1.12}$$

In non-linear and non-Gaussian state-space models, the likelihood integral in equation (1.12) is generally analytically intractable. This is because the dimension of the latent state $h_{1:T}$ grows linearly with the sample size T, while nonlinear or non-Gaussian components in the observation density $f(y_t \mid h_t, \theta)$ and the state transition density $f(h_t \mid h_{t-1}, \theta)$ destroy the linear-Gaussian structure required for closed-form filtering and smoothing. As a consequence, direct likelihood evaluation and classical optimization-based inference become computationally challenging (Durbin and Koopman, 2001). These difficulties are particularly evident in state-space models, motivating simulation-based approaches, in which rather than attempting to evaluate the likelihood analytically, Bayesian methods treat both the model parameters θ and latent states $h_{1:T}$ as unknown quantities and conduct inference through their joint posterior distribution

$$\pi(\theta, h_{1:T} \mid y_{1:T}) \propto f(y_{1:T} \mid h_{1:T}, \theta)\, f(h_{1:T} \mid \theta)\, \pi(\theta). \tag{1.13}$$

The resulting posterior distribution does not admit a closed-form representation. Consequently, posterior summaries must be obtained numerically, most commonly through Monte Carlo methods based on samples drawn from the joint posterior.

Posterior expectations are defined analogously with respect to the joint posterior,

$$\mathbb{E}_\pi[g(\theta, h) \mid y] = \int_\Theta g(\theta, h)\, \pi(\theta, h \mid y)\, d\theta\, dh. \tag{1.14}$$

1.2 Markov Chain Monte Carlo Methods (MCMC)

1.2.1 Basic Idea

MCMC methods provide a general framework for generating dependent draws from a target posterior distribution $\pi(\theta, h_{1:T} \mid y_{1:T})$, which was introduced in the previous section. The basic idea is to construct a time-homogeneous Markov chain whose stationary distribution coincides with the target posterior $\pi(\theta, h_{1:T} \mid y_{1:T})$ (Robert and Casella, 2004). If the chain is irreducible (able to reach all relevant regions of the state-space) and aperiodic (not trapped in deterministic cycles), it will converge to its stationary distribution (Tierney, 1994), allowing the generation of dependent draws $\{(\theta^{(m)}, h_{1:T}^{(m)})\}_{m=1}^{M}$ from the posterior. Under standard regularity conditions, the resulting Markov chain is ergodic, implying that posterior expectations can be approximated through Monte Carlo averages:

$$E[g(\theta, h_{1:T}) \mid y_{1:T}] \approx \frac{1}{M} \sum_{m=1}^{M} g(\theta^{(m)}, h_{1:T}^{(m)}), \tag{1.15}$$

implying that the Monte Carlo average converges to the corresponding posterior expectations as $M \to \infty$ or for large M in practice.

In finite samples, many MCMC algorithms exhibit slow mixing and high autocorrelation, which reduce sampling efficiency and limit exploration of the parameter space (Brooks et al., 2011). The following subsections describe the principal MCMC algorithms used throughout this thesis.

1.2.2 Gibbs Sampling and Metropolis-Hastings Algorithm

To generate draws from the posterior distribution $\pi(\theta, h_{1:T} \mid y_{1:T})$, one commonly used approach is Gibbs sampling, a special case of MCMC, in which model parameters and latent states are updated sequentially by sampling from their full conditional distributions (Gelfand and Smith, 1990).

Let the parameter vector be partitioned as $\theta = (\theta_1, \ldots, \theta_K)$. The Gibbs sampler generates draws by iteratively sampling from the full conditional distributions after fixing some starting values, $\theta_j^{(0)},\ j = 1, \ldots, K$:

$$\theta_1^{(m)} \sim \pi(\theta_1 \mid \theta_2^{(m-1)}, \ldots, \theta_K^{(m-1)}, y), \tag{1.16}$$

$$\theta_2^{(m)} \sim \pi(\theta_2 \mid \theta_1^{(m)}, \theta_3^{(m-1)}, \ldots, \theta_K^{(m-1)}, y), \tag{1.17}$$

$$\vdots$$

$$\theta_K^{(m)} \sim \pi(\theta_K \mid \theta_1^{(m)}, \ldots, \theta_{K-1}^{(m)}, y). \tag{1.18}$$

The resulting sequence forms a Markov chain whose stationary distribution coincides with the target posterior distribution. Consequently, after a sufficiently long run, the simulated draws can be regarded as samples from the posterior distribution of interest. Under the standard regularity conditions, such as irreducibility and aperiodicity, the Markov chain is ergodic, implying that empirical averages computed from the simulated draws converge to the corresponding posterior expectations (Brooks et al., 2011; Tierney, 1994).

While these conditions guarantee convergence in principle, the Gibbs sampler also requires the full conditional distributions to be consistent with the target posterior distribution. This ensures that the chain is able to explore the full support of the target posterior distribution rather than becoming confined to a restricted region. If these conditions are violated, for example, if a full conditional restricts the chain to a subset of the parameter space or if the posterior is not proper, the Markov chain cannot explore the entire space, and in such circumstances, the convergence to the posterior distribution fails (Robert and Casella, 2004).

Even when these theoretical conditions are satisfied, the practical performance of the Gibbs sampler depends on how rapidly the Markov chain approaches its stationary distribution. In practice, convergence may be affected by the choice of starting values and by the strong dependence structure among parameters, which may slow down the movement of the chain. Because the chain is typically initialized away from its stationary distribution, the early draws may not be representative of the target posterior distribution. Consequently, an initial burn-in period is typically discarded before posterior inference is conducted (Gilks et al., 1996).

The effectiveness of Gibbs sampling also depends on the availability of tractable full conditional distributions. When the full conditional distributions have known forms that are easy to sample from, Gibbs sampling provides an efficient and practical method for posterior inference. In contrast, when some full conditionals are not available in closed-form, direct Gibbs updates are no longer feasible. In such cases, the Metropolis-Hastings (MH) algorithm provides a more general approach for sampling from the posterior distribution (Hastings, 1970; Robert and Casella, 2004).

The basic idea of the MH algorithm is to construct a Markov chain by proposing candidate values and accepting or rejecting them according to a probabilistic rule. Given a current value $\theta^{(m-1)}$, a candidate value θ' is proposed from a proposal distribution $q(\theta' \mid \theta^{(m-1)})$. The proposed value is accepted with probability

$$\alpha(\theta^{(m-1)}, \theta') = \min\left\{1, \frac{\pi(\theta' \mid y)\, q(\theta^{(m-1)} \mid \theta')}{\pi(\theta^{(m-1)} \mid y)\, q(\theta' \mid \theta^{(m-1)})}\right\}. \tag{1.19}$$

The acceptance probability corrects for the discrepancy between the proposal distribution and the target posterior distribution. This construction ensures that the resulting Markov chain admits the target posterior distribution as its stationary distribution (Metropolis

et al., 1953; Hastings, 1970).

If the proposal is accepted, the chain moves to $\theta^{(m)} = \theta'$, otherwise it remains at the current state, $\theta^{(m)} = \theta^{(m-1)}$. Unlike Gibbs sampling, the MH algorithm does not require direct sampling from the full conditional distributions. Instead, it only requires evaluation of the posterior density up to a proportionality constant, making it applicable in settings where full conditional distributions do not admit standard forms. This flexibility makes the MH algorithm an important component of many Bayesian samplers (Chib and Greenberg, 1995).

1.2.3 MCMC in Latent Variable Models

In latent-variable models, the posterior distribution typically exhibits strong dependence between the parameters θ and the latent state sequence $h_{1:T}$. This dependence often leads to slow mixing and reduced sampling efficiency, motivating MCMC algorithms that update parameters and latent states in separate blocks (Shephard, 1994):

$$\theta^{(m)} \sim \pi(\theta \mid h_{1:T}^{(m-1)}, y_{1:T}), \quad h_{1:T}^{(m)} \sim \pi(h_{1:T} \mid \theta^{(m)}, y_{1:T}), \tag{1.20}$$

where each step may involve embedded MH updates when closed-form sampling from the full conditional distributions is not available.

To improve convergence and mixing efficiency, modern implementations often combine the Gibbs and Metropolis-Hastings steps with auxiliary variable techniques and simulation smoothing to jointly sample model parameters and latent states (Gelfand and Smith, 1990; Frühwirth-Schnatter, 2006). In addition, centered and non-centered parameterizations (see section 2.5.3) are frequently employed to reduce posterior dependence between parameters and latent states, thereby enhancing the overall efficiency of the sampling algorithm (Robert and Casella, 2004; Yu and Meng, 2011).

1.2.4 Convergence Diagnostics

Regardless of the sampling algorithm employed, convergence diagnostics are necessary to verify whether the Markov chain has converged to the target posterior distribution and whether the simulated draws adequately explore the posterior support (Brooks et al., 2011). Following the recommendations of Cowles and Carlin (1996), a combination of visual inspections and formal statistical tests are employed rather than relying on a single metric as different diagnostics target distinct failure modes.

Graphical diagnostics include trace plots and estimated autocorrelation functions (ACF). In particular, trace plots are used to verify if the chain fluctuates around a stable region without exhibiting persistent trends or structural drifts; whereas ACF evaluates the serial dependence between MCMC draws, and is commonly used to evaluate mixing.

For a scalar parameter ψ, the lag-k autocorrelation is defined as

$$\rho(k) = \mathrm{Corr}\big(\psi^{(m)}, \psi^{(m+k)}\big). \tag{1.21}$$

A well-mixed chain should exhibit rapidly decaying autocorrelations, with $\rho(k)$ approaching zero within a small number of lags, indicating a more effective exploration of the posterior distribution. Slow decay suggests strong serial dependence in the chain and reduces sampling efficiency.

A commonly used formal diagnostic is the Geweke statistic, which compares the means of two non-overlapping segments of the Markov chain (Geweke, 1992). Let $\bar{\psi}_A$ and $\bar{\psi}_B$ denote the sample means of an early segment and a late segment of the chain, respectively. The Geweke statistic is given by:

$$Z_{\mathrm{G}} = \frac{\bar{\psi}_A - \bar{\psi}_B}{\sqrt{\widehat{\mathrm{Var}}(\bar{\psi}_A) + \widehat{\mathrm{Var}}(\bar{\psi}_B)}}, \tag{1.22}$$

where the variances of the sample mean must account for autocorrelation in the MCMC draws, and are given by

$$\mathrm{Var}(\bar{\psi}) = \frac{1}{n}\left(\gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k\right), \tag{1.23}$$

where $\gamma_k = \rho(k)$ defined as above. Under the null hypothesis of convergence, $Z_{\mathrm{G}} \sim \mathcal{N}(0,1)$. Values of $|Z_{\mathrm{G}}|$ exceeding conventional thresholds indicate that the chain has not yet reached stationarity.

Another important quantity is the effective sample size (ESS), which measures the amount of approximately independent information contained in the correlated MCMC draws. For a Markov chain of length M, the ESS is commonly approximated by

$$\mathrm{ESS} = \frac{M}{1 + 2\sum_{k=1}^{\infty}\rho(k)}, \tag{1.24}$$

where $\rho(k)$ denotes the lag-k autocorrelation of the chain as above. The ESS may be interpreted as the number of independent draws that would provide the same amount of information as the correlated MCMC sample. High autocorrelation increases the denominator and therefore reduces the ESS, implying less efficient posterior simulation (Brooks et al., 2011).

1.3 Posterior Predictive Inference

1.3.1 Posterior Predictive Distribution

Bayesian inference naturally extends from parameter estimation to the prediction of future observations. Rather than conditioning on a single point estimate of the parameter θ, posterior predictive inference integrates over the entire posterior distribution, thereby accounting for parameter uncertainty. In state-space models, posterior predictive distributions are useful both for comparing alternative model specifications and for forecasting future volatility dynamics.

The posterior predictive distribution of a future observation is obtained by integrating over the posterior distribution of the parameters and latent states (Geweke and Amisano, 2010). Given observed data $y = \{y_1, \ldots, y_T\}$, the posterior predictive distribution y_{T+1} is:

$$f(y_{T+1} \mid y) = \int_{\Theta} \int_{\mathbb{R}} f(y_{T+1} \mid \theta, h_{T+1})\, p(\theta, h_{T+1} \mid y)\, dh_{T+1}\, d\theta, \tag{1.25}$$

where $p(\theta, h_{T+1} \mid y)$ is the joint predictive distribution of parameter and future state. This expression shows that predictive inference averages the likelihood of future observations with respect to the posterior distribution of the parameters, thereby incorporating both model intrinsic uncertainty (uncertainty about latent variables) and parameter uncertainty. In latent variable models, this predictive distribution does not have a closed-form expression. Instead, predictive quantities are typically computed using Monte Carlo methods based on draws from the posterior distribution.

Given posterior draws $\{(\theta^{(m)}, h^{(m)})\}_{m=1}^{M}$, predictive quantities can be approximated by simulating future observations from the conditional model. The predictive density can be approximated as

$$f\big(y_{T+1} \mid y_{1:T}\big) \approx \frac{1}{M} \sum_{m=1}^{M} f\big(y_{T+1} \mid \theta^{(m)}, h_{T+1}^{(m)}\big), \tag{1.26}$$

where $h_{T+1}^{(m)}$ is drawn from the state transition equation conditional on $(\theta^{(m)}, h_T^{(m)})$. This approximation is consistent with $M \to \infty$.

1.3.2 Predictive Return Distributions in Latent Variable Models

For state-space models, posterior predictive inference is implemented by propagating the latent volatility process forward through the state equation. Given posterior draws of the latent state and model parameters, future volatility states can be simulated according to

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_{t+1}. \tag{1.27}$$

Since the latent state h_{t+1} represents the logarithm of the conditional variance, it determines the scale of future return innovations. Under the observation equation

$$y_{t+1} = \beta + \exp\left(\frac{h_{t+1}}{2}\right)\varepsilon_{t+1}, \tag{1.28}$$

the conditional variance of future returns is

$$\mathrm{Var}(y_{t+1} \mid h_{t+1}) = \exp(h_{t+1}). \tag{1.29}$$

Consequently, uncertainty about future volatility translates into uncertainty about into future returns. Simulating both h_{t+1} and the corresponding return disturbance ε_{t+1} therefore yields draws from the posterior predictive distribution of y_{t+1}.

1.3.3 Predictive Likelihood and Log Predictive Score

Given the predictive return distribution, predictive performance can be assessed through the posterior predictive density evaluated at the realized value of particular interest (Geweke and Amisano, 2010). Let $\mathcal{F}_t = \{y_1, \ldots, y_t\}$ denote the information set available at time t. At each time t, the one-step-ahead predictive likelihood is given by

$$f(y_{t+1} \mid \mathcal{F}_t), \tag{1.30}$$

Taking logarithms yields the log predictive score,

$$\log f(y_{t+1} \mid \mathcal{F}_t). \tag{1.31}$$

Because predictive performance is typically evaluated over a sequence of observations rather than a single forecast origin, the one-step-ahead log predictive scores are aggregated sequentially over the evaluation period. For an evaluation sample of size H, the cumulative log predictive score is defined as

$$\log f(y_{T_0+1}, \ldots, y_{T_0+H} \mid \mathcal{F}_{T_0}) = \sum_{h=1}^{H} \log f(y_{T_0+h} \mid \mathcal{F}_{T_0+h-1}), \tag{1.32}$$

where $T_0 + 1$ denotes the starting point of the period.

The cumulative log predictive score measures how much the posterior predictive distribution agrees with the realized of observations on $\{T_0 + 1, \ldots, T_0 + H\}$. As such, it provides a criterion for comparing alternative model specifications, with higher values indicating a better ability to provide reliable predictions.

To evaluate predictive densities over a period, a sequential estimation scheme is commonly employed. At each forecast origin t, the model is estimated using all

observations available up to time t, and the one-step-ahead predictive density for y_{t+1} is constructed. Once the realization y_{t+1} becomes available, it is incorporated into the dataset, and the model is re-estimated using the expanded sample. The predictive density for the next observation y_{t+2} is then computed. This procedure is repeated over the entire out-of-sample period, ensuring that forecasts are generated using only information available at the time they are made.

1.3.4 Multi-step Forecasting

The predictive performance above are primarily used to evaluate the relative performance of alternative model specifications for model comparison purposes. The state-space formulation, however, facilitates multi-step forecasting beyond one-step-ahead density evaluation.

For forecasting purposes, the posterior draws obtained from the MCMC output are propagated forward through the state-space representation of the model. In particular, for each retained posterior draw, future latent volatility states are generated recursively according to

$$h_{T+1}^{(m)} = \mu^{(m)} + \phi^{(m)} \left(h_T^{(m)} - \mu^{(m)} \right) + \sigma^{(m)} \eta_{T+1}^{(m)}, \tag{1.33}$$

followed by $h_{T+2}^{(m)}, \ldots, h_{T+H}^{(m)}$.

Conditionally on the simulated volatility path, future returns are generated from the observation equation. For the Gaussian SV specification,

$$y_{T+h}^{(m)} = \exp\left(\frac{h_{T+h}^{(m)}}{2} \right) \varepsilon_{T+h}^{(m)}, \tag{1.34}$$

while for the heavy-tailed SVt model,

$$y_{T+h}^{(m)} = \exp\left(\frac{h_{T+h}^{(m)}}{2} \right) t_{\nu}^{(m)}. \tag{1.35}$$

The leverage specifications additionally account for dependence between return and volatility innovations.

Repeating this procedure across all posterior draws produces a Monte Carlo samples from the predictive distributions $\{f(y_{T+1} \mid y_{1:T}), \ldots, f(y_{T+H} \mid y_{1:T})\}$. For each forecast horizon h, the collection $\left\{ y_{T+h}^{(1)}, y_{T+h}^{(2)}, \ldots, y_{T+h}^{(M)} \right\}$ forms a sample from the posterior predictive distribution.

Chapter 2

Stochastic Volatility Models

The stochastic volatility (SV) models considered in this chapter is an application of the latent-variable state-space framework mentioned in chapter 1. In SV models, volatility is treated as an unobserved latent process that evolves over time and drives the distribution of observed returns. This representation allows the Bayesian inference and MCMC methods introduced in the previous chapter to be applied in a unified manner across the alternative model specifications considered below.

The chapter begins with the baseline specification, and then considers extensions with heavy-tailed disturbances and leverage effects. Finally, a multivariate factor stochastic volatility (FSV) model is introduced to capture the joint dynamics of multiple return series.

2.1 Baseline Stochastic Volatility Model (SV)

The baseline SV model provides the fundamental building block for the more elaborate specifications (Taylor, 1986; Kim et al., 1998) considered later in this chapter. It combines an observation equation for returns with a state equation for the latent log-volatility process.

2.1.1 Model Specification

Let y_t denote the continuously compounded return at time t. For $t = 1, \ldots, T$, the baseline SV model is defined by the state-space representation:

$$y_t = \beta + \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1), \tag{2.1}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2), \tag{2.2}$$

where h_t denotes the latent log-volatility process.

Conditionally on the latent state h_t, returns are Gaussian with conditional distribution

$$y_t \mid h_t \sim \mathcal{N}\left(\beta, e^{h_t}\right), \tag{2.3}$$

so that the conditional variance evolves stochastically over time according to the latent autoregressive process. The observation equation links returns to the latent volatility state, while the state equation governs the persistence and evolution of latent volatility process over time.

2.1.2 Parameter Interpretation

The parameter vector $\theta = (\beta, \mu, \phi, \sigma_\eta^2)$ characterizes the dynamics of the volatility process:

- β denotes the constant mean return in the observation equation
- μ represents the unconditional mean level of the log-volatility process,
- $\phi \in (-1, 1)$ controls the persistence of volatility,
- σ_η^2 determines the variability of the volatility innovations.

In practical applications, estimates of ϕ are often close to unity, indicating highly persistent volatility dynamics while still remaining within the stationary region.

2.1.3 Stationarity and Unconditional Moments

The log-volatility process $\{h_t\}$ follows a Gaussian AR(1) specification as in Kim et al. (1998). Under standard conditions for autoregressive processes, the process is covariance-stationary if $|\phi| < 1$. Under stationarity, the unconditional distribution is

$$h_t \sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right) \tag{2.4}$$

The persistence induced by the AR(1) dynamics implies that periods of high and low volatility tend to cluster over time, a key stylized feature of financial return series. Furthermore, integrating over the latent volatility distribution produces a heavy-tailed, non-Gaussian unconditional return distribution. Consequently, the baseline SV model is able to capture both volatility persistence and excess kurtosis through the latent volatility process (see section B.3.2).

2.2 Extensions of the Univariate Stochastic Volatility Model

While the baseline SV model captures time-varying and persistent volatility, it relies on the assumption of conditionally Gaussian returns and does not account for asymmetric responses of volatility to shocks. This section introduces extensions of the SV model that address these features.

2.2.1 SV Model with Heavy-Tailed Errors (SVt)

To account for excess kurtosis in financial returns, in which extreme movements in financial returns occur more frequently than predicted by the normal distribution, the Gaussian assumption on the observation disturbance can be relaxed by introducing heavy-tailed errors. A common approach is to assume that the return disturbances follow a Student-t distribution (Jacquier et al., 2004):

$$y_t = \beta + \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \quad \varepsilon_t \sim t_\nu(0,1), \tag{2.5}$$

where $t_\nu(0,1)$ denotes the standardized Student-t distribution, and the degrees of freedom parameter ν governs the tail behavior of the Student-t distribution. Small values of ν imply heavier tails, allowing for a higher probability of extreme observations. While the model is well-defined for any $\nu > 0$, the condition $\nu > 2$ is required for the variance of the Student-t distribution to be finite.

In this specification, the latent log-volatility process still evolves according to the AR(1) dynamics

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2), \tag{2.6}$$

Relative to the Gaussian specification, the Student-t probability distribution allocates greater probability mass to extreme return realizations, thereby improving robustness to large shocks and accommodating the excess kurtosis commonly observed in financial returns.

While the Student-t specification accommodates heavy-tailed return disturbances, Bayesian estimation is greatly facilitated by its Gaussian scale-mixture representation (Kim et al., 1998):

$$\varepsilon_t \mid \lambda_t \sim \mathcal{N}(0, \lambda_t^{-1}), \quad \lambda_t \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \tag{2.7}$$

where λ_t denotes an auxiliary latent scaling variable. Integrating out the latent scaling

variable λ_t recovers the marginal Student-t distribution for ε_t

$$f(\varepsilon_t) = \int_0^\infty f(\varepsilon_t \mid \lambda_t)\, f(\lambda_t)\, d\lambda_t = \int_0^\infty \mathcal{N}(\varepsilon_t \mid 0, \lambda_t^{-1})\, \text{Gamma}\left(\lambda_t \middle| \frac{\nu}{2}, \frac{\nu}{2}\right) d\lambda_t, \tag{2.8}$$

which evaluates to a Student-t density with ν degrees of freedom, $\varepsilon_t \sim t_\nu$.

This representation expresses the Student-t distribution as a Gaussian variance mixture with random scaling, allowing heavy-tailed behavior to be incorporated through latent scale variables while retaining conditional Gaussianity for Bayesian computation. Consequently, conditionally on the latent scales λ_t, the density $f(y_t \mid h_t)$ becomes conditionally Gaussian, enabling efficient data augmentation and Gibbs-type sampling methods within the MCMC estimation procedure.

2.2.2 SV Model with Leverage Effect (SVl)

Financial markets often exhibit an asymmetric relationship between returns and volatility, commonly referred to as leverage effect: negative returns tend to increase future volatility more than positive returns of the same magnitude.

To capture this feature, the SV model can be extended by allowing correlation between return disturbance ε_t and the volatility innovation η_t (Harvey and Shephard, 1996). The leverage SV model is specified as

$$y_t = \beta + \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \tag{2.9}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \tag{2.10}$$

where the innovations satisfy the joint distribution

$$\begin{pmatrix}\varepsilon_t \\ \eta_t\end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix}0 \\ 0\end{pmatrix}, \begin{pmatrix}1 & \rho\sigma_\eta \\ \rho\sigma_\eta & \sigma_\eta^2\end{pmatrix}\right). \tag{2.11}$$

The leverage parameter $\rho \in (-1, 1)$ controls the contemporaneous dependence between return shocks and volatility shocks. A negative value of ρ implies that negative return shocks are associated with increases in future volatility, thereby generating asymmetric volatility responses consistent with empirical findings in equity markets.

The two extensions described above can be combined to account for both heavy tails and leverage effects. The resulting SVtl model incorporates Student-t errors in the observation equation and allows for correlation between return and volatility innovations (Jacquier et al., 2004; Omori et al., 2007; Nakajima and Omori, 2012).

This specification provides a relatively flexible framework capable of capturing both heavy tails and leverage effects.

2.3 Multivariate Factor Stochastic Volatility Models (FSV)

While univariate SV models describe the dynamics of a single return series, multivariate applications are required to model the joint dependence structure across multiple assets. Direct modeling of a full time-varying covariance matrix becomes increasingly difficult in high-dimensional settings because the number of covariance parameters grows quadratically with the dimension of the system.

Factor stochastic volatility (FSV) models address this issue by introducing a low-dimensional latent factor structure (Pitt and Shephard, 1999; Aguilar and West, 2000). Instead of modeling the entire covariance matrix directly, co-movements between asset returns are explained through a small number of latent common factors with stochastic volatility dynamics. This approach substantially reduces dimensionality while preserving flexible time-varying dependence structures.

2.3.1 Model Specification

Let $y_t = (y_{1t}, \ldots, y_{Nt})'$ denote an N-dimensional vector of asset returns at time t. The FSV model is given by:

$$y_t = \beta + \Lambda f_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_{\varepsilon,t}), \quad f_t \sim \mathcal{N}(0, \Sigma_{f,t}), \tag{2.12}$$

where $\beta \in \mathbb{R}^N$ is a vector of intercept terms capturing the unconditional mean of the return series. The matrix $\Lambda \in \mathbb{R}^{N \times K}$ denotes the $N \times K$ matrix of factor loadings, where $K \ll N$ is the number of latent factors. The latent vector f_t captures the common dynamic movements shared across assets, while the idiosyncratic component ε_t represents asset-specific shocks (Harvey et al., 1994a; Jacquier et al., 1995). To allow both sources of variation to exhibit time-varying volatility, the covariance matrices of the factor and idiosyncratic innovations are specified as

$$\Sigma_{f,t} = \operatorname{diag}\left(e^{h_{f,1,t}}, \ldots, e^{h_{f,K,t}}\right), \tag{2.13}$$

and

$$\Sigma_{\varepsilon,t} = \operatorname{diag}\left(e^{h_{\varepsilon,1,t}}, \ldots, e^{h_{\varepsilon,N,t}}\right), \tag{2.14}$$

where both covariance matrices are diagonal and evolve according to stochastic volatility processes. Combining the factor and idiosyncratic components yields the conditional covariance matrix of returns,

$$\operatorname{Var}(y_t \mid h_t) = \Lambda \Sigma_{f,t} \Lambda' + \Sigma_{\varepsilon,t}. \tag{2.15}$$

The term $\Lambda\Sigma_{f,t}\Lambda'$ represents the common volatility structure induced by the latent factors. Since the number of factors K is typically much smaller than the number of asset N, the resulting matrix is low-rank, with rank at most equal to K. This component captures the co-movement between assets generated by common economic or financial shocks, such as market-wide or sector-specific shocks. The matrix $\Sigma_{\varepsilon,t}$ captures idiosyncratic asset-level volatility and accounts for variation that is not explained by the common factors. In practical applications, this component is often assumed to be diagonal, implying that shocks are conditionally independent across assets after controlling for the factor structure Pitt and Shephard (1999); Aguilar and West (2000).

Since both covariance components, $\Lambda\Sigma_{f,t}\Lambda'$ and $\Sigma_{\varepsilon,t}$, are time-varying, the conditional covariance matrix $\mathrm{Var}(y_t \mid h_t)$ evolves dynamically over time, allowing the model to jointly capture time-varying variances and correlations. Moreover, the low-dimensional factor structure provides a parsimonious representation of high-dimensional covariance dynamics, substantially reducing the number of parameters relative to unrestricted multivariate volatility models. This reduction mitigates the curse of dimensionality and makes estimation feasible even in large financial systems (Kastner et al., 2017).

To complete the specification, the latent log-volatility process underlying $\Sigma_{f,t}$ and $\Sigma_{\varepsilon,t}$ are modeled as independent AR(1) processes with persistence parameters satisfying $|\phi_{f,k}| < 1$ and $|\phi_{\varepsilon,i}| < 1$ to ensure stationarity. For factor volatilities:

$$h_{f,k,t} = \phi_{f,k} h_{f,k,t-1} + \sigma_{f,k}\eta_{f,k,t} \tag{2.16}$$

for $k = 1, \ldots, K$ where $\eta_{f,k,t} \sim \mathcal{N}(0, 1)$.

Similarly, idiosyncratic volatilities evolve as

$$h_{\varepsilon,i,t} = \mu_{\varepsilon,i} + \phi_{\varepsilon,i}(h_{\varepsilon,i,t-1} - \mu_{\varepsilon,i}) + \sigma_{\varepsilon,i}\eta_{\varepsilon,i,t}. \tag{2.17}$$

for $i = 1, \ldots, N$ where $\eta_{\varepsilon,i,t} \sim N(0, 1)$ is independent of $\eta_{f,k,t}$ for all i and k.

