Model Based Inference of Stochastic Volatility via Iterated Filtering

WEIZHE SUN

Thesis Supervisor: Prof. Edward Ionides
Graduate Student Mentor: Jesse Wheeler

Department of Statistics, University of Michigan

April 2024

Abstract

The Heston stochastic volatility model is one of the most widely studied stochastic volatility models, in which the variance follows a Cox–Ingersoll–Ross process. Estimating this model under the physical measure is challenging, as the likelihood function involves high-dimensional integral. While an approximate analytical solution for the likelihood function exists, the task of maximizing the function remains difficult in practice. Furthermore, these approximate solutions are invalid if any modifications or extensions of the Heston model are considered, such as extending the model to higher dimensions. Being full-information, plug-and-play, and frequentist, iterated filtering algorithms are adopted to estimate the volatility process of the Heston model. We use the S&P500 index as an example, estimating model parameters and their confidence intervals. The results indicate that the estimated volatility of the S&P500 index matches the pattern of the VIX index. An application in options pricing is also given. We then demonstrate the benefit of iterated-filtering methods by extending to a multi-dimensional panel of Heston models, estimating the volatility processes of four emerging market indices. The results illustrate that the volatility processes of these emerging market indices may share the same rate of reversion and the same sensitivity to their corresponding price processes with a 95% confidence level.

Keywords: sequential Monte Carlo, iterated filtering, stochastic volatility

An honors thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Science (Honors Statistics) at the University of Michigan, 2024. Code files and data sets corresponding to this thesis are available at https://github.com/wzsun21/Honor_Thesis.

Contents

1. Introduction

Volatility measures the variation of the trading price of an underlying asset. It plays a crucial role across various financial domains, including risk management, portfolio optimization, and options pricing. Black and Scholes firstly proposed their famous Black-Scholes model (1973), in which they assumed the asset price follows a diffusion process with constant volatility. They embedded the option price in a corresponding parabolic partial differential equation (PDE). This model is widely adopted in the financial industry due to the simplicity of the analytical solution of the PDE and the relative ease of calibrating the model to data. One common critique of the Black-Scholes model, however, is that the assumption of constant volatility in stock price is inconsistent with real-world phenomena. This deficiency leads to the consideration of stochastic volatility models. Modeling stochastic volatility was largely based on the Ornstein-Uhlenbeck (OU) process proposed by Uhlenbeck and Ornstein (1930), which is a Gaussian-Markov model that allows mean reversion. Vasicek (1977); Cox, Ingersoll, and Ross (1985); and Hull and White (1987) modeled the interest rate as mean reverting OU processes, and later these models were widely generalized to model stochastic volatility. Scott (1987) modeled the volatility of stock price as a generalization of the Vasicek model, and Heston (1993) modeled the volatility as a generalization of Cox-Ingersoll-Ross model (CIR process). In the Heston model, the asymmetric effect is incorporated by allowing a correlation between the Wiener processes of the stock price and volatility. Due to the important role stochastic volatility plays in financial models, other stochastic volatility models have also been developed. Examples include the model proposed by Barndorff-Nielsen and Shephard (2001), in which the volatility dynamic as an OU process driven by a positive Levy process without the Gaussian component, and the model proposed by Asai and McAleer (2011), in which alternative asymmetric effects are introduced based on leverage and size effects. More recent examples include the models proposed by Jiao et al. (2021) and He and Chen (2020). The Alpha-Heston model proposed by Jiao et al. addresses a key limitation of the Heston model, which struggles to accurately capture extreme volatilities during crises, while the model introduced by He and Chen extends the Heston model by incorporating a stochastic long-term mean reversion level of volatility.

Many techniques have been applied to calibrate the parameters of stochastic volatility models under the risk-neutral measure. The most straightforward method is to minimize the loss function with respect to the options prices in the market. There are typically two types of loss functions, which are squared error (SE) and implied volatility squared error (IVSE). For example, Bakshi, Cao, and Chen (1997) employed the sum of squared error (SSE) as the loss function, and Bams, Lehnert, and Wolff (2009) compared the uncertainty of pricing errors by using relative mean squared error (RMSE) loss function and implied volatility relative mean squared error (IVRMSE). Kilin (2007) and Christoffersen et al. (2009) proposed methods to speed up the loss function minimization. Storn and Price (1997) proposed a more robust and reliable global optimization algorithm, known as the differential evolution algorithm, and this is applied in the Heston model by Vollrath and Wendland (2009). Even though these improvements can make optimization more reliable and efficient, loss function methods require the existence of a liquid and efficient derivative market, otherwise, the market prices of derivatives may not reflect a fair price. Also, the loss-minimizing methods may

not be statistically efficient, meaning some information may be lost at the expense of focusing on a single loss criterion. Since derivatives are priced under the risk-neutral probability measure, these calibrated parameters cannot be used to characterize the market under the physical measure, and they can only be used to price contingent claims.

In contrast, estimating the parameters and volatility process using historical asset prices is a more challenging problem, as we cannot observe volatility and model parameters directly from the data. However, estimating volatility is an important task and has wide applications not only in equity derivatives pricing but also in risk control and portfolio optimization. Early research focused on various methods, such as moment matching (Taylor, 2007; Andersen and Sørensen, 1996; Andersen et al., 1999) and quasi-likelihood methods by linear filtering (Ruiz, 1994; Harvey and Shephard, 1996). Later, research efforts focused on estimating the volatility with numerical Bayesian methods, including Markov chain Monte Carlo (MCMC) (examples include Jacquier et al., 1994; Chib et al., 2002 and 2006; Cappuccio et al., 2004; Nugroho and Morimoto, 2014). Recently, the Heston model was estimated using MCMC (Gruszka and Szwabiński, 2023). Another numerical Bayesian method extensively adopted in recent years is Particle Markov chain Monte Carlo (PMCMC) (Andrieu et al., 2010), which incorporated particle filtering into MCMC algorithms. PMCMC was also used in estimating Value-at-Risk (VaR) based on the stochastic volatility model (Yang et al., 2017). Based on PMCMC, Chopin et al. (2013) proposed the SMC 2 algorithm and tested it in Lévy-driven stochastic volatility models. However, Bayesian methods may be problematic due to the need to specify prior distributions, which can be difficult for stochastic volatility models because of the correlations between model parameters. Another perspective is to analytically solve the problem instead of using numerical simulations. Atiya and Wall (2009) derived the analytical approximation of the likelihood function for the Heston model conditioned on the observed time series of stock prices. While optimizing over this analytical likelihood approximation may be efficient, it does not permit any modifications or extensions of the model, for instance, considering a higher-dimensional version of the Heston model that considers various stock prices simultaneously. Also, the derivation of the analytical likelihood function involves approximating techniques like Taylor expansion, which might be problematic in practice. Aside from the methods we mentioned above, recently some new methods for volatility estimation based on the Heston model (or extended Heston model) were proposed, including polynomial filtering (Cacace et al., 2019) and trinomial tree method (Clayton, 2020), each suffering from similar flexibility issues.

In this thesis, we estimate the one-dimensional Heston model by iterated filtering, which is a relatively new method. First proposed by Ionides et al. (Ionides et al., 2006, 2011), it is an efficient method to search for a maximum likelihood estimator of state space models. It was improved by Ionides et al. (2015), and the new algorithm is known as IF2. This algorithm is based on iterations of particle filters (Arulampalam et al., 2002) which includes a random walk for model parameters that in theory converges to the maximum likelihood estimate. The iterated filter has been applied in estimating stochastic dynamics in finance: Bretó (2014) employed IF1 in the stochastic leverage model and Szczepocki (2020) employed IF2 in the Barndoff-Nielsen and Shephard (BN-S) stochastic volatility model proposed in 2001. To our knowledge, the work in this thesis contains the first com-

plete analysis of a stock index using the Heston model and iterated filtering. We then demonstrate the usefulness of using iterated filtering for the Heston model by performing a similar analysis on a collection of stock indices with a high-dimensional extension of the Heston model via panel Iterated Filter (PIF) proposed by Bretó et al. (2020). This is a very new method, which is still in active development. Moreover, While Bretó (2014) and Szczepocki (2020) considered the sliced likelihood in their research, they did not provide confidence intervals for parameter estimates. In this thesis, we perform the profile likelihood search, based on which we construct confidence intervals for each parameter in the model.

2. Heston Model Background

Heston (1993) proposed the following stochastic volatility model to describe the dynamics of asset prices:

$$dS_t = \mu S_t\, dt + \sqrt{\nu_t} S_t\, dW_t^{s,\mathbb{P}} \tag{1}$$

$$d\nu_t = \kappa(\theta - \nu_t)\, dt + \xi\sqrt{\nu_t}\, dW_t^{\nu,\mathbb{P}}, \tag{2}$$

where S_t is the asset price at time t, μ is the asset return on average, $\sqrt{\nu_t}$ is the volatility (standard deviation) of the underlying asset, κ is the mean rate of reversion in variance, θ is the mean reversion level for the variance, ξ is the volatility coefficient of volatility (standard deviation), W_t^s is a Wiener process in the asset price, W_t^ν is a Wiener process in the volatility, and the correlation between W_t^s and W_t^ν is defined as ρ, i.e., $\rho\, dt = \mathbb{E}^{\mathbb{P}}[dW_t^{s,\mathbb{P}} dW_t^{\nu,\mathbb{P}}]$.

The dynamics (1) and (2) are defined under the physical probability measure $\mathbb{P}$ in continuous time. The volatility should be non-negative. If $2\kappa\theta \geq \xi^2$, the volatility process starting from a positive starting point stays strictly positive, which is known as Feller's condition (Cox et al., 1985). In particular, we only have financial data at discretized time points, which requires a discretized version of the Heston model. We can integrate (1) and (2) from t to $t + dt$ to obtain (3) and (4):

$$S_{t+dt} = S_t + \int_t^{t+dt} \mu S_t\, dt + \int_t^{t+dt} \sqrt{\nu_t} S_t\, dW_t^{s,\mathbb{P}} \tag{3}$$

$$\nu_{t+dt} = \nu_t + \int_t^{t+dt} \kappa(\theta - \nu_t)\, dt + \int_t^{t+dt} \xi\sqrt{\nu_t}\, dW_t^{\nu,\mathbb{P}}. \tag{4}$$

There are several ways to approximate the integrals. The simplest method is to Euler scheme, (i.e., evaluate the integrals using the left-point rule), described in Definition 1.

Definition 1. *The discretized Heston model based on Euler scheme is*

$$\tilde{S}_{t+dt} = \tilde{S}_t + \mu\tilde{S}_t\, dt + \sqrt{\tilde{\nu}_t}\tilde{S}_t W_{dt}^1; \tag{5}$$

$$\tilde{\nu}_{t+dt} = \tilde{\nu}_t + \kappa(\theta - \tilde{\nu}_t)\, dt + \xi\sqrt{\tilde{\nu}_t} W_{dt}^2, \tag{6}$$

where $W_{dt}^1 \backsim \mathcal{N}(0, dt), W_{dt}^2 \backsim \mathcal{N}(0, dt)$ *are normal random walk with correlation* ρ.

Here, $\tilde{S}$ represents the discretized process S. Let Δt be the time step between two discretized time points. As Δt goes to 0, $\tilde{S} \to S$. Under the discretized Heston model with Euler scheme, the asset price satisfies

$$\tilde{S}_{t+\Delta t} = \tilde{S}_t \exp\left(\left(\mu - \frac{1}{2}\nu_t\right) dt + \sqrt{\nu_t} W_{\Delta t}^1\right),$$

where $W_{\Delta t}^1 \backsim \mathcal{N}(0, \Delta t)$. The derivation is shown below:

Derivation. Define $Z_t := \ln S_t$. Apply Ito's lemma to equation (1),

$$\begin{aligned} d\ln(S_t) = dZ_t &= \left[\frac{\mu \cdot S_t}{S_t} + \frac{1}{2}\nu_t S_t^2 \cdot (-S_t^{-2})\right] dt + \frac{\sqrt{\nu} S_t}{S_t} dW_t^1 \\ &= \left(\mu - \frac{1}{2}\nu_t\right) dt + \sqrt{\nu_t}\, dW_t^1. \end{aligned} \tag{7}$$

Write (7) in integral form:

$$\ln(S_{t+\Delta t}) = \ln(S_t) + \int_t^{t+\Delta t} \left(\mu - \frac{1}{2}\nu_t\right) dt + \int_t^{t+\Delta t} \sqrt{\nu_t}\, dW_t^1. \tag{8}$$

Discretize (8) using Euler's scheme:

$$\ln(\tilde{S}_{t+\Delta t}) = \ln(\tilde{S}_t) + \left(\mu - \frac{1}{2}\nu_t\right)\Delta t + \sqrt{\nu_t}(W_{t+\Delta t}^1 - W_t^1). \tag{9}$$

Define $W_{\Delta t}^1 := W_{t+\Delta t}^1 - W_t^1$. Since W_t^1 is a standard Wiener process, $W_{\Delta t}^1 \backsim \mathcal{N}(0, \Delta t)$. Then (9) becomes

$$\ln\left(\frac{\tilde{S}_{t+\Delta t}}{\tilde{S}_t}\right) = \left(\mu - \frac{1}{2}\nu_t\right)\Delta t + \sqrt{\nu_t} W_{\Delta t}^1 \tag{10}$$

and taking exponential on both sides will finish the derivation.

