Research Article

Portfolio Selection Based on Bayesian Theory

Daping Zhao [iD],[1] Yong Fang [iD],[2] Chaoliang Zhang,[3] and Zongrun Wang[4]

[1]*School of Finance, Capital University of Economics and Business, Beijing 100070, China*
[2]*Institute of Systems Science, Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing 100190, China*
[3]*Founder Fubon Fund Company, Beijing 100037, China*
[4]*Business School, Central South University, Changsha, Hunan Province 410083, China*

Correspondence should be addressed to Yong Fang; yfang@amss.ac.cn

The traditional portfolio selection model seriously overestimates its theoretic optimal return. Aiming at this problem, two portfolio selection models are proposed to modify the parameters and enhance portfolio performance based on Bayesian theory. Firstly, a Bayesian-GARCH(1,1) model is built. Secondly, Markov Chain is applied to curve the parameters' state transfer, and a Bayesian Markov regime-Switching-GARCH(1,1) model is constructed. Both the two models can handle the overestimation problem and can obtain self-financing portfolios. In the numerical experiments, both the models are examined with data from China stock market, and their performances are compared and analyzed. The results show that BMS-GARCH(1,1) model is superior to the Bayesian-GARCH(1,1) model.

1. Introduction

Portfolio optimization and diversification have been instrumental in the development and understanding of financial markets and financial decision-making. The major breakthrough came in 1952 with the publication of Harry Markowitz theory of portfolio selection. The theory is now popularly referred to as modern portfolio theory. According to this theory, both the return and risk of a security can be quantified, using the statistical measures of its expected return and standard deviation. Then, the investors should consider return and risk together and determine the allocation of funds among investment alternatives on the basis of their return-risk trade-off. The idea that sound financial decision-making is a quantitative trade-off between return and risk was revolutionary. And, it has had a major impact on academic research and the financial industry as a whole.

However, there are still some concerns with the theory, and one of the concerns is that the original mean-variance model depends on estimated parameters which are uncertain and may be mistaken. Specifically, return-risk optimization can be very sensitive to changes in the inputs, especially when the return and risk estimates are not well aligned or when the problem formulation uses multiple, interacting constraints. As a result, many practitioners consider the output of risk-return optimization to be opaque, unstable, and/or unintuitive. Besides, estimation errors in the forecasts significantly impact the resulting portfolio weights. For example, it is well-known that, in practical applications, equally weighted portfolios often outperform mean-variance portfolios [1–4]. The classical framework has to be modified when used in practice in order to achieve reliability, stability, and robustness with respect to model and estimation errors [5, 6]. In other words, the portfolio selection models face estimation risks. Therefore, how to minimize the estimation risk to portfolio draws much attention in academia.

Recently, more and more parameter uncertainties have been analyzed in Bayesian Framework. The Bayesian method has advantages: first of all, it can fully use prior information; secondly, it explains the uncertainties of estimated risks and models; thirdly, it makes the computation in simulating complicated variables easy. In this way, investors firstly set the parameter prior; then they use Bayesian rules to get the posterior. As a result, Bayesian methods are gradually used in the portfolio selection analysis. Avramov and Zhou [7]

reviewed the development of Bayesian portfolio in the following respect: (1) the returns are i.i.d; (2) the returns are predicable; (3) the mean follows regime switching and the volatility is stochastic. They mainly discussed the Bayesian portfolio in parameter uncertainties as well as with the model uncertainties. In Fabozzis review of "60 Years of Portfolio Optimization: Practical Challenges and Current Trends," they also speak high of the application of Bayesian methods in portfolio optimization [8]. In this paper, we take the financial data with different lengths as the application background, mainly focus on the uncertainties of parameters in portfolio selection, and build two models based on Bayes and modern portfolio theory: a Bayesian-GARCH(1,1) model and a Bayesian-Markov Regime Switching-GARCH(1,1) model. Using the Bayesian optimization algorithm, both the models do not overestimate their theoretic optimal returns, and they can handle the overestimation problem.

The paper is organized as follows. In Section 2, the background studies are briefly introduced, including the literature review and the relative theories and methods, such as Bayesian theory and Markov regime switching. In Section 3, a Bayesian-GARCH(1,1) model and a Bayesian-Markov Regime Switching-GARCH(1,1) model are constructed. In Section 4, the two models are taken into application, and the numerical results are illustrated and compared. In Section 5, the paper is concluded with summary comments and pointers to the future work.

2. Background Studies

2.1. Literature Review.
It is well known that the parameters in portfolio selection may be uncertain or mistaken, and the return-risk optimization can be very sensitive to changes in the inputs. In early studies, most researchers assumed that the risky asset returns are i.i.d and focused on the effect of the uncertainty of the mean and variance on the portfolio selection. Williams [9] discussed how the parameter uncertainty affected the portfolio selection. Bawa et al. [10] firstly consider the uncertainty of the parameters of the return distribution in static portfolio selection. Gennotte [11] analyzed portfolio selection when the expected return was uncertain. Kandel and Stambaugh [12] showed that when the returns were predictable, investors who ignored the parameter uncertainties preferred risky stocks in a short period. Zenios et al. [13] develop multiperiod dynamic models for fixed-income portfolio management under uncertainty, using multistage stochastic programming with recourse. The multiperiod models outperform classical models based on portfolio immunization and single-period models. Brennan [14] studied the dynamic portfolio selection with the assumptions that the expected risk premium was uncertain in discrete time. He thought the assumptions of financial models should be that all the invest opportunities depended on a series of unobservable variables, and the parameters were not certain. Barberis [15] got similar results with the long-period investors. Pástor and Stambaugh [16] discussed how the uncertainties of the mispricing affected the portfolio selection. Pástor [17] treated the model uncertainty using the investors' opinions to the asset pricing as

prior information. Bai et al. [18, 19] develop a bootstrap-corrected estimator to correct the overestimation in portfolio selection and further extend the theory to obtain self-financing portfolios. Besides, in order to improve the portfolio selection efficiency, researchers tend to extend the models with dynamic programming [13, 20, 21], with continuous time framework [22, 23], using Markovian regime-switching methods [24], and so on.

Currently, most of the parameter uncertainties are analyzed in Bayesian Framework. As we mentioned before, the Bayesian method has advantages: first of all, it can fully use prior information; secondly, it explains the uncertainties of estimated risks and models; thirdly, it makes the computation in simulating complicated variables easy. In this way, investors firstly set the parameter prior; then they use Bayesian rules to get the posterior. As a result, Bayesian methods are gradually used in the portfolio selection analysis. One of the implications is that the Bayesian theory can be mainly used to estimate the uncertain returns or other factors (such as a shrinkage estimator and so on) in portfolio [15, 25–29]. In addition, for the parameter uncertainty of portfolio selection, in other words, for the estimation risk evaluation or the estimation error correction, researchers also tend to choose Bayesian methods [4, 30–34]. Besides, the Bayesian theory is also used in model analysis [35], risk measurement or return-risk trade off [36], portfolio optimization [37, 38], and so on.

2.2. Modern Portfolio Theory.
According to Markowitz's portfolio theory, there are N assets. R_t is the excess return of asset N in time t, $R_t = (R_{1t}, \ldots, R_{Nt})'$, and they follow a normal distribution $p(R_t \mid \mu, \Sigma) = N(\mu, \Sigma)$, μ is a $N \times 1$ matrix, and Σ is a $N \times N$ matrix. The weights are denoted as $w = (w_1, \ldots, w_N)$. The return of the portfolio is

$$R_p = \sum_{i=1}^{N} w_i R_{it} = w' R_t. \tag{1}$$

And the expected return and variance are

$$\mu_p \sum_{i=1}^{N} w_i \mu_i = w' \mu,$$
$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \text{cov}(R_i, R_j) = w' \sum w, \tag{2}$$

where μ_i is the return of asset i and $\text{cov}(R_i, R_j)$ is the covariance of assets i and j.