This specification allows both common market-wide volatility movements and asset-specific volatility dynamics to evolve stochastically over time (Kastner et al., 2017).

2.3.2 Identification Problems

Rotational Invariance and Sign Indeterminacy

Despite the flexibility of the FSV specification, the models are not identified without additional restrictions because the likelihood depends only on the product Λf_t. In particular, the factor structure is subject to rotational invariance: for any orthogonal matrix $P \in \mathbb{R}^{K\times K}$, satisfying $PP' = P'P = I$, the transformation $\Lambda^* = \Lambda P'$ and $f_t^* = Pf_t$ generate exactly the same conditional covariance matrix $\Lambda\Sigma_{f,t}\Lambda' = \Lambda^*\Sigma_{f,t}^*(\Lambda^*)'$,

where $\Sigma^*_{f,t} = P\Sigma_{f,t}P'$. Therefore, the likelihood depends only on the product Λf_t rather than on the individual factor loadings and factors. As a result, many different pairs of (Λ, f_t) are observationally equivalent, implying that the factor loadings and latent factors are not uniquely identified from the data (Aguilar and West, 2000; Chib et al., 2006).

An additional problem is sign indeterminacy, in which the signs of the latent factors and factor loadings are not uniquely determined. Without positivity constraints, the transformation

$$\Lambda \to -\Lambda, \quad f_t \to -f_t \tag{2.18}$$

leaves the likelihood unchanged. In other words, reversing the signs of both the factors and their corresponding loadings produces exactly the same model fit.

To restore identification, additional restrictions are required on the factor loading matrix and the factor covariance structure. A commonly used approach is to restrict the upper $K \times K$ block of the factor loadings matrix $\Lambda \in \mathbb{R}^{N\times K}$ to be lower-triangular with strictly positive diagonal elements, while the remaining $(N-K)$ rows are left unrestricted:

$$\Lambda = \begin{pmatrix} \lambda_{11} & 0 & 0 & \cdots & 0 \\ \lambda_{21} & \lambda_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \lambda_{K1} & \lambda_{K2} & \lambda_{K3} & \cdots & \lambda_{KK} \\ \lambda_{K+1,1} & \lambda_{K+1,2} & \lambda_{K+1,3} & \cdots & \lambda_{K+1,K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{N1} & \lambda_{N2} & \lambda_{N3} & \cdots & \lambda_{NK} \end{pmatrix}, \qquad \lambda_{kk} > 0. \tag{2.19}$$

The lower triangular structure removes rotational ambiguity by imposing an ordering on the latent factors. Under the triangular specification, the first factor is introduced through the first variable via Λ_{11}, the second factor through the second variable, and so forth. As a result, this creates an ordered factor structure in which each factor contributes new variation not already explained by previous factors. The positivity constraints on the diagonal elements further eliminate sign indeterminacy. As a result, posterior inference for the latent factors and factor loadings becomes well-defined and admits an economically interpretable representation.

Scale Identification

Although lower-triangular restrictions on the factor loading matrix provide an orientation convention for the latent factors and eliminate much of the rotational indeterminacy inherent in factor models, they do not by themselves fully identify a FSV model. In particular, lower-triangular restrictions determine the ordering of the factors and, when combined with positive diagonal elements, resolve sign ambiguity. However,

they do not identify the scale of the latent factors.

Consider a positive diagonal scaling matrix $S = \text{diag}(s_1, \ldots, s_K), s_j > 0$, and define the transformation $\Lambda^* = \Lambda S, \quad f_t^* = S^{-1} f_t$. If the unconditional levels of the factor log-volatility processes are freely estimated, the factor variances can be rescaled simultaneously so that $D_t^* = S^{-1} D_t S^{-1}$. Under this transformation, $\Lambda^* D_t^* \Lambda^{*\prime} = \Lambda S(S^{-1} D_t S^{-1}) S\Lambda' = \Lambda D_t \Lambda'$, implying that the factor component of the covariance matrix remains unchanged. Consequently, the likelihood cannot distinguish between the original and rescaled representations. Since multiplying a loading column by a positive constant preserves the lower-triangular zero pattern, triangular restrictions alone are insufficient for scale identification (Pitt and Shephard, 1999; Chib et al., 2006).

Therefore, an additional normalization is required. In FSV models, scale identification is typically achieved either by fixing selected loadings, for example through constraints such as $\lambda_{ii} = 1, i = 1, \ldots, K$, or by fixing the unconditional levels of the factor log-volatility processes. Consider the factor log-volatility process

$$h_{f,j,t} = \mu_{f,j} + \phi_{f,j}\big(h_{f,j,t-1} - \mu_{f,j}\big) + \sigma_{f,j}\, \eta_{f,j,t}, \quad \eta_{f,j,t} \sim \mathcal{N}(0,1). \tag{2.20}$$

Under stationarity, the unconditional expectation of the process is

$$\mathbb{E}(h_{f,j,t}) = \mu_{f,j}. \tag{2.21}$$

Therefore, $\mu_{f,j}$ represents the unconditional level, or long-run mean, of the factor log-volatility process. Since the factor variance is given by $\exp(h_{f,j,t})$, the parameter $\mu_{f,j}$ determines the unconditional variance level of the corresponding latent factor and hence its scale. If $\mu_{f,j}$ is freely estimated, the scale of a loading column can be increased while the factor volatility level is shifted downward by an offsetting amount, leaving the implied covariance structure unchanged. Consequently, the model remains scale unidentified, implying that the likelihood cannot distinguish between the original and rescaled parameterizations. Since column-wise rescaling preserves the lower-triangular zero pattern, triangular restrictions alone are insufficient for scale identification. An additional normalization is therefore required, typically by fixing $\mu_{f,j} = 0$ for all factors or by imposing constraints on selected factor loadings. Such restrictions remove the scale indeterminacy by assigning the factor scale either to the latent volatility process or to the loading matrix.

2.3.3 Computational Challenges

Identification issues should also be distinguished from computational issues. Even after an identified parameterization has been imposed, estimation of FSV models remains computationally demanding. The main difficulty arises from the strong posterior

dependence among factor loadings, latent factors, volatility states, and model parameters.

Because different combinations of these latent quantities can often provide a similar fit to the observed data, the posterior distribution may contain highly correlated regions that are difficult for MCMC algorithms to explore efficiently. Consequently, posterior draws tend to exhibit substantial autocorrelation, slow mixing, and reduced effective sample sizes.

These computational challenges become more pronounced as model complexity increases. In particular, larger factor dimensions introduce additional dependence among latent factors and loadings, while highly persistent volatility processes generate strong temporal dependence in the latent states. Both features increase the complexity of posterior exploration and may substantially slow convergence of the MCMC algorithm (Kim et al., 1998; Aguilar and West, 2000).

Part of this computational burden is also related to the identification structure adopted in FSV models. In unrestricted factor models, the likelihood is invariant to rotations, permutations, and sign changes of the latent factors and factor loadings. To obtain an identified specification, lower-triangular restrictions with positive diagonal elements are commonly imposed on the loading matrix. While these constraints remove rotational and sign indeterminacies, they introduce an ordering dependence into the model. Since factors are defined sequentially through the ordering of the observed variables, alternative variable orderings may lead to different factor decompositions and different posterior dependence structures. This issue becomes increasingly relevant in higher-dimensional settings or when common variation is distributed relatively evenly across assets, making individual factors less clearly separated (Gelman et al., 2013; Hosszejni and Kastner, 2021).

The computational difficulties associated with posterior dependence motivate the use of specialized Bayesian estimation techniques. In SV applications, methods such as block updating, simulation smoothing, and reparameterization techniques are therefore essential for obtaining stable posterior simulation. In particular, centered and non-centered parameterizations may exhibit substantially different mixing behavior depending on the persistence of the latent volatility processes. The ASIS helps mitigate this issue by alternating between multiple parameterizations in order to reduce posterior dependence and improve sampling efficiency (Kastner and Frühwirth-Schnatter, 2017).

These computational considerations play an important role in practical estimation of FSV models and partly explain why model complexity must be balanced against numerical stability and posterior simulation efficiency in high-dimensional Bayesian volatility models.

2.4 Prior Specification

Bayesian inference requires the specification of prior distributions for the unknown model parameters. The priors considered in this thesis are weakly informative and are designed to respect essential structural constraints, such as stationarity, positivity, and parameter identifiability, while providing mild shrinkage to improve numerical stability without imposing overly restrictive assumptions on the posterior distribution.

2.4.1 Prior Specification for Univariate SV Models

The prior specifications adopted in this thesis follow the Bayesian SV literature, particularly the framework by Kastner and Frühwirth-Schnatter (2017) and Kastner et al. (2017).

For the baseline SV models, prior distributions are assigned to the parameter vector $\theta = (\beta, \mu, \phi, \sigma_\eta^2)$.

The unconditional mean of observation equation $\beta \in \mathbb{R}$ is assigned a Gaussian prior:

$$\beta \sim \mathcal{N}(b_\beta, B_\beta), \tag{2.22}$$

where b_β and B_β denote the prior mean and variance of the constant mean return parameter.

The unconditional mean of latent log-volatility process $\mu \in \mathbb{R}$ is assigned a Gaussian prior:

$$\mu \sim \mathcal{N}(b_\mu, B_\mu), \tag{2.23}$$

where typically b_μ denotes the prior mean and B_μ controls prior dispersion. In practice, b_μ is assumed to be zero, reflecting the absence of strong prior information about the unconditional level of log-volatility, and B_μ is typically chosen sufficiently large to reflect weak prior information.

To ensure stationarity of the latent AR(1) process, the persistence parameter is restricted to the interval $\phi \in (-1, 1)$ through a reparameterization based on the Beta distribution. Directly specifying a prior on ϕ is inconvenient because standard distributions on the real line do not naturally enforce bounded support. To address this, a transformed variable, $u = \frac{\phi+1}{2}$, which maps $\phi \in (-1, 1)$ to $u \in (0, 1)$. A Beta prior is then assigned as

$$u \sim \text{Beta}(a_\phi, b_\phi), \tag{2.24}$$

which implies $\phi = 2u - 1$. This construction ensures that all prior draws of ϕ automatically satisfy the stationarity constraint. The hyperparameter a_ϕ and b_ϕ control the implied prior distribution persistence, allowing flexible modeling of different degrees of dependence.

The volatility innovation parameter σ_η governs the magnitude of shocks driving the stochastic volatility process, and therefore constrained to be strictly positive, and is assigned a half-normal prior

$$\sigma_\eta \sim \mathcal{N}^+(0, B_\sigma), \tag{2.25}$$

where $\mathcal{N}^+$ denotes the Gaussian distribution with mean zero and variance B_σ, truncated to the positive real line. This prior enforces the positivity constraint on σ_η, while remaining weakly informative when B_σ is chosen sufficiently large. Equivalently, this can be expressed as a prior on the variance parameter σ_η^2, which admits a Gamma representation:

$$\sigma_\eta^2 \sim \text{Gamma}\left(\frac{1}{2}, \frac{1}{2B_\sigma}\right), \tag{2.26}$$

under the shape-rate parameterization.

For models with heavy-tailed errors (SVt and SVtl), the degrees of freedom parameter ν controls the tail thickness of the conditional return distribution, and is restricted to $\nu > 2$ to ensure finite variance. To enforce this constraint while maintaining flexibility, a reparameterization based on exponential prior is adopted:

$$\nu - 2 \sim \text{Exp}(\lambda_\nu), \tag{2.27}$$

which ensures that ν automatically satisfies the positivity constraint above 2. This specification places higher prior mass on moderate values of ν, which correspond to empirically plausible degrees of heavy tails, while avoiding an undue preference for either nearly Gaussian or extremely heavy-tailed behavior.

For models with the leverage effect (SVl and SVtl), the correlation parameter ρ governs the dependence between return disturbances and volatility innovations, and is restricted to the interval $\rho \in (-1, 1)$. To enforce this constraint, a transformed Beta distribution is adopted:

$$\frac{\rho + 1}{2} \sim \mathcal{B}(a_\rho, b_\rho), \tag{2.28}$$

which ensures that all prior draws of ρ lie within the valid correlation range.

Overall, these prior specifications are weakly informative while imposing essential structural constraints required by the model, such as stationarity of autoregressive process and positivity of scale parameters, necessary for a well-defined SV framework. These priors act as a mild form of Bayesian shrinkage by reducing prior mass assigned to inadmissible or extreme parameter values, thereby improving numerical stability during estimation and preventing pathological parameter values configurations. At the same time, the priors remain sufficiently diffuse so that the likelihood largely dominates the posterior, in the sense that changes in the prior within a reasonable range have only a minor impact on posterior summaries.

2.4.2 Prior Specification for FSV Model

For FSV model, prior distributions are specified for the intercept vector β the factor loadings Λ, the latent factors f_t, and the associated stochastic volatility processes.

In particular, the intercept parameters $\beta = (\beta_1, \ldots, \beta_N)'$ are assigned independent Gaussian priors:

$$\beta_i \sim \mathcal{N}(b_i, B_i), \quad i = 1, \ldots, N, \tag{2.29}$$

where b_i denotes the prior mean and and B_i controls the prior variance. In the absence of strong prior information regarding the unconditional return means, the prior variances are chosen sufficiently large so that the priors remain weakly informative.

The unrestricted elements of the factor loadings matrix are assigned independent normal priors. In particular, for unrestricted loadings,

$$\Lambda_{ij} \sim \mathcal{N}(0, \tau_{ij}^2), \tag{2.30}$$

where the variance parameters τ_{ij}^2 control the degree of shrinkage imposed on the factor loadings. Smaller values of τ_{ij}^2 encourage loadings to be concentrated near zero, whereas larger values allow stronger factor effects. In high-dimensional settings, many factor loadings may contribute little to the dependence structure among assets. Shrinkage priors address this issue by pulling negligible loadings toward zero while allowing the most important factor relationships to remain largely unaffected, resulting in a more parsimonious and stable representation of the covariance structure (Carvalho et al., 2008; Bhattacharya and Dunson, 2011).

In addition to the factor loadings, prior specifications are required for the latent factors and idiosyncratic components. These are modeled through conditionally Gaussian stochastic volatility processes:

$$f_t \sim \mathcal{N}(0, \Sigma_{f,t}), \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_{\varepsilon,t}), \tag{2.31}$$

where both covariance matrices $\Sigma_{f,t}$ and $\Sigma_{\varepsilon,t}$ are diagonal and evolve through latent stochastic volatility dynamics.

As equation (2.14) and equation (2.13) presented in section 2.3, the conditional variances of the factors and idiosyncratic components are driven by latent log–volatility processes $h_{f,k,t}$ and $h_{\varepsilon,i,t}$. For each latent factor $k = 1, \ldots, K$ and each series $i = 1, \ldots, N$:

$$h_{f,k,t+1} = \mu_{f,k} + \phi_{f,k}(h_{f,k,t} - \mu_{f,k}) + \sigma_{f,k}\eta_{f,k,t}, \tag{2.32}$$

$$h_{\varepsilon,i,t+1} = \mu_{\varepsilon,i} + \phi_{\varepsilon,i}(h_{\varepsilon,i,t} - \mu_{\varepsilon,i}) + \sigma_{\varepsilon,i}\eta_{\varepsilon,i,t}. \tag{2.33}$$

The associated persistence and volatility parameters are assigned priors analogous to those used in the univariate stochastic volatility specification.

These prior structures improve estimation stability in high-dimensional covariance models by mitigating the influence of weakly identified parameters such as ϕ, σ_η, and Λ, while preserving sufficient flexibility to capture complex dynamic dependence patterns in multivariate financial returns.

2.5 Bayesian Computation and MCMC Estimation

The posterior distributions of the SV models introduced in this chapter are approximated using MCMC methods. Following the Bayesian framework outlined in chapter 1, inference is conducted jointly on the model parameters and the latent volatility states.

The following subsections describe the posterior simulation schemes employed for the different SV specifications considered in this thesis.

2.5.1 Auxiliary Mixture-of-Normals Approximation

For simplicity, consider the observation equation with zero-mean:

$$y_t = \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1), \tag{2.34}$$

so that log-square returns satisfy

$$\log(y_t^2) = h_t + \log(\varepsilon_t^2), \quad \varepsilon_t \sim \mathcal{N}(0,1). \tag{2.35}$$

The transformed error term $\log(\varepsilon_t^2)$ follows a log-chi-square distribution with one degree of freedom, which is highly non-Gaussian and heavy-tailed. Consequently, the resulting observation is no longer conditionally Gaussian, preventing direct application of standard linear Gaussian state-space (see section A.2) simulation methods based on Kalman filtering. To overcome this issue, Kim et al. (1998) approximate the distribution of $\log(\varepsilon_t^2)$ by a finite mixture of normal distributions:

$$\log(\varepsilon_t^2) \approx \sum_{j=1}^{J} \pi_j \mathcal{N}(m_j, s_j^2), \tag{2.36}$$

where $\{\pi_j, m_j, s_j^2\}_{j=1}^J$ are fixed mixture weights, means, and variances, respectively. The finite mixture representation approximates the non-Gaussian log-chi-square distribution by a mixture of a small number of normal distributions, thereby enabling the use of Gaussian state-space simulation methods. Introducing a latent mixture indicators $s_t \in$

$\{1, \ldots, J\}$ with $Pr(s_t = k) = \pi_k$, allows the transformed error term to be represented as

$$\log(\varepsilon_t^2) = m_{s_t} + s_{s_t}\eta_t, \quad \eta_t \sim N(0, 1). \tag{2.37}$$

Substituting this representation into the transformed observation equation yields

$$\log(y_t^2) = h_t + m_{s_t} + s_{s_t}\eta_t, \quad \eta_t \sim \mathcal{N}(0, 1). \tag{2.38}$$

Conditionally on the mixture indicators $s_{1:T}$, the SV model becomes a Gaussian state-space model, allowing efficient simulation of the latent volatility states $h_{1:T}$ using standard Gaussian state-space methods. Bayesian estimation then proceeds by alternating between draws of the mixture indicators $s_{1:T}$ and the latent volatility states $h_{1:T}$ within the MCMC algorithm.

The derivation above follows the original formulation of Kim et al. (1998), which is presented for the zero-mean specification. In the models considered in this thesis, a constant mean parameter β is included in the observation equation. In this case, the transformation is applied to the demeaned observations $y_t - \beta$, yielding

$$\log\left((y_t - \beta)^2\right) = h_t + \log\left(\varepsilon_t^2\right). \tag{2.39}$$

Since β is time-invariant, its inclusion does not affect the latent volatility dynamics or the mixture-of-normals approximation, and the same estimation framework remains applicable.

2.5.2 Sampling Latent States via Kalman Filter-Based Methods

Given the mixture indicators and model parameters, the latent states $h_{1:T}$ are sampled jointly using simulation smoothing methods for linear Gaussian state-space models. The Gaussian structure of the augmented model allows the application of Kalman filter-based methods, such as Forward Filtering Backward Sampling (FFBS), originally described in Frühwirth-Schnatter (1994) and Carter and Kohn (1994).

The algorithm proceeds in two steps. First, Kalman filtering recursions are used to compute the sequence of filtering distributions (forward pass)

$$f(h_t \mid y_{1:t}, \theta), \quad t = 1, \ldots, T \tag{2.40}$$

These distributions summarize the information available up to time t. Second, the latent volatility path is sampled recursively from the smoothing distributions (backward pass)

$$f(h_t \mid h_{t+1}, y_{1:T}, \theta), \tag{2.41}$$

starting from $t = T$ and moving backward to $t = 1$.

Sampling the latent states jointly substantially improves mixing efficiency relative to single-state updating schemes because the full latent trajectory is updated simultaneously while accounting for the strong temporal dependence induced by volatility persistence.

The FFBS draw corresponds to an exact Gibbs update, since it samples from the full conditional posterior:

$$f(h_{1:T} \mid y_{1:T}, \theta, s_{1:T}). \tag{2.42}$$

2.5.3 Ancillarity-Sufficiency Interweaving Strategy (ASIS)

While the auxiliary mixture representation and FFBS provide efficient sampling of the latent states, strong posterior dependence between the volatility states and model parameters may still reduce MCMC efficiency, particularly when the persistence parameter ϕ approaches unity.

To address this issue, the Ancillarity-Sufficiency Interweaving Strategy, proposed by Yu and Meng (2011), alternates between centered and non-centered parameterization of the latent state process.

Under the centered parameterization, the latent volatility process is written as

$$h_t = \mu + \phi(h_{t-1} - \mu) + \sigma\eta_t. \tag{2.43}$$

The corresponding non-centered parameterization is obtained by defining

$$\tilde{h}_t = \frac{h_t - \mu}{\sigma}, \tag{2.44}$$

which yields the transformed state equation

$$\tilde{h}_t = \phi\tilde{h}_{t-1} + \eta_t. \tag{2.45}$$

The two parameterizations describe the same latent volatility process but induce different posterior dependence structures between the latent states and model parameters. ASIS exploits this feature by alternating between the centered and non-centered representations within each MCMC iteration. After updating the model in one parameterization, the latent states are transformed to the alternative representation and the parameter updates are repeated. This interweaving combines the strengths of both parameterizations: the centered form is typically more efficient when σ is small, whereas the non-centered form tends to mix better when ϕ is close to unity (Yu and Meng, 2011; Kastner and Frühwirth-Schnatter, 2017).

In estimation procedure, ASIS is applied after sampling the latent states $h_{1:T}$ via simulation smoothing and before updating the model parameters. This additional

reparameterization step substantially reduces posterior autocorrelation in the Markov chain and improves convergence without altering the target posterior distribution.

2.5.4 Overall MCMC Scheme

A single iteration of the MCMC algorithm proceeds as follows:

1. Sample mixture indicators $s_{1:T}$ from their conditional posterior distribution using Gibbs (see section C.1 for explicit expressions).
2. Sample latent states $h_{1:T}$ jointly via Forward Filtering Backward Smoothing (FFBS), a Kalman Filter-based simulation algorithm.
3. Update parameters:
   - update β via Gibbs,
   - update μ via Gibbs,
   - update (ϕ, σ_η) via Metropolis-Hastings in the centered parameterization.
4. Apply ASIS by transforming to the non-centered parameterizations and repeating parameter updates.

For the multivariate FSV specification, the same computational principles are employed within a block MCMC framework. Mixture indicators and latent volatility states are sampled for both factor and idiosyncratic volatility processes, while the factor loading matrix Λ, latent factors $f_{1:T}$, intercept vector β, factor volatilities $h^f_{1:T}$, idiosyncratic volatilities $h^e_{1:T}$, and their associated persistence and innovation variance parameters are updated from their conditional posterior distributions.

Technical derivations for all parameter blocks, including the full conditional distributions and the corresponding MCMC updating equations schemes, are provided in Appendix C.

Together, these data augmentation, state sampling, and reparameterization steps provide an efficient computational framework for Bayesian inference in SV models.

2.6 Forecasting Procedure

To compare alternative model specifications, predictive performance is evaluated using an expanding-window forecasting scheme. The sample is split into an initial estimation period and an evaluation period.

At each forecast origin t within the evaluation period, the model is estimated using observations up to time t, and a one-step-ahead predictive density for y_{t+1} is computed,

denoted by $f(y_{t+1} \mid \mathcal{F}_t)$. This procedure is repeated sequentially as new observations become available. Here, $f(y_{t+1} \mid y_{1:t})$ denotes the posterior predictive density of the next observation conditional on the information available at time t.

As new observations become available, the estimation sample is expanded sequentially and the predictive density is recomputed. The realized observation y_{t+1} is then evaluated under the corresponding predictive distribution. Models that assign higher probability density to subsequently realized observations are regarded as providing better predictive descriptions of the data.

2.7 Multivariate Model Selection

For the FSV model, several model specifications corresponding to different choices of the number of latent factors K are considered. In particular, a small set of candidate values for K is examined, and their performance is evaluated in order to identify the most suitable factor structure.

For each specification, the model is estimated sequentially using the expanding-window scheme described above, and predictive performance is assessed through the cumulative log predictive score.

The purpose of this exercise is not to obtain forecasts itself, but rather to compare competing factor structures. The specification that assigns the highest predictive density to the realized observations over the evaluation period is selected as the preferred model. Among the candidate specifications considered in this thesis, the model with the largest cumulative log predictive score is taken to provide the most adequate factor representation.

Chapter 3

Application to Financial Data

This chapter presents an illustrative application of the SV models introduced in the previous chapters. The analysis focuses on latent volatility estimation, predictive behavior, and the computational properties of Bayesian SV models in practical applications.

All empirical analyses were conducted in R (R Core Team, 2025). The univariate stochastic volatility specifications were estimated using the `stochvol` package (Kastner, 2016), while the multivariate factor stochastic volatility models were estimated using the `factorstochvol` package (Hosszejni and Kastner, 2021). Data preparation, visualization and forecasting exercises were implemented using additional R packages where appropriate.

3.1 Description of The Dataset

3.1.1 Data Description

The dataset consists of daily financial return series for a collection of iShares STOXX Europe sector ETFs together with the iShares STOXX Europe 600 UCITS ETF, which serves as a proxy for the broad European market index. The sectoral ETFs represent major segments of the European economy, including financials, healthcare, industrials, energy, materials, consumer discretionary, consumer staples, technology, utilities, real estate and telecommunications. All price series were obtained from Yahoo Finance. The sample covers the period from January 2015 to December 2024. The inclusion of multiple sectors provides a suitable environment for investigating both individual volatility dynamics and common dependence structures across assets.

Let $P_{i,t}$ denote the adjusted closing price of asset i at time t. The observed return series are constructed using continuously compounded log-returns,

$$r_{i,t} = 100(\log P_{i,t} - \log P_{i,t-1}), \tag{3.1}$$

where returns are expressed in percentage form for numerical convenience. The log-return transformation removes long-term trends in price levels and produces series that are more appropriate for SV modeling.

Several standard pre-processing procedures are applied before estimation. Since the ETFs are traded on different exchanges and may contain missing observations due to holidays or non-trading days, the price series are first merged by trading date and incomplete observations are removed to obtain a balanced multivariate dataset with synchronized return series across all assets. The final dataset consists of $m = 12$ return series observed over $T = 2,541$ trading days, resulting in a balanced multivariate return matrix of dimension $T \times m$. No additional smoothing or filtering is performed, since the objective of SV models is precisely to capture the time-varying variance structure directly from the observed data.

3.1.2 Descriptive Statistics

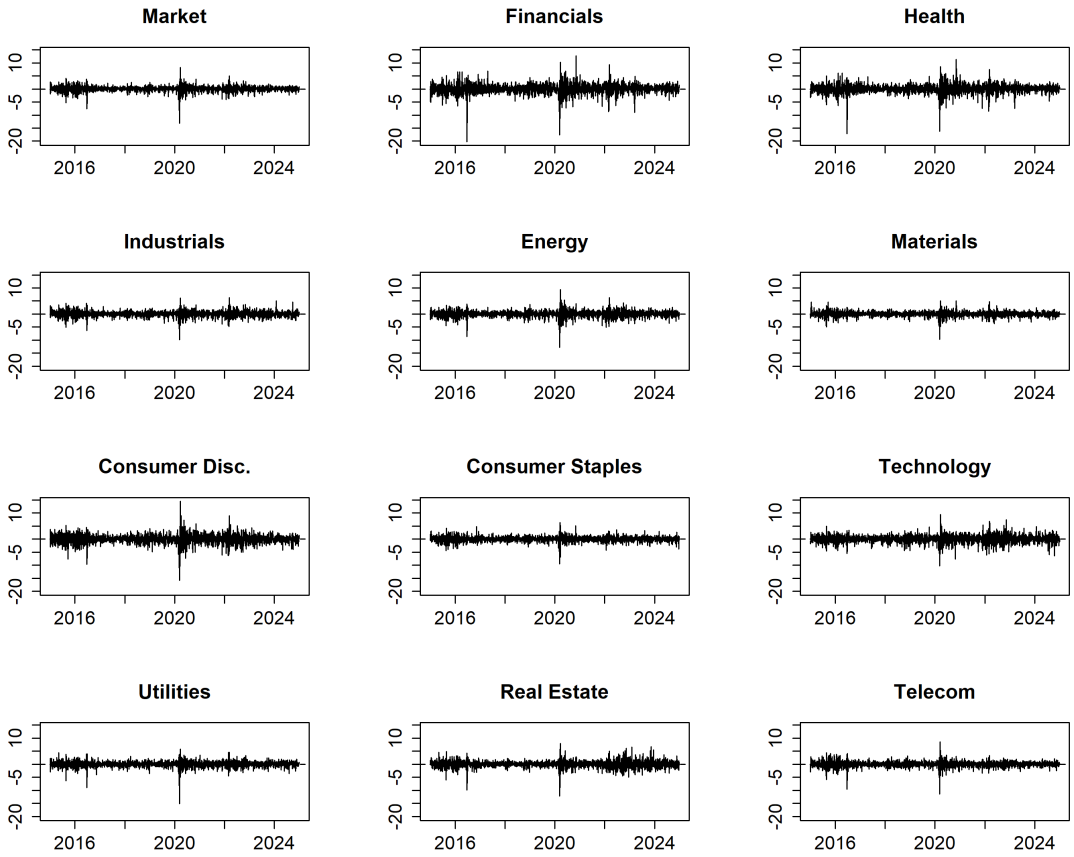


Figure 3.1: Daily return series by sector

Figure 3.1 illustrates the key stylized facts commonly documented in financial return series. First, the sector-level returns fluctuate around zero without exhibiting persistent

deterministic trends. Second, periods of relatively calm market behavior alternate with episodes of elevated volatility, indicating the presence of volatility clustering. Third, large return shocks frequently occur simultaneously across multiple sectors, suggesting the existence of common latent factors affecting market-wide uncertainty, which is fully consistent with the theoretical motivation for multivariate specifications of SV models.

