Hindawi
Discrete Dynamics in Nature and Society
Volume 2019, Article ID 8904162, 15 pages
https://doi.org/10.1155/2019/8904162

Research Article

Research on the Value at Risk of Basis for Stock Index Futures Hedging in China Based on Two-State Markov Process and Semiparametric RS-GARCH Model

Liang Wang, Tingjia Xu, Longhao Qin, and Chenge Liu

School of Economics and Business Administration, Xi'an University of Technology, Xi'an 710048, China

Correspondence should be addressed to Liang Wang; wangliang@xaut.edu.cn

Received 8 February 2019; Revised 30 April 2019; Accepted 9 May 2019; Published 2 June 2019

Academic Editor: Luca Pancioni

This article aims to investigate the Value at Risk of basis for stock index futures hedging in China. Since the RS-GARCH model can effectively describe the state transition of variance in VaR and the two-state Markov process can significantly reduce the dimension, this paper constructs the parameter and semiparametric RS-GARCH models based on two-state Markov process. Furthermore, the logarithm likelihood function method and the kernel estimation with invariable bandwidth method are used for VaR estimation and empirical analysis. It is found that the three fitting errors (MSE, MAD, and QLIKE) of conditional variance calculated by semiparametric model are significantly smaller than that of the parametric model. The results of Kupiec backtesting on VaR obtained by the two models show that the failure days of the former are less than or equal to that of the latter, so it can be inferred that the semiparametric RS-GARCH model constructed in this paper is more effective in estimating the Value at Risk of the basis for Chinese stock index futures. In addition, the mean value and standard deviation of VaR obtained by the semiparametric RS-GARCH model are smaller than that of the parametric method, which can prove that the former model is more conservative in risk estimation.

1. Introduction

As the economic globalization and financial innovation are intensifying the income fluctuation of financial markets, Value at Risk (Morgan, 1996) [1] has become one of the important tools to invest and operate for financial institutions and to conduct market supervision by regulators. At the same time, with the launch of stock index futures of CSI 300, SSE 50, and CSI 500 in China, hedging has increasingly become a hot issue for practitioners and regulators. The basic principle of hedging is to utilize the high correlation between the spot market and the futures market, establishing opposite positions in the two markets, and use one market's profit to offset the losses in the other market, thereby achieving the purpose of assets hedge. However, in this process, the risk caused by the basis volatility is the key factor for the hedging effect (Working, 1953) [2], which can be measured by the VaR method.

The traditional theory holds that hedging can reduce the risk because the basis will not change greatly during this process, but, in the actual transaction, the nonsynchronization of the volatility in the futures market and the spot market will also cause basis volatility and generate risks. Hence, many foreign scholars have explored this issue. Working (1953) [2] first defined the concept of basis, which is the spot price minus the futures price, and he believed that the best hedging is to keep the basis remain unchanged. Fraser & Mckaig (2000) [3] conducted a research by using data of Financial Times Stock Index Futures, the basis of U.S. Treasury futures, the difference between three-month Treasury yields in the United States and the United Kingdom. It is found that macroeconomic factors and the investors' expectations to the market will significantly cause basis volatility. Gulley & Tilton (2014) [4] found that when the basis is positive, the demand of investors in futures market will also affect spot and futures prices, but the basis volatility is smaller. Zheng & Huang (2013) [5] constructed a macroeconomic factor model for the basis of commodity futures, and the empirical results showed that factors such as market interest rate, equity risk premium, and futures price change will have significant impacts on the basis, and the impact in the bull market is stronger than the bear market. Fama & French (2015) [6] found that the interest rate change, storage costs, and opportunity costs have different effects on the basis volatility for commodity futures. Broll, Welzel & Wong (2015) [7] pointed out that if the random price is negatively correlated with the expected value of the basis risk, then partial hedging is optimal; if there is a positive correlation, then excessive hedging is optimal, and the optimal position will be uncertain. Zhuang et al. (2016) [8] explored the basis risk for the hedging in the steel futures market and analyzed the impact of macroeconomic factors and micromarket factors on the basis risk empirically, and it was found that the VaR of basis provides a foundation for hedging the risks with parametric, semiparametric, and nonparametric GARCH methods.