Suppose investors hold the portfolio for τ and their target is to maximize the value at time $T + \tau$, where T is the time when the portfolio is built. The mean-variance model can be expressed as the following optimization problem:

$$\min_{w} \quad \sigma_p^2 = \min_{w} w' \sum_{T+\tau} w$$
$$\text{s.t.} \quad w' \mu_{T+\tau} \geq \mu^*$$
$$w' 1 = 1, \tag{3}$$

where μ^* is the permitted minimum return.

2.3. Bayesian Theory.

2.3. Bayesian Theory. Bayes firstly presented Bayesian formula in his essay. Later on, Laplace put the new approach into applications. Then statisticians developed it as a systematic statistical method, the Bayesian theory. With the new theory, each unknown variable can be seen as a stochastic variable and can be described with a probability distribution. The distribution is the prior information before sampling, and hence it is called prior distribution, or prior.

Traditional statistical estimation depends on the given distributions and data. However, prior is the base of the Bayesian estimation. Posterior is the key to Bayesian estimation. For an unknown parameter θ, its posterior $\pi(\theta \mid x)$ is a conditional distribution of θ under sampler x and it contains all the information that is available.

2.4. GARCH Model. Engle presented the ARCH (autoregressive conditional heteroskedasticity) model in 1982. Then the method is popular in the volatility empirical test. However, in real stock market, the conditional variance depends not only on one or two previous variables but also on more alternating quantities, which account for several lags and parameters. Then, Bollerslev presented the GARCH model:

$$x_t = \beta_0 + \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p} + \varepsilon,$$
$$\sigma_t^2 = E\left(\mu_t^2\right) = \alpha_0 + \alpha_1 \mu_{t-1}^2 + \cdots + \alpha_q \mu_{t-q}^2 \qquad (4)$$
$$+ \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2,$$

where $\alpha_0 > 0$, $\alpha_i \geq 0 \, (i = 1, \ldots, q)$ and $\beta_j \geq 0 \, (j = 1, \ldots, p)$. In general, the real market can be characterized well when $p = q = 1$.

2.5. MCMC Simulation. MCMC sampling is proposed by Metropolis et al. [36], and then Hastings [40] generalized it. MCMC simulation depends on a Markov chain, and the main methods are Gibbs sampler and Metropolis–Hastings. Derin et al. [41] presents Gibbs sampler in a statistics article, which is now the most popular sampler in MCMC. According to Gibbs sampler, it is easy to figure out the total conditional probability density, which makes the sampler not that difficult. However, sometimes the Gibbs sampler cannot solve complicated Bayesian problems with combined probability density. In these cases, the Metropolis–Hastings algorithm is applied. Take a random variable θ as an example, for its symmetric proposal density function $q(\theta^{[j]} \mid \theta^* = q(\theta^* \theta^{[j]}))$, the acceptance probability with M-H algorithm is

$$\min \left\{ \frac{p(\theta^* \mid y)}{p(\theta^{[j]} \mid y)}, 1 \right\}. \qquad (5)$$

In this way, the acceptance probability is free from calculation. It should be mentioned that the Bayesian parameter estimation based on simulation depends on the astringency of Markov chain. In this way, the sample generating from the simulation follows the posterior distribution of expectation. In the posterior simulation, the aim is to generate a fully ergodic Markov chain.

3. Portfolio Selection Models Based on Bayes

3.1. Bayesian-GARCH(1,1) Model. In this section, the residuals of the return are assumed to follow t-distribution, and the posterior is simulated by Metropolis–Hastings and Gibbs sampler. Suppose the return of asset follows the GARCH(1,1) model as belows:

$$R_{i,t} = \mu_{i,t} + \sigma_{i,t|t-1} \varepsilon_{i,t}, \quad t = 1, \ldots, T, \qquad (6)$$

$$\sigma_{i,t|t-1}^2 = w_i + \alpha_i \mu_{i,t-1}^2 + \beta_i \sigma_{i,t-1|t-2}^2. \qquad (7)$$

Equations (6) and (7) show the process that the asset return and volatility follow. Equation (6) is the mean equation, and (7) is the volatility equation, where $R_{i,t}$ is the return of asset i at time t, $\sigma_{i,t|t-1}^2$ is the variance (volatility) of the return of asset i between time $t-1$ and t, and $\mu_{i,t-1} = R_{i,t} - \mu_{i,t}, w_i > 0, \alpha_i \geq 0, \beta_i \geq 0$. For convenience, assume $\mu_{i,t}$ is unchangeable. If the variance is consistent, then (6) turns to a linear regression. For the parameters w_i, α_i, and β_i, they will be estimated with Bayesian methods.

3.2. Likelihood Function and Prior of Parameters

3.2.1. Likelihood Function. With the assumptions of assets and parameters above, denote the parameter variable of the model as $\theta = (w, \alpha, \beta, v, \mu)$. Suppose $\varepsilon_{i,t}$ follows a t-distribution with v_i degrees of freedom, the likelihood function can be written as follows:

$$L\left(\theta \mid R_{1t}, R_{2t}\right) = \prod_{i=1}^{N} L\left(\theta_i \mid R_{i,t}\right) \propto \prod_{i=1}^{N_1} \prod_{t=1}^{T} \left[\sigma_{i,t|t-1}^{-1} \left(1 + \frac{1}{v} \frac{\left(R_{i,r} - \mu_i\right)^2}{\sigma_{i,t|t-1}^2} \right) \right]$$
$$\times \prod_{j=1}^{N_2} \prod_{t=s}^{T} \left[\sigma_{j,t|t-1}^{-1} \left(1 + \frac{1}{v} \frac{\left(R_{j,r} - \mu_i\right)^2}{\sigma_{j,t|t-1}^2} \right)^{-(v+1/2)} \right], \qquad (8)$$

where θ_i is a parameter variable of asset i and $\sigma_{i,t|t-1}^2$ is its conditional variance. The initial conditional covariance $\sigma_{i,0}^2$ is a constant, and $\varepsilon_{i,t}$ follows a t-distribution with degrees of freedom v_i. Its conditional variance at time t is $v_i / ((v_i - 2)\sigma_{i,t}^2), v_i > 2$.

3.2.2. Prior of Parameters. In many cases, people's priors are vague and thus difficult to translate into an informative prior. We therefore want to reflect the uncertainty about the model parameters without substantially influencing the posterior parameter inference. The so-called noninformative priors, also called vague or diffuse priors, are employed to that end. In this section, we suppose that there is a noninformation prior for the conditional variance parameters of asset i:

$$\pi(w_i, \alpha_i, \beta_i) \propto 1 I_{i,\theta_G}, \qquad (9)$$

where $1 I_{i,\theta_G}$ is an indicator function for constraints of parameters as follows:

$$1 I_{i,\theta_G} = \begin{cases} 1, & w_i > 0, \ \alpha_i > 0, \ \beta_i > 0, \\ 0, & \text{others.} \end{cases} \qquad (10)$$

Geweke (1993) proposed that the prior of v_i can be index prior distribution, and its density function is

$$pi(v_i) = \lambda \exp(-v_i\lambda). \tag{11}$$

Based on the distribution, the mean is $1/\lambda$, and λ can be determined by the degree of freedom of v_i. Several empirical experiments show that in the financial markets, the degree of freedom is typically less than 20. Hence, set the upper bound K as 20. According to Bauwens et al. [42], if there is a diffusion prior for v_i on $[0, \infty)$, its posterior will not be appropriate. So, here we use the index prior as (11) shows. Then, assume the prior of μ_i follows normal distribution:

$$\mu_i \sim N(\mu_{i0}, \sigma_{i0}^2). \tag{12}$$

3.3. Parameter Estimation with Bayesian-GARCH(1,1).

Assume there are N assets, of which N_1 assets are of T period. Denote the return of asset i as $R_{1,t}$ ($t = 1, 2, \ldots, T$), and all the T variables are i.i.d, following the same multivariate normal distribution. In addition, the rest of the N_2 ($N_2 = N - N_1$) assets are of S periods, and denote their return as $R_{1,t}$ ($t \geq T - S + 1$).