To provide a preliminary description of the dataset, table 3.1 reports summary statistics for each return series, including the sample mean, standard deviation, minimum, median, maximum, skewness, and kurtosis. The descriptive statistics highlight several pronounced departures from Gaussianity. In particular, all sectors exhibit substantial excess kurtosis, indicating heavy-tailed return distributions with a higher frequency of extreme movements than implied by a normal one. Notably, the market index displays both one of the highest kurtosis values – reflecting the amplification of tail events through cross-sector co-movement during periods of market stress – and markedly stronger negative skewness than most individual sectors, indicating that large downside shocks occur more frequently and with greater magnitude at the aggregate level (see section B.4). This combination of heavy tails and asymmetric return behavior provides further motivation for employing SV models with heavy-tailed innovations and leverage effects, which are specifically designed to capture such empirical features.

	Mean	SD	Min	Median	Max	Skewness	Kurtosis
Market	0.03	1.06	-13.08	0.08	8.19	-1.20	17.84
Financials	0.02	1.85	-20.19	0.07	12.71	-1.01	15.76
Health	0.02	1.62	-17.14	0.09	11.30	-1.07	15.82
Industrials	0.03	1.10	-9.89	0.06	6.36	-0.41	8.31
Energy	0.04	1.24	-12.81	0.09	9.34	-0.75	12.01
Materials	0.01	0.94	-9.75	0.02	5.12	-0.64	10.92
Consumer Disc.	0.02	1.65	-15.77	0.06	14.54	-0.45	11.87
Consumer Staples	0.02	1.02	-9.45	0.04	6.38	-0.44	9.01
Technology	0.04	1.45	-10.29	0.11	9.41	-0.34	6.69
Utilities	0.02	1.11	-15.18	0.05	5.81	-1.51	21.06
Realestate	0.00	1.32	-12.23	0.01	7.90	-0.46	10.26
Telecom	0.00	1.08	-11.39	0.03	8.56	-0.74	14.38

Table 3.1: Descriptive statistics of return series

While the univariate characteristics reported in figure 3.1 and table 3.1 provide useful information about the distributional properties of individual return series, they do not capture the dependence structure across sectors. To investigate the extent of cross-sectional comovement, figure 3.2 reveals substantial positive dependence among all sector returns. Correlations are generally high, ranging from approximately 0.39 to 0.97 ,

with particularly strong associations observed between the market index and most sector indices. The widespread presence of strong correlations suggests that a considerable proportion of return variation is driven by common underlying factors. This evidence provides direct motivation for the use of multivariate FSV models.

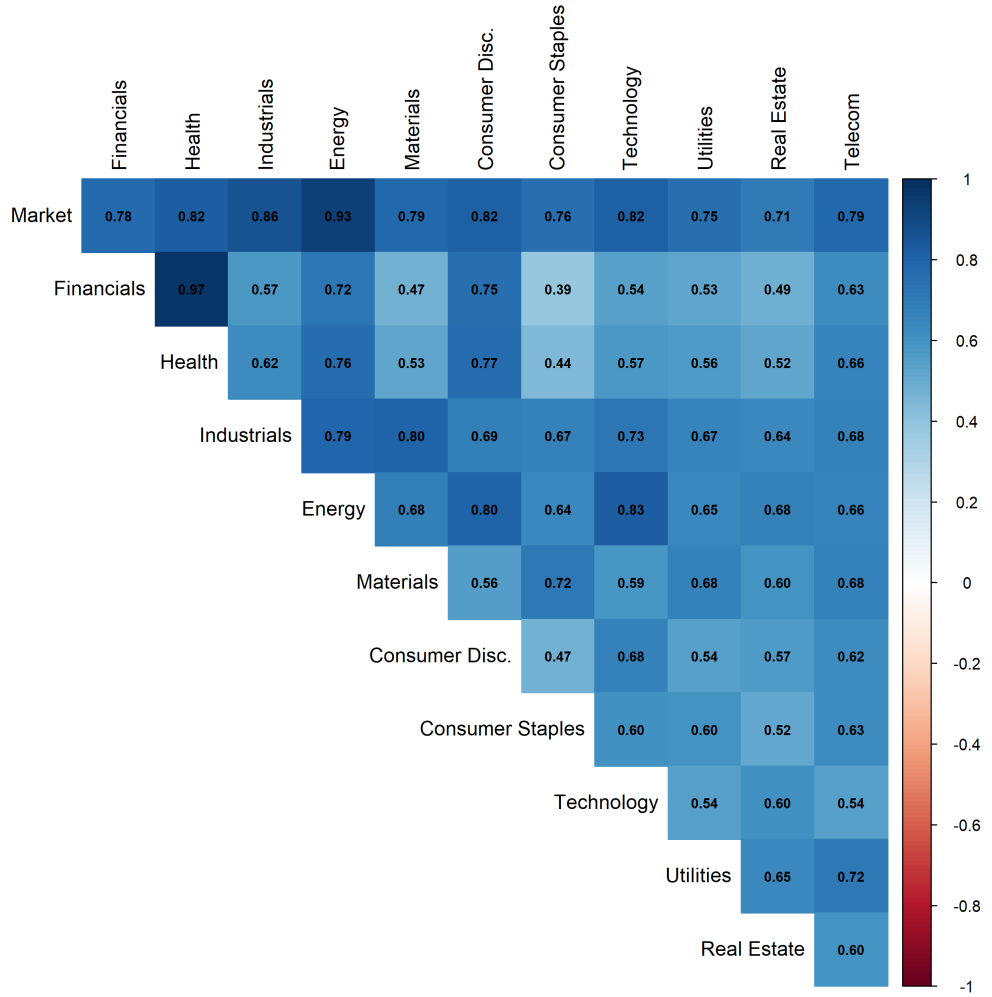


Figure 3.2: Cross-sectional correlation matrix of the market and sectoral return series

3.2 Univariate SV Analysis

3.2.1 Estimation Setup

The SV models are estimated within a fully Bayesian framework using MCMC methods. For each return series, four univariate specifications are considered: a Gaussian SV model (baseline SV), a heavy-tailed variant with Student-t disturbances (SVt), a leverage specification allowing for asymmetric volatility responses (SVl), and a combined heavy-tailed-leverage model (SVtl). The prior distribution follows the standard specifications introduced in chapter 2 and implemented in the `stochvol` package

Parameter	Prior	Hyperparameter Values
β	$\mathcal{N}(b_\beta, B_\beta)$	$b_\mu = 0,\ B_\mu = 10000^2$
μ	$\mathcal{N}(b_\mu, B_\mu)$	$b_\mu = 0,\ B_\mu = 100^2$
$\frac{\phi+1}{2}$	$Beta(a_\phi, b_\phi)$	$a_\phi = 5,\ b_\phi = 1.5$
σ_η^2	Gamma $\left(\frac{1}{2}, \frac{1}{2B_\sigma}\right)$	$B_\sigma = 1$
$\nu - 2$	Exp(λ_ν)	$\lambda_\nu = 0.1$
$\frac{\rho+1}{2}$	$Beta(a_\rho, b_\rho)$	$a_\rho = 4,\ b_\rho = 4$

Table 3.2: Hyperparameter specification for univariate SV models

(Kastner, 2016). The corresponding hyperparameter values employed in the analysis are reported in table 3.2. The selected hyperparameter values reflect commonly adopted choices in the SV literature and correspond to the default prior specifications implemented in the `stochvol` package (Kastner, 2016). For example, the large prior variance $B_\mu = 100^2$ reflects substantial prior uncertainty regarding the unconditional volatility level, while the specification for the persistence parameter $a_\phi = 5$, and $b_\phi = 1.5$ places most prior mass on highly persistent volatility dynamics, a common feature of financial return series. The remaining hyperparameters are selected to provide sufficiently flexible prior support while allowing the data to play the dominant role in posterior inference.

Posterior inference is based on simulation from the joint posterior distribution of the static parameters and latent volatility states. The MCMC algorithm iteratively alternates between updating the static parameters and sampling the latent volatility trajectory conditional on the observed data.

For each model specification, the Markov chain is run for a total of $M = 30{,}000$ iterations. An initial portion of $B = 10{,}000$ iterations is discarded as burn-in, leaving 20,000 post-burn draws available for analysis. To reduce short-range autocorrelation, thinning is applied by retaining every tenth draw. The resulting posterior samples are then used for latent volatility estimation, posterior predictive inference, and forecasting evaluation. Mixing behavior and convergence are assessed using trace plots, autocorrelation functions (ACF), effective sample size (ESS) and Geweke diagnostics for the main static parameters.

3.2.2 Posterior Estimation Results

This section illustrates the behavior of the univariate SV specifications introduced in chapter 2 using daily sectoral equity return data. To avoid repetitive discussion across

all return series, the analysis focuses primarily on the market index as a representative example. The remaining sectoral indices exhibit qualitatively similar patterns, and the corresponding posterior summaries are reported in section D.1.

Although the primary focus of this study is on volatility dynamics, the intercept parameter β is estimated jointly with the remaining model parameters and is therefore included in the posterior summaries reported in table 3.3. Since the estimated intercepts are generally small, the subsequent discussion focuses primarily on the volatility parameters.

	SV	SVtl	SVt	SVl
β	0.085 [0.062, 0.108]	0.049 [0.027, 0.072]	0.086 [0.064, 0.109]	0.039 [0.018, 0.062]
μ	-0.464 [-0.673, -0.260]	-0.399 [-0.545, -0.243]	-0.438 [-0.647, -0.224]	-0.417 [-0.559, -0.261]
ϕ	0.947 [0.926, 0.962]	0.942 [0.923, 0.957]	0.954 [0.936, 0.969]	0.940 [0.923, 0.954]
σ_η	0.320 [0.271, 0.374]	0.302 [0.262, 0.352]	0.287 [0.240, 0.340]	0.315 [0.274, 0.359]
ν		20.302 [12.168, 35.033]	20.007 [12.412, 37.016]	
ρ		-0.533 [-0.609, -0.448]		-0.606 [-0.686, -0.522]

Table 3.3: Posterior medians and 90% credible intervals of the model parameters of market series.

The posterior estimates reported in table 3.3 indicate that all four specifications imply highly persistent volatility dynamics. The persistence parameter ϕ governing the latent log-volatility process is coherently estimated close to one across all specifications, with posterior medians between 0.94 and 0.96 within the market series, suggesting that volatility shocks dissipate only gradually over time. Such behavior is typical of financial return series and confirms the suitability of SV models for capturing volatility clustering.

The posterior estimates of the latent-volatility innovation scale parameter σ_η remain moderate across all specifications, suggesting that the latent log-volatility evolves gradually rather than through abrupt shifts. Compared to the Gaussian SV specification, the heavy-tailed incorporated specifications SVt and SVtl generally produce slightly smaller posterior estimates of σ_η, reflecting the ability of Student-t innovations to absorb extreme observations directly through the heavy-tailed distribution,thereby reducing the need for large shocks to the latent volatility dynamics.

Further evidence against Gaussian return innovations is provided by the posterior estimates of the degrees-of-freedom parameter ν in the SVt and SVtl specifications. The posterior medians of around 20 imply moderately heavy tails in the return disturbances, while the relatively wide credible intervals for ν indicate some remaining uncertainty regarding the precise degree of tail heaviness. In general, the results suggest that allowing for heavy tails provides additional flexibility in capturing extreme market movements.

For the leverage incorporated specifications SVl and SVtl, the posterior estimates of the asymmetry parameter ρ are substantially negative with the corresponding 90% credible intervals lying entirely below zero, indicating a pronounced leverage effect in the market series. This result is consistent with asymmetric volatility responses in which negative return shocks are associated with larger subsequent increases in volatility than positive shocks of comparable magnitude.

3.2.3 Latent Volatility Estimation

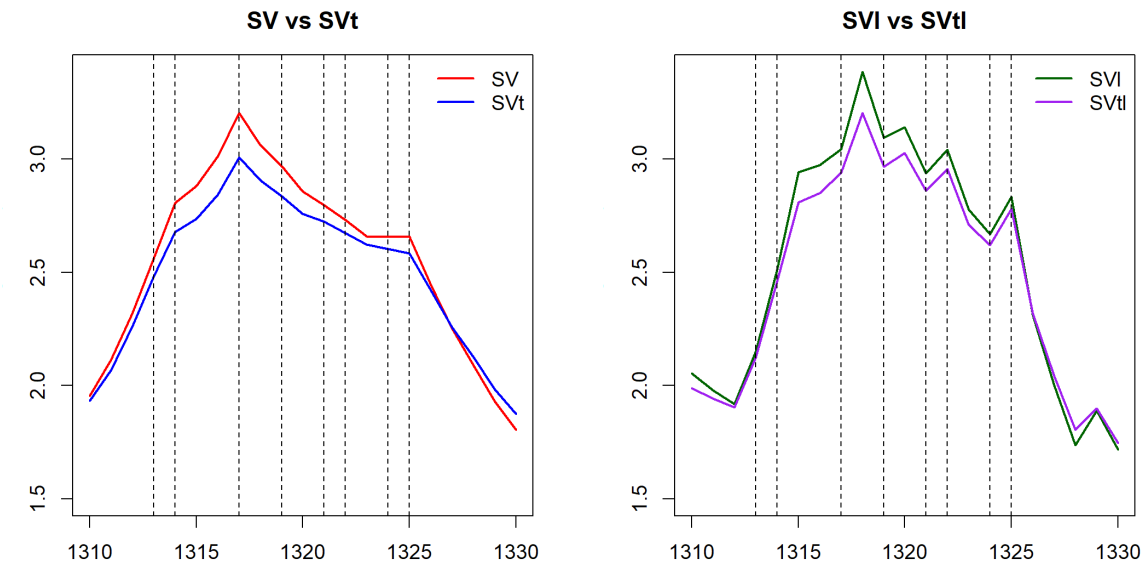


Figure 3.3: Latent volatility for market series around extreme return events
The vertical dashed lines indicate returns exceeding the 99th percentile of the absolute return distribution.

The implications of the parameters related to volatility estimated above can be further examined through the estimated latent volatility trajectories. Figure 3.3 illustrates the posterior mean latent volatility paths for the market series together with the timing of extreme return observations. Although these extreme observations occur frequently throughout the sample period, the estimated volatility trajectories h_t do not exhibit a corresponding sharp spike for every extreme return realization. Instead, the models primarily capture the persistent nature of volatility and its gradual evolution during periods of market stress. While the overall volatility patterns remain qualitatively

similar across specifications, noticeable differences emerge in the local magnitude and persistence of the estimated volatility levels. In particular, the estimated volatility paths under the SVt and SVtl models appear somewhat more robust around several extreme return episodes than those produced by the Gaussian SV specification. This difference arises because, under the Student-t distribution, exceptionally large positive or negative returns are less unusual than under the Gaussian distribution. Consequently, observations that may be regarded as outliers under the Gaussian specification can be accommodated directly through the heavy-tailed return distribution, rather than being attributed entirely to sudden increases in latent volatility.

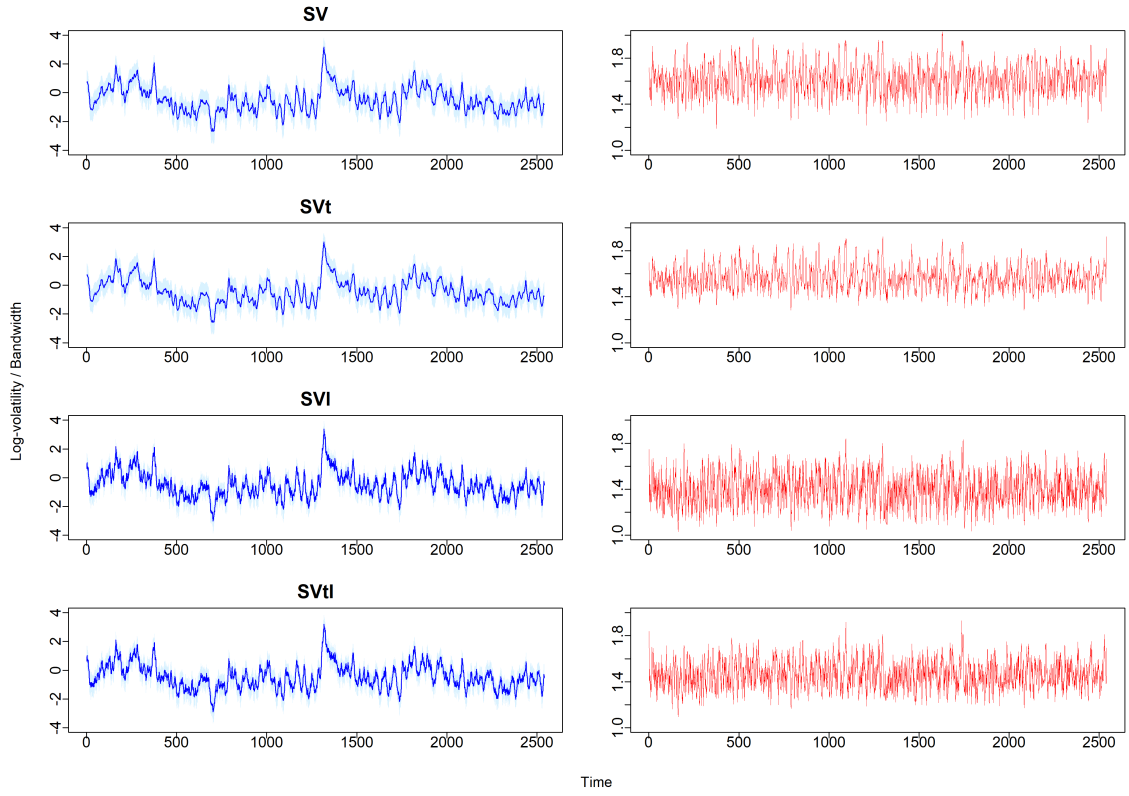


Figure 3.4: Posterior median latent volatility estimates and 90% credible intervals for the market return series under univariate specifications

The blue curves represent the posterior median latent volatility, while the shaded areas denote the corresponding 90% *posterior credible intervals. The red curves represent the bandwidth of the* 90% *credible intervals, computed as the difference between the* 95% *and* 5% *posterior quantiles at each time point.*

To further illustrate posterior uncertainty in the latent volatility process, figure 3.4 presents the posterior median latent volatility together with the 90% credible intervals for the four univariate specifications. While the estimated volatility paths exhibit several episodes of elevated volatility, the widths of credible intervals remain relatively stable over time, indicating that uncertainty surrounding the latent volatility states is

reasonably controlled and does not vary substantially over time. Moreover, the estimated volatility trajectories are qualitatively similar across specifications, indicating that the main features of the estimated latent volatility process h_t are robust to the inclusion of heavy-tailed innovations and leverage effects.

3.2.4 Model Diagnostics

Although the estimated latent volatility paths appear reasonable, further diagnostic checks are required to assess whether the fitted models provide an adequate representation of the volatility dynamics. To this end, the standardized residuals

$$\hat{z}_t = \frac{y_t - \hat{\beta}}{\exp(\hat{h}_t/2)} \tag{3.2}$$

are examined, where $\hat{h}_t$ denotes the posterior mean of the latent log-volatility at time t. Under a correctly specified model, standardized residuals should approximate a white noise process with zero-mean, unit variance and serial dependence. Figure 3.5 presents the residuals diagnostics for the market series under the four univariate SV specifications. All four models produce standardized residuals fluctuating around zero, and the ACF of the squared residuals is essentially zero at all lags, indicating that the main volatility clustering has been captured. Therefore, the residual autocorrelation diagnostics do not provide strong evidence in favor of one specification over another.

More informative differences emerge from the QQ plots. The Gaussian SV model exhibits noticeable departures from the reference line in the tails, whereas the heavy-tailed specifications provide a closer fit to the empirical residual distribution. Among the four models, the SVtl specification shows the closest agreement with the theoretical quantiles, particularly in the tails, suggesting that the combination of heavy-tailed returns and leverage effects provides the most adequate description of the market return series. Consequently, SVtl is the preferred univariate specifications for the market series.

Taken together, the posterior parameter estimates, latent volatility trajectories, and residual diagnostics suggest that all four univariate SV specifications provide a adequate description of the representative return dynamics. However, the heavy-tailed models generally offer a better fit to the empirical return distribution, with the SVtl specification providing the most satisfactory overall performance among the four models.

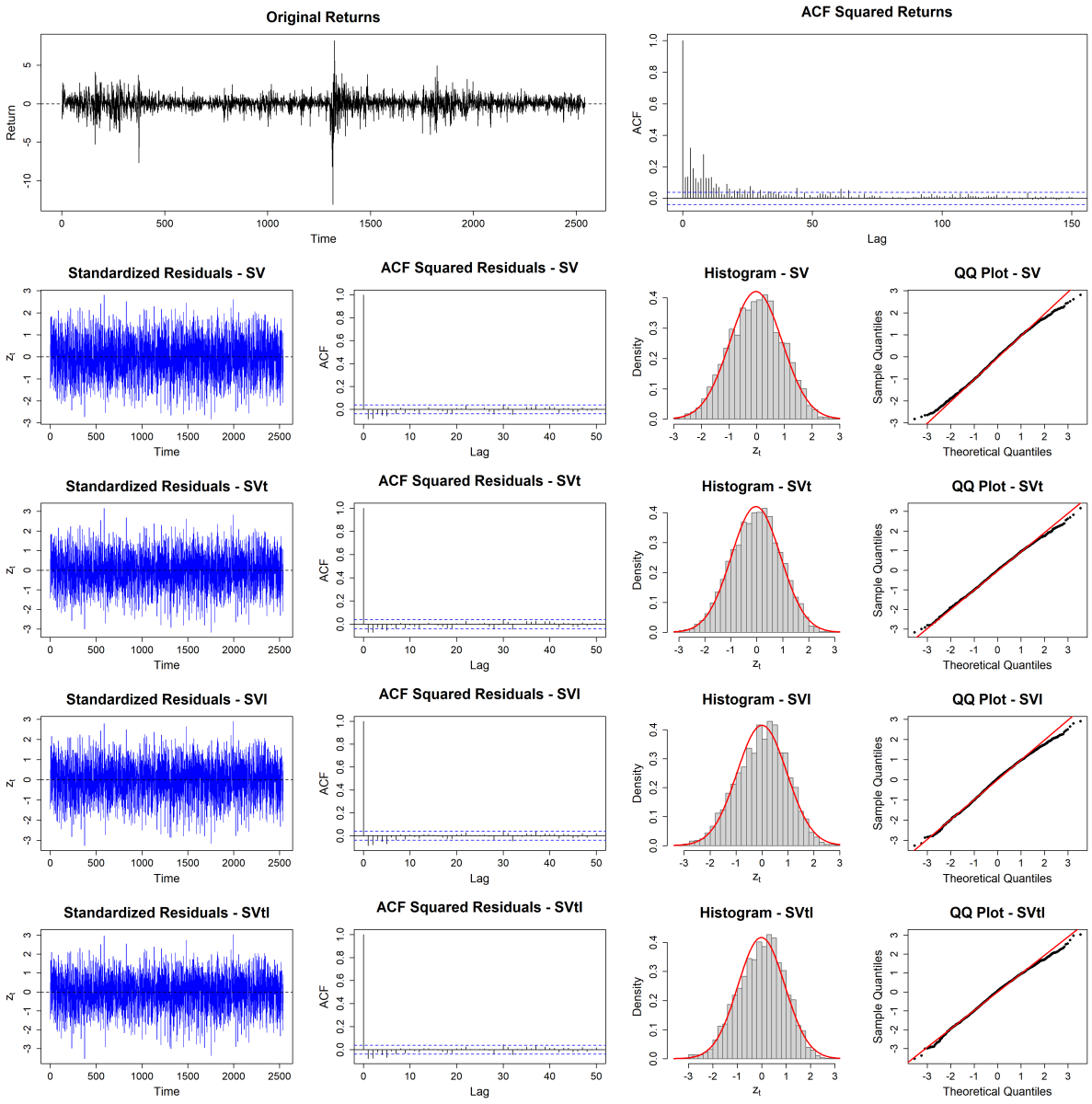


Figure 3.5: Residual diagnostics for the market return series under the univariate SV specifications

3.2.5 MCMC Diagnostics

While the residual diagnostics and QQ plots provide evidence regarding the adequacy of the competing SV specifications, these results are meaningful only if the underlying MCMC estimation procedure has converged and explored the posterior distribution sufficiently well. Therefore, it is necessary to assess the convergence and mixing behavior of the Markov chains. The following diagnostics examine whether the sampled chains exhibit stable behavior and provide reliable inference for the model parameters.

Trace Plots and Mixing Behavior

Trace plots are examined first, as they provide an immediate visual assessment of the mixing behavior of the Markov chains and help identify potential issues such as poor exploration of the posterior distribution or prolonged periods of sticking. Representative MCMC trace plots for the market series are presented in figure 3.6. Across all univariate specifications, the sampled parameter trajectories fluctuate around stable posterior regions throughout the retained iterations, with no visible trends, structural breaks, or evidence that the chains become trapped in particular regions of the parameter space following the burn-in phase. The degrees-of-freedom parameter ν in the SVt and SVtl models exhibits more persistent movements than the remaining parameters, but the chains continue to explore a broad range of posterior values.