Structural state transition is common in the immature financial market. The implementation of major economic policies and the change of financial supervision system may induce the economic state transition. Therefore, it is necessary to apply the volatility model with state transition to estimate and predict the volatility of China financial market and improve the fitting and estimating accuracy of volatility, while the Regime-Switch GARCH (RS-GARCH) model can better estimate the structural transition (Hamilton & Susmel, 1994; Francq & Zakoan, 2005) [9, 10]. Although the GARCH model has prominent effect in measuring the risk of financial market (Engle, 1982; Bollerslev, 1986) [11, 12], it cannot describe the state transition of financial sequences (Hamilton & Susmel, 1994; Francq & Zakoan, 2005) [9, 10]. Many scholars have introduced state transition GARCH model (RS-GARCH) for research. Lamoureux & Lastrapes (1990) [13] found that, due to state changes of return series, the GARCH model will overestimate the volatility. Hamilton & Susmel (1994) [9] and Gray (1996) [14] combined the Markov state transition process with the GARCH model and developed the maximum likelihood estimation method for parameter estimation. Sajjad, Coakley & Nankervis (2008) [15] examined whether the Bayesian MS-GARCH models with two regimes improve the forecasting volatility of VaR model by comparing with their single-regime counterpart. The empirical results showed that Bayesian two-regime MS-GJR-GARCH model with a GED distribution has the best fitting effect to the data based on DIC. Elenjical et al. (2016) [16] assessed the forecasting performance of popular GARCH-based volatility models in the context of VaR estimation and conducted a cross-regime analysis between time periods whereby market conditions experience a shift. Yang & Zhang (2013) [17] applied RS-GARCH and RS-APGARCH models to estimate and forecast the volatility of return series for Shanghai Composite Index and Shenzhen Component Index with Markov regime switching. Peng & Chen (2015) [18] combined the Markov state transition process with the DCC-GARCH model and found that the introduction of Markov state transition process and the range rate of return can effectively improve the estimation accuracy of the hedge ratio.

Due to the short history of stock index futures in China, there are relatively few literatures on the basis risk. In view of this, this paper takes the CSI 300, SSE 50, and CSI 500 stock index futures as the research object, analyzes the formation process of the basis for stock index futures hedging, and further proposes the method to measure the VaR of the basis for long and short hedges. Considering that the RS-GARCH model can effectively solve the structural transition problem of variance in VaR estimation and the Markov process can describe the characteristics of state transition with periodicity for financial variables, this paper intends to calculate the conditional variance of the basis based on Markov process and parametric RS-GARCH model.

Although the multivariate nonparametric regression model can effectively avoid the incorrect setting of the variable distribution in the parametric model, the former also has certain limitations. When there are more explanatory variables, the situation of “dimensional disaster” is prone to occur, such as the sharp increase of variance, the rapid decline of convergence rate of kernel estimation, and local linear estimation. However, the semiparametric model can effectively solve these problems and flexibly deal with the problem of unknown probability density function of variables and disobedient parameter distribution of samples. Furthermore, this paper develops a semiparametric RS-GARCH model based on two-state Markov process to obtain the conditional variance of the basis and uses the log-likelihood function along with the kernel estimation with invariable bandwidth approach to solve the model. Finally, an empirical study is carried out on the VaR of the basis for long and short hedges in three stock index futures.

This paper is organized as follows. Section 1 puts forward the issues concerned in this paper with literature review. Section 2 constructs the parametric and semiparametric RS-GARCH models based on two-state Markov process respectively, and the solution methods of the two models are given. Section 3 analyzes the empirical findings and Section 4 concludes. The innovations of this paper are as follows: (1) It constructs the parametric and semiparametric RS-GARCH models based on two-state Markov process, proposes the measurement method of VaR for long and short hedges, and adopts both the log-likelihood function and kernel estimation with invariable bandwidth approach for model solution; (2) it selects the CSI 300, SSE 50, and CSI 500 stock index futures in China for empirical analysis, and the result shows that the fitting errors (MSE, MAD, and QLIKE) of conditional variance obtained by semiparametric RS-GARCH model are significantly lower than that of the parametric model. Through the Kupiec backtesting on VaR, it is further found that the semiparametric model established in this paper is better in estimating the VaR of the basis.

2. Methodology

This section first analyzes the basis of hedging for stock index futures and then proposes the method of measuring VaR for long and short hedges. Next, the parametric and semiparametric RS-GARCH models are constructed based on two-state Markov process, respectively, and the solution methods of the two models are presented.

2.1. Analysis of the Basis for Stock Index Futures Hedging. The basis is the spread between the spot and the futures price during the hedging process, which plays an important role in the price discovery and information transmission of the futures market. The most perfect hedge is that the basis remains changeless. However, in the real market, the price fluctuations of spot and futures are not synchronized, and the optimal hedge is to ensure the minimization of the basis risk. For the hedge with a hedge ratio of 1, the basis is defined as Eq. (1).

$$B = S - F \tag{1}$$

where B is the basis and S and F denote the spot price and the futures price, respectively. In general, hedge can be divided into short and long hedges. The former refers to buying spot while selling futures at the time of opening a position, and the latter refers to selling spot and buying futures when opening a position. The return of short hedge is

$$S - S_0 + F_0 - F = B - B_0 \tag{2}$$

The return of long hedge is

$$S_0 - S + F - F_0 = B_0 - B \tag{3}$$

In Eq. (2) and Eq. (3), F_0 and S_0 denote the price of futures and spot when opening a position, respectively. For short hedgers, the increase of the basis when closing a position means that the hedge is successful and profitable, while the long hedge is just the opposite.