According to the prior assumptions, the posterior of θ_i is

$$\prod_{i=1}^{N} p(\theta_i \mid R_{i,t}) \propto \prod_{i=1}^{N_1} \prod_{t=1}^{T} \left[\sigma_{i,t|t-1}^{-1} \left(1 + \frac{1}{v_i} \frac{(R_{i,t} - \mu_i)^2}{\sigma_{i,t|t-1}^2} \right)^{-(v_i+1/2)} \right]$$

$$\times \prod_{j=1}^{N_2} \prod_{t=s}^{T} \left[\sigma_{j,t|t-1}^{-1} \left(1 + \frac{1}{v_i} \frac{(R_{j,t} - \mu_i)^2}{\sigma_{j,t|t-1}^2} \right)^{-(v_i+1/2)} \right]$$

$$\times \prod_{i=1}^{N} \exp(-v_i\lambda)$$

$$\times \prod_{i=1}^{N} \prod_{i=1}^{N} \exp\left(\left(\frac{-1}{2}(\mu_i - \mu_{i,\gamma}) \right)' \sum_{i,\gamma}^{-1} (\mu_i - \mu_{i,\gamma}) \right)$$

$$\times \prod_{i=1}^{N} I_{i,\theta_G}. \tag{13}$$

Obviously, there is no analytic solution for the joint posterior density function. Therefore, it is acceptable to replace t−distribution with normal distribution before sampling. Suppose the return of asset i follows t−distribution, there is

$$R_{i,t} \mid \sigma_{i,t|t-1}^2 \sim t(\mu_{i,t}, \sigma_{i,t|t-1}^2, v_i). \tag{14}$$

It is equivalent to the normal distribution as follows:

$$R_{i,t} \mid \sigma_{i,t|t-1}^2, \eta_t \sim N\left(\mu_{i,t}, \frac{\sigma_{i,t|t-1}^2}{\eta_t} \right), \tag{15}$$

where η_t is called mixed variables, following i.i.d. Gamma distribution is

$$\eta_t \mid v_i \sim \text{gamma}\left(\frac{v_i}{2}, \frac{v_i}{2} \right), \quad t = 1, \ldots, T. \tag{16}$$

This substitution helps to make the calculation of the posterior easy. After the replacement, the parameter variable of asset i becomes $\theta_i = (w_i, \alpha_i, \beta_i, v_i, \mu_i', \eta_t')$, and its logarithm likelihood function turns to normal logarithm likelihood function:

$$\lg(L(\theta_i \mid R_{1,t}, R_{2,t})) = \text{const} - \frac{1}{2} \sum_{t=t_0}^{T} \left[\lg(\sigma_{i,t|t-1}^2) - \lg(\eta_t) + \frac{\eta_t(R_{i,t} - \mu_i)^2}{\sigma_{i,t|t-1}^2} \right]. \tag{17}$$

According to the equation, when $i = 1, \ldots, N_1, t_0 = 1$ and when $i = 1, \ldots, N_2$, $t_0 = Y - S + 1$. Meantime, the posterior of θ_i contains other mixed variables. The logarithm posterior is

$$\lg(L(\theta_i \mid R_{1,t}, R_{2,t})) = \text{const} - \frac{1}{2} \sum_{t=t_0}^{T} \left[\lg(\sigma_{i,t|t-1}^2) - \lg(\eta_t) + \frac{\eta_t(R_{i,t} - \mu_i)^2}{\sigma_{i,t|t-1}^2} \right]$$

$$- \frac{1}{2}(\mu_i - \mu_{i,\gamma})' \sum_{i,\gamma}^{-1} (\mu_i - \mu_{i,\gamma}) + \frac{Tv_i}{2} \lg \frac{v_i}{2}$$

$$- T \lg\left(\Gamma\left(\frac{v_i}{2} \right) \right) + \left(\frac{v_i}{2} - 1 \right) \sum_{t=1}^{T} \lg(\eta_t)$$

$$- \frac{v_i}{2} t = 1 \sum_{t=1}^{T} \eta_t - v_t\lambda, \tag{18}$$

where $w_i > 0, \alpha_i \geq 0, \beta_i \geq 0$.

3.3.1. Conditional Posterior Distributions

(1) The Conditional Posterior Distribution of μ_i. It can be proved that the full conditional posterior density of μ_i is a normal distribution:

$$p(\mu_i \mid \theta_{i,G}, \eta, v_i, R_{1,t}, R_{2,t}) = N(\mu^*, \Sigma^*). \tag{19}$$

(2) The Conditional Posterior Distribution of η_t. It can be proved that the full conditional posterior density of η_t ($t = 1, \ldots, T$) is a gamma distribution:

$$p(\eta \mid \theta_{i,G}, \mu_i, v_i, R_{i,t}, R_{2,t}) = \text{gamma}\left(\frac{v_i}{2}, \frac{(R_{i,t} - \mu_i)^2}{2\sigma_{i,t|t-1}^2} + \frac{v_i}{2} \right). \tag{20}$$

(3) The Conditional Posterior Distribution of v_i. As there is no standard form for the conditional posterior density of v_i and the degree of freedom (11), the kernel of the posterior distribution is

$$p(v_i \mid \theta_{i,G}, \mu_i, \eta, R_{1,t}, R_{2,t}) \propto \Gamma\left(\frac{v_i}{2} \right)^{-T} \left(\frac{v_i}{2} \right)^{Tv_i/2} \exp(v_i\lambda^*), \tag{21}$$

where $\lambda^* = (1/2)\sum_{t=1}^{T}(\lg(\eta_t) - \eta_t) - \lambda$.

Here, we put sampling with the Griddy Gibbs sampler algorithm into use. The kernel of the conditional logarithm posterior distribution of v_i is

$$\lg\left(p\left(v_i \,\middle|\, \theta_{-v_i}, R_{1,t}, R_{2,t}\right)\right) = \text{const} + \frac{Tv_i}{2}\log\left(\frac{v_i}{2}\right) - T\lg\left(\Gamma\left(\frac{v_i}{2}\right)\right)$$
$$+ \left(\frac{v_i}{2} - 1\right)\sum_{t=1}^{T}\lg(\eta_t) - \frac{v_i}{2}\sum_{t=1}^{T}\eta_t - v_i\lambda. \tag{22}$$

Set the reasonable range of v_i as (2, 30). Divide v_i equally, and denote the grid interval as $(\text{upsilon}_{i1}, \ldots, \text{upsilon}_{ij})$. Sample from the time m, and the other parameters are denoted as $\theta_{-v_i}^{(m-1)}$.

(4) Conditional Posterior Destiny of $\theta_{i,G}$. The posterior distribution kernel of $\theta_{i,G}$ is

$$\lg\left(p\left(\theta_{i,G}\,\middle|\,\theta_{-i,\theta_G}, R_{1,t}, R_{2,t}\right)\right) = \text{const} - \frac{1}{2}\sum_{t=t_0}^{T}\left[\log\left(\sigma_{i,t|t-1}^2\right) + \frac{\eta_t\left(R_{i,t} - \mu_i\right)^2}{\sigma_{i,t|t-1}^2}\right], \tag{23}$$

where $w_i > 0$, $\alpha_i \geq 0$, and $\beta_i \geq 0$. Here, we set a proposal density function for $\theta_{i,G}$, and it follows t−distribution.