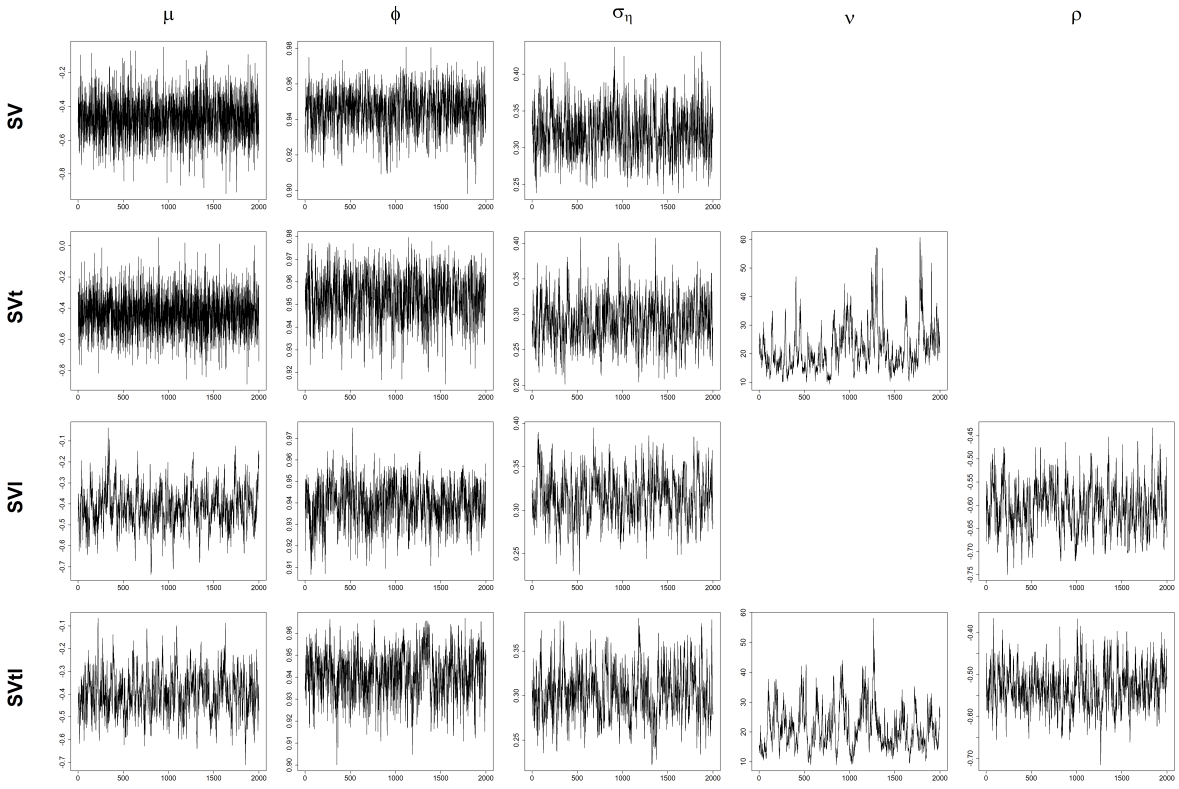


Figure 3.6: Trace plot for the market return series under the univariate SV specifications

To complement the visual inspection of the trace plots, the ACFs of the MCMC draws in figure 3.7 are examined to assess chain dependence and mixing efficiency. For the Gaussian and heavy-tailed SV models, the ACFs of the level parameter μ is essentially zero beyond lag 0, indicating very weak serial dependence and good mixing of the corresponding chains. In contrast, the leverage specifications (SVl and SVtl) exhibit somewhat slower decay in the ACF of μ, reflecting the additional complexity introduced by the leverage structure. Nevertheless, the magnitude of the autocorrelations is not indicative of problematic MCMC behavior. Similar patterns are observed for

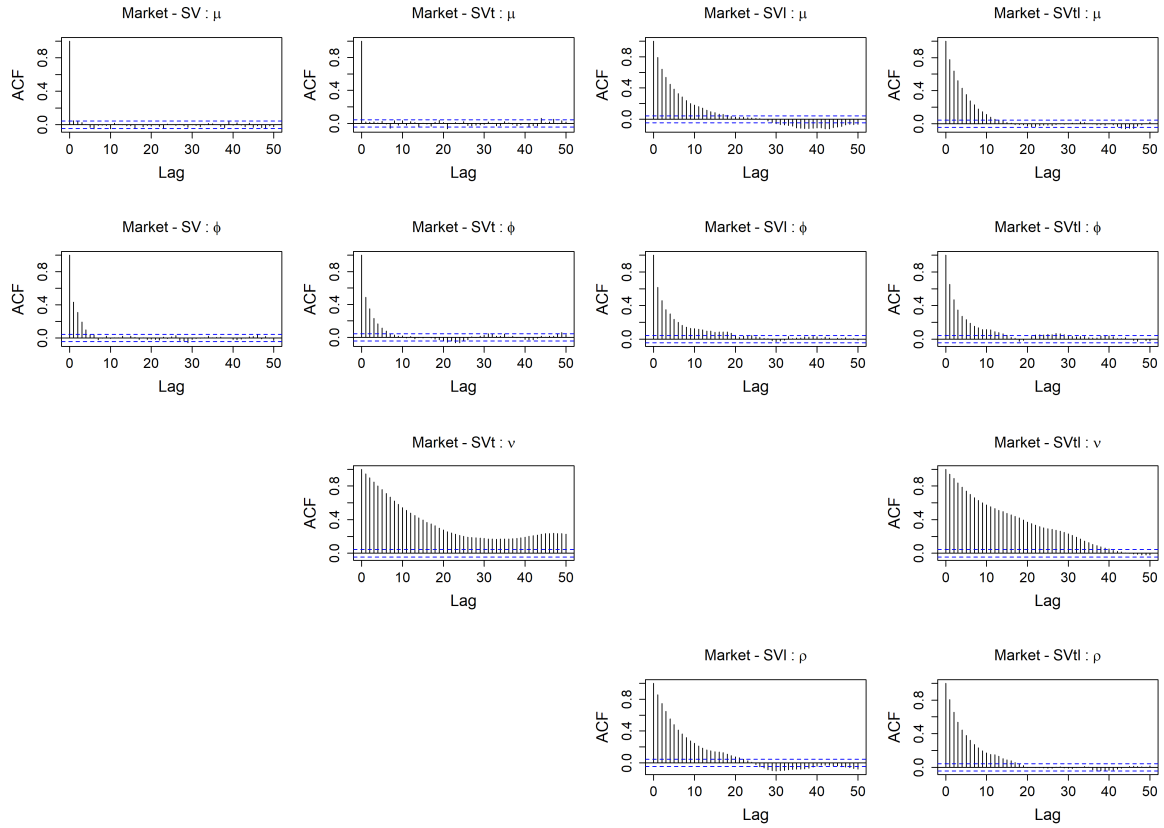


Figure 3.7: ACF of selected MCMC chains for the market univariate SV specifications

the persistence parameter ϕ, whose ACFs decay more gradually across all specifications, particularly in the leverage models. The degrees-of-freedom parameter ν in the SVt and SVtl models and the leverage parameter ρ in the SVl and SVtl models also display noticeable serial dependence, which is common for these parameters and reflects their more challenging posterior geometry. Overall, the autocorrelation diagnostics are consistent with the trace plots and suggest satisfactory mixing behavior for all four univariate specifications.

Convergence Diagnostics

Convergence diagnostics were additionally assessed using Geweke statistics for static parameters to assess whether the retained draws have converged to the target distribution.

The Geweke statistics for the market series across all four univariate SV specifications, reported in table 3.4, fall within the $[-1.96, 1.96]$ acceptance bounds, providing no evidence against convergence at the 5% significance level.

The effective sample sizes (ESS) reported in table 3.5 further characterize sampling efficiency across specifications. The Gaussian SV and SVt models yield comparatively high ESS values for μ (1691 and 2000, respectively), indicating efficient exploration of the

Parameter	SV	SVt	SVl	SVtl
μ	0.18	0.67	-1.44	-0.93
ϕ	-0.73	-0.35	-1.76	-0.02
σ_η	0.20	0.02	0.39	1.12
ν		-1.96		-0.06
ρ			1.10	0.30

Table 3.4: Geweke convergence diagnostics for the market univariate SV specifications

posterior for the level parameter. In contrast, the leverage specifications SVl and SVtl show substantially lower ESS across all parameters, with μ dropping to approximately 197-210 and ϕ to around 291-296, consistent with the slower mixing observed in the ACF plots for these models. The degrees-of-freedom parameter ν and leverage parameter ρ return the lowest ESS values across their respective specifications (as low as 60 for ν in SVtl), reflecting the additional sampling difficulty introduced by heavy tails and asymmetric volatility responses.

Parameter	SV	SVt	SVl	SVtl
μ	1691.0	2000.0	197.3	209.7
ϕ	587.5	512.6	295.6	291.8
σ_η	409.7	290.0	178.6	199.6
ν		69.7		60.1
ρ			140.4	214.0

Table 3.5: Effective sample sizes (ESS) for the market univariate SV specifications

Overall, the Geweke and ESS diagnostics jointly support the conclusion that the retained draws are reliable, though the degrees-of-freedom parameter ν in the heavy-tailed specifications warrants more cautious interpretation given its substantially lower sampling efficiency.

3.2.6 Predictive Performance and Practical Forecasting Application

Predictive Performance

Predictive performance is evaluated using the expanding-window forecasting procedure described in section 1.3. The final 100 observations in the dataset are reserved

for the one-step-ahead forecasting evaluation period. At each forecast origin, the model is re-estimated using all observations available up to time t, and a predictive density for the subsequent observation y_{t+1} is constructed, and predictive performance is assessed using cumulative log predictive likelihood scores in equation (1.32), computed over the evaluation sample.

The resulting scores for the four univariate SV specifications are reported in table 3.6. Since higher values of scores indicate better predictive performance, the preferred model for each series is the specification attaining the largest cumulative score.

	SV	SVt	SVl	SVtl	Winning Model
Market	-105.49	-105.64	-105.43	-105.33	SVtl
Financials	-153.35	-152.28	-154.23	-153.13	SVt
Health	-142.15	-141.49	-142.52	-141.94	SVt
Industrials	-141.23	-141.12	-141.90	-141.18	SVt
Energy	-133.87	-133.99	-134.35	-133.99	SV
Materials	-108.83	-109.18	-108.19	-108.37	SVl
Consumer Disc.	-165.08	-165.19	-166.04	-165.42	SV
Consumer Staples	-131.81	-130.69	-130.35	-129.27	SVtl
Technology	-176.63	-175.12	-177.88	-174.98	SVtl
Utilities	-114.83	-113.83	-114.79	-113.69	SVtl
Realestate	-150.81	-150.93	-149.97	-150.60	SVl
Telecom	-106.40	-105.33	-105.73	-104.93	SVtl

Table 3.6: Cumulative log predictive scores across univariate SV model specifications

The results indicate that no single specification uniformly dominates across all sectoral return series. Nevertheless, the extended SV models generally outperforms the Gaussian benchmark, suggesting that allowing for heavy-tailed return innovations and leverage effects improves the ability of the predictive distribution to accommodate future observations. Among the competing specifications, the SVtl model is selected most frequently, providing the highest predictive scores for five of the twelve series. The heavy-tailed SVt model also performs strongly, particularly for the financials, health, and industrials sectors.

At the same time, score differences across specifications are generally moderate rather than dramatic. Several models achieve broadly similar predictive performance, suggesting that the dominant volatility dynamics of the return series are already captured reasonably well by relatively parsimonious SV specifications. In particular, even the Gaussian SV model is capable of generating excess kurtosis through its stochastic

volatility mechanism (see section B.3 and section B.3.2), producing heavier-tailed unconditional return distributions despite Gaussian innovations. This helps explain why the Gaussian specification remains competitive despite the absence of explicitly heavy-tailed innovations.

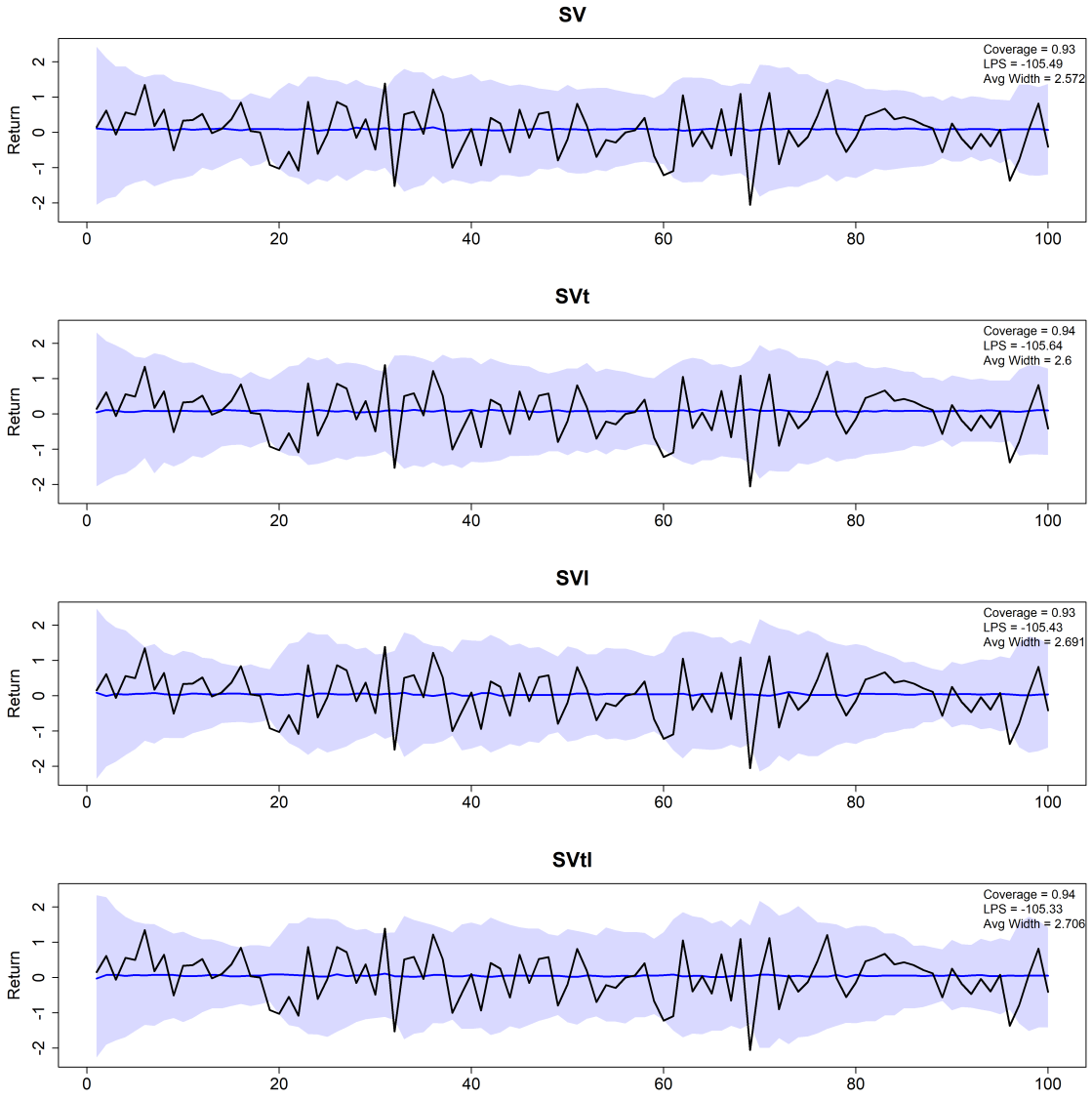


Figure 3.8: One-step-ahead predictive return distributions for the market return series under the four univariate specifications
The black line represents the realized future returns, while the blue line and shaded region denote the predictive median and 90% *credible interval*

Nevertheless, the superior performance of the SVt and SVtl models indicates that directly modeling heavy-tailed return disturbances provides additional flexibility in the tails of the predictive distribution and greater robustness to extreme return realizations. On the other hand, leverage effects provide additional gains in some series, but these improvements are generally smaller than those achieved by incorporating heavy-tailed

innovations. Overall, the results suggest that accounting for heavy tails is the primary source of predictive improvement, whereas leverage effects play a secondary but still beneficial role.

To complement the cumulative log predictive scores reported in table 3.6, figure 3.8 presents the one-step-ahead posterior predictive distributions for the market return series under the four competing specifications. The figure illustrates several features that are common across all specifications. First, the predictive intervals exhibit some variation over time, reflecting changes in the uncertainty implied by the latent volatility process. However, differences in interval width are relatively moderate, indicating that the level of predictive uncertainty remains fairly stable throughout most of the evaluation period. Second, the predictive medians remain close to zero throughout the evaluation period, which is consistent with the well-known difficulty of predicting the direction of financial returns. Finally, differences across specifications are relatively modest, as the predictive intervals largely overlap and achieve similar coverage rates. This observation is consistent with the relatively small differences in cumulative log predictive scores reported in table 3.6.

Taken together, the results of model comparison between univariate SV specifications suggest that increasing distributional flexibility through heavy-tailed innovations yields more consistent predictive gains than introducing leverage effects, while the overall similarity of predictive scores across specifications indicates that the core stochastic volatility mechanism already captures a substantial portion of the dynamics present in the return series.

Practical Forecasting Application

While the preceding analysis focuses on one-step-ahead predictive performance, multi-step forecasts provide a useful illustration of how predictive uncertainty evolves over increasing forecast horizons. The multi-step forecasts in figure 3.9. In contrast to the one-step-ahead predictive shown in figure 3.8, the multi-step forecasts in figure 3.9 are generated from a fixed information set at time T. As the forecast horizon increases, uncertainty accumulates through repeated propagation of the latent volatility process, leading to a gradual widening of the predictive intervals. However, this increase is relatively moderate, and the interval width becomes increasingly stable latter. This behavior reflects the mean-reverting nature of the latent log-volatility process, which causes the influence of the current volatility state to diminish as the forecast horizon increases.

Another notable difference is that the one-step-ahead forecasts display substantial variation across calendar time, whereas the multi-step forecasts evolve smoothly across horizons. The former primarily reflects changes in market conditions during the

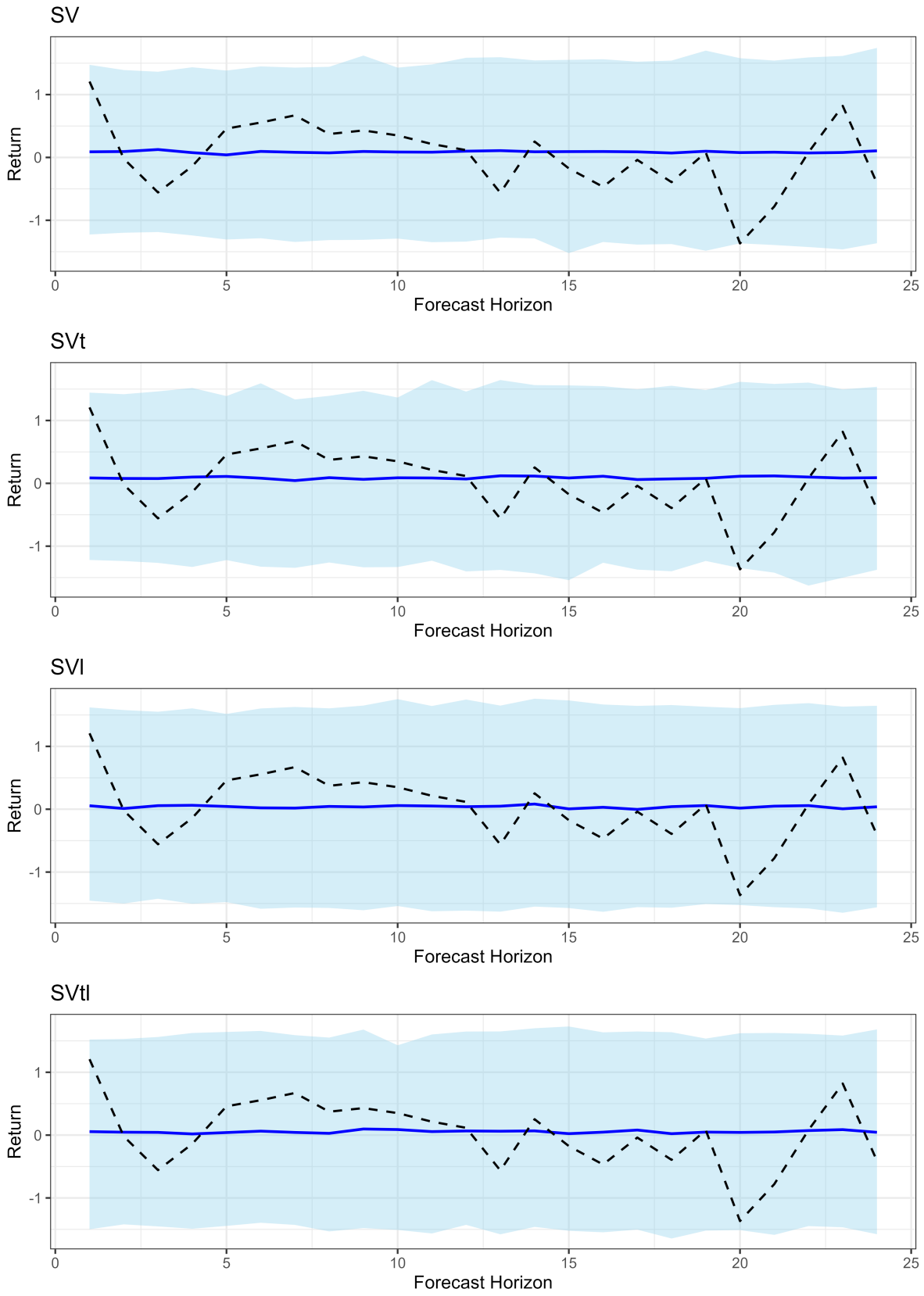


Figure 3.9: Forecast of 24-step-ahead posterior predictive distributions for the market return series under the four univariate SV specifications

The blue line represents the posterior predictive median, the shaded region corresponds to the 90% *credible interval obtained from the posterior predictive distribution, and the dashed black line is the realized future returns.*

evaluation period, while the latter illustrates how predictive uncertainty develops when forecasting progressively further into the future from a single forecast origin.

Compared with the one-step-ahead forecasts discussed previously, the multi-step predictive distributions exhibit a more stable evolution across forecast horizons. In the one-step-ahead setting, changes in the predictive intervals are driven by updates in the information set and the estimated latent volatility state as new observations become available. In contrast, the multi-step forecasts are generated from a fixed information set, so the gradual widening of the predictive intervals reflects the accumulation of uncertainty as the latent volatility process is propagated forward in time. Nevertheless, both forecasting exercises convey a similar message: the predictive medians of return remain close to zero and the four specifications produce highly similar predictive distributions, suggesting that the core stochastic volatility mechanism captures the dominant features of return uncertainty.

3.3 Multivariate FSV Analysis

This section illustrates the behavior of the FSV framework introduced in chapter 2 using the multivariate sectoral return data. In contrast to the univariate SV specifications, the FSV framework jointly models the dynamic dependence structure across return series through a small K number of latent common factors together with series-specific idiosyncratic volatility components. The analysis focuses primarily on the statistical behavior of the model, including latent factor dynamics, volatility propagation mechanisms, posterior uncertainty, and the implied multivariate dependence structure.

To evaluate the contribution of common latent factors, FSV models with K latent factors, $K \in \{1, 2, 3\}$, are considered. The factor dimension is restricted to be relatively small to provide a parsimonious representation of cross-sectional dependence while maintaining computational tractability.

3.3.1 Estimation Setup

For the FSV specifications, the unrestricted factor loadings are assigned the default row-wise Normal-Gamma shrinkage prior implemented in the `factorstochvol` package (Hosszejni and Kastner, 2021). Conditionaly on a local variance parameter τ_{ij}^2, each factor loading follows a Gaussian priors $\lambda_{ij} \mid \tau_{ij}^2 \sim \mathcal{N}(0, \tau_{ij}^2)$, where the variance parameters τ_{ij}^2 determines the amount of shrinkage applied to the corresponding loading. Small values of τ_{ij}^2 imply that the loading is strongly concentrated around zero, whereas larger values permit more substantial factor exposures. The default shrinkage specification encourages weakly supported factor loadings to shrink towards zero while allowing important loadings to remain sizeable when supported by the data (Hosszejni and Kastner, 2021; Kastner

et al., 2017).

Other stochastic volatility parameters associated with both the latent factor volatility processes and the idiosyncratic volatility processes are assigned the same prior distributions and hyperparameter values as those used in the univariate SV models, shown in table 3.2. Each FSV specification is estimated and assessed using the same framework as in the univariate case.

3.3.2 Posterior Estimation Results

The latent factor volatility processes $h_{f,t}$ exhibit persistent dynamics and pronounced volatility clustering. These characteristics are evident from both the posterior parameter estimates reported in table 3.7 and the estimated latent factor volatility trajectories $h_{f,t}$ shown in figure 3.10. As is shown in table 3.7, the posterior estimates of the persistence parameters $\phi_{f,k}$ remain consistently high across all factor dimensions, with posterior medians close to 0.9, implying that common volatility shocks dissipate only gradually over time. At the same time, the posterior estimates of the factor volatility innovation parameters $\sigma_{f,k}$ are comparatively moderate across all specifications. Since $\sigma_{f,k}$ governs the magnitude of innovations in the latent volatility process, these estimates suggest that the latent factor volatility trajectories $h_{f,t}$ are sufficiently flexible to adapt to changing market conditions while retaining substantial persistence. Together, the high values of $\phi_{f,k}$ and the moderate values of $\sigma_{f,k}$ indicate a balance between adaptability and persistence in the evolution of latent market volatility.

	Factor 1	Factor 2	Factor 3
$\phi_{f,k}$	0.897 [0.594, 0.979]	0.881 [0.583, 0.979]	0.884 [0.608, 0.978]
$\sigma_{f,k}$	0.321 [0.132, 0.665]	0.333 [0.141, 1.061]	0.358 [0.150, 1.563]

Table 3.7: Posterior medians and 90% credible intervals for the persistence parameters $\phi_{f,k}$ and volatility innovation parameters $\sigma_{f,k}$ of the $K = 3$ FSV specification

These parameter estimates are complemented by the estimated latent factor volatility trajectories presented in figure 3.10. The three-latent-factor volatility process ($K = 3$ FSV specification) exhibits persistent fluctuations and occasional periods of elevated volatility, consistent with the high persistence implied by the posterior estimates of $\phi_{f,k}$. The associated credible intervals indicate that posterior uncertainty differs across factors.

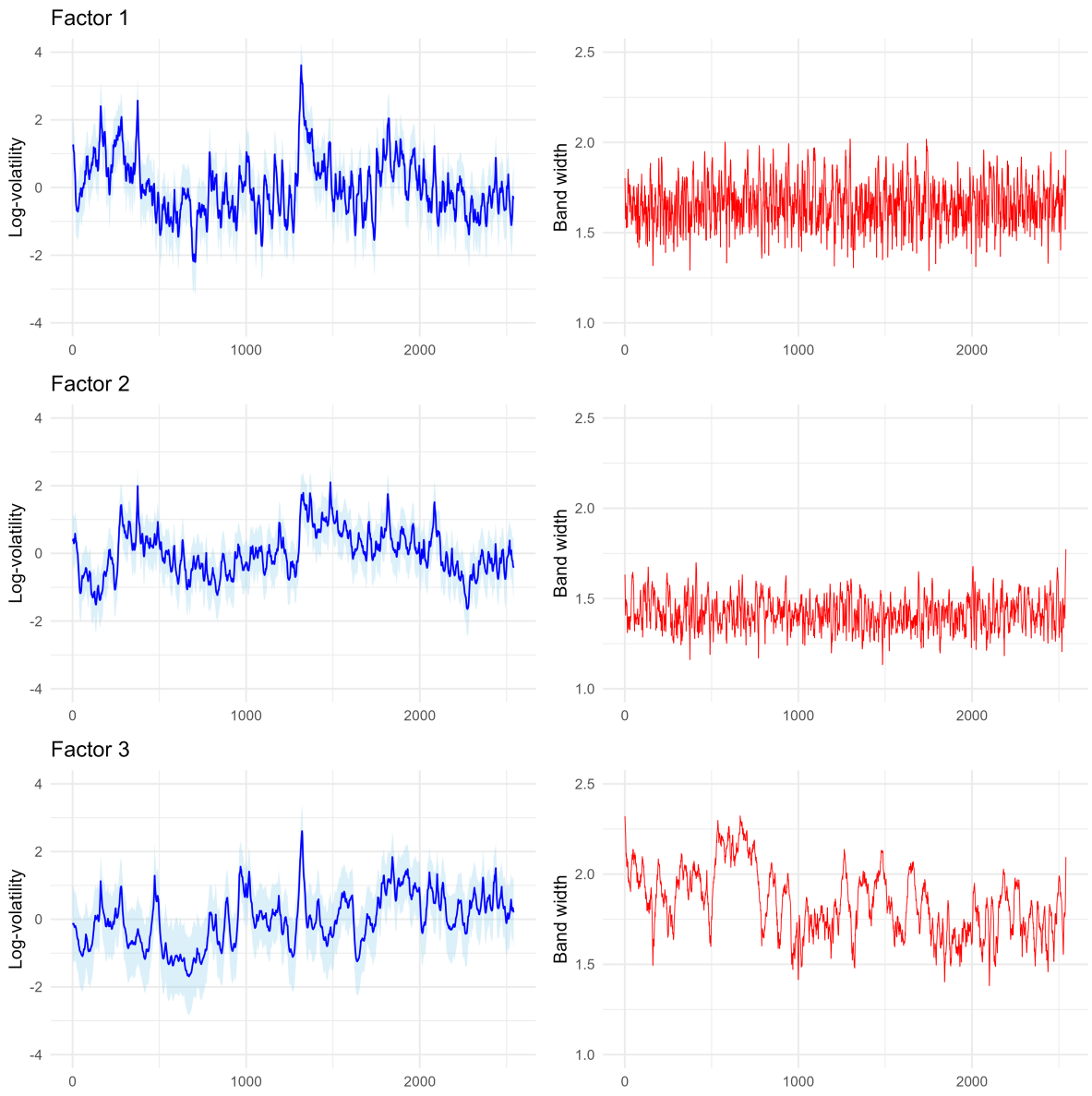
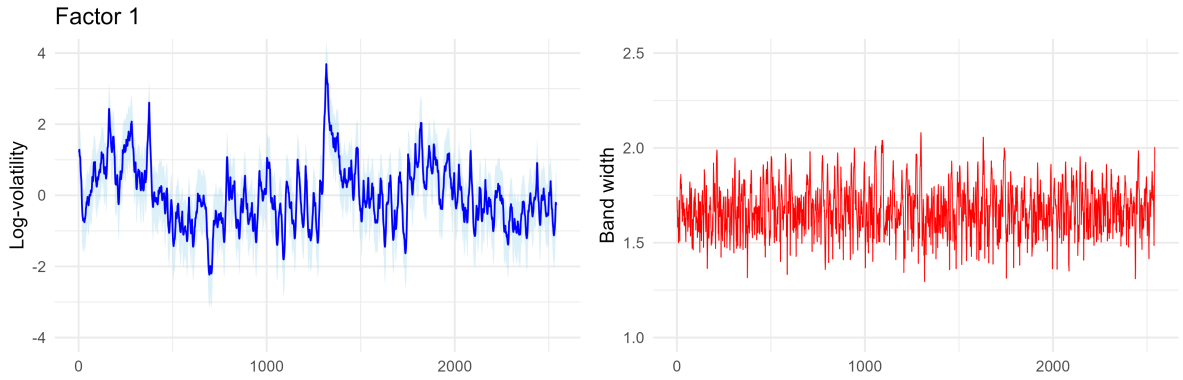


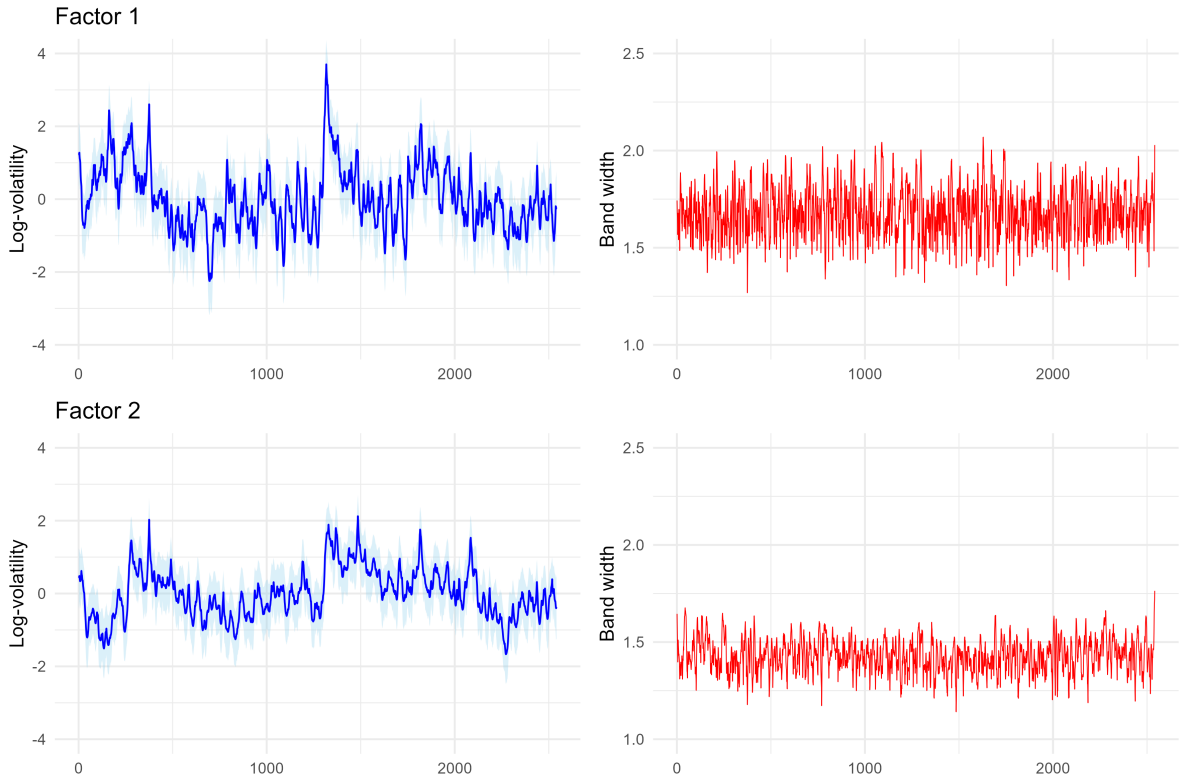
Figure 3.10: Posterior mean latent factor log-volatility processes $h_{f,t}$ with 90% credible intervals under the three-latent-factor model ($K = 3$ specification)
The blue curves represent the posterior median latent volatility, while the shaded areas denote the corresponding 90% *posterior credible intervals. The red curves represent the bandwidth of the* 90% *credible intervals, computed as the difference between the* 95% *and* 5% *posterior quantiles at each time point.*

In particular, the first two factors exhibit relatively stable uncertainty over time, suggesting that the latent factor volatilities are estimated with a fairly consistent degree of precision throughout the sample period. In contrast, the third factor exhibits noticeably wider and more variable credible intervals, indicating greater posterior uncertainty regarding the volatility trajectory of the third factor. Nevertheless, the overall pattern of persistent volatility dynamics remains broadly consistent across all latent factors, with each factor displaying gradual changes in volatility and periods of elevated volatility

clustering.