2.2. The VaR Model for the Basis in Stock Index Futures Hedging

2.2.1. The VaR Model for Long and Short Hedges. Risk management has received much attention from practitioners and regulators in the last few years, with Value at Risk (VaR) emerging as one of the most popular tools. Morgan (1996) [1] proposed the parametric method for VaR, and the zero-value of VaR can be expressed as Eq. (4).

$$VaR = W_0 - W^* = -W_0 R^* = -W_0 \left(\mu_R - \alpha\sigma_R\right) \tag{4}$$

where W^* is the final value of the asset at a certain confidence level, R^* is the minimum yield, μ_R and σ_R denote the mean value and standard deviation of the yield, and α is the quantile. For a long hedge, when the basis becomes larger, the return may be negative, so the VaR of the basis can be expressed as Eq. (5) (Jorion, 2010) [19].

$$VaR_L = B^* - B_0 = \left(\mu + \alpha\sigma\right) - B_0 \tag{5}$$

For a short hedge, when the basis becomes smaller, the return may be negative, so the VaR of the basis can be expressed as Eq. (6).

$$VaR_S = B_0 - B^* = B_0 - \left(\mu - \alpha\sigma\right) \tag{6}$$

In Eq. (5) and Eq. (6), B^* is the basis of the quantile at a certain confidence level, and μ and σ denote the mean value and standard deviation, respectively. It can be seen that how to obtain the standard deviation in Eq. (5) and Eq. (6) is the key step to get VaR. However, for the financial time series, the standard deviation changes with time and has the phenomenon of structural transition. Hence, this paper will develop the parametric and semiparametric RS-GARCH models for investigation.

2.2.2. The Kupiec Backtesting of VaR. The backtesting of VaR is the coverage degree of the model results to the actual loss, and one feasible method is the failure rate test introduced by Kupiec (1995) [20]. Assuming that the estimation of VaR is time-independent, if the actual loss exceeds the VaR, this situation is recorded as failure; otherwise, it is recorded as success. Therefore, the binomial results of failure observation represent a series of independent Bernoulli experiment, and the expected probability of failure is $p^* = 1-\alpha$. In addition, it is assumed that the total number of backtesting days is T, and the number of failure days is N; then the failure probability is $P = N/T$. The null hypothesis is $p = p^*$, and the statistic of the likelihood ratio test for null hypothesis can be expressed as Eq. (7):

$$\begin{aligned} LR = &-2\ln\left[\left(1-p^*\right)^{T-N} p^{*N}\right] \\ &+ 2\ln\left[\left(1-\frac{N}{T}\right)^{T-N}\left(\frac{N}{T}\right)^{N}\right] \end{aligned} \tag{7}$$

where LR obeys the χ^2-distribution when the degree of freedom is 1.

2.3. The Estimation of Conditional Variance Based on Two-State Markov Process and Parametric RS-GARCH Model

2.3.1. Analysis of State Transition for Conditional Variance Based on Two-State Markov Process. For Markov process (1907), the future change of the variable only depends on its current state, not on its past situation. According to this characteristic, the problem of the state transition path for variance in RS-GARCH model can be solved. Assuming that h_t is the conditional variance of the basis at time t, we can obtain

$$\begin{aligned} h_t\left(b_t\right) &= E\left(b_t^2\right) - \left(E\left(b_t\right)\right)^2 \\ &= p_t\left(\mu_{1,t}^2 + h_{1,t}\right) + \left(1-p_t\right)\left(\mu_{2,t}^2 + h_{2,t}\right) \\ &\quad - \left(p_t\mu_{1,t} + \left(1-p_t\right)\mu_{2,t}\right)^2 \\ &= F\left(p_t, \mu_{1,t}, \mu_{2,t}, h_{1,t}, h_{2,t}\right) \end{aligned} \tag{8}$$

Without using Markov process, the variance change is shown in the left diagram of Figure 1, where h_0 denotes the conditional variance when $t = 0$, and $h_{t=1|s_1=2}$ indicates the conditional variance when $t = 1$ with State 2, $s_1 = 1$ is the State 1 when $t = 1$, and so on, $h_{t=2|s_1=2\longrightarrow s_2=1}$ indicates the conditional variance that the State 2 when $t = 1$ converts to the State 1 when $t = 2$. The right diagram in Figure 1 shows the path changes of variance by two-state Markov process. It can be seen that the variance change is converted from the original 2^t paths to only two paths ($t = 1, 2\ldots, N$), which avoids the geometrical growth of the paths and reduces the computational dimension of the model.

In this section, we use two-state Markov process to analyze the probability of the state transition. The first column

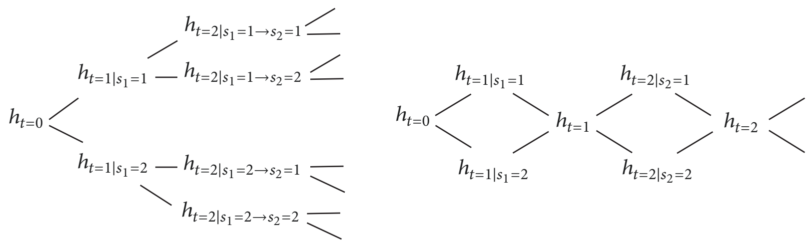

Figure 1: The path change of conditional variance.