Now we define the discretized log return $\tilde{R}_t := \ln\left(\frac{\tilde{S}_{t+\Delta t}}{\tilde{S}_t}\right)$. By (10), we can write the original $(\tilde{S}_t, \tilde{\nu}_t)$ system (5 and 6) as $(\tilde{R}_t, \tilde{\nu}_t)$ system:

$$\tilde{R}_{t+1} = \left(\mu - \frac{1}{2}\tilde{\nu}_t\right)\Delta t + \sqrt{\tilde{\nu}_t \Delta t}\epsilon_1, \tag{11}$$

$$\tilde{\nu}_{t+1} = [\kappa\theta + (1-\kappa)\tilde{\nu}_t]\Delta t + \xi\sqrt{\tilde{\nu}_t \Delta t}\epsilon_2, \tag{12}$$

where ϵ_1 and ϵ_2 are two standard normal variables with correlation ρ. For sampling purposes, if we want to generate $\epsilon_1, \epsilon_2 \backsim \mathcal{N}(0, \Delta t)$ with correlation ρ, we only need to generate two independent standard normal variables u_1, u_2 and let $\epsilon_1 = u_1$ and $\epsilon_2 = \rho u_1 + \sqrt{1-\rho^2} u_2$. In other words, (11)

and (12) can be written as

$$\tilde{R}_{t+1} = \left(\mu - \frac{1}{2}\tilde{\nu}_t\right)\Delta t + \sqrt{\tilde{\nu}_t \Delta t}\, u_1 \tag{13}$$

$$\tilde{\nu}_{t+1} = [\kappa\theta + (1-\kappa)\tilde{\nu}_t]\Delta t + \xi\sqrt{\tilde{\nu}_t \Delta t}(\rho u_1 + \sqrt{1-\rho^2}u_2) \tag{14}$$

where u_1 and u_2 are two independent standard random variables.

Directly from equation (10), we can derive the following lemma:

Lemma 2. *The discretized log return defined by* $\tilde{R}_t := \ln\left(\frac{\tilde{S}_{t+\Delta t}}{\tilde{S}_t}\right)$ *conditionally follows the normal distribution below:*

$$\tilde{R}_t | \nu_t = \tilde{\nu}_t \backsim \mathcal{N}\left(\left(\mu - \frac{1}{2}\tilde{\nu}_t\right)\Delta t, \tilde{\nu}_t \Delta t\right).$$

The proof is directly from (11), in which the only undetermined part is $\sqrt{\nu_t} W^1_{\Delta_t}$ which is a normal random variable itself conditioning on ν_t.

Notice that everything above is based on the physical measure $\mathbb{P}$. The parameters under the physical measure can be used to characterize the market, but for pricing purposes we need processes under the risk-neutral probability $\mathbb{Q}$, assuming the market is arbitrage-free. Here is a brief explanation of the risk-neutral probability measure: Consider the stochastic discount factor (SDF, or pricing kernel) model of asset pricing $P_t = \mathbb{E}_t[M_{t+1}P_{t+1}]$, where M is the SDF and P is the price of the asset. The SDF takes both time value of money and risks of future cash flows into consideration, and is always stochastic. The intuition of the risk-neutral probability measure $\mathbb{Q}$ is to incorporate the SDF into the expectation by changing the probability measure from the physical measure $\mathbb{P}$ to $\mathbb{Q}$, i.e.,

$$P_t = \mathbb{E}^{\mathbb{P}}_t[M_{t+1}P_{t+1}] = e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}[P_{t+1}].$$

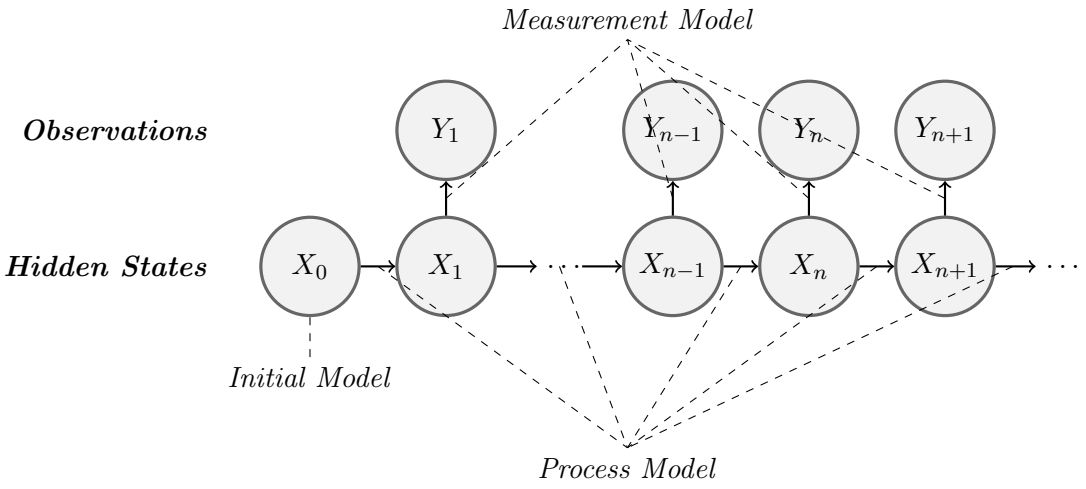
The benefit is, under $\mathbb{Q}$ we can directly price the claim without specifying the SDF. The risk-neutral probability measure $\mathbb{Q}$ is corresponding to a world in which all investors are risk-neutral, i.e., they will not require to be compensated for taking risks, and we can simply use the risk-free rate r to discount the future risky cash flows. In other words, if we define B_t as the price of the risk-free asset at time t, then $\frac{P_t}{B_t}$ is a martingale under $\mathbb{Q}$. Mathematically, SDF is the discounted Radon-Nykodim derivative: $M_t = e^{-r\Delta t}\frac{d\mathbb{Q}}{d\mathbb{P}}$.

3. One-Dimensional Example with IF2

3.1 Partially Observed Markov (POMP) Models

A POMP model consists of a noisy and incomplete observed process conditionally independent of a latent Markov process (which is unobserved) (King et al., 2016). Suppose we observe the noisy and incomplete process $Y(t)$ as $y^*_1, \ldots, y^*_n$ at time points $t_1 < \cdots < t_N$. Let $\{X(t) : t > t_0\}$ be the hidden Markov process defined on all continuous time points greater than t_0, conditioning on

which the process $Y(t)$ is independent. We denote the value of the latent process at measurement time t_n as $X_n = X(t_n)$ and write $X_{0:N} = (X_0, \ldots, X_N)$. Since the process $X(t)$ is defined to be Markov, it evolves by the one-step *transition density (or process model)* $f_{X_{n+1}|X_n}(x_{n+1}|x_n;\theta)$ where θ is a vector of parameters specifying the model in the parameter space $\Theta \in \mathbb{R}^p$. We also need the *measurement density (or measure model)* $f_{Y_n|X_n}(y_n|x_n;\theta)$, which describes how observations an the system are made. Here Y_i and Y_j are conditionally independent given the hidden states for any $i \neq j$. Finally, the *initial density (or initialization model)* $f_{X_0}(x_0;\theta)$ is required to initialize the hidden process. The figure below demonstrates the structure of an arbitrary POMP model:



3.2 Particle Filter and Iterated Filter

Given the initial density, the transition density, the measurement density and the model parameter $\theta \in \Theta$, we can specify the joint density of $X_{0:N}$ and $Y_{0:N}$:

$$f_{X_{0:N},Y_{0:N}}(x_{0:N}, y_{0:N};\theta) = \underbrace{f_{X_0}(x_0;\theta)}_{\text{initial density}} \prod_{i=1}^{N} \underbrace{f_{X_n|X_{n-1}}(x_n|x_{n-1};\theta)}_{\text{transition density}} \underbrace{f_{Y_n|X_n}(y_n^*|x_n;\theta)}_{\text{measurement density}}. \tag{15}$$

Based on (15), the marginal density of $Y_{1:N}$ evaluated at the observed time series $y^*_{1:N}$ is

$$\mathcal{L}(\theta) = f_{Y_{1:N}}(y^*_{1:N};\theta) = \int_{\mathbb{R}^{N+1}} f_{X_0}(x_0;\theta) \prod_{i=1}^{N} f_{X_n|X_{n-1}}(x_n|x_{n-1};\theta) f_{Y_n|X_{n-1}}(y_n^*|x_{n-1};\theta)\, dx_{0:N}. \tag{16}$$

The integral in (16) is high-dimensional and is extremely hard to evaluate analytically in the general case. A natural idea is to write (16) in the expectation form and then use Monte Carlo simulations to approximate. The problem is that we need to simulate the whole time series to approximate the expectation, and once the simulated trajectory deviates from the true pattern, it is hard to get back. For long time series, almost all trajectories would deviate from the true pattern, making the Monte Carlo approximation unreliable (Doucet et al., 2001).

Particle filters can be used to obtain a Monte Carlo estimation of the likelihood function using

a different decomposition of the integral. It sequentially estimates the integral with a two-step recursion of the *prediction formula* (18) and the *filtering formula* (19) to prevent the particles from deviating from the true pattern, using normalized measurement density of the observed data given the simulated prediction particles as weights to resample states. (17) to (19) are specified below:

$$\mathcal{L}(\theta) = \prod_{i=1}^{N} \int_{\mathbb{R}} f_{Y_n|X_n}(y_n^*|x_n;\theta) f_{X_n|Y_{1:n-1}}(x_n|y_{1:n-1}^*;\theta)\, dx_n \tag{17}$$

$$f_{X_n|Y_{1:n-1}}(x_n|y_{1:n-1}^*;\theta) = \int_{\mathbb{R}} f_{X_n|X_{n-1}}(x_n|x_{n-1};\theta) f_{X_{n-1}|Y_{1:n-1}}(x_{n-1}|y_{1:n-1}^*;\theta)\, dx_{n-1} \tag{18}$$

$$f_{X_n|Y_{1:n}}(x_n|y_{1:n}^*;\theta) = \frac{f_{Y_n|X_n}(y_n^*|x_n;\theta) f_{X_n|Y_{1:n-1}}(x_n|y_{1:n-1}^*;\theta)}{\int f_{Y_n|X_n}(y_n^*|u_n;\theta) f_{X_n|Y_{1:n-1}(u_n|y_{1:n-1}^*;\theta)}\, du_n}. \tag{19}$$

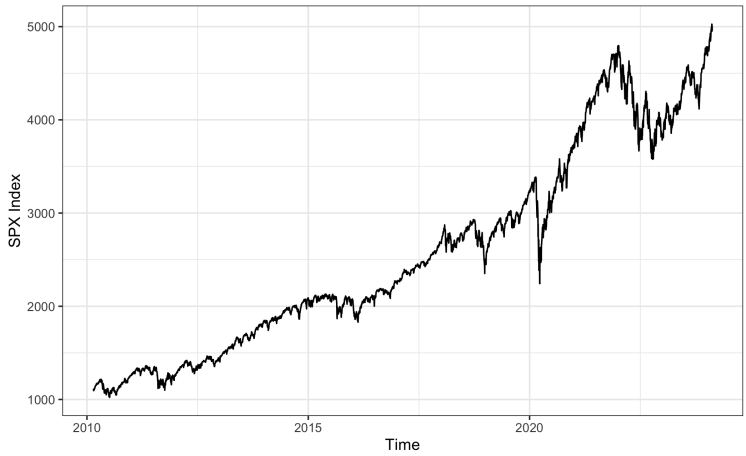
Because the particle filter provides only a stochastic estimation of the model likelihood, black-box numeric optimization algorithms often fail to properly maximize model likelihoods. The particle filter can be extended, however, to efficiently maximize model likelihoods by using techniques known as iterated filtering (Ionides et al., 2006, 2015). Specifically, the IF2 algorithm of Ionides et al. (2015) considers an extended version of the model where model parameters are treated as latent states that perform carefully constructed random walks between time points. Over time, the random walk standard deviation goes to zero, and it can be shown that the parameter vector converges to the maximum likelihood estimator (MLE). Each iteration of IF2 consists of one particle filter with perturbed parameter vectors, and at the end of the time series, the updated parameter vector is recycled to the next iteration. Theoretically, the sequence of estimated parameter vectors converges to the maximum likelihood estimator.