(a) $K = 1$ FSV specification



(b) $K = 2$ FSV specification

Figure 3.11: Posterior mean latent factor log-volatility processes $h_{f,t}$ with 90% credible intervals under the $K = 1$ and $K = 2$ FSV specification

The blue curves represent the posterior median latent volatility, while the shaded areas denote the corresponding 90% *posterior credible intervals. The red curves represent the bandwidth of the* 90% *credible intervals, computed as the difference between the* 95% *and* 5% *posterior quantiles at each time point.*

Together with figure 3.10, figure 3.11 further illustrates the stability of the estimated latent factor volatility processes across alternative model dimensions. In particular, the first latent factor exhibits remarkably similar volatility dynamics under the $K = 1$, $K = 2$, and $K = 3$ specifications, with the timing and magnitude of major volatility episodes

remaining largely unchanged. This suggests that the dominant common volatility component is robustly identified regardless of the number of latent factors included in the model. A similar pattern is observed for the second latent factor, whose volatility trajectories remain highly comparable between the $K = 2$ and $K = 3$ specifications. These results suggest that the primary sources of common volatility are robustly identified and largely insensitive to the choice of the number of latent factors included in the model. The stability of the estimated factor trajectories across model dimensions suggests that the dominant sources of common volatility are robust to the precise choice of factor dimension. This finding provides additional support for the interpretation of the latent factors as persistent drivers of broad market uncertainty rather than artifacts of a particular model specification.

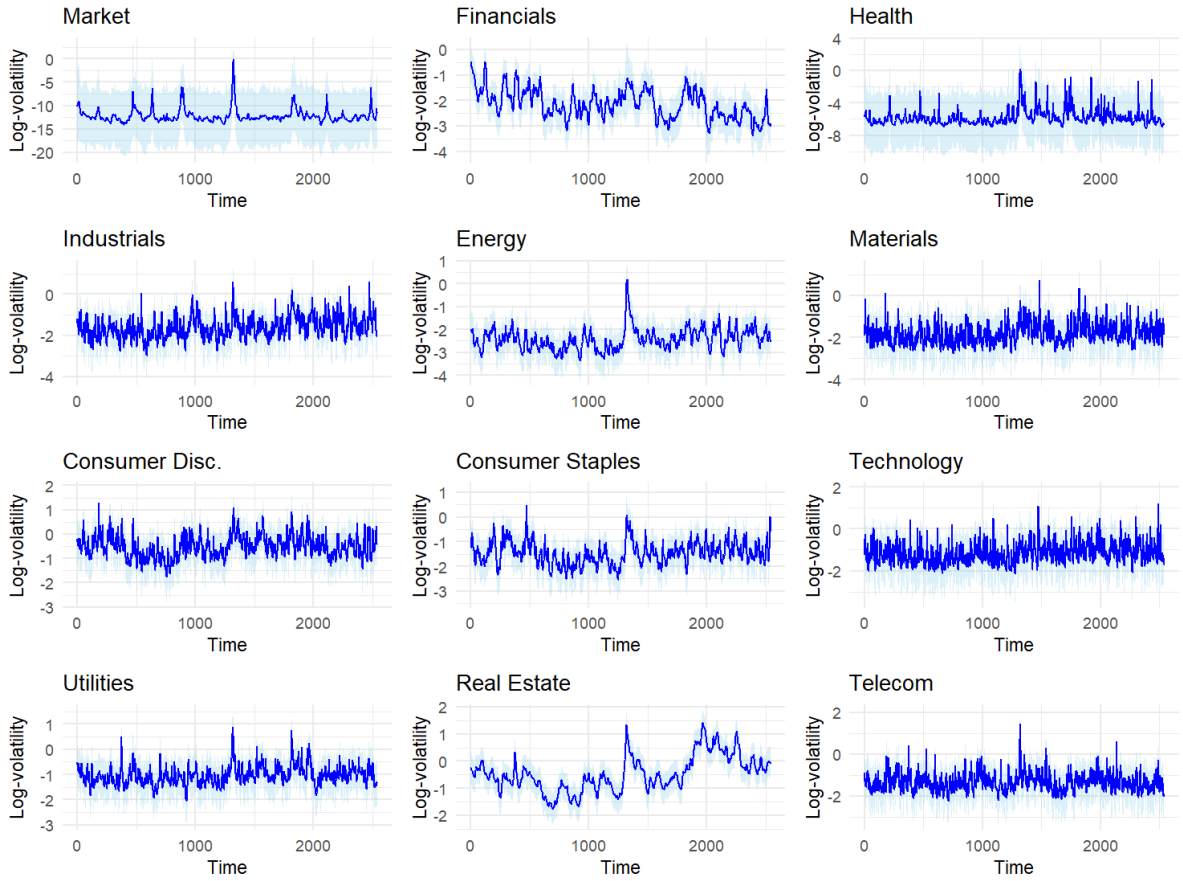


Figure 3.12: Posterior mean of idiosyncratic log-volatility processes $h_{i,t}$ with 90% credible intervals under $K = 3$ FSV specification

In addition to the common latent factor processes, the FSV framework also estimates series-specific idiosyncratic volatility components that capture residual volatility dynamics unique to individual return series after controlling for the common multivariate factors. Figure 3.12 shows that the estimated idiosyncratic volatility trajectories $h_{\varepsilon,i,t}$ are generally more irregular and localized than the latent factor volatility processes. Several sectors, such as financials and real estate, exhibit relatively persistent

idiosyncratic volatility dynamics characterized by recurring volatility waves over extended periods without extremely sharp local spikes. In contrast, series such as market, health and energy, exhibit more pronounced isolated volatility spikes and abrupt short-lived increases in volatility, suggesting the presence of a more localized sector-specific shocks.

Together, the factor and idiosyncratic volatility dynamics jointly determine the time-varying covariance structure $\text{Var}(y_t \mid h_t)$ (equation (2.15)) implied by the FSV framework: the former driving common co-movement across sectors, and the latter capturing sector-specific variation. Beyond identifying the sources of uncertainty, the latent volatility processes describe how these common and sector-specific components evolve over time. The relative contribution of each component is governed by the factor loading matrix Λ. Examination of the posterior factor loadings therefore provides insight into how the common volatility factors are transmitted across sectors and how strongly individual sectors are connected to the underlying latent structure. The posterior mean of the loading structure, presented in figure 3.13, indicates that a relatively small number of latent factors explains a substantial proportion of the common volatility dynamics.

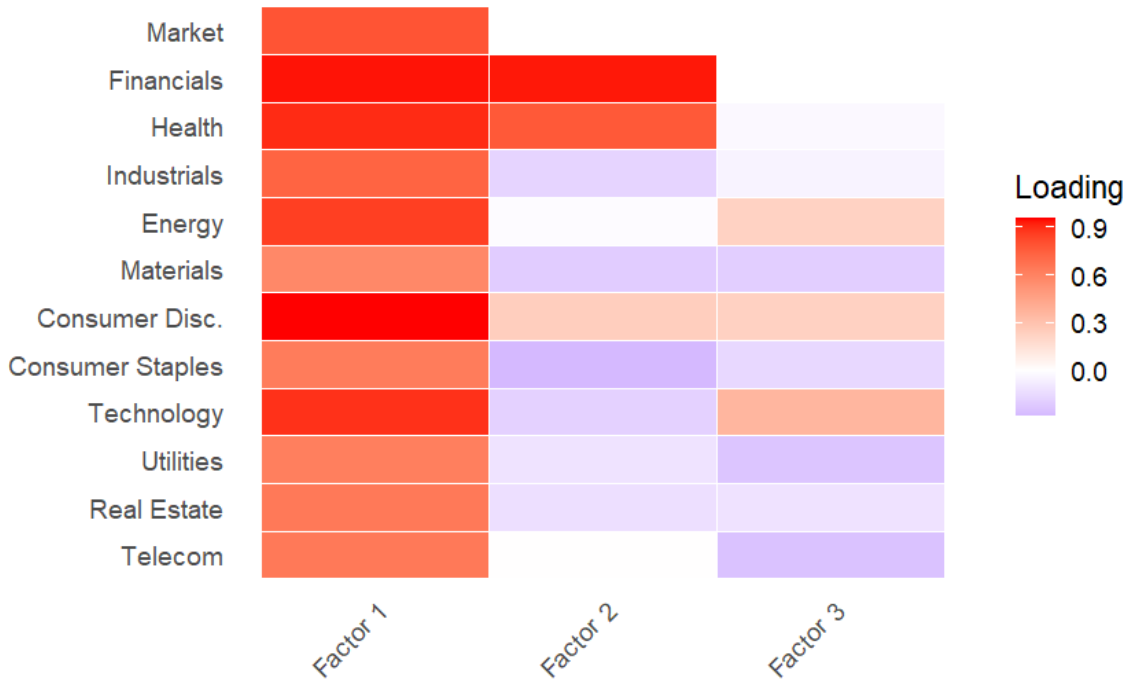


Figure 3.13: Posterior mean of factor loading matrix for the $K = 3$ FSV specification

In particular, the first latent factor loading column in Λ exhibits consistently large positive loading across most sectors, indicating that the first factor contributes substantially to the variation of nearly all return series and captures a pervasive source of common co-movement. This suggests that the first factor captures broad

market-wide volatility movements shared across sectors, consistent with the dominant and highly persistent volatility dynamics observed previously in the latent factor volatility trajectories $h_{f,t}$. The higher-order factors exhibit comparatively smaller and more heterogeneous loading patterns across sectors. Certain sectors, including financials, healths, and consumer discretionary, display relatively stronger exposures to the additional latent factors, suggesting the presence of secondary covariance structures beyond the dominant market-wide component. In contrast, several sectors exhibit loadings close to zero for the higher-order factors and therefore rely more heavily on their idiosyncratic volatility components. The estimated loading matrix therefore demonstrates that the FSV framework learns a low-rank representation of the multivariate covariance structure. Rather than estimating a fully unrestricted time-varying covariance matrix, the model concentrates most of the cross-sectional dependence into a relatively small number of latent common factors, confirming the parsimony of the factor structure. This result is particularly important from a modeling perspective. A fully unrestricted covariance specification would require a substantially larger number of parameters, especially in higher dimensions. By summarizing the dominant dependence structure through only a few latent factors, the FSV framework achieves a considerable reduction in model complexity while preserving the key features of the observed co-movement across sectors.

3.3.3 Dynamic Covariance and Correlation Estimation

An important advantage of the FSV framework is its ability to estimate the entire time-varying covariance structure of the asset returns. Unlike univariate SV models, which focus solely on marginal volatilities, the FSV specification produces posterior draws of the conditional covariance matrix $\mathrm{Var}(y_t \mid h_t)$ (equation (2.15)). This decomposition allows the covariance structure to evolve dynamically over time, capturing both common market-wide shocks through the latent factors and asset-specific variation through the idiosyncratic components. Posterior draws of Σ_t obtained from the MCMC output provide a flexible estimate of the conditional variance-covariance matrix for all sector returns.

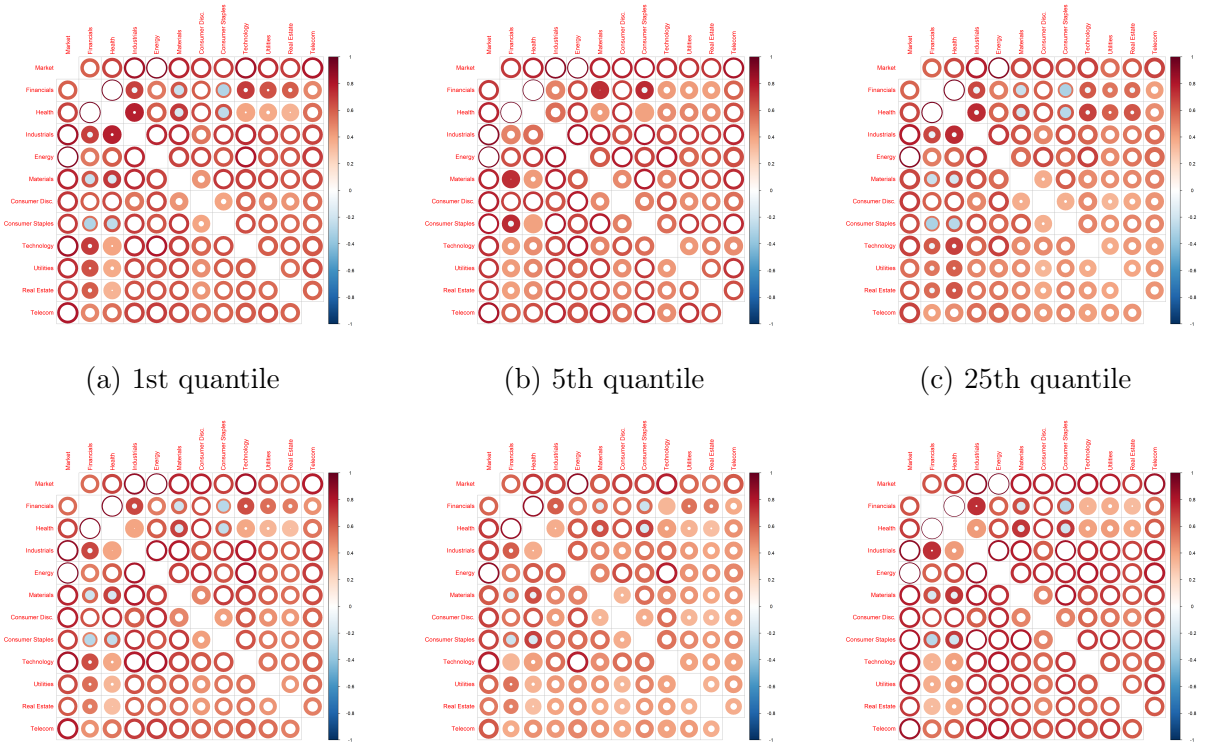
For interpretation, the covariance matrices can be transformed into correlation matrices using

$$\rho_{ij,t} = \frac{\Sigma_{ij,t}}{\sqrt{\Sigma_{ii,t}\Sigma_{jj,t}}}, \tag{3.3}$$

where $\rho_{ij,t}$ denotes the conditional correlation between sectors i and j at time t. The resulting correlation matrices provide a standardized measure of dependence structure between sectors implied by the FSV model and facilitate comparisons across different market conditions.

Figure 3.14 examines the dependence structure across different volatility regimes, represented by conventional quantile levels of the STOXX Europe 600 index volatility

distribution.



(a) 1st quantile (b) 5th quantile (c) 25th quantile

(d) 50th quantile (e) 75th quantile (f) 95th quantile

Figure 3.14: Pairwise correlation matrices at selected market volatility quantiles under $K = 3$ FSV specification

Pairwise correlations are visualized using circles. The inner and outer radii of each circle correspond to the posterior mean minus and plus two standard deviation, respectively. Circle colors indicate the sign and magnitude of the posterior mean correlations, with red (blue) representing positive (negative) values (Hosszejni and Kastner, 2021). Consequently, the width of the band reflects the uncertainty associated with the estimated correlation, with wider bands indicating greater posterior uncertainty and thinner bands indicating more precise estimates.

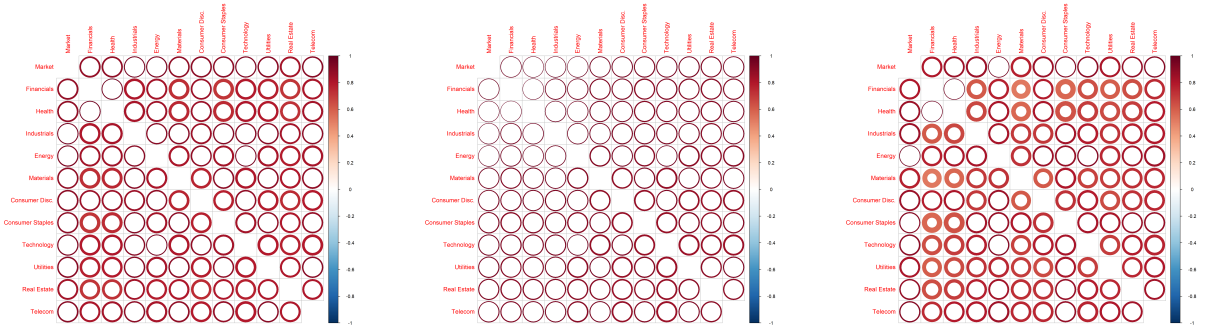
The correlation matrices can be broadly classified into three volatility regimes: low volatility (1st–5th percentiles), intermediate volatility (25th–75th percentiles), and extreme volatility (95th percentile). Although each matrix corresponds to a single representative date and therefore reflects some idiosyncratic features, several common patterns emerge across the different regimes.

In the low-volatility regime, sectoral correlations are generally moderate and exhibit greater heterogeneity. While positive correlations remain dominant, several sector pairs display noticeably weaker dependence, suggesting that sector-specific factors play a more important role when overall market uncertainty is subdued. As volatility increases to intermediate levels, the correlation structure becomes relatively stable, with most sector pairs maintaining moderate positive correlations and only limited changes in the overall dependence pattern. This stability indicates that the sectoral dependence structure is

largely preserved during normal market conditions.

A more pronounced shift is observed in the extreme-volatility regime represented by the 95th percentile. Relative to the low- and intermediate-volatility regimes, correlations become more uniformly positive across sectors, and stronger co-movement emerges among a larger number of sector pairs. This finding suggests that common market-wide shocks increasingly dominate sector-specific influences during periods of elevated uncertainty, leading to a more synchronized market structure. In addition to stronger correlations, the uncertainty bands surrounding the posterior correlation estimates become noticeably narrower during extreme-volatility episodes. Since the width of the band reflects posterior uncertainty, this suggests that sectoral correlations are estimated more precisely when market volatility is elevated.

The evidence from the 95th-percentile matrix is further reinforced by the correlation matrices corresponding to the three highest realizations of latent factor volatility shown in figure 3.15. During these extreme episodes, positive correlations become widespread across nearly all sector pairs. Particularly strong dependence is observed among Financials, Materials, Industrials, and Technology, suggesting that these sectors tend to move closely together during periods of severe market stress.



(a) Top 1: COVID-19 market crash (24 Mar 2020)

(b) Top 2: China global selloff (25 Aug 2015)

(c) Top 3: Oil-price shock (9 Mar 2022)

Figure 3.15: Pairwise sectoral correlation matrices during the three most extreme market volatility episodes

The correlation matrices discussed above summarize posterior dependence at particular dates. In contrast, figure 3.16 displays one-step-ahead posterior predictive draws generated from the fitted $K = 3$ FSV specification, which describe the expected joint behavior of future sector returns after integrating over posterior uncertainty in model parameters, latent factors, and volatility states.

Overall, the predictive dependence among sectors is relatively modest. With the exception of the Financials–Health pair, which exhibits the strongest positive predictive correlation $\rho = 0.51$, most sector pairs display correlations below 0.30 in absolute value. Several pairs involving the Energy sector show moderate negative dependence, including

energy–technology $\rho = 0.47$, energy–materials $\rho = -0.43$, or energy–health $\rho = -0.34$. These predictive distributions indicate that future sector returns exhibit only moderate dependence in most cases, despite sharing common sources of risk.

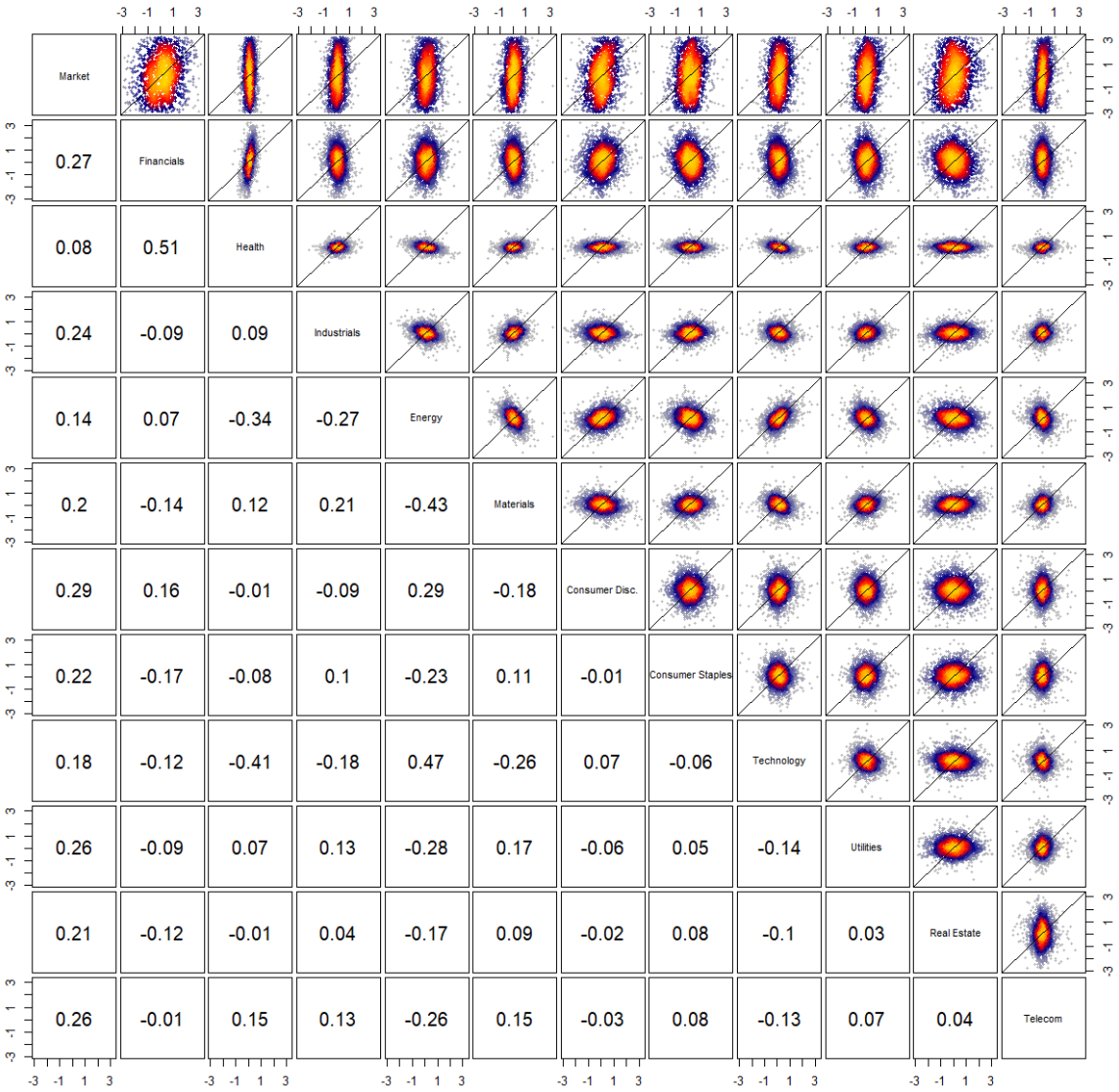


Figure 3.16: Posterior predictive draws for sectoral returns under the $K = 3$ FSV specification

The figure is based on one-step-ahead posterior predictive draws generated from the fitted $K = 3$ FSV specification rather than on the observed returns. The upper triangular panels visualize the joint predictive distributions for each pair of sectors, whereas the lower triangular panels report the corresponding predictive correlations.

The scatterplots provide a visual confirmation of these findings. Most predictive distributions exhibit an approximately elliptical shape centered near zero, which is consistent with the Gaussian specification adopted in the multivariate FSV model. The orientation of the major axis reflects the sign of predictive dependence, with ellipses

aligned along the positive diagonal indicating positive correlations and ellipses aligned along the negative diagonal indicating negative correlations. More elongated ellipses correspond to stronger predictive relationships, whereas nearly circular clouds indicate weak predictive dependence.

Taken together, the predictive distributions imply that future sector returns are neither independent nor excessively synchronized. Moderate predictive correlations coexist with substantial sector-specific variation, suggesting that diversification benefits remain available even after accounting for common sources of risk.

3.3.4 Model Diagnostics

Having established the estimated factor structure and the implied covariance decomposition, the adequacy of the FSV specification is assessed through standardized residual diagnostics to validate the findings discussed above. Specifically, standardized residuals

$$\hat{z}_t = D_t^{-1/2}(y_t - \hat{\beta} - \hat{\Lambda}\hat{f}_t) \tag{3.4}$$

are examined, where $\hat{\Lambda}$ denotes the posterior mean factor loading matrix, $\hat{f}_t$ is the posterior mean latent factor vector, and $D_t^{-1/2} = \hat{\Sigma}_{\varepsilon,t} = \text{diag}\left(e^{\hat{h}_{\varepsilon,1,t}}, \ldots, e^{\hat{h}_{\varepsilon,N,t}}\right)$ is the diagonal matrix of posterior mean idiosyncratic variances. Under a well-specified FSV models, these standardized residuals should exhibit approximately white-noise behavior, implying little remaining serial dependence or volatility clustering after accounting for both the common factor structure and time-varying idiosyncratic volatility. The diagnostics shown in figure 3.17 provide evidence that the FSV specification offers an adequate fit to the data: the standardized residuals fluctuate around zero with relatively stable variance, and the ACFs of squared standardized residuals are generally close to zero. The QQ plots, however, exhibit moderate deviations from the theoretical normal quantiles in the tails, suggesting the probable presence of residual heavy-tailed behaviour.

The residual correlation matrix in figure 3.18, reveals non-negligible remaining cross-sectional dependence across most sector pairs. This implies that while the latent factors account for a substantial proportion of the common co-movement, some residual correlation persists - consistent with the inherent parsimony of restricting the common component to $K = 3$ latent factors rather than reflecting model misspecification. Taken together, the diagnostics indicate that the FSV model provides an adequate representation of the main temporal and cross-sectional dynamics of the return series, despite evidence of some remaining cross-sectional dependence.