Table 1: Two-state Markov process.

t	$t+1$ State 1	State 2
State 1	P	1-P
State 2	1-Q	Q

and the first row in Table 1 give the state of variance at time t and t+1, respectively. If the probability that the variance remains at State 1 from time t to time t+1 is P, then the probability of converting to State 2 at time t+1 is 1-P. Similarly, the probability of the variance remaining at State 2 from time t to time t+1 is Q, and the probability of transforming to State 1 at time t+1 is 1-Q.

After multiple transitions, Markov process will tend to be stable with the steady-state probability. Assuming that the steady-state probability is π_1 and π_2 in State 1 and State 2, respectively, then we can obtain Eq. (9) (Markov, 1907):

$$\begin{aligned}\pi_1 &= \pi_1 \times P + \pi_2 \times (1-Q)\\ \pi_2 &= \pi_1 \times (1-P) + \pi_2 \times Q\end{aligned} \tag{9}$$

From Eq. (9), π_1 and π_2 can be formulated as

$$\pi_1 = 1 - \pi_2, \quad \pi_2 = \frac{1-P}{(1-Q)\times(1-P) - P \times Q + 1} \tag{10}$$

Let $\widetilde{b}_t$ denote the basis of the time series from time 1 to time t, and let $s_t = 1$ and $s_t = 2$ represent State 1 and State 2 at time t, respectively. Since the transition of state is affected by $\widetilde{b}_t$, we have

$$\begin{aligned}&p\left(s_t = 1 \mid \widetilde{b}_{t-1}\right)\\ &= \sum_{a=1}^{2} p\left(s_{t-1} = 1 \mid s_{t-1} = a, \widetilde{b}_{t-1}\right) \times p\left(s_{t-1} = a \mid \widetilde{b}_{t-1}\right)\end{aligned} \tag{11}$$

Further, there is

$$\begin{aligned}p\left(s_t = 1 \mid \widetilde{b}_{t-1}\right) = P \times p\left(s_{t-1} = 1 \mid \widetilde{b}_{t-1}\right) + (1-Q)\\ \times \left(1 - p\left(s_{t-1} = 1 \mid \widetilde{b}_{t-1}\right)\right)\end{aligned} \tag{12}$$

With Figure 1, we can get $p(s_{t-1} = 1 \mid \widetilde{b}_{t-1}) = p(s_{t-1} = 1 \mid b_{t-1}) = p(s_{t-1} = 1)$ at time t-1. By substituting it into Eq. (12), we can develop Eq. (13) according to Bayesian formula:

$$\begin{aligned}p\left(s_{t-1} = 1 \mid \widetilde{b}_{t-1}\right) &= p\left(s_{t-1} = 1 \mid b_{t-1}, \widetilde{b}_{t-2}\right)\\ &= \frac{f\left(b_{t-1} \mid s_{t-1} = 1, \widetilde{b}_{t-2}\right) p\left(s_{t-1} = 1 \mid \widetilde{b}_{t-2}\right)}{\sum_{a=1}^{2} f\left(b_{t-1} \mid s_{t-1} = a, \widetilde{b}_{t-2}\right) p\left(s_{t-1} = a \mid \widetilde{b}_{t-2}\right)}\end{aligned} \tag{13}$$

Hence, the probability of State 1 at time t is

$$\begin{aligned}p_t = (1-Q)\left[\frac{f_{2,t-1}\left(1-p_{t-1}\right)}{f_{1,t-1}p_{t-1} + f_{2,t-1}\left(1-p_{t-1}\right)}\right]\\ + P\left[\frac{f_{1,t-1}p_{t-1}}{f_{1,t-1}p_{t-1} + f_{2,t-1}\left(1-p_{t-1}\right)}\right]\end{aligned} \tag{14}$$

where $f_{1,t-1}$ and $f_{2,t-1}$ are the probability densities of State 1 and State 2 at time t-1, respectively.

2.3.2. The Parametric RS-GARCH Model Based on Two-State Markov Process. According to the study of Hamilton & Susmel (1994) [9], we estimate the conditional variance of basis sequence with two-state Markov process and parametric RS-GARCH model and obtain Eq. (15):

$$\begin{aligned}b_t &= \begin{cases}\mu_{1,t} + \sigma_{1,t}v_t & \text{with probability } p_t\\ \mu_{2,t} + \sigma_{2,t}v_t & \text{with probability } 1-p_t\end{cases}\\ &\qquad v_t \sim i.i.d\ N(0,1)\end{aligned} \tag{15}$$