3.3 Data and Model Setup under Market Physical Measure

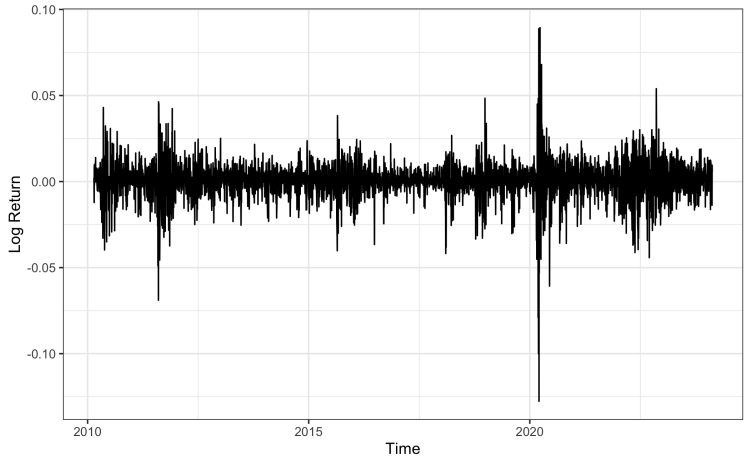
We estimate the Heston models under the physical probability measure and the risk-neutral probability measure for the Standard & Poor's 500 index (S&P500, SPX), which is an index that tracks the performance of 500 of the largest companies listed in stock exchanges in the United States. We used the daily data for the period from Feb 18, 2010, to Feb 16, 2024. The United States has the largest and most financially advanced stock market in the world, making it reasonable to assume it is arbitrage-free and complete. Figure 1 shows the time series plots of the SPX index (a) and log return of SPX index (b) during the time period specified above.

Our estimate under the market physical measure is based on the Euler-discretized Heston model (5, 6). Instead of using the original index data, we fit the model using the log-return time series data of SPX index as the observations to use measurement model based on the solution of geometric Brownian motion (Lemma 3). Here, the discretization Δt is set to 1 to represent 1 trading day. The processes pair (R_t, ν_t) constructs a POMP model, where R_t is the process of SPX log-return and ν_t is the squared-volatility process which is unobservable from the market.

One concern that arises is that the transition of the latent variable from ν_t to ν_{t+1} depends on the observable variable R_t if $\rho \neq 0$, which violates the structure of the POMP model. To address



(a) SPX daily Index



(b) SPX daily log return

Figure 1: Time series plots of the SPX index and its log return. Plot (a) and (b) consist of 3523 and 3522 observations respectively. Plot (a) uses values of the daily adjusted-closing prices of the SPX index. Plot (b) uses the log return values calculated based on the time series of daily adjusted-closing prices of the SPX index. *Data Source: Yahoo Finance*

this concern, we consider an extended model to get rid of the violation: Let $X_n = (R_n, \nu_n)$ be the state variable of the extended POMP model, where R_n is perfectly observed with r_n^*. In this case, there are no more violations of the POMP model structure, but the process R_n has zero measurement error. When the latent state is continuous and there is no measurement error, the basic particle filter fails since all prediction particles have infinite weights. Thus we need a modification of sequential Monte Carlo. For filtered particle j at time $t-1$, the state variable is $X_{j,n-1}^F = (r_{n-1}^*, \nu_{j,n-1})$. The prediction particle j is then

$$\nu_{j,n}^P \backsim f_{\nu_n|\nu_{n-1},R_{n-1}}(v_n|\nu_{j,n-1}^F, r_{n-1}^*) \tag{20}$$

with corresponding resampling weights

$$w_{j,n} = f_{R_n|\nu_n}(r_n^*|\nu_{j,n}^P).$$

Then we can resample the probabilities proportional to the corresponding weights. This gives the Sequential Monte Carlo (SMC) representation of the filtering distribution at time t, as derived by Ionides (2021).

Based on the extended POMP model and extended particle filter, there is no randomness in $\Delta W_{\Delta t}^1$ of equation (5) in filtering given the hidden states, as (20) is conditioned on the observed data (log-returns). From (10), we can obtain

$$W_{\Delta t}^1 = \frac{R_t - \mu + \frac{1}{2}\nu_t}{\sqrt{\nu_t}}.$$

The distribution specified in Lemma 2 is used as the measurement distribution. Since the squared-volatility process is unobservable, we treat the initial squared-volatility ν_0 as a parameter to be estimated. Thus we have 6 parameters to be estimated in total under the physical measure: let γ be the parameter vector, then $\gamma = (\mu, \theta, \kappa, \xi, \rho, \nu_0)$. We perform monotonic transformations to ensure that our parameter estimates are in valid regions of the parameter space. Specifically, θ, κ, ξ and ν_0 are transformed using a log-transformation to enforce $\hat{\theta}, \hat{\kappa}, \hat{\xi}, \hat{\nu}_0 \in \mathbb{R}^+$. Since $\rho \in [-1, 1]$, we transform (and transform back) ρ with

$$f(\rho) = \frac{\rho + 1}{1 - \rho}, \quad f^{-1}(\tilde{\rho}) = -1 + \frac{2}{1 + \exp(-\tilde{\rho})}.$$

Also, we apply Feller's condition when setting up the initial parameter space to reduce the chances of negative squared volatility. Due to the discretization of the latent process, however, Feller's condition alone does not guarantee the process to be positive at all time points. Specifically, based on equation (6) we have

$$\nu_{t+\Delta t} \backsim \mathcal{N}\left(\kappa\theta + (1 - \kappa\Delta t)\nu_t, \xi^2\nu_t\Delta t\right)$$

and thus

$$\mathbb{P}(\nu_{t+\Delta t} < 0) = \Phi\left(\frac{0 - \kappa\theta - (1 - \kappa\Delta t)\nu_t}{\xi\sqrt{\nu_t\Delta t}}\right) = \Phi\left(\frac{-\kappa\theta - (1 - \kappa\Delta t)\nu_t}{\xi\sqrt{\nu_t\Delta t}}\right)$$

where $\Phi(\cdot)$ is the cumulative distribution function of standard normal random variable. We set the squared volatility at that time point to be 1×10^{-32} if it is negative.

3.4 Profile Likelihoods

We conduct a profile likelihood search for each of the parameters. We first profile over μ and obtain its smoothed profile MLE $\mu^* = 3.71 \times 10^{-4}$. Then, when profiling over other parameters, we fix $\mu = \mu^*$. The resulting profile traces are shown in Figure 2, and values of MLEs and confidence intervals based on Monte Carlo Adjusted Profiles (MCAP; Ionides et al., 2017) are demonstrated in Table 1.

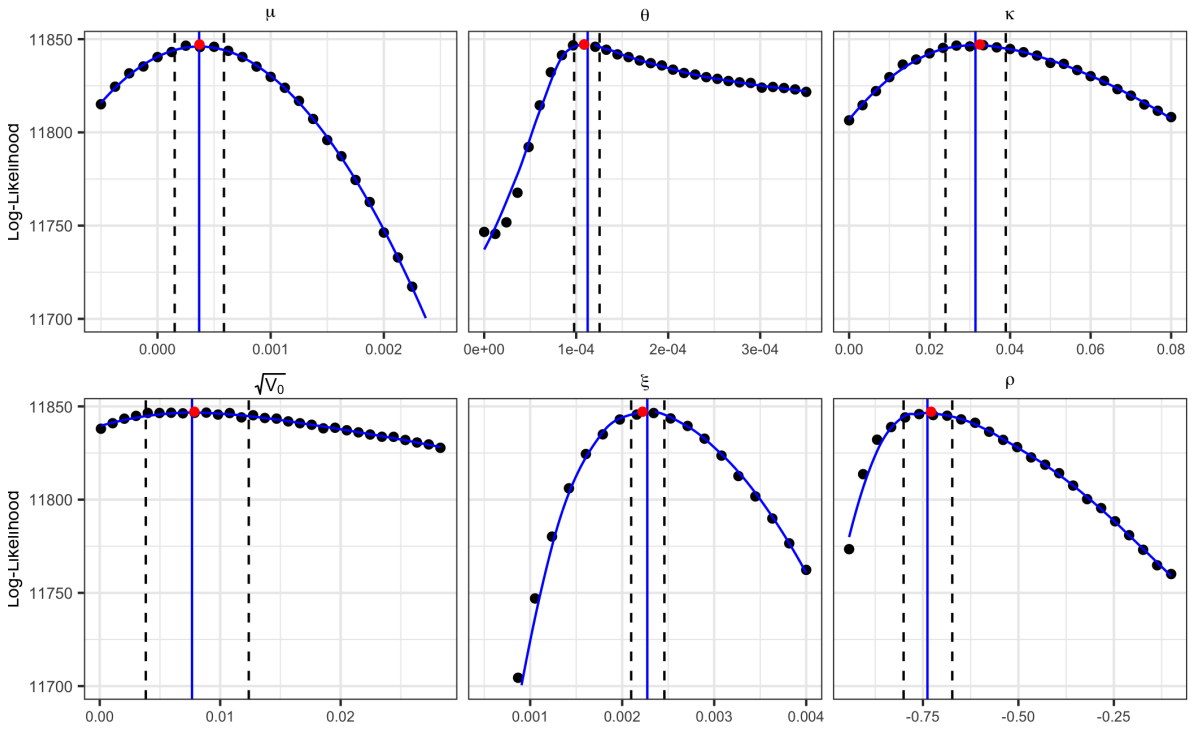


Figure 2: Profile traces for each of the parameters (fixing μ at its smoothed profile MLE), with 200 IF2 iterations and 2500 particles. In each of the plots, the red point is the profile MLE, the blue curve is the fitted loess-smoothed curve based on profile searches, the blue vertical line is the smoothed profile MLE for each parameter, and the interval between two vertical dashed lines is the confidence interval for the parameter. All confidence intervals are calculated using MCAP.

In a standard profile search, parameters that are not being profiled are typically not fixed, as was done here. However, our initial investigations revealed that fixing μ results in an easier optimization problem, resulting in profile likelihood curves that are comparable with those from much heavier computations. Theoretically, the profiles must pass through the same maximum level, so the results obtained while not fixing μ demonstrate that we do not obtain a proper estimate of the profile likelihood (see Appendix A1). In the first plot of Appendix A1, as the blue smoothed curve fits the black points almost perfectly, it is reasonable to fix μ at its MLE to simplify the optimization problem. After we fixed μ at its MLE, all estimates of the profile curves increased (Figure 2). Whether or not these new curves represent the true profile is questionable, but what is certain is that they are better estimates of the profile than the standard approach since they lie above the estimate obtained when μ is not fixed.