Overall, these results indicates that the FSV specification provides a coherent and parsimonious representation of multivariate volatility dynamics. The latent factor processes capture persistent sources of common uncertainty, while the idiosyncratic volatility components account for sector-specific fluctuations. Together with the

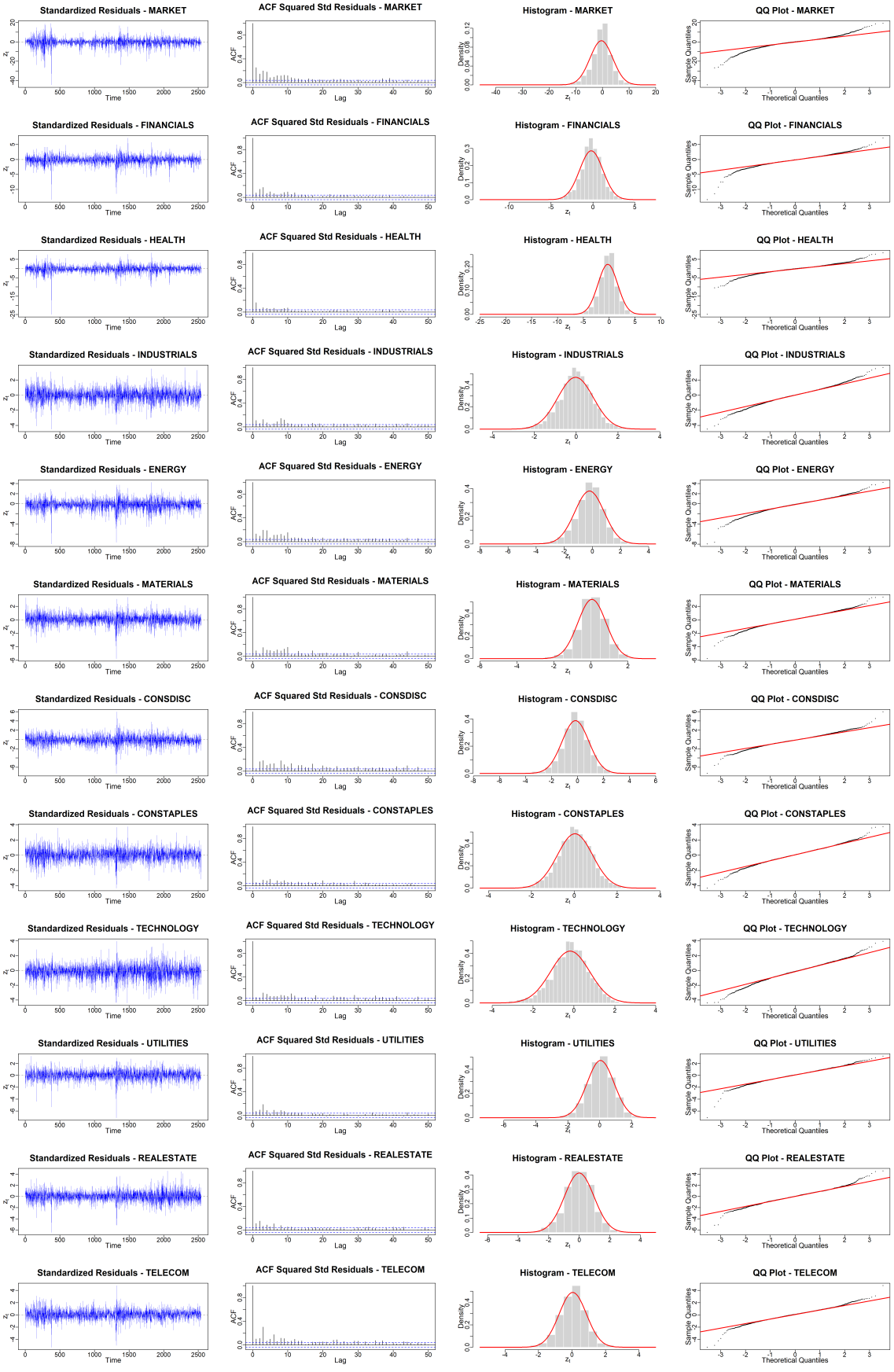


Figure 3.17: Standardized residual diagnostics for the $K = 3$ FSV specification

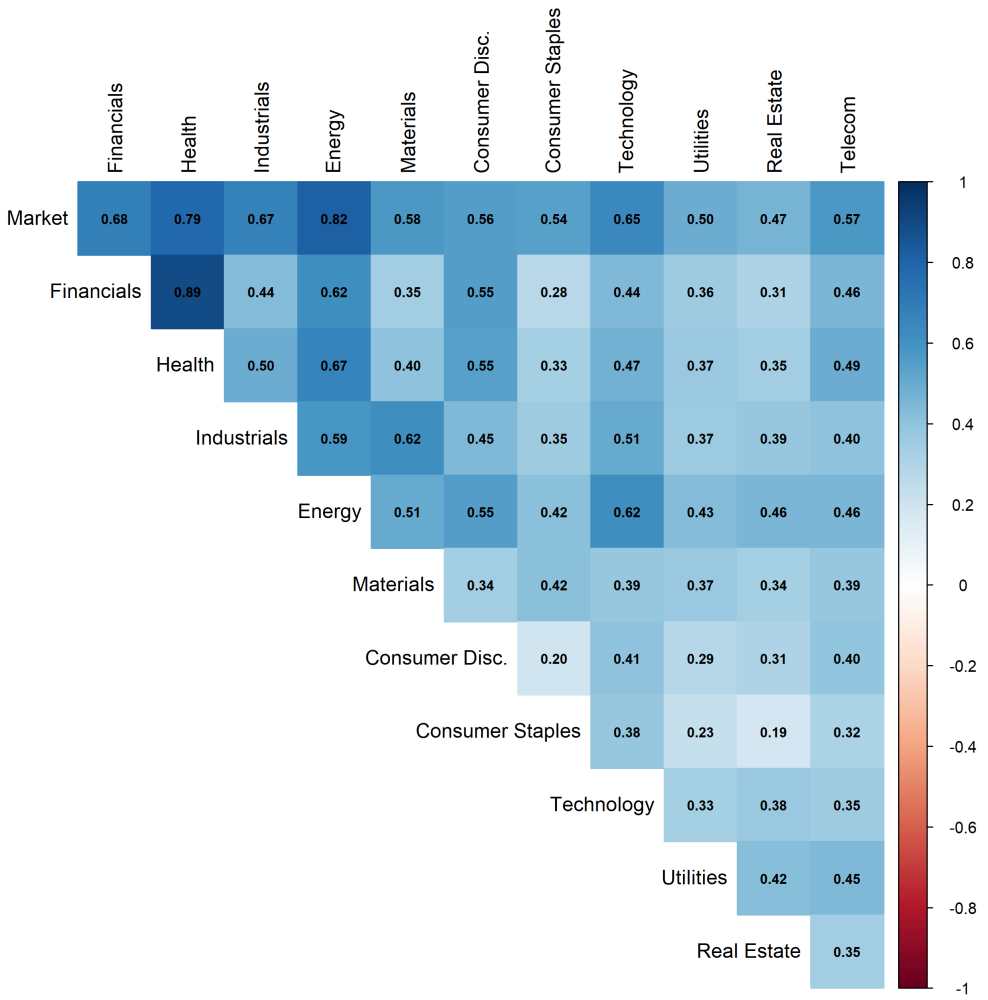


Figure 3.18: Correlation matrix of standardized residuals $\hat{z}_t$ implied by the $K = 3$ FSV specification

estimated factor loadings, these elements generate a flexible time-varying covariance structure that captures the main features of the cross-sectional dependence observed in the data, although some residual dependence remains.

3.3.5 MCMC Diagnostics

Trace Plots and Mixing Behavior

Trace plots of the FSV specifications in figure 3.19, and figure 3.20 indicate that mixing behavior varies across series and specifications. The chains for most idiosyncratic series and all latent factors fluctuate around stable levels throughout the MCMC chain, with no systematic drifts or explosive movements, suggesting adequate mixing for these components. However, the chains of μ, ϕ, and σ for market series exhibit pronounced gradual drifts in the $K = 1$ specification. In the higher-dimensional $K = 2$ and $K = 3$ specifications, mixing becomes noticeably slower for the market and health volatility parameters. Their trace plots exhibit persistent autocorrelation and gradual shifts across

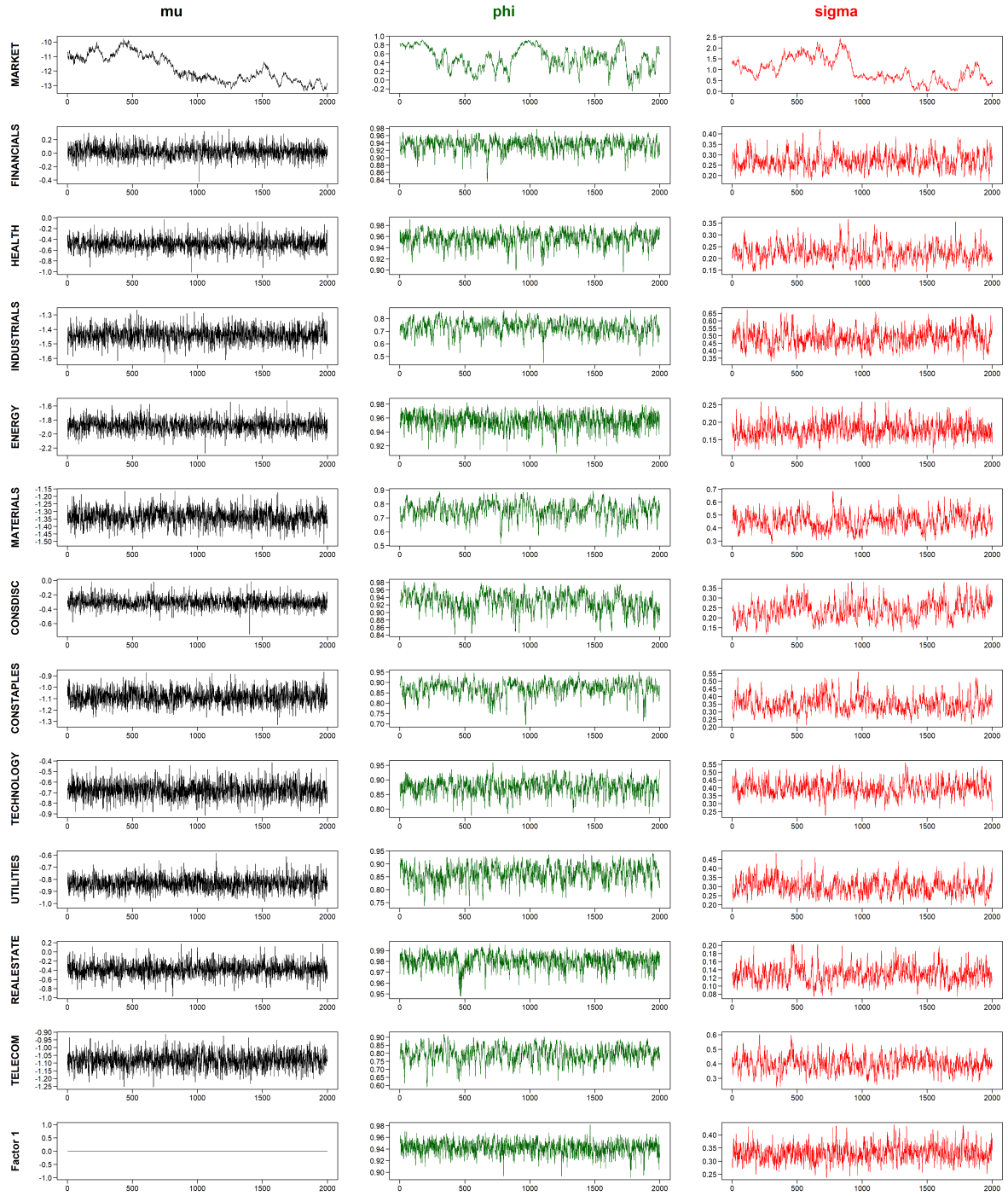


Figure 3.19: Trace plots of posterior draws of the volatility parameters in the $K = 1$ FSV specification

Sector-labelled rows correspond to idiosyncratic volatility processes, while Factor 1 denotes the latent common volatility factor capturing co-movements across sectors.

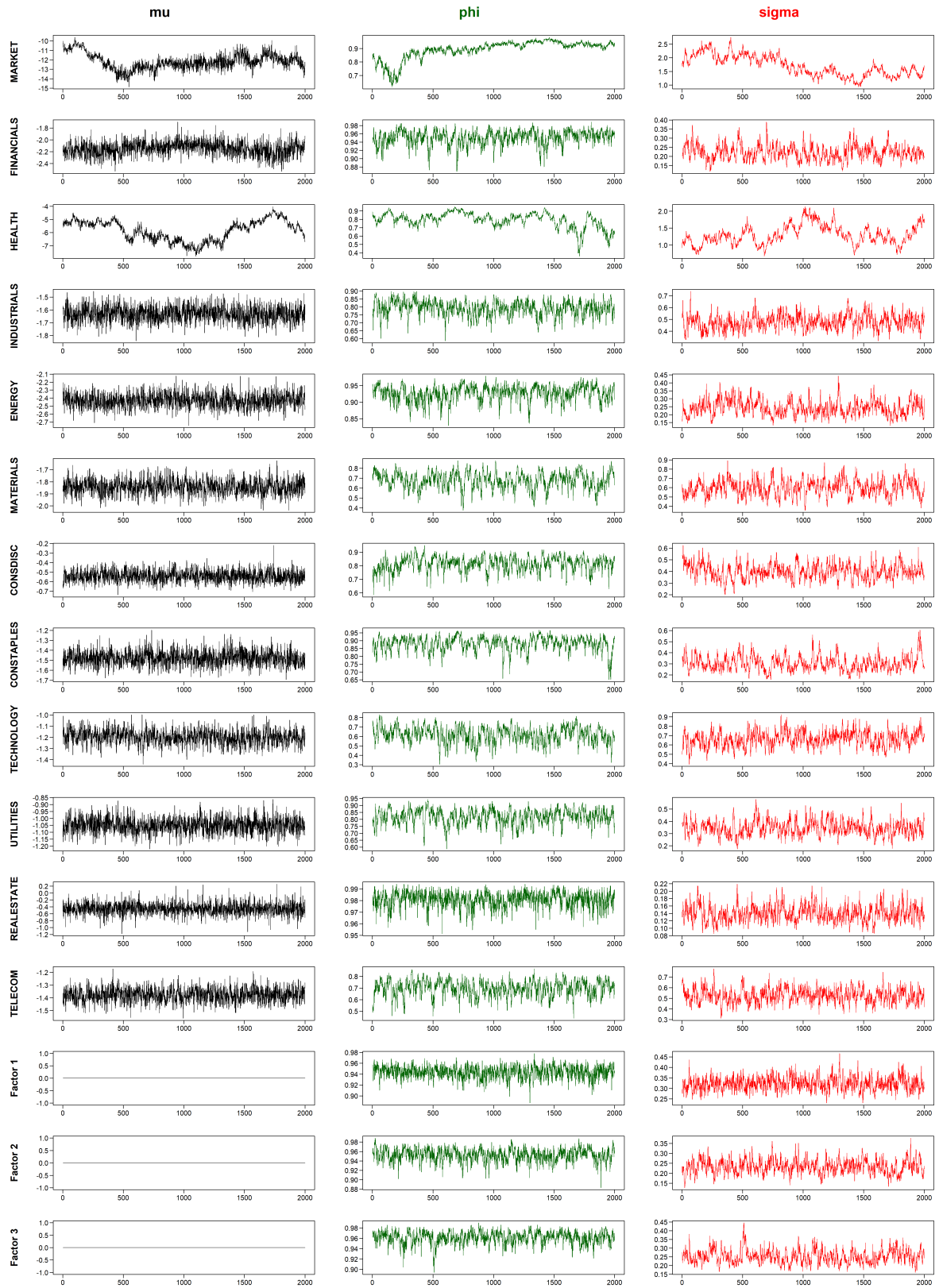


Figure 3.20: Trace plots of posterior draws of the volatility parameters in the $K = 3$ FSV specification

Sector-labelled rows correspond to idiosyncratic volatility processes, while Factor 1 – Factor 3 denotes the latent common volatility factor capturing co-movements across sectors.

the MCMC iterations, suggesting less efficient exploration of the posterior distribution than for the remaining series. These patterns suggest that the sampler mixes poorly for the volatility parameters of these two sectors when additional factors are introduced, and interpretations of the corresponding estimates should be made with caution.

The ACFs of selected chains, shown in figure 3.21, reveal moderate serial dependence in some parameters but generally satisfactory mixing for others. For both the $K = 1$ and $K = 3$ models, the ACFs of representative factor loadings Λ and latent factor states f are negligible, indicating good mixing of these components. The persistence parameter ϕ_f displays more persistent autocorrelations, with the higher-dimensional specification $K = 3$ showing a somewhat slower decay than the lower-dimensional $K = 1$, reflecting the additional complexity introduced by the richer factor structure. Nevertheless, the autocorrelations decline steadily across lags and remain consistent with satisfactory mixing behaviour.

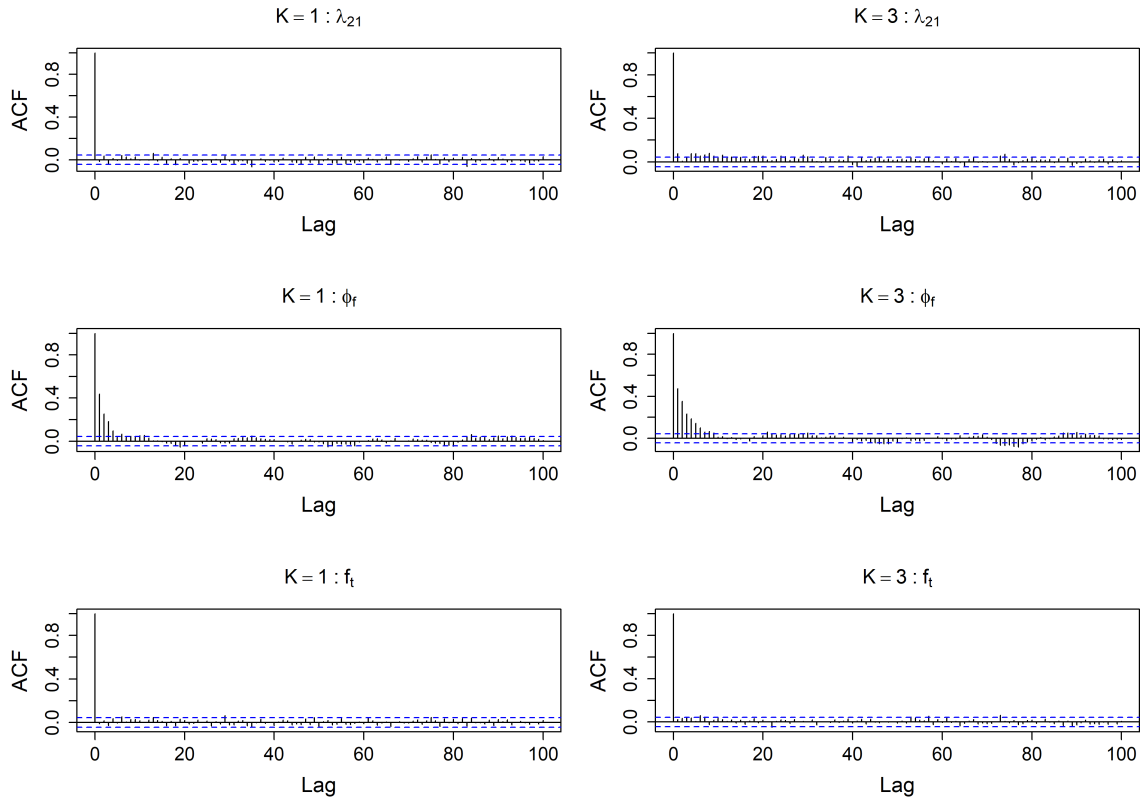


Figure 3.21: ACF of selected MCMC chains for the FSV specifications

Figure 3.22 presents trace plots of the factor loading parameters under the $K = 3$ specification. The MCMC chains for the first factor loadings The chains remain stable throughout the sampling period and fluctuate around well-defined posterior regions, providing evidence of satisfactory mixing for the loading parameters. The first-factor

loadings (black lines) are consistently positive and relatively large across most sectors, supporting the interpretation of the first latent factor as the dominant source of common market-wide volatility variation. In contrast, the second- and third-factor loadings (red and green lines) are generally smaller and exhibit greater cross-sectional heterogeneity, with several loadings centered close to zero while others remain clearly distinct from zero, suggesting that the higher-order factors capture more localized covariance structures affecting only subsets of sectors. The absence of noticeable drifts or regime-switching behavior in the chains indicates that the estimated loading structure is reasonably stable and well identified under the selected specification.

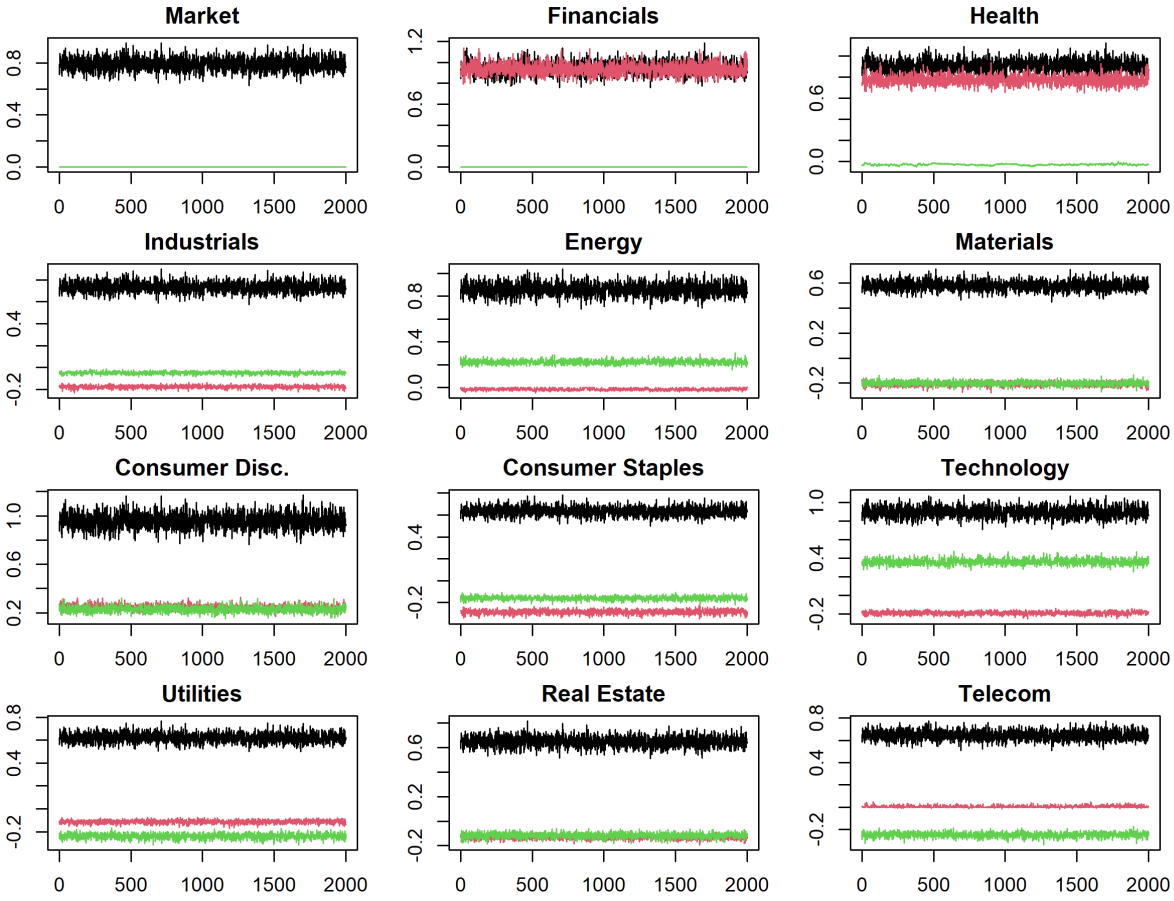


Figure 3.22: Trace plots of factor loading parameters under the $K = 3$ FSV specification

The slower mixing observed for some volatility parameters likely is consistent with increased dimensionality and dependence structure of the FSV models, particularly as the number of latent factors increases. While most parameters exhibit stable trace plots and moderate autocorrelation levels, a subset of volatility parameters shows weaker mixing behaviour. Therefore, the estimation results for these parameters should be interpreted with some caution, and the diagnostics suggest that posterior exploration becomes more challenging in the richer model specifications.

Convergence Diagnostics

For the FSV models, convergence becomes increasingly challenging as the number of latent factors increases. The higher-dimensional specifications exhibit stronger posterior dependence and greater serial correlation, which reduces sampling efficiency and slows posterior exploration. The Geweke statistics and ESS reported in table 3.8 and table 3.9 broadly support the conclusions obtained from the trace plot and ACF diagnostics. Most sector-specific volatility parameters exhibit Geweke statistics within the conventional $[-1.96, 1.96]$ interval together with moderate to high effective sample sizes, indicating satisfactory convergence and mixing behavior for the majority of series. In particular, sectors such as industrials, energy, materials, consumer discretionary, utilities, real estate, and telecommunication display relatively large ESS values, suggesting efficient posterior exploration despite the high dimensionality of the FSV specification.

	μ	ϕ	σ
Market	8.35	-5.88	6.92
Financials	-1.20	-2.28	1.93
Health	1.17	1.70	-2.30
Industrials	0.29	0.94	-0.86
Energy	0.24	0.43	-0.54
Materials	0.80	1.48	-1.51
Consumer Disc.	-0.63	-3.64	3.14
Consumer Staples	0.33	-0.30	0.57
Technology	2.65	2.12	-1.85
Utilities	0.53	0.10	-0.10
Realestate	0.05	-0.79	0.78
Telecom	-1.50	-0.34	0.76

Table 3.8: Geweke z-statistics for static parameters under the $K = 3$ FSV specifications

In contrast, the market and health series exhibit noticeably weaker convergence diagnostics. For the market series, the Geweke statistics for μ, ϕ, and σ substantially exceed the conventional acceptance bounds. The corresponding ESS values are exceptionally low and among the smallest observed across all estimated parameters, indicating substantial autocorrelation and inefficient posterior exploration for these parameters. For the health series, only the volatility innovation parameter σ exceeds the conventional Geweke threshold, whereas the diagnostics for μ and ϕ remain within acceptable limits. Nevertheless, all three volatility parameters within the health series

	μ	ϕ	σ
Market	7.60	5.20	9.30
Financials	174.10	127.40	130.30
Health	3.70	19.50	17.30
Industrials	1108.60	178.60	189.50
Energy	947.90	112.20	92.10
Materials	717.10	118.40	118.70
Consumer Disc.	957.80	151.90	131.70
Consumer Staples	957.80	94.40	92.90
Technology	583.30	164.90	173.60
Utilities	884.50	153.60	150.90
Realestate	2000.00	334.10	212.20
Telecom	807.20	175.70	197.10

Table 3.9: Effective sample sizes for the $K = 3$ FSV specification

exhibit comparatively low ESS values, suggesting persistent autocorrelation and slow mixing. These findings are consistent with the irregular drifts and slower mixing behavior observed in the trace plots and ACF diagnostics for the market and health series under the $K = 3$ specification.

The convergence diagnostics indicate that posterior exploration becomes increasingly challenging as the dimensionality of the FSV specification increases. Although most parameters exhibit satisfactory convergence behavior, selected parameters for the market and health series display slower or even poor mixing and lower ESS. These findings highlight the computational challenges associated with estimating high-dimensional latent state-space models. In particular, the increased model complexity appears to reduce sampling efficiency for a subset of volatility parameters, suggesting that longer MCMC runs may be required to obtain more reliable posterior summaries.

3.3.6 Density Forecast Evaluation

The predictive evaluation considered in this section is primarily intended for model comparison rather than for practical forecasting. Specifically, the models are assessed using one-step-ahead log predictive scores as equation (1.32), which measure how much probability mass each model assigns to observations that were not included in the estimation sample. A model receiving a higher predictive density for the realized one-step-ahead observations is regarded as providing a more accurate description of

the underlying data-generating process. Consequently, the log predictive scores serve as a criterion for comparing competing specifications on the basis of their predictive distributions. The objective is not to generate long-horizon forecasts, but rather to determine which model provides the most plausible probabilistic description of near future observations among the candidate specifications considered in this study.

For the FSV specifications, the introduction of latent common factors allows the models to capture dynamic cross-sectional dependence and common volatility movements across sectors. The cumulative log predictive scores indicate improved predictive performance as the factor dimension increases from $K = 1$ to $K = 2$. In particular, the $K = 1$ specification produces a cumulative log predictive score of approximately $-1{,}284.39$ while the $K = 2$ specification improves to approximately $-1{,}156.56$. Additional gains are obtained from the $K = 3$ specification, which achieves the highest predictive score of approximately $-1{,}122.69$, although the improvement relative to $K = 2$ is considerably smaller than that observed between $K = 1$ and $K = 2$. These results suggests diminishing marginal predictive gains as additional latent factors are introduced. Within the range of specifications considered in this thesis, a relatively small number of latent factors appears sufficient to capture most of the predictive information contained in the joint return dynamics. Consequently, the $K = 3$ FSV specification is preferred according to the predictive model comparison criterion adopted here. However, this conclusion should be interpreted relative to the set of candidate models examined, rather than as evidence that three factors are universally optimal.