In Eq. (15), at time t, b_t is the basis; $\mu_{i,t}$ and $\sigma_{i,t}$ represent the conditional mean and standard deviation, where $i = 1, 2$ indicates State 1 and State 2; p_t and $1-p_t$ are the probability of the variance transforming to State 1 and State 2, respectively; v_t is a random variable. According to Eq. (15), the equations of the mean value and conditional variance are

$$\begin{aligned}\mu_{1,t} &= x_1 + x_2 b_{t-1}\\ \mu_{2,t} &= x_3 + x_4 b_{t-1}\end{aligned} \tag{16}$$

$$\begin{aligned}h_{1,t} &= x_5 + x_6\varepsilon_{t-1}^2 + x_7 h_{t-1}\\ h_{2,t} &= x_8 + x_9\varepsilon_{t-1}^2 + x_{10} h_{t-1}\end{aligned} \tag{17}$$

In Eq. (16) and Eq. (17), x_i is the parameter to be estimated, and $h_{1,t}$ and $h_{2,t}$ are the conditional variance of

State 1 and State 2 at time t, where ε_{t-1} is the random error at time t-1. Moreover,

$$\varepsilon_t = b_t - \left[p_t \mu_{1,t} + (1 - p_t)\, \mu_{2,t}\right] \tag{18}$$

and the variance of the basis b_t can be written as Eq. (19):

$$\begin{aligned} h_t(b_t) &= E\left(b_t^2\right) - \left(E(b_t)\right)^2 \\ &= p_t\left(\mu_{1,t}^2 + h_{1,t}\right) + (1 - p_t)\left(\mu_{2,t}^2 + h_{2,t}\right) \\ &\quad - \left(p_t \mu_{1,t} + (1 - p_t)\, \mu_{2,t}\right)^2 \end{aligned} \tag{19}$$

Assuming that the probability densities $f_{1,t}$ and $f_{2,t}$ obey the normal distribution, we have

$$\begin{aligned} f_{1,t} &= \frac{1}{\sqrt{2\pi h_{1,t}}} e^{-(b_t-\mu_{1,t})^2/2h_{1,t}} \\ f_{2,t} &= \frac{1}{\sqrt{2\pi h_{2,t}}} e^{-(b_t-\mu_{2,t})^2/2h_{2,t}} \end{aligned} \tag{20}$$

Let γ_t denote the daily volatility at time t, that is, the percentage of the increase or decrease of the price at time t compared with the time t-1, and $\overline{\gamma}$ is the mean value of γ_t.

Furthermore, we derive the parameters in Eq. (16) and Eq. (17) with the following steps:

Step 1. Use the basis sequence b_t in the sample period to obtain the daily volatility γ_t and the mean value $\overline{\gamma}$. When $\gamma_t > \overline{\gamma}$, let the state at time t be $s_t = 1$; otherwise, $s_t = 2$. Hence, the homogeneous Markov chain is derived.

Step 2. Let $t = 1$.

Step 3. According to equations (16)-(20) and the basis sequence in the sample period, we get the estimated value of the parameters: $\widehat{\mu}_{1,t}$, $\widehat{\mu}_{2,t}$, $\widehat{h}_{1,t}$, $\widehat{h}_{2,t}$, $\widehat{h}_t$, $\widehat{\varepsilon}_t$, $\widehat{f}_{1,t}$, $\widehat{f}_{2,t}$ and $\widehat{p}_t$.

Step 4. Substitute the estimated value in Step 3 into the log likelihood function

$$L(x) = -\sum_{t=1}^{N} \ln\left(f_{1,t}(\mathbf{x})\, p_t + f_{2,t}(\mathbf{x})\,(1 - p_t)\right) \tag{21}$$

Step 5. Use the "fminunc" function in Matlab; when $L(x)$ reaches the minimum value at time t, the estimated value of the parameter vector $\mathbf{x} = (x_1, x_2, \cdots, x_{10})$ can be obtained.

Step 6. Substitute the estimated value of $\mathbf{x} = (x_1, x_2, \cdots, x_{10})$ in Step 5 into equations (16)-(20), and recalculate the estimated value of the parameters: $\widehat{\mu}_{1,t}$, $\widehat{\mu}_{2,t}$, $\widehat{h}_{1,t}$, $\widehat{h}_{2,t}$, $\widehat{h}_t$, $\widehat{\varepsilon}_t$, $\widehat{f}_{1,t}$, $\widehat{f}_{2,t}$ and $\widehat{p}_t$.

Step 7. Let $t = t + 1$.

Step 8. Substitute the estimated value in Step 6 into the log likelihood function in Step 4; repeat Steps 5–8.

Step 9. When $t = N$, stop the loop, and get the final estimated value $\widehat{\mathbf{x}} = (\widehat{x}_1, \widehat{x}_2, \cdots, \widehat{x}_{10})$ of the parameter vector $\mathbf{x} = (x_1, x_2, \cdots, x_{10})$.

Step 10. According to the estimated value of the parameters at time $t = 1, 2, ..., N$, the conditional variance sequence can be obtained from Eq. (19), and the sequence of Value at Risk (VaR) of the basis can be further confirmed by Eq. (5) and Eq. (6).