To examine the effect of fixing μ, we performed a profile of ρ with 1000 IF2 iterations without fixing μ at its MLE. The result is shown in Appendix A2 (b). Appendix A2 is a good demonstration that our profile is a good approximation of the true profile because our profile trajectory over ρ (Appendix A2 (b)) is very comparable to that with μ unfixed (Appendix A2 (a)). We may refer to the plot in Appendix A3, which shows that 200 iterations may not be sufficient, but the

Table 1: Profile MLE and Confidence Interval for Each Parameter

Parameter	MLE[1]			Confidence Interval[1,3]	
	Profile MLE	Smoothed Profile MLE	Annualized MLE[2]	2.5%	97.5%
μ	3.68×10^{-4}	3.71×10^{-4}	9.27%	1.52×10^{-4} (3.83%)	5.87×10^{-4} (14.8%)
θ	1.12×10^{-4}	1.09×10^{-4}	16.6%	9.78×10^{-5} (15.7%)	1.25×10^{-4} (17.7%)
κ	3.14×10^{-2}	3.25×10^{-2}	8.19	2.39×10^{-2} (6.02)	3.89×10^{-2} (9.80)
$\sqrt{\nu_0}$	7.66×10^{-3}	7.86×10^{-3}	–	3.82×10^{-3}	1.23×10^{-2}
ξ	2.27×10^{-3}	2.22×10^{-3}	3.52%	2.10×10^{-3} (3.33%)	2.46×10^{-3} (3.91%)
ρ	-7.38×10^{-1}	-7.29×10^{-1}	–	-8.01×10^{-1}	-6.73×10^{-1}

Note: when profiling over other parameters, we fix μ at its smoothed profile MLE.
[1] Based on MCAP.
[2] Annualized by 252 trading days of a year based on the smoothed profile MLE.
[3] The numbers in the parentheses are corresponding to the annualized MLE based on the smoothed profile MLE.

log-likelihood converges well after 1000 iterations. However, profiling with much more iterations is extremely computationally exhaustive in this case. Based on the plots in Appendix A2 and A3, we feel fairly confident that the large-computation version (with 1000 IF2 iterations) gives us a good profile, and it seems to match very closely to the fixed-μ profile, so this suggests that fixing μ does not result in much bias for the profile of rho, and it is not unreasonable to suspect that other parameters should have a similar situation. One possible explanation is that the IF2 procedure involves a random walk for μ, which may confuse the dynamic relationship with the random walk in equation (5). Thus, the model converges only when the perturbation of μ becomes negligible. This issue may be solved by an automatic differentiable particle filter (ADPF; Tan, 2023) as it does not introduce random perturbations to the parameters, however, ADPF is still under active research.

The approximate quadratic behavior of the MCAP curve near the marginal MLEs, coupled with the relatively small deviations from the MCAP curve and Monte Carlo evaluations of the profiles, suggest that the IF2 algorithm has converged to the MLE. In essence, we see that we obtain consistent estimates of model parameters, even though each search started from diverse regions of the parameter space and is subject to Monte Carlo variability in both the maximization and evaluation procedures. While these results make us confident that we have maximized the model likelihood, the log-likelihood of a model only provides a relative measure of goodness-of-fit compared to other measures. Therefore to better understand how well the model fits the observed data, we compare this relative measure to a statistical benchmark, as recommended by Wheeler et al.

(2024). In this case, our choice of benchmark models is a generalized autoregressive conditional heteroscedastic (GARCH) model and an exponential general autoregressive conditional heteroskedastic (EGARCH) model, which allows time-varying variance and has previously been used to model stock return and volatility (Alberga et al., 2008; Olowe, 2009; Neokosmidis, 2009). The log-likelihood of the maximized Heston model is 11847.12, which is 146.27 units higher than that of the GARCH(1,1) benchmark model (11700.85) and 88.01 log units higher than that of the EGARCH(1,1) benchmark model (11759.11), demonstrating that our model provides a better statistical fit to the observed data than the GARCH and EGARCH benchmarks.

Table 1 provides some results to characterize the dynamics of the SPX index. We annualize the smoothed profile MLE of μ by 252 (the number of trading days in a year) to give an annualized return of 9.27%. The average annualized return of SPX between Feb 18, 2010, and Feb 16, 2024, is 12.3%, which falls into the 95% confidence interval for μ. The mean rate of reversion, which can be used to calculate the half-life to the mean level of reversion (θ) by half-life $= \frac{1}{\kappa}\ln(2)$(years), is also annualized to give 8.19, indicating that the half-life of the variance (or squared-volatility) is 0.0846 years or equivalently 21.3 trading days. This means that under the smoothed profile MLE of κ, people would expect the variance of the SPX index to travel half the distance toward the mean reversion level of variance (θ) in roughly 1 month. In the past, estimating the rate of mean reversion mostly depends on the GARCH and EGARCH models (Arsalan et al., 2022; Ahmed et al., 2018). Arsalan et al. estimated a half-life of 13.24 days for the NASDAQ index, and Ahmed et al. estimated a half-life of 49.41 days for the S&P500 index and 72.98 days for the NASDAQ index. However, the validity of GARCH and EGARCH models is questionable, because they may not be a perfect fit for the market data. The GARCH and EGARCH models both have far lower log-likelihood than our model concerning the S&P500 historical data, which might be a signal that the GARCH and EGARCH models do not fit the data well. On the other hand, Ahmed et al. found that their GARCH(1,1) model only explained 0.45% of the price variations, indicating that this model may not be good enough.

The smoothed MLE of θ suggests that the long-term mean reversion level of volatility (i.e., square root of variance) is about 16.6%, and the annualized volatility of variance is 3.52% indicated by the smoothed profile MLE of ξ. Moreover, the estimator of ρ (which is the source of asymmetry) is significantly negative, indicating that a negative shock in stock would lead to a positive shock in volatility (and vice versa). There are many existing explanations to the negative correlation, including the leverage ratio effect (Black, 1976; Christie, 1982) and volatility feedback theory (Campbell and Hentschel, 1992; Bekaert and Wu, 2000; Bollerslev et al., 2006; Bae et al., 2007). Figure 3 shows the estimated volatility process. We can see that the filtered mean has a similar pattern to the VIX index, which measures the expected (annualized) implied volatility of SPX options over the next 30 days.

Figure 4 demonstrates the estimated volatility risk premium and variance risk premium, as well as the 95% confidence interval. The volatility risk premium refers to the (typically positive) gap between the implied volatility and realized volatility, reflecting the costs of insuring against equity

volatility fluctuations. The volatility risk premium (as well as variance risk premium) reached an extremely high level in the early period of 2020 when the world market crashed due to the COVID-19 pandemic. Other peaks also represent severe market crashes, including those in May 2010 (2010 flash crash), August 2011 (European debt crisis and US credit rating downgrade), 2018 (cryptocurrency crash), and Jan 2022 (global financial instability due to the unpredictability of COVID-19).

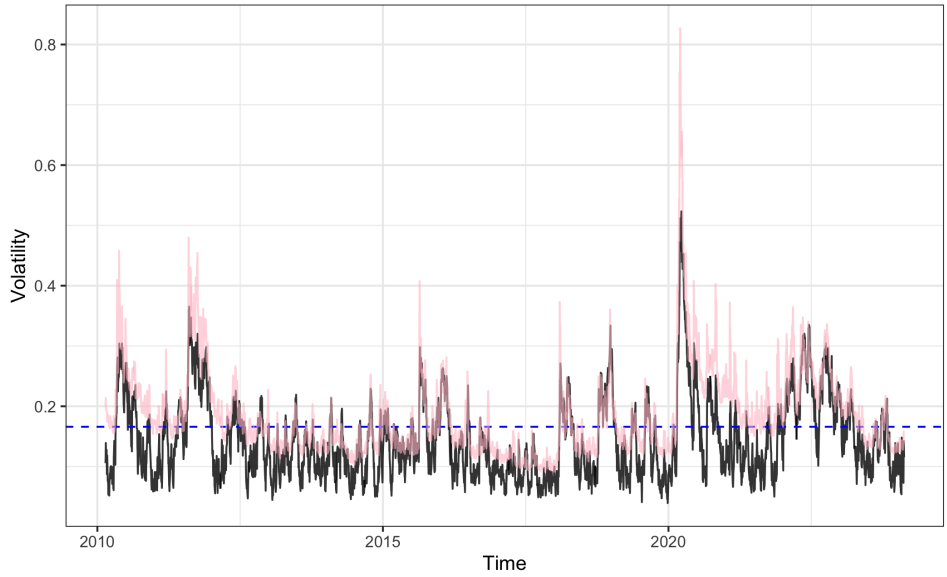
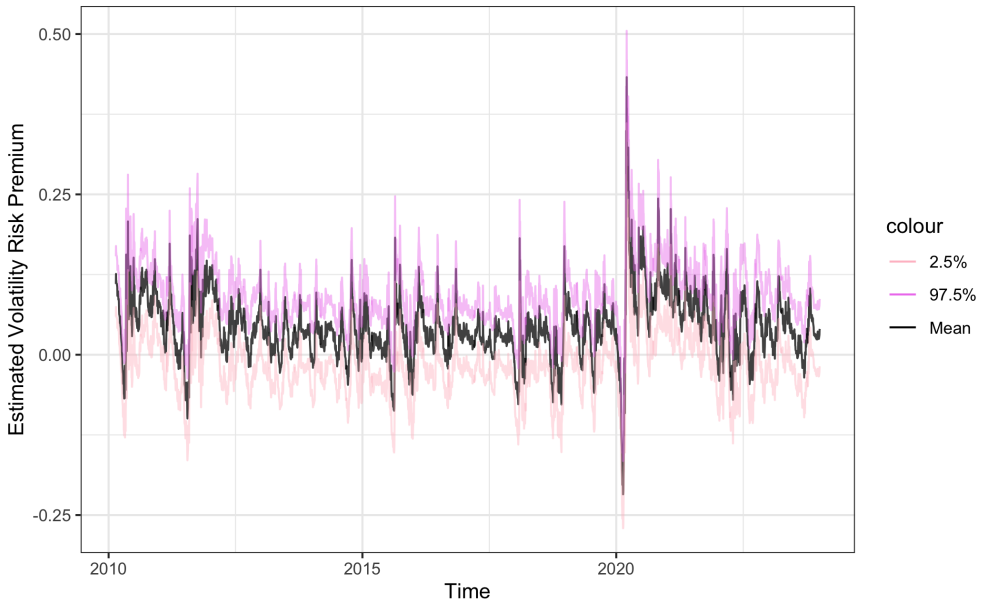


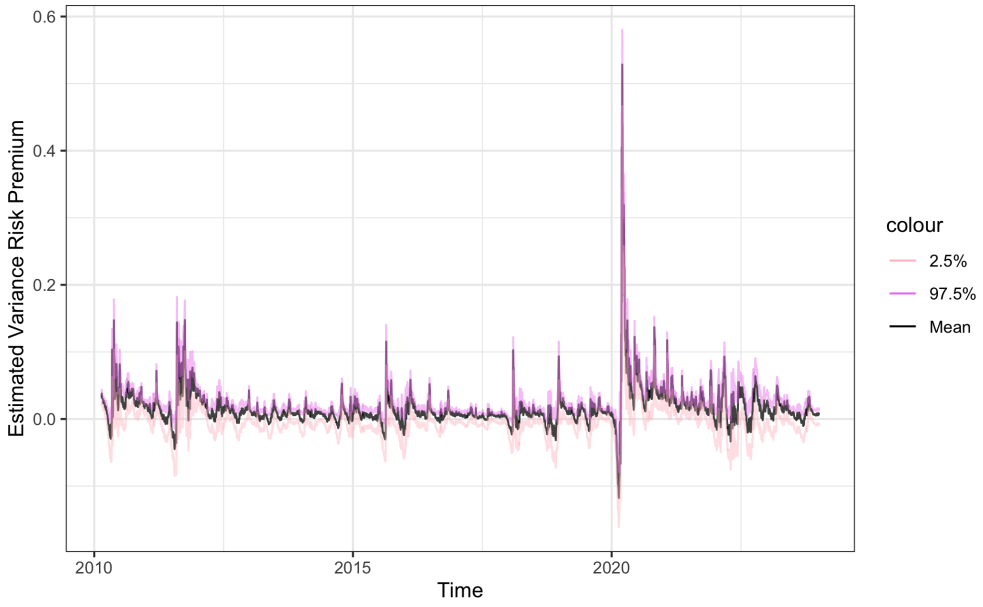
Figure 3: The estimated volatility process (filtering mean with 1000 particles). The blue horizontal dashed line represents the estimated mean reversion level of volatility and the pink process represents the corresponding VIX index as a reference.

3.5 Application: Options Pricing

An option is a financial claim whose payoffs depend on their underlying securities. It is a contract between the seller and buyer, and the buyer has the right to choose if they would exercise the contract. The price of options has to compensate the issuer for the risks they take, and the buyer of the contract has to pay an extra premium for risk reduction. The two primary classes of options are call options and put options. A call option provides the buyer a right to buy the underlying security at a pre-determined strike price at maturity, in contrast, a put option provides the buyer a right to sell the underlying security at a pre-determined strike price. Options are typically classified into two main categories: European options and American options. European options only have one single maturity, and American options have a period available for exercising. An American call (put) option gives the holder the right to buy (sell) the underlying asset at the strike price in a period, while a European call (put) option only provides the holder the right to buy (sell) the underlying asset at a certain time spot.