3.3.2. Sampling Algorithm of Bayesian-GARCH(1,1). We decompose the parameter variable $\theta_i = (\theta_{i,G}, v_i, \mu_i', \eta_t')$ of asset i, where $\theta_{i,G} = (w_i, \alpha_i, \beta_i)$. Then, we can use a coherent sampling strategy to sample from each conditional posterior distribution:

$$p\left(\theta_{i,G}\,\middle|\,\mu_i, \eta, v_i, R_{1,t}, R_{2,t}\right),$$
$$p\left(v_i\,\middle|\,\theta_{i,G}, \mu_i, \eta, R_{1,t}, R_{2,t}\right),$$
$$p\left(\eta\,\middle|\,\theta_{i,G}, \mu_i, v_i, R_{1,t}, R_{2,t}\right),$$
$$p\left(\mu_i\,\middle|\,\theta_{i,G}, \eta, v_i, R_{1,t}, R_{2,t}\right). \tag{24}$$

3.4. Portfolio Selection Based on BMS-GARCH(1,1). Many researches find that the returns are heteroscedastic, but there are still some other empirical experiments which show that the parameters are not fixed and unchangeable. As for the reasons of changes, it might be due to the potential transfer in different mechanisms during data generating processes. For example, business cycle fluctuations can be seen as an endogenous factor. And, in the regime-switching situation, the Markov regime-switching model is good at describing the dynamics of the returns and variance.

The Markov regime-switching model is proposed by Hamilton in 1989, providing great convenience in modeling the volatility between different states. Then, several researchers try to introduce the method in the GARCH process. For instance, the regime-dependent parameter can be integrated in the standard deviation [44]:

$$R_t = \mu_{t|t-1} + \sqrt{g_{S_t}}\,\sigma_{t|t-1}\varepsilon_t, \tag{25}$$

where S_t is the state at time t. Besides, the regime-dependent parameters can be part of the variance equation [45], that is,

$$\sigma_{t|t-1}^2 = \left(w + g_{S_t}\right) + \sum_{i=1}^{p}\alpha_i u_{t-i}^2. \tag{26}$$

Models of (25) and (26) are both based on the dynamics of the conditional variance since the status dependency makes the likelihood function hard to deal with. According to Henneke et al. [43], the conditional mean can be treated as a ARMA(1,1) process in Bayesian framework, where the volatilities will change in different statuses.

3.4.1. Markov Chain. Suppose that the volatility has three states and denote $i = 1, 2, 3$ as the low volatility, normal volatility, and high volatility. Denote $\pi_{i,j}$ as the transfer probability from state to state. The transfer probability matrix is

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix}, \tag{27}$$

where the sum of each row is 1. And, the property of Markov chain can be shown as follows:

$$P\left(S_t\,\middle|\,S_{t-1}, S_{t-2}, \ldots, S_1\right) = P\left(S_t\,\middle|\,S_{t-1}\right). \tag{28}$$

In matrix (27), each row means the conditional probability distribution of transferring from the realized state S_{t-1} to state S_t, where $\{S_t\}_{t=1}^{T}$ is a tridimensional Markov chain with the transfer probability matrix Pi.

Here, we suppose all the parameters in GARCH(1,1) are state-dependent, and for convenience, the calculation below is for asset j ignoring the subscript and the conditional variance equation is

$$\sigma_{t|t-1}^2 = w\left(S_t\right) + \alpha\left(S_t\right)u_{t-1}^2 + \beta\left(S_t\right)\sigma_{t-1|t-2}^2, \tag{29}$$

and in every stage, there is

$$\left(w\left(S_t\right), \alpha\left(S_t\right), \beta\left(S_t\right)\right) = \begin{cases} \left(w_1, \alpha_1, \beta_1\right), & S_t = 1, \\ \left(w_2, \alpha_2, \beta_2\right), & S_t = 2, \\ \left(w_3, \alpha_3, \beta_3\right), & S_t = 3. \end{cases} \tag{30}$$

As for the return, there is $R_t = \mu_{S_t} + \sigma_{t|t-1}\varepsilon_t$, $t = 1, \ldots, t$. If the given states are the same, the conditional variances are the same according to equations (29) and (7).

3.4.2. Prior Information. Denote the parameter variables of asset and Markov chain $\{S_t\}_{t=1}^{T}$ in BMS-GARCH(1,1) model as

$$\theta = \left(v, \mu', \eta', \theta_{G1}, \theta_{G2}, \theta_{G3}, \pi_1, \pi_2, \pi_3, S\right), \tag{31}$$

where $\theta_{Gi} = (w_i, \alpha_i, \beta_i)$, $i = 1, 2, 3$, $\pi_i = (\pi_{i1}, \pi_{i2}, \pi_{i3})$, $i = 1, 2, 3$, and S is the state path in all the periods, $S = (S_1, S_2, \ldots, S_T)$.

For parameters μ_j, η, and v, they have the non-information prior, which means they are not affected by the states of the BMS-GARCH(1,1) model. The following are for the other parameters.

(1) Prior Distribution of θ_{Gi}. For θ_{Gi} $(i = 1, 2, 3)$, there is a normal distribution prior:

$$\theta_{Gi} \sim N\big(\mu_{\theta_{Gi}}, \Sigma_{\theta_{Gi}}\big) I_{\{\theta_{Gi}\}}, \tag{32}$$

where $I_{\{\theta_{Gi}\}}$ is an indicator function.

(2) Prior Distribution of π_i. In a binomial environment, beta distribution can be a proper prior for probability parameters, and it is called Dirichlet distribution in multivariate situation. Therefore, there is

$$\pi_i \sim \text{Dirichlet}\big(a_{i1}, a_{i2}, a_{i3}\big). \tag{33}$$

In order to get the prior parameters a_{ij}, it is necessary to get the prior of every expected value of each transfer probability and figure out the parameters with the equations.

3.4.3. Parameter Estimation with BMS-GARCH(1,1).
BMS-GARCH(1,1) evolves with invisible status variable S_t $(t = 1, \ldots, T)$. Hence, the discrete Markov Chain is called invisible Markov process as well. However, the parameters can be estimated by Bayesian methods: simulate the invisible variable and the parameters at the same time. Here, the distribution of S is multivariable distribution:

$$
\begin{aligned}
p(S \mid \pi) &= \prod_{t=1}^{T-1} p\big(S_{t+1} \mid S_t, \pi\big) \\
&= \pi_{11}^{n_{11}} \pi_{12}^{n_{12}} \cdots \pi_{32}^{n_{32}} \pi_{33}^{n_{33}} \\
&= \pi_{11}^{n_{11}} \pi_{12}^{n_{12}} \big(1 - \pi_{11} - \pi_{12}\big)^{n_{13}} \cdots \pi_{32}^{n_{32}} \big(1 - \pi_{31} - \pi_{32}\big)^{n_{33}},
\end{aligned}
\tag{34}
$$

where n_{ij} is the number of times the chain transferring from status i to status j between time 1 and time T. The first equation in (34) draws the property of a Markov Chain.

According to the GARCH(1,1) model with t−distribution, the invisible Markov process, π_i and θ_{Gi} $(i = 1, 2, 3)$, the joint log posterior distribution of parameter variables in BMS-GARCH(1,1) is shown as follows:

$$
\begin{aligned}
\lg\big(L\big(\theta \mid R_{1,t}, R_{2,t}\big)\big) = {}& \text{const} - \frac{1}{2} \sum_{t=t_0}^{T} \left[\lg\big(\sigma_{t\mid t-1}^2\big) - \lg\big(\eta_t\big) + \frac{\eta_t \big(R_t - \mu\big)^2}{\sigma_{t\mid t-1}^2} \right] \\
& - \frac{1}{2}\big(\mu - \mu_\gamma\big)' \sum_{\gamma}^{-1} \big(\mu - \mu_\gamma\big) \\
& - \frac{1}{2} \sum_{i=1}^{3} \big(\theta_{Gi} - \mu_{\theta_{Gi}}\big)' \sum_{\theta_{Gi}}^{-1} \big(\theta_{Gi} - \mu_{\theta_{Gi}}\big) I_{\{S(t)=i\}} \\
& + \frac{Tv}{2} \lg\Big(\frac{v}{2}\Big) - T \lg\Big(\Gamma\Big(\frac{v}{2}\Big)\Big) \\
& + \Big(\frac{v}{2} - 1\Big) \sum_{t=1}^{T} \lg\big(\eta_t\big) - \frac{v}{2} \sum_{t=1}^{T} \eta_t - v\lambda \\
& + \sum_{i=1}^{3} \sum_{j=1}^{3} \big(a_{ij} + n_{ij} - 1\big) \lg\big(\pi_{ij}\big),
\end{aligned}
\tag{35}
$$

where $w_i > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$ and i is a certain status. The likelihood function depends on the whole sequence of invisible variable S.