The comparison across factor dimensions further illustrates the trade-off between model fit and computational complexity. In the present application, the three-factor specification achieves the strongest predictive performance among the models considered, while the computational diagnostics shows that gains from additional factors come at the cost of more challenging posterior inference and slower MCMC mixing for certain parameters. The results therefore illustrate a trade-off between model flexibility and computational tractability in high-dimensional Bayesian latent state-space models.

Chapter 4

Conclusion

4.1 Summary of Findings

This thesis investigated a range of SV models for financial return series within a Bayesian framework. The analysis considered both univariate and multivariate specifications, including the standard Gaussian SV model, extended to heavy-tailed specification with Student-t and leverage effects, and FSV models for multivariate returns.

Several key findings emerge from the analysis. First, the results confirm that for univariate SV specifications, allowing for heavy-tailed return innovations improves predictive performance relative to the Gaussian benchmark. Although no single specification uniformly dominates across all sectors, the heavy-tailed SVt and SVtl models are selected most frequently according to log predictive scores, indicating that explicit accommodation of extreme return realizations is beneficial for forecasting financial returns.

Second, the predictive differences across specifications are generally moderate rather than dramatic. This finding is consistent with the theoretical result that even the Gaussian SV model generates excess kurtosis through its stochastic volatility mechanism, allowing it to capture part of the heavy-tailed behavior commonly observed in financial data. Nevertheless, the Student-t extensions provide additional flexibility in the tails of the predictive distribution and exhibit greater robustness to extreme observations.

Third, leverage effects contribute additional predictive gains for some sectors, although these improvements are generally smaller than those obtained from introducing heavy-tailed innovations. The results therefore suggest that modeling tail behavior plays a more important role than leverage effects in improving predictive performance for the sectoral return series considered in this study. However, the leverage framework remains valuable because it explicitly captures the asymmetric volatility response to positive and negative return shocks, thereby providing a richer characterization of volatility dynamics.

For the multivariate analysis, the FSV framework provides a parsimonious

representation of the covariance structure by introducing a small number of latent factors. The estimated factor loadings reveal the presence of common sources of variation shared across sectors, while the latent volatility processes generate a flexible time-varying covariance structure capable of capturing the dominant dependence patterns in the data. Although some residual cross-sectional dependence remains, the FSV specification adequately captures the primary features of the multivariate volatility dynamics.

Finally, the MCMC diagnostics highlight the computational challenges associated with estimating higher-dimensional stochastic volatility models. Most factor loadings, latent factor states, and volatility parameters exhibit satisfactory mixing behavior and convergence diagnostics. However, the introduction of additional latent factors increases sampling difficulty for a small subset of volatility parameters, particularly those associated with the market and health series. These parameters display stronger autocorrelation, lower effective sample sizes, and slower mixing behavior and, in the case of the market volatility parameters, evidence of weak convergence diagnostics under the $K = 3$ specification.

Beyond the specific financial application, these findings highlight an important implication for Bayesian inference in multivariate stochastic volatility models. As model complexity increases, gains in flexibility and explanatory power must be balanced against reduced sampling efficiency and greater computational demands, illustrating a fundamental trade-off between model richness and inferential tractability.

Overall, the findings demonstrate that Bayesian SV models provide a flexible framework for modeling financial return dynamics. Heavy-tailed innovations improve predictive performance, leverage effects offer more modest gains. FSV models provide an effective and parsimonious approach to modeling multivariate volatility and dependence structures, but they become computationally challenging as the model dimension increases.

4.2 Limitations and Directions for Future Research

4.2.1 Limitations

Several limitations of the present study should be acknowledged. First, the analysis was primarily intended to illustrate the behavior and estimation of SV models rather than to provide a comprehensive investigation across multiple markets, asset classes, or forecasting environments. Consequently, the empirical conclusions should be interpreted within the context of the specific sectoral return data considered in this thesis.

Second, the forecasting evaluation focused exclusively on density forecasts evaluation through cumulative log predictive scores. Other economically relevant forecasting objectives, such as multi-step forecasting, risk measurement, portfolio allocation, and

Value-at-Risk prediction, were beyond the scope of the present study.

Third, the estimation of higher-dimensional FSV specifications revealed non-negligible computational challenges. Although the overall diagnostic evidence supported the adequacy of the posterior inference, a subset of volatility parameters exhibited slower mixing, stronger autocorrelation, and relatively low effective sample sizes. These findings highlight the substantial computational burden associated with estimating FSV models and suggest that posterior summaries should be interpreted with appropriate caution, particularly for parameters exhibiting weaker MCMC performance.

In addition, a further limitation concerns the FSV specifications. The lower-triangular identification scheme adopted in this thesis implies that the interpretation of the latent factors may depend on the ordering of the observed series. Consequently, different orderings may lead to alternative factor representations, and the estimated factors should therefore be interpreted in the context of the chosen identification structure.

Last but not least, the multivariate analysis is restricted to the Gaussian FSV specification. As a result, the model may not fully capture tail dependence and joint extreme movements across sectors. Consequently, the conclusions regarding multivariate dependence should be interpreted within the context of the Gaussian FSV framework considered in this thesis.

4.2.2 Directions for Future Research

Several directions for future research naturally emerge from the present study. From an empirical perspective, future research could extend the forecasting analysis beyond density forecast evaluation based on log predictive scores to consider additional forecasting objectives, including explicit multi-step forecasting, risk forecasting, portfolio allocation, and Value-at-Risk applications. Such extensions would provide a broader assessment of the practical usefulness of SV models in financial decision-making.

From a methodological perspective, future work may investigate more flexible SV specifications capable of capturing additional features of financial return dynamics. Possible extensions include models with dynamic factor selection, time-varying factor loadings, regime-switching volatility processes, or heavy-tailed multivariate specifications. These approaches may provide a richer representation of dependence structures and changing market conditions.

The computational aspects of Bayesian inference also remain an important area for further development. As model dimension increases, posterior sampling becomes increasingly challenging, particularly in multivariate FSV settings. Future research may therefore explore strategies for improving sampling efficiency, including alternative parameterizations, interweaving schemes, advanced MCMC algorithms, sequential Monte Carlo methods, variational Bayesian approximations, and parallel computing techniques.

More broadly, SV models continue to provide a valuable framework for studying the interaction between Bayesian inference, latent variable modeling, and financial econometrics. Continued developments in scalable Bayesian computation and high-dimensional volatility modeling are likely to further expand the applicability of these methods in both statistical research and quantitative finance.

Appendix A

Theoretical and Methodological Background

A.1 Markov Chains in Bayesian SV Models

The Bayesian estimation procedures considered in this thesis are based on MCMC methods. In this setting, the sequence of simulated draws generated by the sampler forms a Markov chain whose stationary distribution corresponds to the target posterior distribution. The presentation in this section follows the treatment of Markov chains and MCMC methods in Robert and Casella (2004).

A stochastic process $\{X_t\}_{t\geq 0}$ is called a Markov chain if

$$P(X_{t+1} \mid X_t, X_{t-1}, \ldots, X_0) = P(X_{t+1} \mid X_t), \tag{A.1}$$

meaning that the future evolution of the process depends only on the current state and not on the entire past history. In MCMC estimation, the state of the Markov chain consists of all unknown quantities in the model. For the SV models considered in this thesis, the state vector can be written as $x = (\theta, h_{1:T})$, where θ denotes the model parameters and $h_{1:T}$ the latent volatility states. Within each MCMC iteration, these components are updated in blocks according to their conditional posterior distributions, producing a sequence of states $\{x^{(0)}, x^{(1)}, \ldots, x^{(m)}\}$; and this resulting sequence of draws defines a stochastic process on the state space of the unknown quantities. Under the Gibbs and Metropolis-Hastings update rules used in this thesis, the distribution of the next draw depends on the past only through the current state. Consequently,

$$P(x^{(m+1)} \mid x^{(m)}, x^{(m-1)}, \ldots, x^{(0)}) = P(x^{(m+1)} \mid x^{(m)}), \tag{A.2}$$

so that the simulated sequence forms a Markov chain. The objective of MCMC is to construct the transition mechanism of the chain so that its stationary distribution

coincides with the target posterior distribution.

A distribution π is called stationary if it remains invariant under the transition dynamics of the chain, that is,

$$\pi(A) = \int K(x, A)\pi(dx), \tag{A.3}$$

where $K(x, A)$ denotes the transition kernel, which describes the probability of moving from the current state x to a set of states A in one iteration of the chain.

Under standard ergodicity conditions, including irreducibility, aperiodicity, and positive recurrence, the distribution of the Markov chain converges to its stationary distribution as the number of iterations increases. Consequently, posterior expectations can be approximated using Monte Carlo averages computed from the simulated draws.

In finite samples, however, the Markov chain typically undergoes an initial transition phase during which the influence of the starting values gradually diminishes and the simulated draws become increasingly representative of the target posterior distribution. For this reason, an initial burn-in period is discarded before conducting posterior inference. Even after convergence, however, MCMC draws are typically serially correlated, so convergence diagnostics and autocorrelation measures remain important for evaluating the quality of the posterior sample.

These issues are particularly relevant in SV and FSV models, where strong dependence between latent volatility states and persistence parameters may lead to slow mixing and high autocorrelation in the simulated chains.

A.2 Linear Gaussian State-Space Models

State-space models provide a general probabilistic framework for representing dynamic systems with latent variables. The SV models considered in this thesis can also be formulated within the state-space framework. However, unlike the linear Gaussian model presented below, the SV observation equation is nonlinear in the latent volatility and the resulting observation density is non-Gaussian. Consequently, standard Kalman filtering methods cannot be applied directly. Nevertheless, the linear Gaussian case provides useful notation and serves as the foundation for the filtering and simulation-smoothing methods employed after suitable data augmentation. The following review of linear Gaussian state-space models closely follows Durbin and Koopman (2001).

Let y_t denote the observed variable and let α_t denote a latent state vector. A general linear Gaussian state-space model can be written as

$$y_t = Z_t \alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, H_t), \tag{A.4}$$

$$\alpha_t = T_t \alpha_{t-1} + R_t \eta_t, \qquad \eta_t \sim N(0, Q_t), \tag{A.5}$$

where:

- Z_t is the observation matrix,
- T_t is the state transition matrix,
- R_t maps the state innovations into the latent state dynamics,
- H_t and Q_t denote covariance matrices for observation and state disturbances.

Conditional on the latent states, the observation equation is Gaussian, while the latent states evolve according to a first-order Markov process. The linear Gaussian structure implies that the full system remains jointly Gaussian, which allows likelihood evaluation and latent-state inference to be performed efficiently using recursive filtering and smoothing algorithms. This setting differs from the SV models studied in this thesis. In particular, the SV observation equation

$$y_t = \exp\left(\frac{h_t}{2}\right) \varepsilon_t \tag{A.6}$$

is nonlinear in the latent state h_t and the resulting return distribution is non-Gaussian after integrating over the latent volatility process. Consequently, standard Kalman filtering methods cannot be applied directly. Nevertheless, the linear Gaussian framework provides the conceptual foundation for the simulation-smoothing methods used after suitable data augmentation and Gaussian mixture approximations.

Although the original SV model falls outside the linear Gaussian framework, the Kalman filter remains central because it can be applied to conditionally Gaussian representations obtained through data augmentation. For this reason, it is useful to briefly review the main ideas underlying Kalman filtering and smoothing. In particular, the Kalman filter provides a recursive procedure for computing the conditional distribution $f(\alpha_t \mid y_{1:t})$, by sequentially updating predictions of the latent states as new observations become available. The filtering recursions consist of:

- A prediction step based on the state transition equation,
- An updating step using the new observation y_t

In linear Gaussian systems, the filtering distribution remains Gaussian at every time index and is therefore fully characterized by its conditional mean and covariance matrix.

In addition to filtering, smoothing methods are often required to sample or estimate the entire latent state trajectory conditional on all observations, $f(\alpha_{1:T} \mid y_{1:T})$. commonly used approach is the Forward Filtering Backward Sampling (FFBS) algorithm, originally developed by Frühwirth-Schnatter (1994) and Carter and Kohn (1994). FFBS combines Kalman filtering recursions with backward simulation in order to jointly sample the complete latent trajectory. This procedure substantially improves computational efficiency in highly persistent latent-variable models relative to single-state updating methods.

As mentioned above, SV models considered in this thesis are not directly linear Gaussian because the observation equation y_t introduces nonlinear dependence through the exponential volatility term. In addition, after logarithmic transformation, the disturbance term follows a non-Gaussian log-chi-square distribution.

To overcome this issue, the estimation procedure employs the auxiliary mixture-of-normals approximation of Kim et al. (1998), under which the transformed observation equation becomes conditionally Gaussian given latent mixture indicators. Conditional on these auxiliary variables, the SV model admits a linear Gaussian state-space representation. As a result, this conditional Gaussian representation allows simulation smoothing methods such as FFBS to be applied efficiently within the Bayesian MCMC estimation procedure.

Appendix B

Technical Derivations

B.1 Bayes Estimators under Alternative Loss Functions

This appendix derives the Bayes estimators associated with several commonly used loss functions. The following derivations follow standard Bayesian decision theory as presented in Robert (2001).

Let θ denote the unknown parameter and q a decision rule. The Bayes estimator is defined as the minimizer of the posterior risk

$$R(q) = \mathbb{E}_\pi \left[\ell(\theta, q) \mid y_{1:T} \right]. \tag{B.1}$$

B.1.1 Quadratic Loss

Consider the quadratic loss function

$$\ell(\theta, q) = (\theta - q)^2. \tag{B.2}$$

The posterior risk is

$$R(q) = \mathbb{E}_\pi \left[(\theta - q)^2 \mid y_{1:T} \right]. \tag{B.3}$$

Expanding the square gives

$$R(q) = \mathbb{E}_\pi \left[(\theta - \mathbb{E}[\theta \mid y_{1:T}] + \mathbb{E}[\theta \mid y_{1:T}] - q)^2 \mid y_{1:T} \right] \tag{B.4}$$

$$= \mathrm{Var}(\theta \mid y_{1:T}) + \left(\mathbb{E}[\theta \mid y_{1:T}] - q \right)^2. \tag{B.5}$$

Since the posterior variance does not depend on q, the posterior risk is minimized when

$$q^* = \mathbb{E}[\theta \mid y_{1:T}]. \tag{B.6}$$

Therefore, the Bayes estimator under quadratic loss is the posterior mean.

B.1.2 Absolute Loss

Consider the absolute loss function

$$\ell(\theta, q) = |\theta - q|. \tag{B.7}$$

The posterior risk is

$$R(q) = \int |\theta - q| \pi(\theta \mid y_{1:T}) d\theta. \tag{B.8}$$

Since

$$|\theta - q| = \begin{cases} q - \theta, & \theta < q, \\ \theta - q, & \theta \geq q, \end{cases} \tag{B.9}$$

the posterior risk can be written as

$$R(q) = \int_{-\infty}^{q} (q - \theta) \pi(\theta \mid y_{1:T}) d\theta + \int_{q}^{\infty} (\theta - q) \pi(\theta \mid y_{1:T})\, d\theta. \tag{B.10}$$

Differentiating with respect to q and applying Leibniz's rule gives

$$\begin{aligned} R'(q) &= \int_{-\infty}^{q} \pi(\theta \mid y_{1:T}) d\theta - \int_{q}^{\infty} \pi(\theta \mid y_{1:T}) d\theta && \text{(B.11)} \\ &= F(q \mid y_{1:T}) - \big[1 - F(q \mid y_{1:T})\big] && \text{(B.12)} \\ &= 2F(q \mid y_{1:T}) - 1, && \text{(B.13)} \end{aligned}$$

where $F(\cdot \mid y_{1:T})$ denotes the posterior cumulative distribution function.

Setting the derivative equal to zero gives

$$F(q^* \mid y_{1:T}) = \frac{1}{2}, \tag{B.14}$$

showing that the Bayes estimator under absolute loss is the posterior median.

B.1.3 Asymmetric Check Loss

Consider the asymmetric check loss

$$\ell_\alpha(\theta, q) = (\alpha - \mathbf{1}\{\theta < q\})(\theta - q), \quad 0 < \alpha < 1. \tag{B.15}$$

The loss can be written equivalently as

$$\ell_\alpha(\theta, q) = \begin{cases} (1-\alpha)(q-\theta), & \theta < q, \\ \alpha(\theta - q), & \theta \geq q. \end{cases} \tag{B.16}$$

The posterior risk is

$$R(q) = \mathbb{E}_\pi \left[\ell_\alpha(\theta, q) \mid y_{1:T} \right]. \tag{B.17}$$

Substituting the piecewise representation yields

$$R(q) = (1-\alpha) \int_{-\infty}^{q} (q-\theta)\pi(\theta \mid y_{1:T}) d\theta + \alpha \int_{q}^{\infty} (\theta - q)\pi(\theta \mid y_{1:T}) d\theta. \tag{B.18}$$

Differentiating with respect to q gives

$$R'(q) = (1-\alpha) \int_{-\infty}^{q} \pi(\theta \mid y_{1:T}) d\theta - \alpha \int_{q}^{\infty} \pi(\theta \mid y_{1:T}) d\theta \tag{B.19}$$

$$= (1-\alpha) F(q \mid y_{1:T}) - \alpha[1 - F(q \mid y_{1:T})] \tag{B.20}$$

$$= F(q \mid y_{1:T}) - \alpha. \tag{B.21}$$

Therefore, the posterior risk is minimized when

$$F(q^* \mid y_{1:T}) = \alpha. \tag{B.22}$$

Consequently,

$$q^* = Q_\alpha(\theta \mid y_{1:T}), \tag{B.23}$$

where $Q_\alpha(\theta \mid y_{1:T})$ denotes the posterior α-quantile. Thus, the Bayes estimator associated with the asymmetric check loss is the posterior quantile of the corresponding order.

B.2 Derivation of the Posterior Predictive Distribution

This appendix derives the posterior predictive distribution used for forecast evaluation and model comparison.

Let $y_{1:T}$ denote the observed data and y_{T+1} a future observation. In Bayesian inference, uncertainty about future outcomes is described by the posterior predictive distribution. Since both the model parameters θ and the future latent volatility state h_{T+1} are unknown, Bayesian prediction requires averaging over their posterior uncertainty. The posterior predictive distribution is therefore defined as

$$f(y_{T+1} \mid y_{1:T}) = \int f(y_{T+1}, \theta, h_{T+1} \mid y_{1:T}) \, d\theta \, dh_{T+1}. \tag{B.24}$$

Applying the chain rule gives

$$f(y_{T+1} \mid y_{1:T}) = \int f(y_{T+1} \mid \theta, h_{T+1}, y_{1:T})\, p(\theta, h_{T+1} \mid y_{1:T})\, d\theta\, dh_{T+1}. \tag{B.25}$$

For state-space models, the observation density at time $T+1$ depends on the past observations only through the model parameters and latent state h_{T+1}. Consequently,

$$f(y_{T+1} \mid \theta, h_{T+1}, y_{1:T}) = f(y_{T+1} \mid \theta, h_{T+1}), \tag{B.26}$$

so that

$$f(y_{T+1} \mid y_{1:T}) = \int f(y_{T+1} \mid \theta, h_{T+1})\, p(\theta, h_{T+1} \mid y_{1:T})\, d\theta\, dh_{T+1}. \tag{B.27}$$

Equation (B.27) shows that the predictive density is obtained by averaging the conditional predictive distribution over the joint posterior distribution of the unknown parameters and future latent volatility.

In practice, the integral in equation (B.27) is analytically intractable for stochastic volatility models and is approximated using posterior draws generated by MCMC. Let $\left\{\theta^{(m)}, h_{T+1}^{(m)}\right\}_{m=1}^{M}$ denote samples from the joint posterior distribution. The posterior predictive density can then be approximated by

$$f(y_{T+1} \mid y_{1:T}) \approx \frac{1}{M} \sum_{m=1}^{M} f\left(y_{T+1} \mid \theta^{(m)}, h_{T+1}^{(m)}\right). \tag{B.28}$$

This Monte Carlo approximation forms the basis for the one-step-ahead predictive densities used throughout the forecasting and model comparison exercises in this thesis.

B.3 Unconditional Moments of the SV Model

This appendix derives several unconditional moments implied by the baseline SV model and provides the theoretical motivation for the moment expansions reported in the main chapters.

Consider the stochastic volatility model

$$y_t = \exp\left(\frac{h_t}{2}\right) \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1), \tag{B.29}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \sigma\eta_t, \quad \eta_t \sim \mathcal{N}(0,1). \tag{B.30}$$

Under the stationarity condition $|\phi| < 1$, the latent log-volatility process has the unconditional distribution

$$h_t \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{1-\phi^2}\right). \tag{B.31}$$

Define

$$V_h = \frac{\sigma^2}{1-\phi^2}. \tag{B.32}$$

B.3.1 Mean and Variance of Returns

Since ε_t is independent of h_t, the unconditional mean can be obtained as

$$\mathbb{E}(y_t) = \mathbb{E}\left[\exp\left(\frac{h_t}{2}\right)\right]\mathbb{E}(\varepsilon_t) = 0 \tag{B.33}$$

Hence, the unconditional mean return is zero.

For the second moment,

$$\mathbb{E}(y_t^2) = \mathbb{E}\left[\exp(h_t)\right]\mathbb{E}(\varepsilon_t^2). \tag{B.34}$$

Since $h_t \sim \mathcal{N}(\mu, V_h)$, $\exp(h_t)$ follows a lognormal distribution. Therefore,

$$\mathbb{E}[\exp(h_t)] = \exp\left(\mu + \frac{V_h}{2}\right), \tag{B.35}$$

which gives

$$\mathbb{E}(y_t^2) = \exp\left(\mu + \frac{V_h}{2}\right). \tag{B.36}$$

Hence,

$$\mathrm{Var}(y_t) = \exp\left(\mu + \frac{V_h}{2}\right). \tag{B.37}$$

B.3.2 Higher-Order Moments

More generally,

$$\mathbb{E}(y_t^{2m}) = \mathbb{E}[\exp(mh_t)]\mathbb{E}(\varepsilon_t^{2m}) \tag{B.38}$$

where $\mathbb{E}(\varepsilon_t^{2m}) = (2m-1)!!$ for a standard Gaussian random variable, and

$$\mathbb{E}[\exp(mh_t)] = \exp\left(m\mu + \frac{m^2 V_h}{2}\right) \tag{B.39}$$

since $h_t \sim \mathcal{N}(\mu, V_h)$. Therefore,

$$\mathbb{E}(y_t^{2m}) = (2m-1)!! \exp\left(m\mu + \frac{m^2 V_h}{2}\right). \tag{B.40}$$

For example,

$$\mathbb{E}(y_t^4) = 3\exp\left(2\mu + 2V_h\right). \tag{B.41}$$

Using $\mathbb{E}(y_t^2) = \exp(\mu + V_h/2)$, the unconditional kurtosis is

$$\kappa = \frac{\mathbb{E}(y_t^4)}{\mathbb{E}(y_t^2)^2} = \frac{3\exp(2\mu + 2V_h)}{\left[\exp\left(\mu + \frac{V_h}{2}\right)\right]^2} = 3\exp(V_h). \tag{B.42}$$

Since $V_h > 0$, it follows that $\kappa > 3$, showing that the SV model generates excess kurtosis even when the innovations ε_t are Gaussian.

B.4 Moment Expansions for Aggregate Returns

This section derives the third- and fourth-order central moments of an aggregate market return constructed as a weighted sum of sector-level returns.

Let $R_{j,t}$ denote the return of sector $j = 1, \ldots, N$ at time t, and let $w_j > 0$ be its portfolio weight with $\sum_{j=1}^{N} w_j = 1$. The market return is defined as

$$R_{M,t} = \sum_{j=1}^{N} w_j R_{j,t}. \tag{B.43}$$

Let $\mu_j = \mathbb{E}[R_{j,t}]$ and $\mu_M = \mathbb{E}[R_{M,t}] = \sum_{j=1}^{N} w_j \mu_j$. The centered variables are defined as

$$X_{j,t} = R_{j,t} - \mu_j, \tag{B.44}$$

$$X_{M,t} = R_{M,t} - \mu_M = \sum_{j=1}^{N} w_j X_{j,t}. \tag{B.45}$$

The third central moment of the market return is

$$\mu_3(M) = \mathbb{E}[X_M^3]. \tag{B.46}$$

Substitute $X_{M,t} = \sum_{j=1}^{N} w_j X_{j,t}$, we obtain

$$\mathbb{E}[X_M^3] = \sum_{j=1}^{N} w_j^3 \mathbb{E}[X_j^3] + 3\sum_{i \neq j} w_i^2 w_j \mathbb{E}[X_i^2 X_j] + 6\sum_{i<j<k} w_i w_j w_k \mathbb{E}[X_i X_j X_k] \tag{B.47}$$

The skewness of the market return is defined as the third standardized moment:

$$\gamma_3(M) = \frac{\mathbb{E}[X_M^3]}{\sigma_M^3}, \tag{B.48}$$

where $\sigma_M^2 = \mathbb{E}[X_M^2]$. As the expansion includes cross-moments such as $\mathbb{E}[X_i^2 X_j]$ and $\mathbb{E}[X_i X_j X_k]$, the skewness of the aggregate return is not a simple weighted average of individual skewness, but is driven by systematic coskewness, which measures how individual sectors co-move with the volatility of the aggregate market.

In addition, the fourth central moment of the market return is

$$\mu_4(M) = \mathbb{E}[X_M^4]. \tag{B.49}$$

Substitute $X_{M,t} = \sum_{j=1}^{N} w_j X_{j,t}$, we obtain

$$\begin{aligned}\mathbb{E}[X_M^4] = &\sum_{j=1}^{N} w_j^4 \mathbb{E}[X_j^4] + 6 \sum_{i<j} w_i^2 w_j^2 \mathbb{E}[X_i^2 X_j^2] + 4 \sum_{i \neq j} w_i^3 w_j \mathbb{E}[X_i^3 X_j] \\ &+ 12 \sum_{i<j<k} w_i^2 w_j w_k \mathbb{E}[X_i^2 X_j X_k] + 24 \sum_{i<j<k<\ell} w_i w_j w_k w_\ell \mathbb{E}[X_i X_j X_k X_\ell].\end{aligned}$$

In the fourth-moment expansion of the aggregate return, the joint tail behavior is governed by the higher-order cross-moments that link extreme realizations across assets. The term $\mathbb{E}[X_i^2 X_j^2]$ captures pairwise tail dependence: it increases when two assets simultaneously experience large movements, particularly in the left tail, and therefore reflects bivariate co-kurtosis. The asymmetric component $\mathbb{E}[X_i^3 X_j]$ represents co-skewness and becomes large when an extreme shock in one series is systematically associated with a directional response in another, generating joint downside asymmetry. Higher-order interactions such as $\mathbb{E}[X_i^2 X_j X_k]$ describe three-way tail co-movements, while the four-way term $\mathbb{E}[X_i X_j X_k X_\ell]$ captures systemic tail events in which multiple assets jointly realize extreme outcomes. These cross-moment terms dominate the behavior of the aggregate fourth moment and explain why the market portfolio exhibits heavier tails and stronger negative skewness than its individual constituents: idiosyncratic skewness diversifies away, whereas joint tail dependence, especially the pairwise and systemic components, remains and amplifies downside risk at the aggregate level.

B.5 Scale-Mixture Representation of the Student-t Distribution

This appendix derives the Gaussian scale-mixture representation of the Student-t distribution used in subsection 2.2.1. The key idea is that a Student-t random variable can be represented as a Gaussian random variable whose variance is itself random. This representation is particularly useful for Bayesian computation because it transforms heavy-tailed observation errors into conditionally Gaussian observations.

The following representation follows the data-augmentation approach commonly used in Bayesian SV models by Kim et al. (1998); Jacquier et al. (2004).

Consider the hierarchical specification

$$\varepsilon \mid \lambda \sim \mathcal{N}(0, \lambda^{-1}), \tag{B.50}$$

$$\lambda \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \tag{B.51}$$

where the Gamma distribution is parameterized in terms of shape and rate parameters. The conditional density of ε given λ is

$$f(\varepsilon \mid \lambda) = \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda \varepsilon^2}{2}\right), \tag{B.52}$$

while the prior density of λ is

$$f(\lambda) = \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma(\nu/2)} \lambda^{\nu/2-1} \exp\left(-\frac{\nu}{2}\lambda\right). \tag{B.53}$$

The marginal density of ε is obtained by integrating out the latent scale variable:

$$f(\varepsilon) = \int_0^\infty f(\varepsilon \mid \lambda)\, f(\lambda)\, d\lambda. \tag{B.54}$$

Substituting the two densities yields

$$f(\varepsilon) = \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\sqrt{2\pi}\Gamma(\nu/2)} \int_0^\infty \lambda^{(\nu+1)/2-1} \exp\left[-\frac{\nu + \varepsilon^2}{2}\lambda\right] d\lambda. \tag{B.55}$$

Using the Gamma integral identity

$$\int_0^\infty x^{a-1} e^{-bx} dx = \frac{\Gamma(a)}{b^a}, \quad a, b > 0, \tag{B.56}$$

with $a = \frac{\nu+1}{2}$ and $b = \frac{\nu+\varepsilon^2}{2}$, gives

$$f(\varepsilon) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{\varepsilon^2}{\nu}\right)^{-(\nu+1)/2}. \tag{B.57}$$

This is precisely the density of a Student-t distribution with ν degrees of freedom $\varepsilon \sim t_\nu(0, 1)$.