(a) Volatility Risk Premium



(b) Variance Risk Premium

Figure 4: Estimated Volatility Risk Premium (a) and Variance Risk Premium (b) based on the filtered volatility (ν_t). We estimate Volatility Risk Premium$_t$ by $\text{VIX}_t - \frac{1}{30}\sum_{i=t+1}^{t+30}\hat{\nu}_t$, and estimate Variance Risk Premium$_t$ by $\text{VIX}_t^2 - \frac{1}{30}\sum_{i=t+1}^{t+30}\hat{\nu}_t{}^2$.

The parameters under the physical measure can be used to evaluate the characteristics of the market. For pricing, however, we need to estimate the model under risk-neutral measure based on the following proposition:

Proposition 3. *(Rouah, 2013). The discretized risk-neutral processes of Heston model with Euler scheme is*

$$d\hat{S}_t = r\hat{S}_t\,dt + \sqrt{\nu_t}\hat{S}_t\,d\hat{W}_t^1 \tag{21}$$

$$d\hat{\nu}_t = \kappa^*(\theta^* - \hat{\nu}_t)\,dt + \xi\sqrt{\hat{\nu}_t}\,d\hat{W}_t^2 \tag{22}$$

where $\hat{W}_t^1$ *and* $\hat{W}_t^2$ *are Wiener processes under the risk-neutral probability measure and* $\mathbb{E}^{\mathbb{Q}}[\hat{W}_t^1\hat{W}_t^2] = \rho\,dt$.

Here, we use κ^* and θ^* to distinguish these two parameters from the Heston parameters under the physical measure, while ξ and ρ are the same in the discretized model under both the physical measure (5 and 6) and the risk-neutral measure (21 and 22). Similar to what we did before, we can write this $(\hat{S}, \hat{\nu})$ system as a $(\hat{R}, \hat{\nu})$ system where $\hat{R}$ represents the discretized log return. The general pricing formula of contingent claim $\mathcal{X}$ with maturity T is

$$\Pi(t; \mathcal{X}) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\Phi(\mathcal{X}_T)|\mathcal{F}_t] \tag{23}$$

where $\Pi(t; \mathcal{X})$ is the price of the claim $\mathcal{X}$ at time t, $\Phi(\cdot)$ is the payoff function of the contingent claim, and $\mathcal{F}_t$ is the filtration representing information up to time t. The expectation could be estimated by Monte Carlo simulation given the risk-neutral parameters. In particular, there are only two parameters that need to be adjusted from the physical measure to the risk-neutral measure: κ and θ. These parameters can be adjusted based on the following proposition, using the variance risk premium derived in the previous subsection:

Proposition 4. *(Heston, 1993). Suppose both the physical measure and the risk-neutral measure exist. Let* $(\bar{S}_t, \bar{\nu}_t)$ *represents the continuous processes under the risk-neutral measure. Define function* $\lambda(\bar{S}, \bar{\nu}_t, t)$ *that solves*

$$d\bar{\nu}_t = [\kappa(\theta - \bar{\nu}_t) - \lambda(\bar{S}_t, \bar{\nu}_t, t)]\,dt + \sigma\sqrt{\bar{\nu}_t}\,d\bar{W}_t^2,$$

where

$$\bar{W}_t^2 = \left(W_t^{\nu,\mathbb{P}} + \frac{\lambda(\bar{S}_t, \bar{\nu}_t, t)}{\xi\sqrt{\bar{\nu}_t}}t\right).$$

Here, $\bar{W}_t^2$ *is the Brownian motion under the risk-neutral measure, and* $W_t^{\nu,\mathbb{P}}$ *is the Brownian motion under the physical measure (defined in (1)). Let* $\lambda(\bar{S}_t, \bar{\nu}_t, t) = \lambda\bar{\nu}_t$, *where* λ *is a constant. Then,*

$$\kappa^* = \kappa + \lambda, \quad \theta^* = \frac{\kappa\theta}{\kappa + \lambda}.$$

Proposition 4 allows us to convert the Heston parameters under physical measure to Heston parameters under risk-neutral measure. Here, λ corresponds to the variance risk premium (VRP; as a convention, $-\lambda$ stands for the premium). The uncertainty of investing in a stock includes two parts: one is the uncertainty about the return (which depends on the return variance), and another is the return variance itself (which depends on the variance of variance). The second uncertainty

is compensated by VRP as representative investors in the real world are risk-averse. Typically, $\text{VRP}_t = \mathbb{E}[\text{Implied Volatility}_{t+1}] - \mathbb{E}[\text{Realized Volatility}_{t+1}]$, where implied volatility indicates the extent to which the market expects the corresponding asset price to change in the future, and the realized volatility is the actual volatility under the physical measure.

Then, Monte Carlo methods can be adopted to simulate stochastic differential equations to get the approximated expectation in the pricing formula (23). We choose options with relatively frequent tradings as examples, and we may assume the markets of selected options are efficient so that the market prices reflect all the information we have. The results in Table 2 show that the prices simulated by the Heston model with the Euler scheme recover the market price very well. This suggests that our model works well in this example. One interesting point is that the VRP does not affect the price significantly. This might be good because not every index has a corresponding index for its volatility like VIX, making VRP hard to estimate in practice. While calibrating stochastic volatility models using market options prices has been well studied, our approach may still be valuable in options pricing, as it is not restricted by the liquidity and efficiency of a specific option market. However, if the expiration is much longer than 30 days, the compensation for VRP may increase dramatically, making the adjustment much more important, and leading to a deficiency of the method.

4. Multi-Dimensional Example with Panel Iterated Filter

4.1 Panel Iterated Filter Background

The particle filter suffers from the "curse of dimensionality", where the quality of the likelihood estimation deteriorates exponentially as the size of the dynamic system increases (Daum and Huang, 2003). As such, the IF2 algorithm cannot readily be used in a higher-dimensional setting, preventing us from considering panel data which consists of time series of several assets. Research efforts for high-dimensional stochastic volatility models have focused on numerically solving high-dimensional model-corresponding PDEs to get the price for embedded contingent claims. These methods typically need to transform the high-dimensional PDEs to stochastic differential equations (SDEs) using the Feynman-Kac formula, then apply deep learning techniques to solve the corresponding SDE system (Ee et al., 2017). While these methods are efficient, it is not possible to estimate stochastic volatility models under the physical measure based on high-dimensional PDEs or Stochastic PDEs (SPDEs). While this is still an active current research topic, some numerical methods under Bayesian structures may be of use (Hoffman et al., 2013; Chen et al., 2016), but they are still subject to the limitations of Bayesian models. An extension of iterated filtering in a high-dimensional setting known as the Panel Iterated Filter (PIF) (Bretó et al., 2020) provides us with a new method to extend iterated filter for panel POMP models, inheriting all advantages of iterated filter.

A panel POMP model is a high-dimensional extension of the POMP model. We label the panel units as $u \in \{1, \ldots, U\}$, written as $u \in 1 : U$. For each unit $u \in 1 : U$, there is an independent

Table 2: Empirical 30-Day (22-Trading-Day) SPX Call Options Pricing

Strike	Market Price[1]	Heston Price (History VRP)[2,4]	Heston Price (Zero VRP)[2,3]	BS-Price (History Vol.)[5]
4900	155.25	155.55 (0.30)	155.43 (0.18)	175.78 (20.53)
4960	111.23	112.08 (0.85)	111.14 (-0.09)	137.94 (26.71)
5010	80.92	80.39 (-0.53)	80.09 (-0.83)	110.54 (29.62)
5060	54.73	53.85 (-0.88)	53.84 (-0.89)	86.95 (32.22)
5100	38.90	36.74 (-2.16)	36.66 (-2.24)	68.16 (29.26)
5200	10.97	10.68 (-0.29)	10.57 (-0.40)	40.03 (29.06)

Note: The numbers in parentheses represent the difference between the estimated price and the market price (i.e., $\hat{P} - P$). We only test the 30-day call options because the VIX index (used in estimating the variance risk premium) is the expected implied volatility over the next 30 days, and the call options are more frequently traded than put options. Pricing options that will expire in less than 30 days should be more accurate than the result above as the simulated trajectories are less likely to deviate from the true process. The closing price of the SPX index on Feb 16, 2024, is 5055.57 (USD).

[1] We use 'last price' as market price. Here, we only pick frequently-traded options whose last trade happened just around 4 p.m. on Feb 16, 2024. In this case, the 'last price' can better represent the market price at that time spot.

[2] We use the Euler scheme (which is consistent with our discretization scheme) to simulate the Heston model with smoothed profile MLE (and adjustment for VRP) to approximate the expectation in the pricing formula (23). Each estimated price is based on 100,000 simulated trajectories. Specifically, we use the filtered mean of the ending variance 5.16×10^{-5} as ν_0 instead of directly employing the estimated ν_0.

[3] We use zero variance risk premium ($-\lambda = 0$). In this case, the parameters under risk-neutral pricing is exactly the same as the parameters under the physical measure.

[4] We use the average variance risk premium $-\lambda = 5.13 \times 10^{-5}$. This number is derived by the same data for Figure 4.

[5] We take the historical volatility into the Black-Schole pricing formula. The historical volatility is estimated by $\hat{\sigma} = \sqrt{\frac{S_\xi^2}{\Delta t}}$, where $S_\xi^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\xi_i - \bar{\xi})^2$, $\xi_i = \ln(\frac{S(t_i)}{S(t_{i-1}}})$ for $i = 1, \ldots, N$, $\bar{\xi}$ is the mean of log-return, and N is the number of log-return data points. Here $\Delta t = 1$ as in the model.

corresponding POMP model, with latent process $\{X_u(t) : t > t_0\}$ and noisy and incomplete process $Y_u(t)$ observed at $t_1 < \ldots, t_N$ as $y^*_{u,1}, \ldots y^*_{u,N}$. One necessary assumption for panel POMP models is the interaction between dynamics of different composite units should be negligible, i.e., the elements in set $\{(X_{u,1:N_u}, Y_{u,1:N_u}) : u \in 1 : U\}$ are pairwise independent, where N_u is the number of observations of unit u, and $1 : N_u$ stands for the sequence $1 \ldots, N_u$. We can split all

the parameters into two categories, which are the unit-specific parameters (i.e., are different among all units) and the shared parameters (i.e., are shared among all units). In mathematical form, $\theta = \theta_{\text{shared}} \cup \{\theta_{u,\text{specific}} : u \in 1 : U\}$. If we have A shared parameters and B unit specific parameters, the parameter space is of dimension $A + UB$, i.e., $\theta \in \Theta \in \mathbb{R}^{A+UB}$. The ability to split parameters as shared or unit-specific provides an opportunity to formally test if dynamic systems have similar characteristics.

Like the iterated filter algorithm mentioned in section 3, PIF explores the parameter space by introducing perturbations between time points. Over time, the perturbation diminishes to zero as in the iterated filter, achieving the convergence of MLE. To achieve faster maximization of model likelihood, we use an extension of the PIF algorithm known as block panel iterated filtering (BPIF) (Bretó et al., 2023). There exist empirical and heuristic arguments that favor the use of the BPIF algorithm for likelihood maximization over the PIF algorithm, though the theoretical justification of BPIF is still currently an open research problem.

4.2 Panel POMP Model for Multiple Emerging Market Indices

We apply the PIF to a 4-dimensional panel POMP model of emerging market indices, which are BVSP (São Paulo Bovespa Index, Brazil), BSESN (S&P Bombay Stock Exchange Sensitive Index, India), JKSE (Jakarta Stock Exchange Composite Index, Indonesia), and MXX (Mexican Indice de Precios y Cotizaciones Index, Mexico). These indices reflect the performance of their corresponding stock markets. As emerging markets, these stock markets only have regional influences and are less likely to influence the performance of other stock markets directly. This is concordant with the required independence between the dynamics of the units involved. To fit the panel POMP model, we use daily time series data of the 4 indices from Apr 10, 2014, to Apr 5, 2024. Figure 5 shows the normalized prices (a) and log returns (b) of these 4 indices.