(1) The Conditional Posterior Distribution of π_i. For transfer probability π_i $(i = 1, 2, 3)$, its conditional posterior distribution of π_i is as follows:

$$\lg\big(p\big(\pi_i \mid R_t, \theta_{-\pi_i}\big)\big) = \text{const} + \sum_{j=1}^{3} \big(a_{ij} + n_{ij} - 1\big) \lg\big(\pi_{ij}\big), \tag{36}$$

where $\theta_{-\pi_i}$ means the variable of all the other parameters except π_i. Equation (36) can be seen as the log of kernel of Dirichlet distribution, a_{ij} can be set with a prior, and n_{ij} is the times of $\{S_t\}_{t=1}^{T}$ shifting from status i to j.

(2) The Conditional Posterior Distribution of S. With the three states, the number of potential possible status paths is 3^T. Hence, it is impossible to get the whole variable at a time. But, part of the parameters can be sampled out. In other words, we can do sampling from the full conditional posterior density of S_t. And, the full conditional posterior density is

$$p\big(S_t = i \mid R_t, \theta_{-S}, S_{-t}\big), \tag{37}$$

where θ_{-S} is the variable with S removed and S_{-t} is the status path without the state at time t. With Bayes theorem, $p(S_t = i \mid R_t, \theta_{-S}, S_{-t})$ becomes

$$
\begin{aligned}
p\big(S_t = i \mid R_t, \theta_{-S}, S_{-t}\big) &= \frac{p\big(S_t = i, R_t, S_{-t} \mid \theta_{-S}\big)}{p\big(R_t, S_{-t} \mid \theta_{-S}\big)} \\
&= \frac{p\big(R_t \mid \theta_{-S}, S_{-t}, S_t = i\big) p\big(S_t = i, S_{-t} \mid \theta_{-S}\big)}{p\big(R_t, S_{-t} \mid \theta_{-S}\big)},
\end{aligned}
\tag{38}
$$

where $p(R_t \mid \theta_{-S}, S_{-t}, S_t = i)$ is the likelihood function with $S_t = i$ according to equation (38). Considering Markov properties, there is

$$p\big(S_t = i, S_{-t} \mid \theta_{-S}\big) = p\big(S_t = i, S_{t-1} = l, S_{t+1} = k \mid \theta_{-S}\big) = \pi_{li} \pi_{ik}. \tag{39}$$

The denominator is

$$p\big(R_t, S_{-t} \mid \theta_{-S}\big) = \sum_{s=1}^{3} p\big(S_t = s, R_t, S_{-t} \mid \theta_{-S}\big). \tag{40}$$

Combine equation (38)–(40), then it is possible to figure out the conditional posterior distribution of S_t:

$$p\big(S_t = i \mid R_t, \theta_{-S}, S_{-t}\big) = \frac{p\big(R_t \mid \theta_{-S}, S_{-t}, S_t = i\big) \pi_{li} \pi_{ik}}{\sum_{s=1}^{3} p\big(R_t \mid \theta_{-S}, S_{-t}, S_t = s\big) \pi_{li} \pi_{ik}},$$
$$i = 1, 2, 3. \tag{41}$$

In addition, the proposal density of θ_{Gi} is similar to that in single statuses with GARCH(1,1) and t−distribution.

3.4.4. Sampling Algorithm of Parameter Estimation with the BMS-GARCH(1,1) Model. The sampling procedures are as follows. In the iteration,

(1) Get $\pi_i^{(m)}$ from the posterior density according to equation (36), $i = 1, 2, 3$

(2) Get $S^{(m)}$ with equation (41)

(3) Get $\eta^{(m)}$ by equation (20)

(4) Get $v^{(m)}$ according to equation (21)

(5) Get $\mu_j^{(m)}$ with equation (19)

(6) Get θ_{Gi} from the proposal distribution of θ_{jGi} $(i = 1, 2, 3)$

(7) Check that if each element is satisfying the constraint or continue sampling until all the parameter constraints are satisfied

(8) Calculate the acceptable probability mentioned in Section 2; then decide whether or not θ_{Gi}^* $(i = 1, 2, 3)$ can be accepted

With the new parameters from sampling, update the variable θ. Repeat all the procedures for a certain time until the Markov Chain is convergent.

According to the parameters and Bayes formula, figure out the posterior distribution $p(\mu, V \mid R_{1,t}, R_{2,t})$. The forecast distribution of return R_{T+1} at time $T + 1$ is

$$p(R_{T+1} \mid R_{1,t}, R_{2,t}) \propto \int p(R_{T+1} \mid \mu, V) p(\mu, V \mid R_{1,t}, R_{2,t}) d\mu \, dV. \tag{42}$$

In most cases, there is not an accurate analytic solution to the formula above. But fortunately, we can find a numerical solution by simulation and construct a portfolio with the returns from sampling. Besides, the models can obtain self-financing portfolios if we set $w'1 = 0$ instead of 1 in model (3) as well as removing the restriction of $w_i > 0$ in equation (10) and others. In this way, self-financing portfolios are constructed and the models can be used to estimate the parameters and optimize the portfolios.

4. Numerical Examples

4.1. Data. We select 30 stocks from China stock markets in total, 10 stocks are picked up from GEM board market, and they are from the very beginning of the market, from Nov 6, 2009, to Oct 21, 2011. Besides, 10 stocks are selected from Shanghai stock exchange and 10 from Shenzhen stock exchange. These 20 stocks have a longer history from Jan 7, 2000, to Oct 21, 2011. All the data are weekly. Considering the interest market situation of China, we choose monthly rate of interest that is converted with the three months deposit rate as risk-free rate because the interest rate does not follow the market principle yet. In the following analysis, the sample is taken between Nov 6, 2009, and Oct 21, 2011.

4.2. Parameter Estimation of Bayesian-GARCH(1,1)

4.2.1. Parameter Estimation. We set the freedom degree $v = 5$, which means $\lambda = 0.2$. The initial variance σ_0^2 can be

Table 1: Sharpe ratios with Bayesian-GARCH(1,1).

	Short-selling	No short-selling
Sharpe ratio	0.411	0.097

treated as the variance of residuals. We make 10000 iterations and choose the last 5000 as the posterior. According to the Bayesian methods and models stated above, the uncertain parameters are estimated.

4.2.2. Efficient Frontiers. According to the mean-variance model and with the expected return and covariance matrix, the efficient frontiers are shown as follows (risk-averse parameter $A = 3$). And, the Sharpe ratios are shown in Table 1.

In Figure 1, it can be seen that the efficient frontier without short-selling is below the frontier with short-selling, and the Sharpe ratio is smaller. This suggests short-selling can hedge most nonsystematic risks.

4.3. Parameter Estimation of BMS-GARCH(1,1)

4.3.1. Parameter Estimation. In this model, we suppose that with Dirichlet distribution, the prior parameter $a_{ij} = 1$, $(i, j = 1, 2, 3)$, which means the prior of the transfer probability follows uniform distribution. The prior mean of the parameters in conditional variance is set as $\theta_{G1} = (0.0002 \, 0.1 \, 0.6)'$, $\theta_{G2} = (0.2 \, 0.3 \, 0.4)'$, and $\theta_{G3} = (2 \, 0.6 \, 0.1)'$.