Therefore, the Student-t distribution can be represented as a Gaussian scale mixture in which the conditional precision (inverse variance) is governed by a latent Gamma-distributed scale variable. This representation is particularly useful for Bayesian computation because, conditional on λ, the observation equation remains Gaussian, allowing efficient Gibbs-type data augmentation methods to be employed.

Appendix C

Conditional Posterior Distributions

C.1 Mixture Indicators $s_{1:T}$

Under the mixture-of-normals approximation of Kim et al. (1998), auxiliary mixture indicators $s_{1:T}$ are introduced so that the transformed observation equation becomes conditionally Gaussian. Conditional on the latent volatility and model parameters, the mixture indicators are independent across time and admit tractable full conditional distributions, allowing direct Gibbs updates within the MCMC algorithm.

Given latent volatility $h_{1:T}$ and model parameters θ, the full conditional distribution of s_t is multinomial over $\{1, \ldots, M\}$ with probabilities

$$\Pr(s_t = m \mid h_t, y_t, \theta) \propto \pi_m \exp\left\{-\frac{1}{2v_m}\left(\log y_t^2 - h_t - c_m\right)^2\right\}, \qquad m = 1, \ldots, M, \quad \text{(C.1)}$$

where π_m, c_m, and v_m denote the mixture weights, means, and variances of the normal approximation to the $log(\chi_1^2)$ distribution. Normalization over m yields the Gibbs sampling probabilities used to update the mixture indicators.

C.2 Intercept Parameter β

When an intercept parameter β is included in the observation equation, the univariate SV model is given by

$$y_t = \beta + \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1). \qquad \text{(C.2)}$$

Conditional on the latent volatility states $h_{1:T}$, the observation equation implies

$$y_t \mid h_t, \beta \sim \mathcal{N}\left(\beta, e^{h_t}\right) \qquad \text{(C.3)}$$

Assume β is Gaussian prior

$$\beta \sim \mathcal{N}(b_\beta, B_\beta), \tag{C.4}$$

the conditional likelihood is therefore

$$f(y_{1:T} \mid h_{1:T}, \beta) \propto \exp\left[-\frac{1}{2}\sum_{t=1}^{T} e^{-h_t}(y_t - \beta)^2\right] \tag{C.5}$$

Combining the likelihood with the Gaussian prior yields

$$f(\beta \mid h_{1:T}, y_{1:T}) \propto \exp\left[-\frac{1}{2}\left(A_\beta \beta^2 - 2B^*_\beta \beta\right)\right], \tag{C.6}$$

where

$$A_\beta = \sum_{t=1}^{T} e^{-h_t} + \frac{1}{B_\beta}, \tag{C.7}$$

and

$$B^*_\beta = \sum_{t=1}^{T} y_t e^{-h_t} + \frac{b_\beta}{B_\beta}. \tag{C.8}$$

Completing the square implies that the conditional posterior distribution is Gaussian:

$$\beta \mid h_{1:T}, y_{1:T} \sim \mathcal{N}(\bar{\beta}, V_\beta), \tag{C.9}$$

with posterior variance $V_\beta = A_\beta^{-1}$, and posterior mean $V_\beta B^*_\beta$.

Hence, the intercept parameter β admits a closed-form Gibbs update within the MCMC algorithm.

For the multivariate FSV specification,

$$y_t = \beta + \Lambda f_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_{\varepsilon,t}), \quad f_t \sim \mathcal{N}(0, \Sigma_{f,t}), \tag{C.10}$$

where β is an $N \times 1$ vector of intercept parameters.

Conditionally on the factor loadings Λ, latent factors f_t, and the idiosyncratic volatility states, the observation equation implies

$$y_t \mid \beta, \Lambda, f_t, \Sigma_{\varepsilon,t} \sim \mathcal{N}(\beta + \Lambda f_t, \Sigma_{\varepsilon,t}). \tag{C.11}$$

Assume a Gaussian prior $\beta \sim \mathcal{N}(b_\beta, B_\beta)$, the conditional likelihood is therefore

$$f(y_{1:T} \mid f_{1:T}, \Lambda, \Sigma_{\varepsilon,1:T}, \beta) \propto \exp\left[-\frac{1}{2}\sum_{t=1}^{T} (y_t - \Lambda f_t - \beta)' \Sigma_{\varepsilon,t}^{-1} (y_t - \Lambda f_t - \beta)\right]. \tag{C.12}$$

Combining the likelihood with the Gaussian prior yields

$$f(\beta \mid y_{1:T}, f_{1:T}, \Lambda, \Sigma_{\varepsilon,1:T}) \propto \exp\left[-\frac{1}{2}\left(\beta' A_\beta \beta - 2b_\beta^{*'}\beta\right)\right], \tag{C.13}$$

where

$$A_\beta = \sum_{t=1}^{T} \Sigma_{\varepsilon,t}^{-1} + B_\beta^{-1}, \tag{C.14}$$

and

$$b_\beta^* = \sum_{t=1}^{T} \Sigma_{\varepsilon,t}^{-1}\left(y_t - \Lambda f_t\right) + B_\beta^{-1} b_\beta. \tag{C.15}$$

Completing the square implies that the conditional posterior distribution is Gaussian,

$$\beta \mid y_{1:T}, f_{1:T}, \Lambda, \Sigma_{\varepsilon,1:T} \sim \mathcal{N}(\mu_\beta, V_\beta), \tag{C.16}$$

with posterior mean $\mu_\beta = V_\beta b_\beta^*$ and variance $V_\beta = A_\beta^{-1}$.

Hence, for the FSV specification the intercept parameter β also admits a closed-form Gibbs update within the MCMC algorithm.

C.3 Latent Volatility $h_{1:T}$

Conditional on the mixture indicators $s_{1:T}$ and model parameters $\theta = (\beta, \mu, \phi, \sigma_\eta)$, the transformed SV model admits a linear Gaussian state-space representation

$$y_t^* = h_t + \varepsilon_t^*, \quad \varepsilon_t^* \sim \mathcal{N}(0, v_{s_t}^2), \tag{C.17}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2), \tag{C.18}$$

where

$$y_t^* = \log\left((y_t - \beta)^2\right) - m_{s_t}, \tag{C.19}$$

and

$$\varepsilon_t^* = \log(\varepsilon_t^2). \tag{C.20}$$

Since ε_t^* does not follow a Gaussian distribution, the density of $\log(\varepsilon_t^2)$ is approximated using the finite normal mixture of Kim et al. (1998), introducing latent mixture indicators s_t.

The conditional posterior distribution of the latent volatility trajectory is therefore

$$f(h_{1:T} \mid s_{1:T}, \theta, y_{1:T}), \tag{C.21}$$

which is Gaussian conditional on the mixture indicators.

Because the latent states exhibit strong serial dependence, single-state updates lead to inefficient posterior exploration and slow mixing. The latent volatility trajectory is therefore sampled using the FFBS algorithm or equivalent simulation smoothing methods for linear Gaussian state-space systems.

The forward recursion computes the sequence of filtered distributions

$$f(h_t \mid y^*_{1:t}, s_{1:T}, \theta), \tag{C.22}$$

while the backward recursion samples sequentially from

$$f(h_t \mid h_{t+1}, y^*_{1:t}, s_{1:T}, \theta). \tag{C.23}$$

This procedure produces an exact joint draw from the full conditional posterior distribution of $h_{1:T}$.

For the FSV model, the underlying idea is analogous to the univariate SV case. The observation equation is given by

$$y_t = \beta + \Lambda f_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_{\varepsilon,t}), \quad f_t \sim \mathcal{N}(0, \Sigma_{f,t}), \tag{C.24}$$

Conditionally on the latent factors $f_{1:T}$, factor loadings Λ, and intercept vector β, the common factor component can be removed from the observed returns, yielding the idiosyncratic residuals

$$r_t = y_t - \beta - \Lambda f_t = \varepsilon_t. \tag{C.25}$$

Since the idiosyncratic covariance matrix is diagonal, each residual series satisfies

$$r_{i,t} = \exp\left(\frac{h_{it}}{2}\right)\varepsilon_{it}, \quad \varepsilon_{it} \sim \mathcal{N}(0,1), \tag{C.26}$$

independently across i, and has exactly the same form as the observation equation of the univariate SV model.

In addition, the factor volatility processes are treated analogously. Conditional on the current draws of the factor volatility states, the latent factors follow $f_t \sim \mathcal{N}(0, \Sigma_{f,t})$, where $\Sigma_{f,t} = \text{diag}\left(e^{h_{f,1,t}}, \ldots, e^{h_{f,K,t}}\right)$. Since the factor covariance matrix is diagonal, each factor satisfies

$$f_{k,t} \mid h_{f,k,t} \sim \mathcal{N}(0, e^{h_{f,k,t}}) \tag{C.27}$$

Therefore, each factor can be viewed as following a standard univariate SV series and $h_{f,k,t}$ representing the corresponding latent log-volatility process.

Consequently, the Gaussian mixture approximation of Kim et al. (1998) can be applied separately to each idiosyncratic volatility process and each factor volatility process.

C.4 Level Parameters μ

Conditional on the latent volatility states $h_{1:T}$, the persistence parameter ϕ, and the volatility innovation variance σ_η^2, the latent AR(1) state equation is given by

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \tag{C.28}$$

Rearranging the state equation yields

$$h_t - \phi h_{t-1} = (1 - \phi)\mu + \eta_t. \tag{C.29}$$

Define

$$z_t = h_t - \phi h_{t-1}, \tag{C.30}$$

so that

$$z_t = (1 - \phi)\mu + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2). \tag{C.31}$$

The conditional likelihood for μ is therefore

$$f(h_{1:T} \mid \mu, \phi, \sigma_\eta^2) \propto \exp\left(-\frac{1}{2\sigma_\eta^2} \sum_{t=2}^{T} \left[z_t - (1 - \phi)\mu\right]^2\right). \tag{C.32}$$

Combining the likelihood with the Gaussian prior gives

$$f(\mu \mid h_{1:T}, \phi, \sigma_\eta^2) \propto \exp\left(-\frac{1}{2\sigma_\eta^2} \sum_{t=2}^{T} \left[z_t - (1 - \phi)\mu\right]^2\right) \times \exp\left(-\frac{1}{2B_\mu}(\mu - b_\mu)^2\right).$$

Expanding the quadratic terms and collecting terms in μ yields

$$f(\mu \mid \cdot) \propto \exp\left(-\frac{1}{2}\left[A_\mu \mu^2 - 2B_\mu^* \mu\right]\right), \tag{C.33}$$

where

$$A_\mu = \frac{(T-1)(1-\phi)^2}{\sigma_\eta^2} + \frac{1}{B_\mu}, \tag{C.34}$$

and

$$B_\mu^* = \frac{(1-\phi)}{\sigma_\eta^2} \sum_{t=2}^{T} z_t + \frac{b_\mu}{B_\mu}. \tag{C.35}$$

Completing the square implies that the conditional posterior distribution is Gaussian:

$$\mu \mid h_{1:T}, \phi, \sigma_\eta^2 \sim \mathcal{N}(\bar{\mu}, V_\mu), \tag{C.36}$$

with posterior variance $V_\mu = A_\mu^{-1}$, and posterior mean $\bar{\mu} = V_\mu B_\mu^*$.

Hence, the level parameter μ admits a closed-form Gibbs update within the MCMC

algorithm.

In the FSV model, the same derivation applies to the level parameters of both the idiosyncratic and factor volatility processes. Since each volatility process follows an independent AR(1) specification, the corresponding conditional posterior distributions are obtained by applying the univariate SV results to each process separately.

C.5 Persistence and Volatility Parameters (ϕ, σ^2)

Conditional on the latent volatility states $h_{1:T}$ and the level parameter μ, the latent volatility process satisfies

$$h_t - \mu = \phi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \tag{C.37}$$

The conditional likelihood is therefore

$$f(h_{1:T} \mid \mu, \phi, \sigma_\eta^2) \propto (\sigma_\eta^2)^{-(T-1)/2} \tag{C.38}$$

$$\times \exp\left(-\frac{1}{2\sigma_\eta^2} \sum_{t=2}^{T} \left[(h_t - \mu) - \phi(h_{t-1} - \mu)\right]^2\right). \tag{C.39}$$

The persistence parameter is constrained to be stationary $\phi \in (-1, 1)$ with assigned prior

$$\frac{\phi + 1}{2} \sim \text{Beta}(a_\phi, b_\phi), \tag{C.40}$$

which implies

$$f(\phi) \propto \left(\frac{\phi + 1}{2}\right)^{a_\phi - 1} \left(1 - \frac{\phi + 1}{2}\right)^{b_\phi - 1}. \tag{C.41}$$

Similarly, the volatility parameter is assigned the prior

$$\sigma_\eta^2 \sim \text{Gamma}\left(\frac{1}{2}, \frac{1}{2B_\sigma}\right), \tag{C.42}$$

where the Gamma distribution is parameterized in terms of shape and rate.

Combining the likelihood and priors gives the joint conditional posterior

$$p(\phi, \sigma_\eta^2 \mid h_{1:T}, \mu) \propto f(h_{1:T} \mid \mu, \phi, \sigma_\eta^2)\, f(\phi)\, f(\sigma_\eta^2). \tag{C.43}$$

The resulting density does not belong to a standard conjugate family because:

- The persistence parameter ϕ enters nonlinearly through the autoregressive residuals,
- The transformed Beta prior destroys Gaussian conjugacy,
- The support restriction $\phi \in (-1, 1)$ prevents direct Gaussian updating.

Consequently, the posterior simulation for (ϕ, σ_η^2) requires a Metropolis-Hastings update within the Gibbs sampling scheme.

C.6 Factor Loadings Matrix Λ

Consider the observation equation

$$y_t = \beta + \Lambda f_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_{\varepsilon,t}), \quad f_t \sim \mathcal{N}(0, \Sigma_{f,t}). \tag{C.44}$$

Conditional on the latent factors $f_{1:T}$, intercept β, and volatility states $h_{1:T}$, the rows of the factor loading matrix are conditionally independent. Therefore, it is sufficient to derive the conditional posterior distribution for a generic row λ_i.

Let λ_i denote the i-th row of the factor loading matrix Λ. The corresponding observation equation is

$$y_{it} - \beta_i = \lambda_i' f_t + \varepsilon_{it}, \quad \varepsilon it \sim \mathcal{N}(0, e^{h_{it}}). \tag{C.45}$$

Assuming the Gaussian prior $\lambda_i \sim \mathcal{N}(b_\lambda, B_\lambda)$, the conditional likelihood is

$$f(y_{i,1:T} \mid \lambda_i, f_{1:T}, h_{i,1:T}, \beta_i) \propto \exp\left[-\tfrac{1}{2}\sum_{t=1}^{T} e^{-h_{it}}(y_{it} - \beta_i - \lambda_i' f_t)^2\right]. \tag{C.46}$$

Combining the likelihood with the Gaussian prior yields

$$f(\lambda_i \mid y_{i,1:T}, f_{1:T}, h_{i,1:T}, \beta_i) \propto \exp\left[-\tfrac{1}{2}\left(\lambda_i' A_i \lambda_i - 2 d_i' \lambda_i\right)\right], \tag{C.47}$$

where

$$A_i = \sum_{t=1}^{T} e^{-h_{it}} f_t f_t' + B_\lambda^{-1}, \tag{C.48}$$

and

$$d_i = \sum_{t=1}^{T} e^{-h_{it}} f_t (y_{it} - \beta_i) + B_\lambda^{-1} b_\lambda. \tag{C.49}$$

Completing the square implies

$$\lambda_i \mid y_{i,1:T}, f_{1:T}, h_{i,1:T} \sim \mathcal{N}(m_i, V_i), \tag{C.50}$$

with posterior mean $m_i = V_i d_i$ and posterior covariance matrix $V_i = A_i^{-1}$.

Therefore, each row of the factor loading matrix admits a closed-form Gibbs update. The complete loading matrix Λ is obtained by sequentially sampling all rows $\lambda_i, i = 1, \ldots, N$.

Appendix D

Supplementary Results and Additional Analyses

D.1 Posterior Parameter Estimates for Univariate SV Models

This appendix reports posterior summary statistics for the static parameters of all univariate SV specifications. For each return series and model specification, the posterior median together with the corresponding 5% and 95% credible intervals are reported. The parameters include the unconditional mean parameter μ, the persistence parameter ϕ, the volatility innovation parameter σ_η, the degrees-of-freedom parameter ν for heavy-tailed specifications, and the leverage parameter ρ for asymmetric specifications.

	SV	SVtl	SVt	SVl
μ	0.748 [0.536, 0.958]	0.810 [0.636, 1.010]	0.798 [0.545, 1.068]	0.775 [0.591, 0.964]
ϕ	0.959 [0.940, 0.974]	0.955 [0.937, 0.971]	0.974 [0.959, 0.985]	0.953 [0.934, 0.968]
σ_η	0.250 [0.208, 0.303]	0.247 [0.205, 0.293]	0.189 [0.151, 0.237]	0.259 [0.217, 0.305]
ν		13.538 [9.496, 21.351]	11.683 [8.237, 19.551]	
ρ		-0.349 [-0.433, -0.254]		-0.394 [-0.499, -0.295]

Table D.1: Posterior summaries for financial series

	SV	SVtl	SVt	SVl
μ	0.472 [0.252, 0.692]	0.531 [0.328, 0.734]	0.522 [0.251, 0.788]	0.492 [0.311, 0.662]
ϕ	0.962 [0.944, 0.976]	0.956 [0.938, 0.970]	0.977 [0.962, 0.987]	0.955 [0.939, 0.967]
σ_η	0.241 [0.199, 0.292]	0.245 [0.208, 0.290]	0.177 [0.140, 0.223]	0.254 [0.219, 0.295]
ν		14.268 [10.024, 24.068]	10.942 [7.684, 16.709]	
ρ		-0.372 [-0.466, -0.276]		-0.437 [-0.529, -0.342]

Table D.2: Posterior summaries for health series

	SV	SVtl	SVt	SVl
μ	-0.168 [-0.328, -0.010]	-0.081 [-0.264, 0.108]	-0.098 [-0.293, 0.102]	-0.150 [-0.299, -0.005]
ϕ	0.937 [0.907, 0.959]	0.968 [0.948, 0.982]	0.967 [0.945, 0.981]	0.940 [0.912, 0.957]
σ_η	0.281 [0.231, 0.352]	0.171 [0.130, 0.221]	0.187 [0.141, 0.249]	0.265 [0.218, 0.332]
ν		10.131 [7.512, 15.453]	10.115 [7.175, 17.442]	
ρ		-0.325 [-0.420, -0.212]		-0.387 [-0.492, -0.283]

Table D.3: Posterior summaries for industrial series

	SV	SVtl	SVt	SVl
μ	-0.016 [-0.223, 0.199]	0.041 [-0.127, 0.206]	-0.003 [-0.222, 0.238]	0.025 [-0.142, 0.174]
ϕ	0.961 [0.944, 0.975]	0.955 [0.938, 0.968]	0.966 [0.950, 0.978]	0.953 [0.936, 0.967]
σ_η	0.239 [0.200, 0.286]	0.239 [0.199, 0.285]	0.218 [0.179, 0.265]	0.247 [0.210, 0.294]
ν		24.173 [14.612, 53.606]	29.839 [15.244, 58.455]	
ρ		-0.514 [-0.597, -0.422]		-0.579 [-0.662, -0.485]

Table D.4: Posterior summaries for energy series

	SV	SVtl	SVt	SVl
μ	-0.501 [-0.661, -0.333]	-0.441 [-0.607, -0.280]	-0.440 [-0.622, -0.245]	-0.491 [-0.628, -0.352]
ϕ	0.941 [0.912, 0.962]	0.955 [0.932, 0.971]	0.963 [0.940, 0.978]	0.942 [0.915, 0.962]
σ_η	0.268 [0.215, 0.336]	0.216 [0.171, 0.271]	0.197 [0.151, 0.253]	0.259 [0.210, 0.319]
ν		12.830 [8.682, 23.949]	11.452 [7.705, 17.854]	
ρ		-0.328 [-0.425, -0.222]		-0.378 [-0.479, -0.276]

Table D.5: Posterior summaries for material series

	SV	SVtl	SVt	SVl
μ	0.635 [0.386, 0.885]	0.677 [0.400, 1.014]	0.684 [0.351, 1.050]	0.662 [0.457, 0.876]
ϕ	0.976 [0.960, 0.987]	0.982 [0.969, 0.992]	0.988 [0.978, 0.995]	0.973 [0.959, 0.984]
σ_η	0.168 [0.131, 0.218]	0.136 [0.106, 0.173]	0.110 [0.084, 0.147]	0.175 [0.141, 0.223]
ν		10.227 [7.625, 15.080]	9.182 [6.942, 12.938]	
ρ		-0.308 [-0.407, -0.193]		-0.412 [-0.520, -0.292]

Table D.6: Posterior summaries for consumer discretionary series

	SV	SVtl	SVt	SVl
μ	-0.247 [-0.406, -0.074]	-0.211 [-0.358, -0.051]	-0.196 [-0.399, 0.007]	-0.240 [-0.380, -0.105]
ϕ	0.957 [0.930, 0.976]	0.966 [0.944, 0.980]	0.975 [0.959, 0.987]	0.953 [0.924, 0.971]
σ_η	0.199 [0.151, 0.260]	0.161 [0.125, 0.213]	0.138 [0.107, 0.184]	0.198 [0.156, 0.260]
ν		15.368 [9.957, 30.729]	12.319 [8.521, 23.239]	
ρ		-0.375 [-0.473, -0.275]		-0.456 [-0.560, -0.340]

Table D.7: Posterior summaries for consumer staples series

	SV	SVtl	SVt	SVl
μ	0.425 [0.257, 0.608]	0.502 [0.325, 0.661]	0.474 [0.275, 0.680]	0.446 [0.300, 0.590]
ϕ	0.953 [0.930, 0.970]	0.964 [0.944, 0.978]	0.969 [0.951, 0.982]	0.946 [0.925, 0.963]
σ_η	0.237 [0.192, 0.291]	0.183 [0.144, 0.230]	0.180 [0.137, 0.227]	0.236 [0.191, 0.286]
ν		11.843 [8.355, 18.492]	12.132 [8.462, 19.819]	
ρ		-0.345 [-0.441, -0.237]		-0.419 [-0.525, -0.316]

Table D.8: Posterior summaries for technology series

	SV	SVtl	SVt	SVl
μ	-0.192 [-0.342, -0.043]	-0.137 [-0.279, 0.038]	-0.150 [-0.312, 0.023]	-0.172 [-0.321, -0.034]
ϕ	0.937 [0.911, 0.958]	0.946 [0.920, 0.965]	0.956 [0.934, 0.973]	0.939 [0.913, 0.958]
σ_η	0.264 [0.216, 0.321]	0.230 [0.190, 0.279]	0.210 [0.163, 0.262]	0.254 [0.212, 0.303]
ν		17.493 [11.422, 30.611]	15.510 [10.073, 29.348]	
ρ		-0.305 [-0.398, -0.205]		-0.331 [-0.433, -0.217]

Table D.9: Posterior summaries for utilities series

	SV	SVtl	SVt	SVl
μ	0.107 [-0.178, 0.413]	0.205 [-0.065, 0.428]	0.137 [-0.175, 0.459]	0.145 [-0.083, 0.415]
ϕ	0.977 [0.965, 0.987]	0.977 [0.966, 0.985]	0.982 [0.972, 0.990]	0.977 [0.966, 0.985]
σ_η	0.188 [0.155, 0.228]	0.184 [0.153, 0.221]	0.164 [0.131, 0.200]	0.184 [0.155, 0.218]
ν		21.827 [13.548, 42.318]	21.277 [13.614, 37.987]	
ρ		-0.381 [-0.488, -0.261]		-0.425 [-0.526, -0.314]

Table D.10: Posterior summaries for real estate series

	SV	SVtl	SVt	SVl
μ	-0.259 [-0.458, -0.051]	-0.229 [-0.427, -0.016]	-0.216 [-0.444, 0.020]	-0.243 [-0.412, -0.061]
ϕ	0.965 [0.947, 0.979]	0.969 [0.952, 0.981]	0.976 [0.962, 0.986]	0.961 [0.943, 0.975]
σ_η	0.204 [0.166, 0.253]	0.183 [0.146, 0.230]	0.163 [0.126, 0.203]	0.208 [0.172, 0.256]
ν		16.109 [10.402, 32.136]	14.967 [9.417, 28.429]	
ρ		-0.315 [-0.421, -0.190]		-0.355 [-0.473, -0.238]

Table D.11: Posterior summaries for telecommunication series

Bibliography

Aguilar, O. and West, M. (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business and Economic Statistics*, 18(3):338–357.

Bhattacharya, A. and Dunson, D. B. (2011). Sparse Bayesian infinite factor models. *Biometrika*, 98(2):291–306.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the Business and Economic Statistics Section*, pages 177–181.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Brooks, S., Gelman, A., Jones, G. L., and Meng, X.-L. (2011). *Handbook of Markov Chain Monte Carlo*. Chapman and Hall/CRC.

Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553.

Carvalho, C. M., Chang, J., Lucas, J. E., Nevins, J. R., Wang, Q., and West, M. (2008). High-dimensional sparse factor modeling. *Journal of the American Statistical Association*, 103(484):1438–1456.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4):327–335.

Chib, S., Nardari, F., and Shephard, N. (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 134(2):341–371.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2).

Cowles, M. K. and Carlin, B. P. (1996). Markov Chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, 91(434):883–904.

Durbin, J. and Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. Oxford University Press.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4).

Fridman, M. and Harris, L. (1998). A maximum likelihood approach for non-Gaussian stochastic volatility models. *Journal of Business & Economic Statistics*, 16(3):284–291.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer, New York.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC Press, 3 edition.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculating of posterior moments. In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., editors, *Bayesian Statistics 4*, pages 169–193. Oxford University Press.

Geweke, J. and Amisano, G. (2010). Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230.

Gilks, W., Richardson, S., and Spiegelhalter, D. (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378.

Harvey, A., Ruiz, E., and Shephard, N. (1994a). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

Harvey, A. C., Ruiz, E., and Shephard, N. (1994b). Multivariate stochastic variance models. *Review of Economic Studies*, 61(2):247–264.

Harvey, A. C. and Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*, 57(1):97–109.

Hosszejni, D. and Kastner, G. (2021). Modeling univariate and multivariate stochastic volatility in R with stochvol and factorstochvol. *Journal of Statistical Software*, 100(12):1–34.

Jacquier, E., Polson, N. G., and Rossi, P. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business Economic Statistics*, 12(4):371–389.

Jacquier, E., Polson, N. G., and Rossi, P. E. (1995). Models and priors for multivariate stochastic volatility. Working paper, Graduate School of Business, University of Chicago.

Jacquier, E., Polson, N. G., and Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1):185–212.

Kastner, G. (2016). Dealing with stochastic volatility in time series using the R package stochvol. *Journal of Statistical Software*, 69(5):1–30.

Kastner, G. and Frühwirth-Schnatter, S. (2017). Ancillarity-Sufficiency Interweaving Strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics & Data Analysis*, 76:408–423.

Kastner, G., Frühwirth-Schnatter, S., and Lopes, H. F. (2017). Efficient Bayesian inference for multivariate factor stochastic volatility models. *Journal of Computational and Graphical Statistics*, 26(4):905–917.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with arch models. *Review of Economic Studies*, 65(3):361–393.

Lee, K.-M. and Koopman, S. J. (2004). Estimating stochastic volatility models: A comparison of two importance samplers. *Studies in Nonlinear Dynamics & Econometrics*, 8(2):1–23.

Liesenfeld, R. and Jung, R. C. (2000). Stochastic volatility models: Conditional normality versus heavy-tailed distributions. *Journal of Applied Econometrics*, 15(2):137–160.

McCausland, W. J., Miller, S., and Pelletier, D. (2011). Simulation smoothing for state-space models: A computational efficiency analysis. *Computational Statistics & Data Analysis*, 55(1):199–212.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

Nakajima, J. and Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t-distribution. *Computational Statistics & Data Analysis*, 56(11):3690–3704.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2):425–449.

Pitt, M. K. and Shephard, N. (1999). Time varying covariances: A factor stochastic volatility approach (with discussion). *Bayesian Statistics*, 6:547–570.

R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robert, C. P. (2001). *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Springer.

Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer.

Rue, H. (2001). Fast sampling of Gaussian Markov random fields. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):325–338.

Shephard, N. (1994). Comment on "Bayesian analysis of stochastic volatility models" by Jacquier, Polson and Rossi. *Journal of Business & Economic Statistics*, 12(4):406–410.

Taylor, S. J. (1986). *Modelling Financial Time Series*. Wiley, Chichester.

Tierney, L. (1994). Markov Chains for exploring posterior distributions. *The Annals of Statistics*, 22(4):1701–1728.

Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics*, 127(2):165–178.

Yu, Y. and Meng, X.-L. (2011). To center or not to center: That is not the question-an Ancillarity-Sufficiency Interweaving Strategy (ASIS) for boosting MCMC efficiency. *Journal of Computational and Graphical Statistics*, 20(3):531–570.