We use the same model in section 3.4 for each unit. We consider 6 versions of the panel POMP model. The first version has no shared parameters, and the remaining five have only one single shared parameter from the list $(\mu, \kappa, \theta, \rho, \xi)$. The results are shown in Table 3, and the estimations of parameters are shown in Table 5 (Appendix B). For each of the models with one shared parameter, we profile over the shared parameter (see Figure 6). One modification is that instead of considering the overall log-likelihood, we select the best parameters for each specific unit while fixing the shared parameter to improve the log-likelihood, as all units are independent. The modification is justified by the figures in Appendix A4. In each figure of Appendix A4, the trajectory based on our modification (red curve) is higher than the regular trajectory (blue curve), indicating that the model based on our modification fits the data better. The confidence interval for each of the shared parameters is demonstrated in Table 4.

One interesting finding is that the LRT statistic of the model with shared ρ is not statistically significant with a 95% confidence level, indicating that we fail to reject the null hypothesis that the more general model is significantly better than the simpler model, and we may prefer the model with

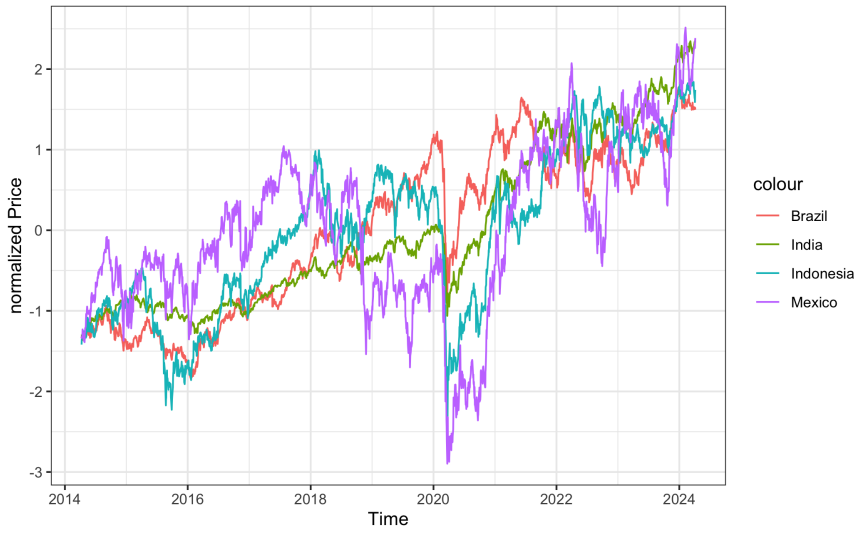


Figure 5: Time series plots of normalized prices from Apr 10, 2014, to Apr 5, 2024. In plot (a), the index prices are normalized using $\frac{S_t - \bar{S}_{1:T}}{\sigma(\bar{S}_{1:T})}$, where $\bar{S}_{1:T}$ and $\sigma(\bar{S}_{1:T})$ are the mean and standard deviation of the full corresponding time series. *Data Source: Yahoo Finance and Bombay Stock Exchange*

Table 3: Log-Likelihood and Likelihood Ratio Test (LRT) Statistic for 1-Shared-Parameter Models

Shared Parameter	Log-Likelihood	Difference[1]	LRT Statistic[2]	p-value
μ	27841.61	4.49	8.98	0.03
κ	27845.09	1.01	2.02	0.57
θ	27820.58	25.52	51.04	4.80×10^{-11}
ρ	27844.15	1.95	3.90	0.27
ξ	27835.82	10.28	20.56	1.30×10^{-4}

Note: The log-likelihood of the model without any shared parameter is 27846.10. This model has the largest number of parameters and is considered as the most flexible model. We use 250 iterations and 2500 particles for all these models.

[1] Difference = Log-Likelihood of the model without any shared parameter − Log-Likelihood of the model a specific shared parameter.

[2] The LRT statistic follows a χ^2_3 distribution.

shared ρ compared with the model with no shared parameter. This implies a potential similarity in the volatility responses among the four emerging markets to price shocks. The estimated shared ρ value is -0.530, with a 95% confidence interval of $[-0.583, -0.461]$ (refer to Table 4). This suggests that a shock in the price processes of these emerging indices would result in a comparatively less negative impact on volatility compared to a shock in the S&P500 index, where the estimated ρ is -0.729. In other words, our results suggest that our selected emerging indices have less negative correlations between the price and volatility, compared to the S&P500 index.

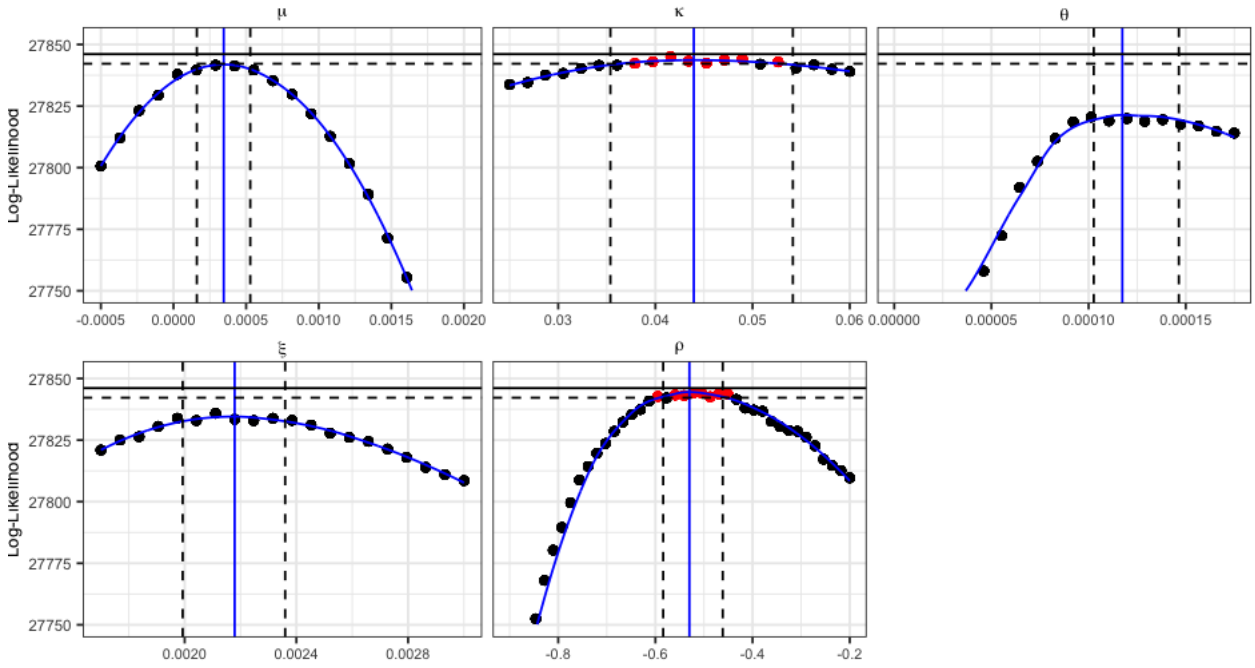


Figure 6: Profile trajectories over each shared parameter, with 250 IF2 iterations and 2500 particles. The horizontal solid line represents the log-likelihood for the model with no shared parameters, and the horizontal dashed line represents the upper boundary of the rejection region for a likelihood ratio test with a 95% confidence level. Any point that does not fall within the rejection region is colored red. The region between two vertical dashed lines is the 95% confidence interval for the shared parameter.

Another interesting finding in Table 4 is that the model featuring a shared κ outperforms the model with no shared parameters in terms of statistical significance. Within the shared-κ model, the annualized estimation of κ is 11.08, suggesting a half-life of 15.8 trading days. This half-life is shorter than that estimated for the S&P500 index, which is 21.3 trading days. This suggests that volatility reverts at a quicker speed in these emerging markets compared to the S&P500 index, which reflects the performance of the world's most developed and efficient market. This is consistent with the findings of Ahmed et al. (2018) based on their GARCH model. Since a greater half-life indicates greater volatility, our results suggest that the US market is even more volatile than some emerging markets. While this may seem contradictory to our intuition, there are several explanations of this result. Bartram et al. (2012) pointed out that higher country risk, better investor protection, greater financial development and openness, more apparent disclosures and noise trading, and more innovation and growth opportunity may all lead to higher volatility of US stocks. Moreover, the extremely active algorithm tradings in the United States may also contribute to the high volatility.

Figure 7 demonstrates the estimated volatility process for each country. Unfortunately, it's challenging to validate the accuracy of these processes directly since volatility is not directly observed, and there are no other readily available metrics like the VIX index we used previously. However, Atiya

Table 4: Profile MLE and Confidence Interval for Each Shared Parameter

Shared Parameter	MLE[1]		Confidence Interval[1,3]	
	Smoothed Profile MLE	Annualized MLE[2]	2.5%	97.5%
μ	3.46×10^{-4}	8.72%	1.61×10^{-4} (4.06%)	5.29×10^{-4} (13.33%)
θ	1.17×10^{-4}	17.20%	1.03×10^{-4} (16.08%)	1.46×10^{-4} (19.21%)
κ	4.40×10^{-2}	11.08	3.54×10^{-2} (8.59)	5.41×10^{-2} (13.63)
ξ	2.18×10^{-3}	3.46%	1.99×10^{-3} (3.16%)	2.36×10^{-3} (3.75%)
ρ	-5.30×10^{-1}	–	-5.83×10^{-1}	-4.61×10^{-1}

Note: only the model with shared κ and the model with shared ρ are better than the model with no shared parameters, up to a 95% confidence level.

[1] Based on MCAP.

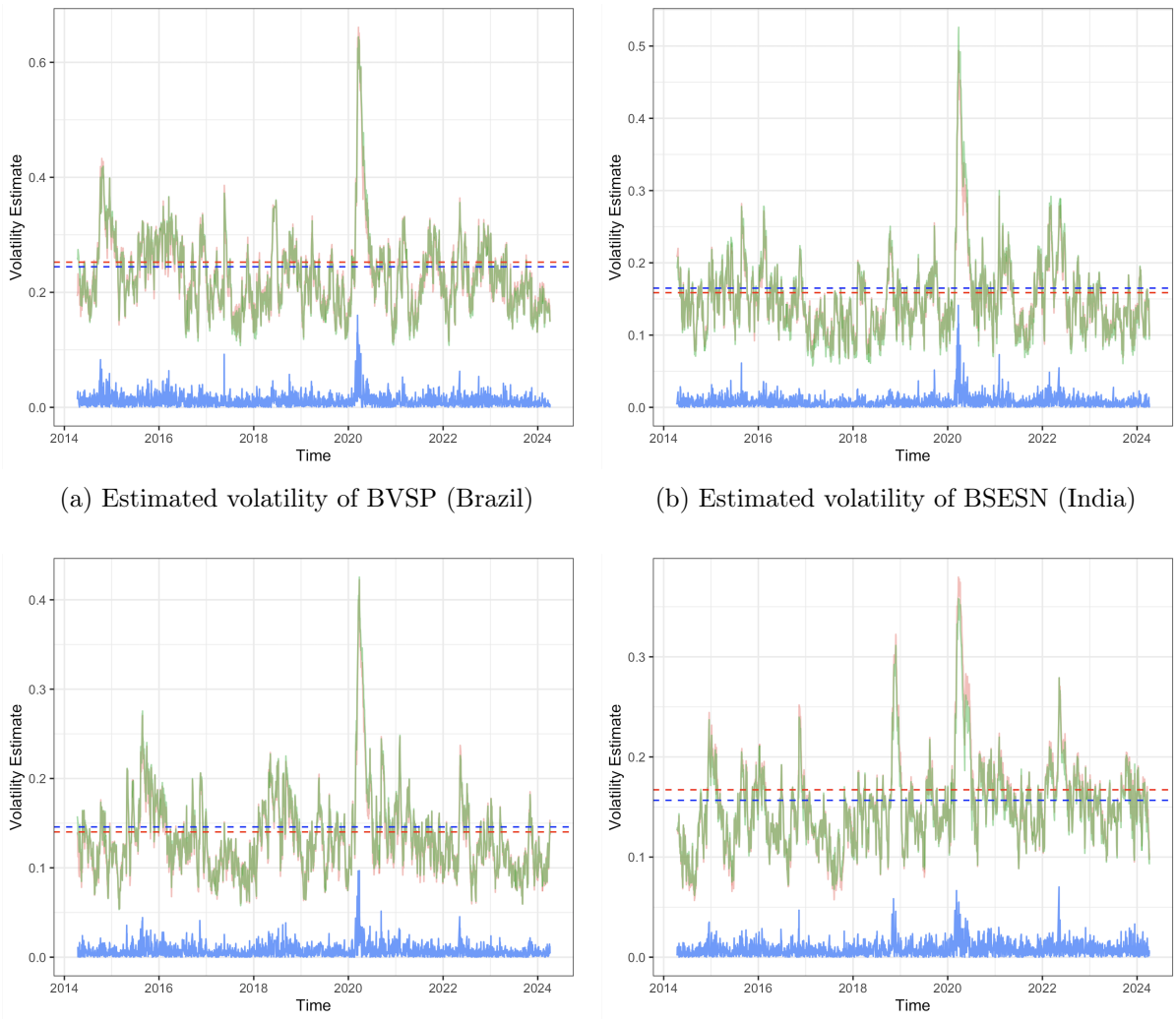
[2] Annualized by 252 trading days of a year based on the smoothed profile MLE.