The parameter priors are selected with the following rule: prior mean reflects the different statuses (status 1, status 2, and status 3) of asset. We suppose the sum of α_i and β_i is fixed, and its value fluctuates with statuses. In the case of high volatility, investors' reaction to the unknown information may be sharper than that in the low volatility case, and then α is greater than β. For simplicity, we suppose the prior covariance matrix of θ_{Gi} $(i - 1, 2, 3)$ is a unit matrix. With M-H algorithm running 10000 iterations, keep the last 5000 as the posterior inference. Then, the modified mean and variance in BMS-GARCH(1,1) can be estimated under the posterior.

According to the estimation results, it is easy to find that the values of posterior mean of conditional variance parameters are nearly the same with prior, but it is only satisfied with status 1 and status 2. In the case of status 3, there is $\alpha_3 + \beta_3 > 1$, implying the status is not steady.

We stochastically pick up one stock with short history data (Hanwei Electronics, A_3) and one stock with long history data (Dazhonggongyong, A_{11}). Their state transition matrixes are shown in Table 2, and the state transfer probability is shown in Figures 2 and 3.

From Table 2 and Figure 2, we can see that, for stock A_3 (Hanwei Electronics), its status 3 is a transfer status compared with statuses 1 and 2. Besides stock A_3 is newly public on GEM board, it is interesting that it gradually transfers from state 3 (high volatility) to a steady one, state 1 (low volatility). Instead, with stock A_{11} (Dazhonggongyong; since its data length is longer than that of A_3, here the status
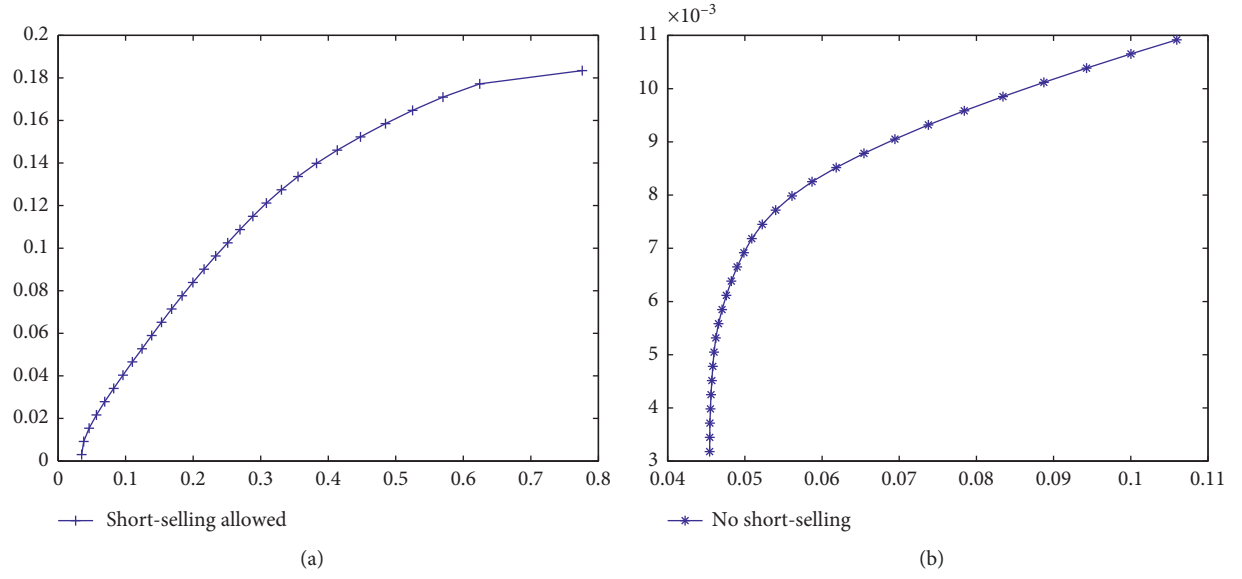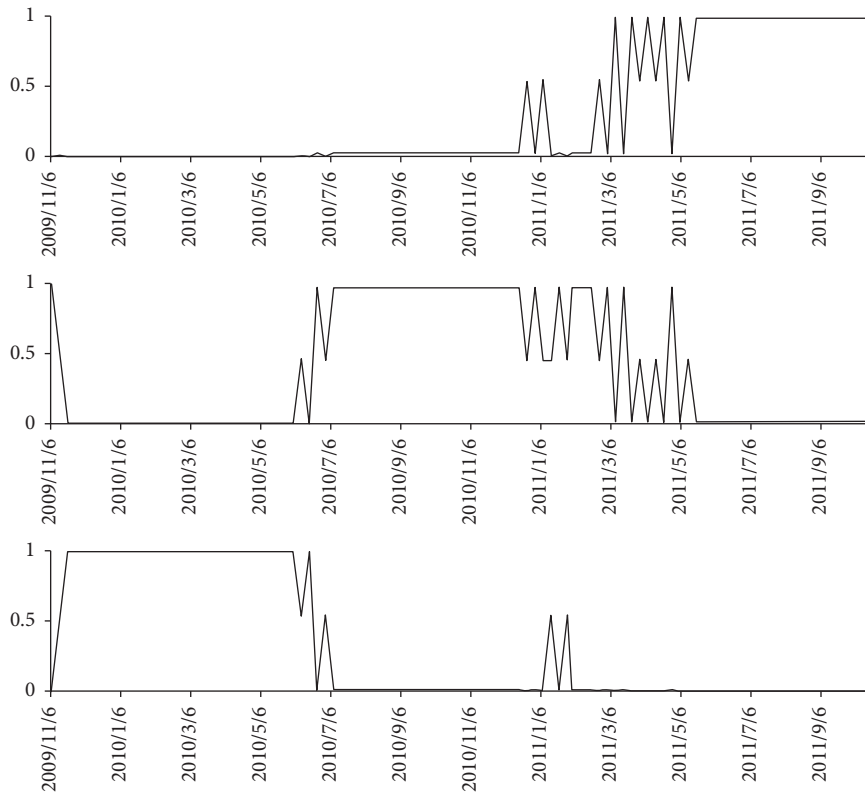
FIGURE 1: Efficient frontiers with (a) short-selling or (b) without short-selling with Bayesian-GARCH(1,1).

TABLE 2: The state transition matrixes of stock A_3 and A_{11}.

	A_3			A_{11}		
	π_{j1}	π_{j2}	π_{j3}	π_{j1}	π_{j2}	π_{j3}
π_{1k}	0.854	0.976	0.013	0.011	0.142	0.004
π_{2k}	0.041	0.029	0.955	0.016	0.936	0.023
π_{3k}	0.103	0.329	0.245	0.426	0.116	0.781

FIGURE 2: Status (statuses 1, 2, and 3) transfer probability of stock A_3.
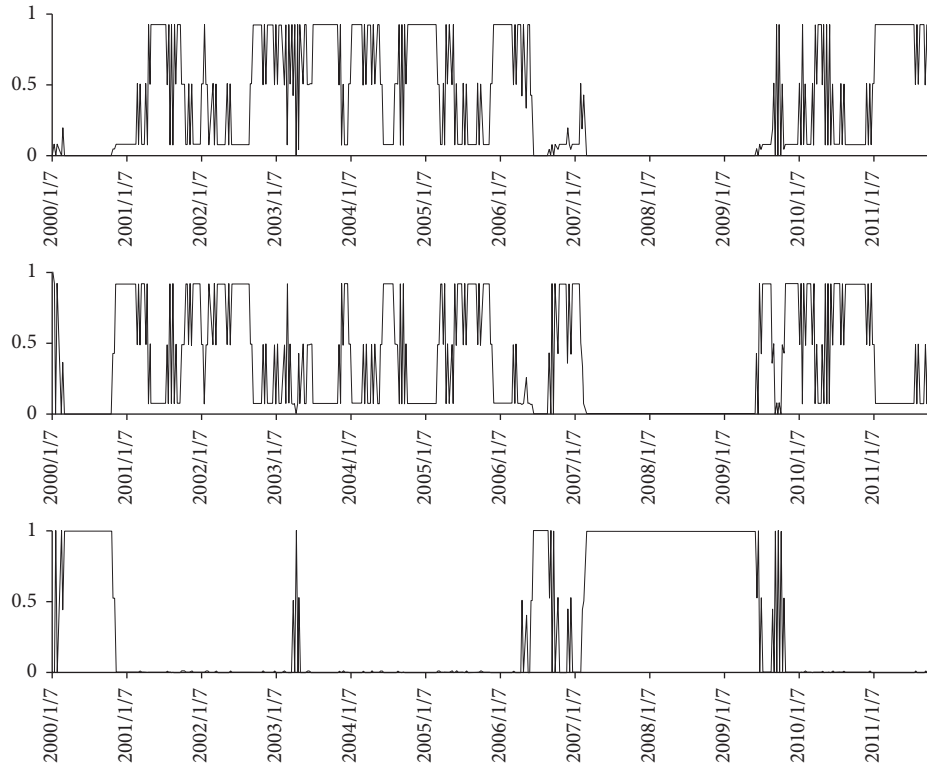
FIGURE 3: Status (statuses 1, 2, and 3) transfer probability of stock A_{11}.