2.3.3. The Semiparametric RS-GARCH Model Based on Two-State Markov Process. The relationship of the economic variables is unknown in reality, and there are specification errors of the traditional econometric model in practical applications, which cannot satisfy the needs of application research in economy and management. The nonparametric regression model assumes that the relationship of economic variables is unknown, so the model is more realistic, compared with linear and nonlinear regression models. The nonparametric regression model has the following characteristics: the form of regression function can be arbitrary, without any constraints, and the distribution of explanatory as well as explained variables is rarely limited. Hence, the model has greater adaptability. Although the nonparametric regression model has a better fitting effect than the classical regression model, in reality, there are several influencing factors for economic phenomena, and when the number of explanatory variables increases, the convergence rate of kernel estimation and local linear estimation of the multivariate nonparametric regression model will slow down. Since the semiparametric model has the characteristics of parametric and nonparametric models, it can improve the convergence rate of model estimation. On the basis of the previous research, this section will construct a semiparametric RS-GARCH model and propose the estimation method.

(1) The Semiparametric RS-GARCH Model Based on Two-State Markov Process. The conditional variance with two states in Eq. (17) is rewritten as the nonparametric form, while the other equations in Eq. (16) to Eq. (20) are unchanged; then the semiparametric RS-GARCH model is obtained:

$$\begin{aligned} h_{1,t} &= \widehat{m}_1\left(\varepsilon_{t-1}, h_{t-1}\right) \\ h_{2,t} &= \widehat{m}_2\left(\varepsilon_{t-1}, h_{t-1}\right) \end{aligned} \tag{22}$$

Here, Eq. (22) is the nonparametric form of conditional variance equation, where $\widehat{m}_1()$ and $\widehat{m}_2()$ are unknown functions. In this paper, the kernel estimation with invariable bandwidth approach in multivariate parametric regression model is used to estimate the semiparametric RS-GARCH model (Ye, 2003) [21]. Let Y be the explained variable, which is a random variable; X is the d-dimensional explanatory variable vector, and the elements of the vector can be either deterministic or random. Given the samples $(X_1, Y_1), \ldots, (X_N, Y_N)$ and assuming that $[Y_t]$ is independent and identically distributed, then a multivariate nonparametric regression model can be established:

$$Y_t = m\left(X_t\right) + u_t, \quad t = 1, \ldots, N \tag{23}$$

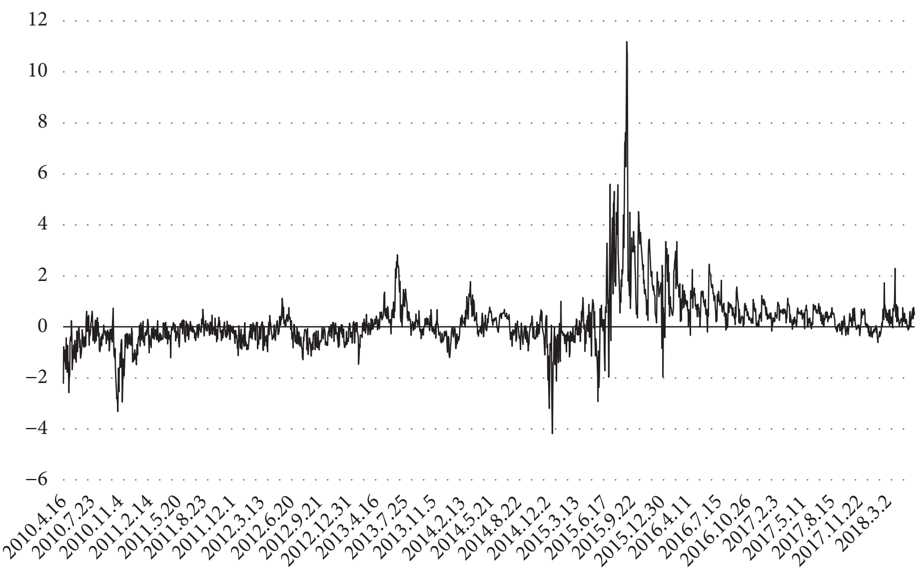

Figure 2: The basis sequence of CSI 300 stock index futures.

where $m()$ is an unknown function; u_t is a random error term, which reflects the influence of other factors on the explained variable as well as the specification error of the model. The estimation of $m()$ can be expressed as

$$\widehat{m}(x, H) = \frac{\sum_{t=1}^{N} K_H (X_t - x) Y_t}{\sum_{t=1}^{N} K_H (X_t - x)} \tag{24}$$

In Eq. (24), H is the bandwidth, $K_H(u) = H^{-d} K(u/H)$, and $K()$ is the d-dimensional density function, where $K(u) \geq 0$ and $\int K(u)du = 1$. According to Ye (2003) [21], this paper uses the following kernel function:

$$K(u) = \frac{d(d+2)}{2s_d}\left(1 - u_1^2 - \cdots - u_d^2\right) \tag{25}$$

where $s_d = 2\pi^{d/2}/\Gamma(d/2)$. Substitute $h_{1,t}$ and $h_{2,t}$ in Eq. (22) into Y_t in Eq. (23), respectively. X is the two-dimensional vector as the form of $(\varepsilon_{t-1}, h_{t-1})$ in Eq. (22). Furthermore, we can use Eq. (23)-Eq. (25) to finish the nonparametric estimation of conditional variance.