[3] The numbers in the parentheses are corresponding to the annualized MLE based on the smoothed profile MLE.

and Wall (2009) suggests that we may use $|y_{t+1} - y_t - \mu\Delta t|$ as a benchmark to assess the plausibility of the estimated volatility processes, where y is the log-price of the corresponding index. (Atiya and Wall used the squared-increment $(y_{t+1} - y_t - \mu\Delta t)^2$ as a reference, commonly used in GARCH modeling of variance, we adopted the absolute increment $|y_{t+1} - y_t - \mu\Delta t|$ to directly compare estimated volatilities.) Within each plot depicted in Figure 6, there is a consistent pattern: when the benchmark process reaches high values, the estimated volatility tends to be correspondingly high; conversely, when the benchmark process reaches low values, the estimated volatility tends to be lower as well. Also, for each index, the estimated volatility processes from the shared-ρ model and the shared-κ model almost overlap. In summary, we may conclude that these estimated volatility processes appear to be aligned with expectations.

5. Discussion

In this thesis, we estimated the Heston model under the physical measure, exploring both one-dimensional and multi-dimensional scenarios. For the one-dimensional Heston stochastic volatility model, we treated it as an extended partially observed Markov Processes (POMP) model and adopted the iterated filtering algorithm (IF2; Ionides et al., 2015) to estimate both the volatility process and the model parameters, along with their respective 95% confidence intervals. The results confirmed the validity of the method and indicated good convergence of parameter estimation. We then used a risk-neutral version of the fitted model to obtain option pricing, which did well in recovering



(a) Estimated volatility of BVSP (Brazil)

(b) Estimated volatility of BSESN (India)

(c) Estimated volatility of JKSE (Indonesia)

(d) Estimated volatility of MXX (Mexico)

Figure 7: Estimated volatility processes of BVSP, BSESN, JKSE and MXX. The estimated volatility processes are the filtered mean with their corresponding smoothed MLE using 1000 particles.

the market prices of frequently-traded S&P500 European call options with 30-day maturities. Extending our analysis, we transitioned from the one-dimensional model to multi-dimensional cases, based on block panel iterated filtering (BPIF; Bretó et al., 2023). We applied this method in a panel data set consisting of 4 emerging market indices, which are BVSP (Brazil), BSESN (India), JKSE (Indonesia), and MXX (Mexico). The findings from this analysis suggested a potential sharing of mean reversion rate and sensitivity between the volatility processes and their corresponding price processes across these market indices. This research is meaningful because it provides a full-information, plug-and-play, and frequentist perspective to stochastic volatility models, while most

previous research in estimating these models focused on a Bayesian structure. Also, due to the plug-and-play nature of our approach, this analysis can easily be extended to other data sets or stochastic volatility models.

However, our research does have some limitations and areas for potential future improvements. First, we only considered the case with time discretization step $\Delta t = 1$ trading day, thus the discretized model may not well approximate the true model. One potential future improvement is to use $\Delta t < 1$. However, implementing this could pose challenges as the corresponding sampling techniques are more complex than those employed in this thesis, as indicated by equation (23). second, While we believe that our profile searching (with fixed μ) in section 3.4 is reliable, the perturbation of μ does hurt the efficiency of the IF2 likelihood optimizer. Efficiency may be improved through the use of automatic differentiable particle filters (ADPF; Tan, 2023) , which do not involve random perturbations of parameters. Another possible extension of this work involves modeling stochastic volatility as an interacting network; in real-world scenarios, markets may be interdependent, violating the core independence assumption of panel POMP models. In this case, we may need spatiotemporal partially observed Markov processes (Spat POMP) models (Asfaw et al., 2023; Asfaw et al., 2024). Finally, our work on the multi-dimensional model only considered models with one shared parameter. An extension for future research could involve exploring models with multiple shared parameters. For instance, in our multi-dimensional example, it may be beneficial to share both ρ and κ across all units.

References

Ahmed, R., Vveinhardt, J., Streimikiene, D., & Channar, Z. (2018). Mean reversion in international markets: Evidence from G.A.R.C.H. and half-life volatility models. *Ekonomska Istraživanja / Economic Research*, *30*, 1198–1217.

Alberga, D., Shalita, H., & Yosef, R. (2008). Estimating stock market volatility using asymmetric GARCH models. *Applied Financial Economics*, *18*(15), 1201–1208.

Andersen, T. G., Chung, H. J., & Sørensen, B. E. (1999). Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics*, (1), 61–87.

Andersen, T. G., & Sørensen, B. E. (1996). GMM estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Business and Economic Statistics*, (3), 328–352.

Andrieu, C., Doucet, A., & Holenstein, R. (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *72*(3), 269–342.

Arsalan, T., Chishty, B. A., Ghouri, S., & Ansari, N. U. H. (2022). Comparison of volatility and mean reversion among developed, developing and emerging countries. *Journal of Economic and Administrative Sciences*.

Arulampalam, M., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *Signal Processing, IEEE Transactions on*, *50*(2), 174–188.

Asai, M., & McAleer, M. (2011). Alternative asymmetric stochastic volatility models. *Econometric Reviews*, *30*(5), 548–564.

Asfaw, K., Park, J., King, A. A., & Ionides, E. (2023). Partially observed Markov processes with spatial structure via the R package spatPomp.

Asfaw, K., Park, J., King, A. A., & Ionides, E. (2024). *spatPomp: Statistical inference for spatiotemporal partially observed Markov processes* [R package, version 0.34.2]. https://github.com/kidusasfaw/spatPomp

Atiya, A., & Wall, S. (2009). An analytic approximation of the likelihood function for the Heston model volatility estimation problem. *Quantitative Finance*, *9*(3), 289–296.

Bae, J., Kim, C. J., & Nelson, C. R. (2007). Why are stock returns and volatility negatively correlated? *Journal of Empirical Finance*, *14*(1), 41–58.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance*, *52*(5), 2003–2049.

Bams, D., Lehnert, T., & Wolff, C. (2009). Loss functions in option valuation: A framework for model selection. *Management Science*, *55*, 853–862.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 167–241.

Bartram, S. M., Brown, G., & Stulz, R. M. (2012). Why are U.S. stocks more volatile? *The Journal of Finance*, *67*(4), 1329–1370.

Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *The Review of Financial Studies*, *13*(1), 1–42.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *The Pricing of Options and Corporate Liabilities*, *81*(3), 885–887.

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

Bollerslev, T., Litvinova, J., & Tauchen, G. (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics*, *4*(3), 353–384.

Bretó, C. (2014). On idiosyncratic stochasticity of financial leverage effects. *Statistics & Probability Letters*, *91*(100), 20–26.

Bretó, C., Ionides, E. L., & King, A. (2020). Panel data analysis via mechanistic models. *Journal of the American Statistical Association*, *115*(531), 1178–1188.

Bretó, C., Ionides, E. L., & King, A. A. (2023). *panelPomp: Inference for panel partially observed Markov processes* [R package, version 1.1].

Cacace, F., Germani, A., & Papi, M. (2019). On parameter estimation of Heston's stochastic volatility model: A polynomial filtering method. *Decisions in Economics and Finance*, (2), 503–525.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Econometrics*, *31*(3), 281–318.

Cappuccio, N., Lubian, D., & Raggi, D. (2004). MCMC Bayesian estimation of a skew-GED stochastic volatility model. *Studies in Nonlinear Dynamics and Econometrics*, *8*.

Chen, Y., Shen, K., Shan, S. O., & Kou, S. (2016). Analyzing single-molecule protein transportation experiments via hierarchical hidden Markov models. *Journal of the American Statistical Association*, *111*, 1–49.

Chib, S., Nardari, F., & Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, *108*, 281–316.

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

Chopin, N., Jacob, P., & Papaspiliopoulos, O. (2013). SMC 2 : An efficient algorithm for sequential analysis of state space models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, (3), 397–426.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics*, (4), 407–432.

Christoffersen, P., Heston, S., & Jacobs, K. (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, *55*(12), 1914–1932.

Clayton, M. (2020). *Time-series Heston model calibration using a trinomial tree.*

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica*, *53*(2), 363–384.

Daum, F., & Huang, J. (2003). Curse of dimensionality and particle filters. *2003 IEEE Aerospace Conference Proceedings*, *4*, 1979–1993.

Doucet, A., de Freitas, N., & Gordon, N. J. (2001). Sequential Monte Carlo methods in practice.

Ee, W., Han, J., & Jentzen, A. (2017). Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *To appear in Communications in Mathematics and Statistics*, *5*(4), 349–380.

Gruszka, J., & Szwabiński, J. (2023). Parameter estimation of the Heston volatility model with jumps in the asset prices. *Econometrics*, *11*(2), 1–26.

Harvey, A. C., & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics*, *14*(4), 429–434.

He, X., & Chen, W. (2020). A closed-form pricing formula for European options under a new stochastic volatility model with a stochastic long-term mean. *Mathematics and Financial Economics*, *15*, 381–396.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, *6*(2), 327–343.

Hoffman, M., Blei, D., Wang, C., & Paisley, J. (2013). Stochastic variational inference. *Journal of Machine Learning Research*, *14*, 1303–1347.

Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, *42*(2), 281–300.

Ionides, E. L., Breto, C., Park, J., Smith, R. A., & King, A. A. (2017). Monte Carlo profile confidence intervals for dynamic systems. *Journal of The Royal Society Interface*, *14*(132), 20170126.

Ionides, E. L., Bretó, C., & King, A. A. (2006). Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, *103*(49), 18438–18443.

Ionides, E., Nguyen, D., Atchadé, Y., Stoev, S., & King, A. (2015). Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(3), 719–724.

Ionides, E. L. (2021). *Lecture 16 notes in analysis of time series (stats531): A case study of financial volatility and a pomp model with observations driving latent dynamics*, University of Michigan. https://ionides.github.io/531w21/16/notes.pdf

Ionides, E. L., Bhadra, A., Atchadé, Y., & King, A. (2011). Iterated filtering. *The Annals of Statistics*, *39*(3), 1776–1802.

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

Jiao, Y., Ma, C., & Scotti, S. (2021). The Alpha-Heston stochastic volatility model. *Mathematical Finance*, *31*(3), 825–1095.

Kilin, F. (2007). Accelerating the calibration of stochastic volatility models. (6).

King, A., Nguyen, D., & Ionides, E. L. (2016). Statistical inference for partially observed Markov processes via the R package pomp. *Journal of Statistical Software*, *69*(12), 1–43.

Neokosmidis, I. (2009). Econometric analysis of realized volatility: Evidence of financial crisis. https://api.semanticscholar.org/CorpusID:10434593

Nugroho, D. B., & Morimoto, T. (2014). Modeling of stochastic volatility to validate IDR anchor currency. *Journal of Japan Statistical Society*, *44*(1), 83–118.

Olowe, R. A. (2009). Stock return, volatility and the global financial crisis in an emerging market: The Nigerian case. *International Review of Business Research Papers*, *5*(15), 426–447.

Rouah, F. D. (2013). *The Heston model and its extensions in Matlab and C#*. Wiley.

Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics*, *63*(1), 289–306.

Scott, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application. *The Journal of Financial and Quantitative Analysis*, *22*(4), 419–438.

Storn, R., & Price, K. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, *11*(12), 341–359.