TABLE 3: Sharpe ratios of BMS-GARCH(1,1).

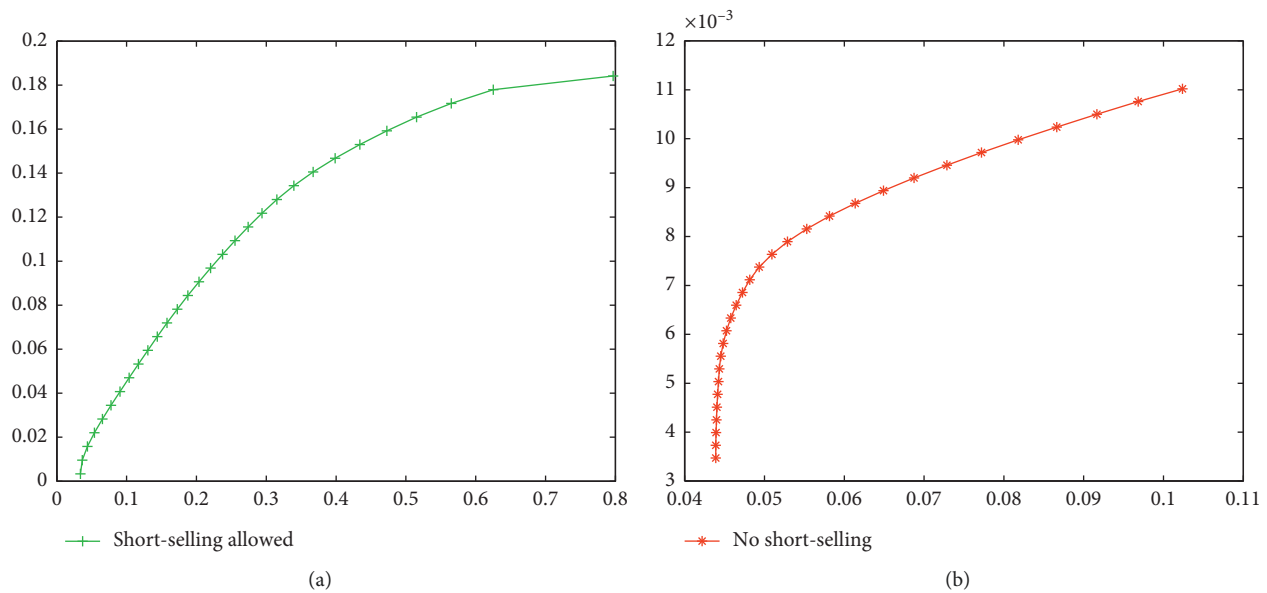|	Short-selling	No short-selling
Sharpe ratio	0.440	0.101

(a)

(b)

FIGURE 4: Efficient frontiers with (a) short-selling or (b) without short-selling with BMS-GARCH(1,1).

transfer sample period is from Jan 7, 2000), when status 3 appears, it will soon transfer to status 1 or status 2. But during the year of 2008, status 3 shows consistency. We believe this is because during the financial crisis, there are more noises in stock market, and the asset is trapped in a state of high volatility.

4.3.2. Efficient Frontier. According to te BMS-GARCH(1,1) model, with the expected return and covariance in last section, the portfolio weights can be calculated (the risk-averse parameter $A = 3$).

Comparing the portfolio weights with BMS-GARCH(1,1) to those with Bayesian-GARCH(1,1), it can be seen that there are obvious differences. We believe that is because there is a state variable in BMS-GARCH(1,1) and the Markov chain contains a lot of information, hence making the results different. Table 3 shows the Sharpe ratios. It can be seen that the Sharpe ratio with BMS-GARCH(1,1) model is much greater than that with Bayesian-GARCH(1,1) and the Sharpe ratios in Section 4.2. It suggests that the state variables overcome the problem of not considering the relationship between long series and short series. In other words, the state variables carry a lot of useful information.

Figure 4 is about the efficient frontier of portfolios under short-selling or without short-selling. It is clear that, in the case of no short-selling, the frontier is below that with short-selling, and the corresponding Sharpe ratio is smaller as well. This also proves that short-selling can hedge some nonsystematic risk, which improves the Sharpe ratio. Comparing Figure 4 with Figure 1, it can be found that the portfolio frontiers based on BMS-GARCH(1,1) are above the others, implying the state variable, the Markov chain, carries useful information and is helpful to the portfolio selection. Therefore, the portfolio selection model based on BMS-GARCH(1,1) performs better than Bayesian-GARCH(1,1).

5. Conclusions

This paper introduces and constructs two models and optimizes portfolios based on them. With the Bayesian-GARCH(1,1) model, the asset returns are assumed to follow t-distribution and the likelihood function can be figured out. Then, with Gibbs sampler, the parameters are estimated and modified.

BMS-GARCH(1,1) can be seen as introducing Markov state variables to the Bayesian-GARCH(1,1). In this model, the parameters are state-dependent and all the states construct a Markov chain. The state changes can be curves with state transfer matrixes, and the parameters can be estimated with the MCMC method. At last, we compare the two models with real stock market data. Numerical results show that portfolios with Bayesian-GARCH(1,1) perform poorer than those with BMS-GARCH(1,1), implying that after introducing state variables, the transferred states carry useful information for the investment portfolio.

People always want to make the optimal financial decision. However, many investors ignore the uncertainties of the parameters and models, which lead to a suboptimal portfolio at last. From this point of view, these models may be of some practical significance and enlightenment. Besides, in the future work, we can try to take other informative priors into consideration, try to expand the models to the multi-stage situation, or even try other frameworks instead of mean-variance framework, such as the utility function, safety-first framework, and so on.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] J. D. Jobson and R. M. Korkie, "Putting Markowitz theory to work," *The Journal of Portfolio Management*, vol. 7, no. 4, pp. 70–74, 1981.

[2] P. Jorion, "International portfolio diversification with estimation risk," *The Journal of Business*, vol. 58, no. 3, pp. 259–278, 1985.

[3] L. Chiodi, R. Mansini, and M. G. Speranza, "Semi-absolute deviation rule for mutual funds portfolio selection," *Annals of Operations Research*, vol. 124, no. 1–4, pp. 245–265, 2003.

[4] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy?," *Review of Financial Studies*, vol. 22, no. 5, pp. 1915–1953, 2009.

[5] F. J. Fabozzi, D. Huang, and G. Zhou, "Robust portfolios: contributions from operations research and finance," *Annals of Operations Research*, vol. 176, no. 1, pp. 191–220, 2010.