(2) The Kernel Estimation with Invariable Bandwidth of Semiparametric RS-GARCH Model. The specific steps for estimating the semiparametric RS-GARCH model by using the method of kernel estimation with invariable bandwidth are as follows.

Step 1. Follow Step 1 to Step 9 in Section 2.3.2; we can obtain the estimated value $\widehat{\mathbf{x}} = (\widehat{x}_1, \widehat{x}_2, \cdots, \widehat{x}_{10})$ of the vector $\mathbf{x} = (x_1, x_2, \cdots, x_{10})$.

Step 2. Substitute $\widehat{\mathbf{x}} = (\widehat{x}_1, \widehat{x}_2, \cdots, \widehat{x}_{10})$ into Eq. (16) and Eq. (17), and get the parameter estimated value $\widehat{\mu}_{1,t}$, $\widehat{\mu}_{2,t}$, $\widehat{\varepsilon}_{t-1}$, $\widehat{h}_{t-1}$, $\widehat{h}_{1,t}$, and $\widehat{h}_{2,t}$.

Step 3. Take $\widehat{\varepsilon}_{t-1}$, $\widehat{h}_{t-1}$, $\widehat{h}_{1,t}$, and $\widehat{h}_{2,t}$ into Eq. (23)-Eq. (25), programming with Matlab (the code is shown in the Appendix) for the kernel estimation with invariable bandwidth, and further have the new parameter estimated value $\widehat{h}_{1,t}$ and $\widehat{h}_{2,t}$.

Step 4. Substitute $\widehat{h}_{1,t}$, $\widehat{h}_{2,t}$ in Step 3 and $\widehat{\mu}_{1,t}$, $\widehat{\mu}_{2,t}$, $\widehat{\varepsilon}_{t-1}$, $\widehat{h}_{t-1}$ in Step 2 into Eq. (19); the conditional variance of the basis b_t can be obtained from the semiparametric RS-GARCH model.

3. Empirical Analysis

3.1. Data Selection. At present, China Financial Futures Exchange (CFFEX) has launched three kinds of stock index futures contracts: CSI 300, SSE 50, and CSI 500, which are taken as the research samples in this paper. Furthermore, due to the limited trading time of futures contracts (seasonal contracts have a relatively long trading time), investors usually choose continuous dominant contract according to the volume of trading. This paper selects the futures contract with the largest daily trading volume as the continuous dominant contract, so that the continuity of the sample can be maintained. The sample range of the CSI 300 stock index futures contract is from April 16, 2010, to May 31, 2018, with a total of 1975 samples, while SSE 50 and CSI 500 are from April 16, 2015, to May 31, 2018, totaling 764 trading days.

3.2. The Descriptive Statistical Analysis. Assume that the closing price of the index is S_t on day t, the closing price of the dominant index contract is F_t, and the basis on day t is $B_t = 100(\ln S_t - \ln F_t)$. Further, we obtain the basis sequence diagrams of CSI 300, SSE 50, and CSI 500, as shown in Figures 2–4. It can be seen that, in Figure 2, before June 2015, the CSI 300 stock index futures contract was basically in the "reverse" market (the basis is less than zero), while it was in the "normal" market (the basis is greater than zero) after June. In Figures 3 and 4, the SSE 50 and CSI 500 stock index futures

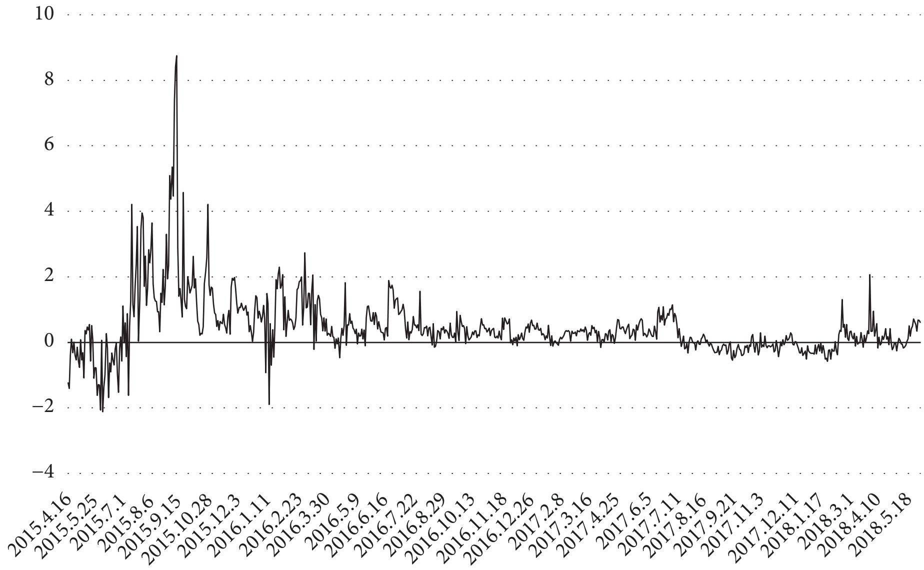

Figure 3: The basis sequence of SSE 50 stock index futures.