Szczepocki, P. (2020). Application of iterated filtering to stochastic volatility models based on non-Gaussian Ornstein-Uhlenbeck process. *Statistics in Transition New Series*, *21*(2), 173–187.

Tan, K. (2023). *Differentiable plug-and-play particle filtering* [unpublished].

Taylor, S. J. (2007). *Modelling financial time series*. World Scientific Publishing Co. Pte. Ltd.

Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the Brownian motion. *Physical review*, *36*(5), 823.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, *5*(2), 177–188.

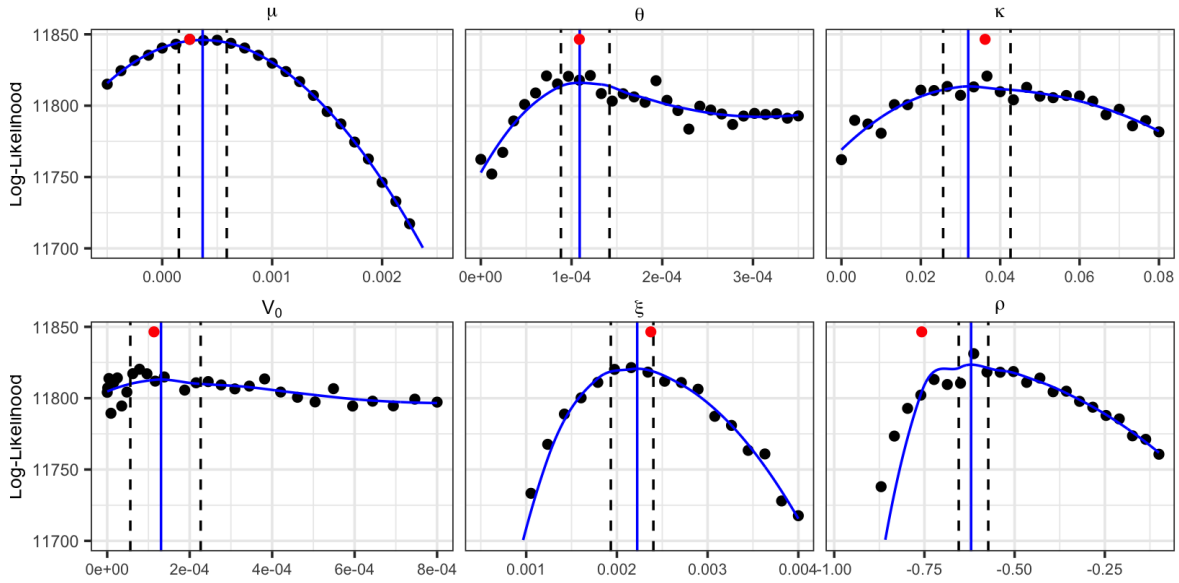
Vollrath, I., & Wendland, J. (2009). Calibration of interest rate and option models using differential evolution. *SSRN Electronic Journal*.

Wheeler, J., Rosengart, A., Jiang, Z., Tan, K., Treutle, N., & Ionides, E. L. (2024). Informing policy via dynamic models: Cholera in Haiti [arXiv:2301.08979].

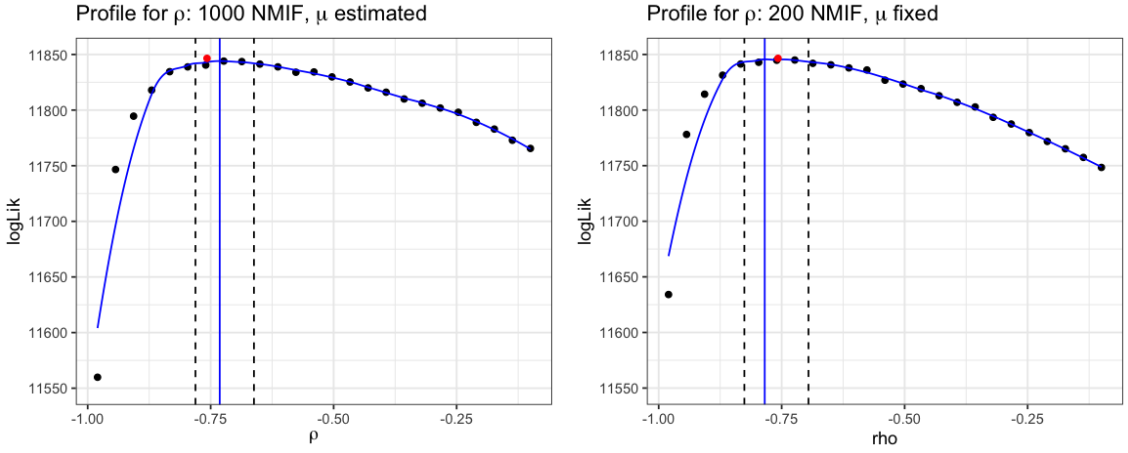
Yang, X., Chatpatanasiri, R., & Sattayatham, P. (2017). Value at risk estimation under stochastic volatility models using adaptive PMCMC methods. *Communications in Statistics - Simulation and Computation*, *46*(9), 7221–7237.

Appendix A. Corresponding Figures

A1



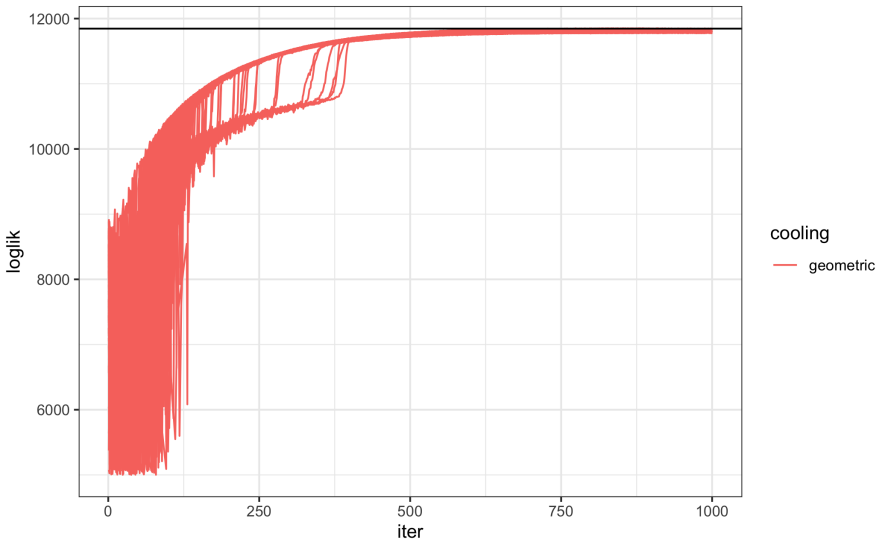
A2



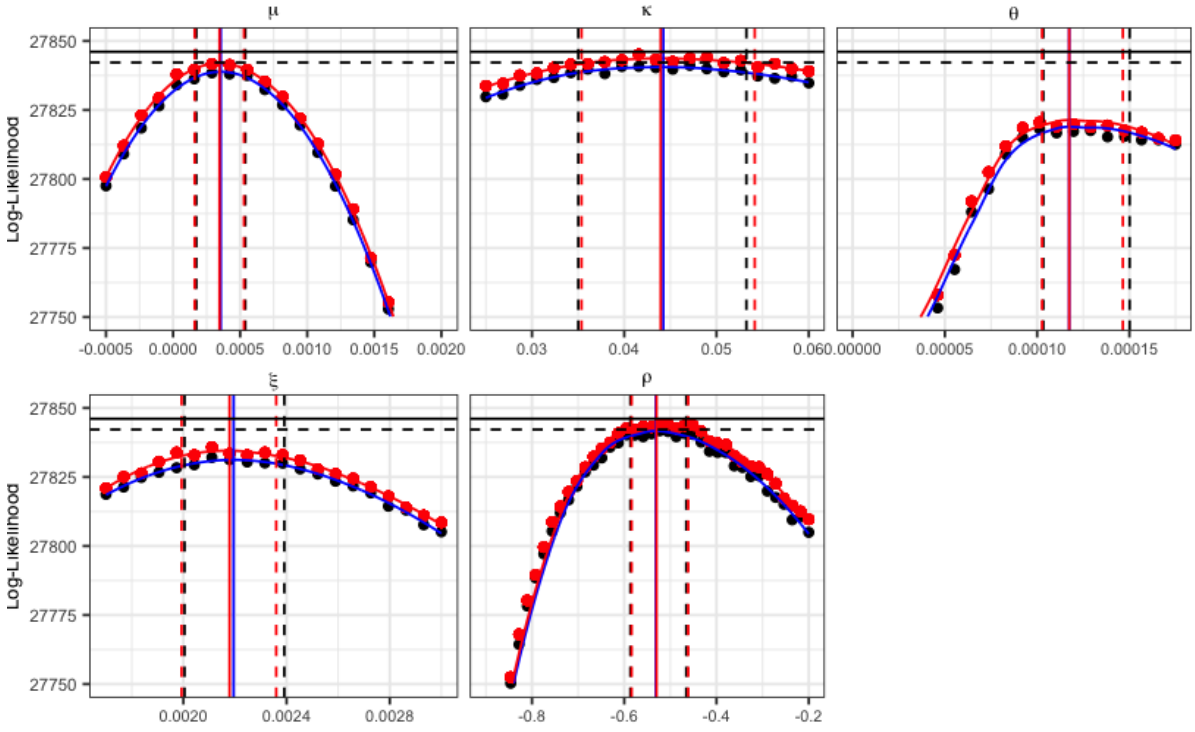
(a) Profile traces of ρ without fixing μ (1000 IF2 iterations)

(b) Profile traces of ρ with fixed μ (200 IF2 iterations)

A3


12000
10000
8000
6000
loglik
0
250
500
750
1000
iter
cooling
geometric

A4


μ
κ
θ
ξ
ρ
Log-Likelihood
27850
27825
27800
27775
27750
-0.0005
0.0000
0.0005
0.0010
0.0015
0.0020
0.03
0.04
0.05
0.06
0.00000
0.00005
0.00010
0.00015
0.0020
0.0024
0.0028
-0.8
-0.6
-0.4
-0.2

Appendix B. Parameter Estimation for the Multi-Dimensional Example

Table 5: Parameter Estimation

Market Index	Model	Parameter Estimation					
		μ	θ	κ	ξ	ρ	ν_0
BVSP	no shared parameter	5.13×10^{-4}	2.39×10^{-4}	4.39×10^{-2}	2.81×10^{-3}	-0.48	1.99×10^{-4}
	shared ρ	1.75×10^{-4}	2.53×10^{-4}	5.86×10^{-2}	3.44×10^{-3}	-0.51	1.40×10^{-4}
	shared κ	7.21×10^{-4}	2.37×10^{-4}	4.16×10^{-2}	2.90×10^{-3}	-0.51	2.70×10^{-4}
BSESN	no shared parameter	7.94×10^{-4}	8.74×10^{-5}	4.98×10^{-2}	2.21×10^{-3}	-0.64	9.11×10^{-5}
	shared ρ	7.79×10^{-4}	1.00×10^{-4}	5.84×10^{-2}	2.29×10^{-3}	-0.51	1.82×10^{-4}
	shared κ	5.10×10^{-4}	1.08×10^{-4}	4.16×10^{-2}	2.18×10^{-3}	-0.61	1.50×10^{-4}
JKSE	no shared parameter	9.39×10^{-7}	8.22×10^{-5}	4.93×10^{-2}	1.87×10^{-3}	-0.51	5.46×10^{-5}
	shared ρ	9.65×10^{-7}	7.80×10^{-5}	5.31×10^{-2}	1.98×10^{-3}	-0.51	7.66×10^{-5}
	shared κ	1.26×10^{-4}	8.43×10^{-5}	4.16×10^{-2}	1.80×10^{-3}	-0.49	1.00×10^{-4}
MXX	no shared parameter	8.40×10^{-8}	1.04×10^{-4}	2.58×10^{-2}	1.52×10^{-3}	-0.47	1.22×10^{-4}
	shared ρ	1.52×10^{-7}	1.11×10^{-4}	2.82×10^{-2}	1.72×10^{-3}	-0.51	6.46×10^{-5}
	shared κ	7.39×10^{-7}	9.74×10^{-5}	4.16×10^{-2}	1.80×10^{-3}	-0.53	7.83×10^{-5}