[6] D. Huang, S. Zhu, F. J. Fabozzi, and M. Fukushima, "Portfolio selection under distributional uncertainty: a relative robust CVaR approach," *European Journal of Operational Research*, vol. 203, no. 1, pp. 185–194, 2010.

[7] D. Fukushima and G. Zhou, "Bayesian Portfolio Analysis," *Annual Review of Financial Economics*, vol. 2, no. 1, pp. 25–47, 2010.

[8] P. N. Kolm, R. Tütüncü, and F. J. Fabozzi, "60 Years of portfolio optimization: Practical challenges and current trends," *European Journal of Operational Research*, vol. 234, no. 2, pp. 356–371, 2014.

[9] J. T. Williams, "Capital asset prices with heterogeneous beliefs," *Journal of Financial Economics*, vol. 5, no. 2, pp. 219–239, 1977.

[10] V. S. Bawa, S. J. Brown, and R. W. Klein, *Estimation Risk and Optimal Portfolio Choice*, North-Holland Publishing Company, Amsterdam, Netherlands, 1979.

[11] G. Gennotte, "Optimal portfolio choice under incomplete information," *The Journal of Finance*, vol. 41, no. 3, pp. 733–746, 1986.

[12] S. Kandel and R. F. Stambaugh, "On the predictability of stock returns: an asset-allocation perspective," *The Journal of Finance*, vol. 51, no. 2, pp. 385–424, 1996.

[13] S. A. Zenios, M. R. Holmer, R. McKendall, and C. Vassiadou-Zeniou, "Dynamic models for fixed-income portfolio

management under uncertainty," *Journal of Economic Dynamics and Control*, vol. 22, no. 10, pp. 1517–1541, 1998.

[14] M. J. Brennan, "The role of learning in dynamic portfolio decisions," *Review of Finance*, vol. 1, no. 3, pp. 295–306, 1998.

[15] N. Barberis, "Investing for the long run when returns are predictable," *The Journal of Finance*, vol. 55, no. 1, pp. 225–264, 2000.

[16] L. Pástor and R. F. Stambaugh, "Costs of equity capital and model mispricing," *The Journal of Finance*, vol. 54, no. 1, pp. 67–121, 1999.

[17] L. Pástor, "Portfolio selection and asset pricing models," *The Journal of Finance*, vol. 55, no. 1, pp. 179–223, 2000.

[18] Z. Bai, H. Liu, and W. K. Wong, "On the Markowitz mean-variance analysis of self-financing portfolios," *Social Science Electronic Publishing*, vol. 1, no. 1, pp. 35–42, 2009.

[19] Z. Bai, H. Liu, and W. K. Wong, "Enhancement of the applicability of Markowitz's portfolio optimization by utilizing random matrix theory," *Mathematical Finance*, vol. 19, no. 4, pp. 639–667, 2010.

[20] S. A. Zenios, *Financial Optimization*, Cambridge University Press, Cambridge, UK, 2002.

[21] N. Topaloglou, H. Vladimirou, and S. A. Zenios, "A dynamic stochastic programming model for international portfolio management," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1501–1524, 2008.

[22] F.-Y. Chen and S.-L. Liao, "Modelling VaR for foreign-asset portfolios in continuous time," *Economic Modelling*, vol. 26, no. 1, pp. 234–240, 2009.

[23] H. Yao, Z. Li, and S. Chen, "Continuous-time mean-variance portfolio selection with only risky assets," *Economic Modelling*, vol. 36, pp. 244–251, 2014.

[24] R. J. Elliott, T. K. Siu, and A. Badescu, "On mean-variance portfolio selection under a hidden Markovian regime-switching model," *Economic Modelling*, vol. 27, no. 3, pp. 678–686, 2010.

[25] Z. Wang, "A shrinkage theory of asset allocation using asset-pricing models," *SSRN Electronic Journal*, vol. 18, no. 2, pp. 673–705, 2001.

[26] L. MacLean, M. E. Foster, and W. T. Ziemba, "Empirical bayes estimation with dynamic portfolio models," *SSRN Electronic Journal*, 2004.

[27] T. L. Lai, H. Xing, and Z. Chen, "Mean-variance portfolio optimization when means and covariances are unknown," *The Annals of Applied Statistics*, vol. 5, no. 2A, pp. 798–823, 2011.

[28] E. Vilkkumaa, J. Liesiö, and A. Salo, "Bayesian evaluation and selection strategies in portfolio decision analysis," *European Journal of Operational Research*, vol. 233, no. 3, pp. 772–783, 2012.

[29] T. Bodnar, S. Mazur, and Y. Okhrin, "Bayesian estimation of the global minimum variance portfolio," *European Journal of Operational Research*, vol. 256, no. 1, pp. 292–307, 2017.

[30] R. Kan and G. Zhou, "Optimal Portfolio Choice with Parameter Uncertainty," *Journal of Financial and Quantitative Analysis*, vol. 42, no. 3, pp. 621–656, 2007.

[31] D. Niedermayer, *Portfolio Optimization under Parameter Uncertainty*, Springer, Boston, MA, USA, 2008.

[32] A. Kourtis, G. Dotsis, and R. N. Markellos, "Parameter uncertainty in portfolio selection: shrinking the inverse covariance matrix," *Journal of Banking & Finance*, vol. 36, no. 9, pp. 2522–2531, 2012.

[33] V. DeMiguel, A. Martín-Utrera, and F. J. Nogales, "Parameter uncertainty in multiperiod portfolio optimization with transaction costs," *Journal of Financial and Quantitative Analysis*, vol. 50, no. 6, pp. 1443–1471, 2015.

[34] A. A. Santos, A. M. Monteiro, and R. Pascoal, *Portfolio Choice under Parameter Uncertainty: Bayesian Analysis and Robust Optimization Comparison*, GEMF-Faculdade de Economia, Universidade de Coimbra, Coimbra, Portugal, 2014.

[35] L. Pastor and P. Veronesi, "Learning in financial markets," *Annual Review of Financial Economics*, vol. 1, no. 1, pp. 361–381, 2009.

[36] J. Li, M. Li, D. Wu, Q. Dai, and H. Song, "A Bayesian networks-based risk identification approach for software process risk: the context of Chinese trustworthy software," *International Journal of Information Technology & Decision Making*, vol. 15, no. 6, pp. 1391–1412, 2016.

[37] W. Bessler, A. Leonhardt, and D. Wolff, "Analyzing hedging strategies for fixed income portfolios: a Bayesian approach for model selection," *International Review of Financial Analysis*, vol. 46, pp. 239–256, 2016.

[38] J. Liu, *Essays on Risk and Uncertainty in Financial Decision Making: Bayesian Inference of Multi-Factor Affine Term Structure Models and Dynamic Optimal Portfolio Choices for Robust Preferences*, Boston University, Boston, MA, USA, 2014.

[39] N. Metropolis and S. Ulam, "A property of randomness of an arithmetical function," *American Mathematical Monthly*, vol. 60, no. 4, pp. 252-253, 1953.

[40] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[41] H. Derin, H. Elliott, R. Cristi, and D. Geman, "Bayes smoothing algorithms for segmentation of binary images modeled by markov random fields," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 6, no. 6, pp. 707–720, 1984.

[42] L. Bauwens, M. Lubrano, and J. F. Richard, *Bayesian Inference in Dynamic Econometric Models*, OUP Oxford, Oxford, UK, 2000.

[43] J. D. Hamilton and R. Susmel, "Autoregressive conditional heteroskedasticity and changes in regime," *Journal of Econometrics*, vol. 64, no. 1-2, pp. 307–333, 1994.

[44] J. Cai, "A Markov model of switching-regime ARCH," *Journal of Business & Economic Statistics*, vol. 12, no. 3, pp. 309–316, 1994.

[45] J. S. Henneke, S. T. Rachev, F. J. Fabozzi, and M. Nikolov, "MCMC-based estimation of Markov Switching ARMA–CGARCH models," *Applied Economics*, vol. 43, no. 3, pp. 259–271, 2011.