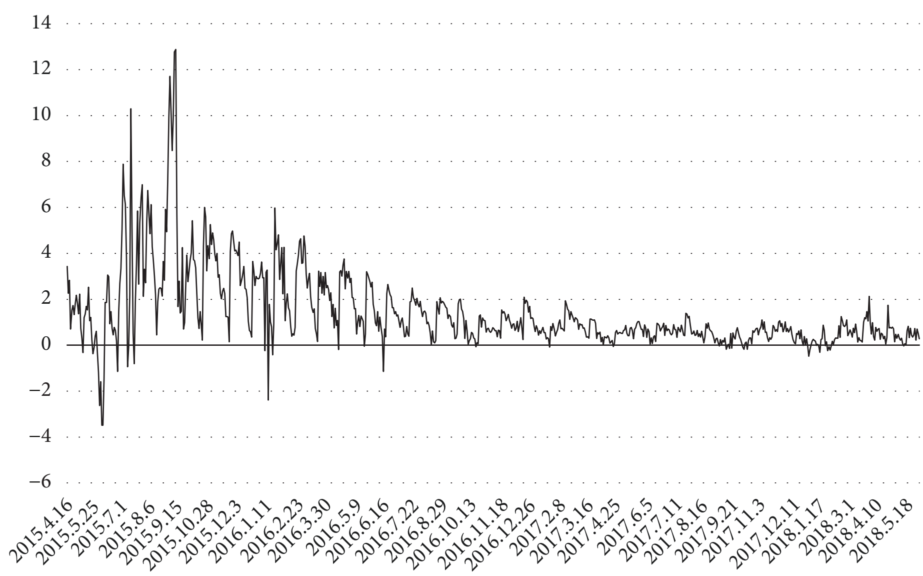

Figure 4: The basis sequence of CSI 500 stock index futures.

contracts were basically in a "normal" market situation after June 2015. This is probably because the stock index futures trading was affected by Chinese stock market crash in 2015. In addition, the basis sequences have an obvious cluster effect in Figures 2–4; that is, a large (small) fluctuation is accompanied by a large (small) fluctuation. Particularly, in the early stage of futures listing and before June 2015, the basis sequences showed strong volatility agglomeration. In addition, the trend of basis sequences for stock index futures has a strong consistency, which showed a tendency of increasing first and then decreasing in 2015. For the convenience of the following study, this paper divides the volatility of basis into two states, high and low volatility, respectively. Specifically, the mean values $\overline{\gamma}$ of the daily volatility γ_t for the basis sequence b_t of three futures are obtained separately, where γ_t refers to the percentage of the increase or decrease of the price at time t compared with the time t-1. When $\gamma_t > \overline{\gamma}$ on day t, we record the state as high volatility; otherwise we set it as low volatility.

Table 2 shows the descriptive statistical results of the basis for the three futures. It can be seen that the mean value of the basis is greater than zero, and the standard deviation is between 0.9 and 2. The JB statistic is larger and the P-value is zero, indicating that the basis sequences disobey the normal distribution. Moreover, the skewness is greater than zero, and the kurtosis is larger, which suggests that the basis sequences have the characteristics of leptokurtosis and fat-tail with right skew. In Table 3, the ADF statistic is lower than the critical value at each significance level, and the P-value is

Table 2: The descriptive statistics of basis sequences for stock index futures.

Stock index futures	Mean value	Standard deviation	Skewness	Kurtosis	JB statistics	P-value
CSI 300	0.205	1.061	2.717	22.22	32816.660	0.000
SSE 50	0.471	0.964	3.135	22.03	12778.254	0.000
CSI 500	1.414	1.727	2.439	12.849	3845.504	0.000

Table 3: The ADF and ARCH-LM test of basis sequences for stock index futures.

Stock index futures	ADF test		ARCH-LM test	
	t-Statistic	Prob	Obs∗R-squared	Prob
CSI 300	-7.866	0.000	110.228	0.000
SSE 50	-6.299	0.000	64.925	0.000
CSI 500	-8.503	0.000	50.588	0.000

Table 4: The state transition statistics of basis sequences for stock index futures.

CSI 300	State 1	State 2	SSE 50	State 1	State 2	CSI 500	State 1	State 2
State 1	801 times	307 times	State 1	136 times	70 times	State 1	174 times	63 times
State 2	306 times	200 times	State 2	69 times	127 times	State 2	63 times	102 times

Table 5: The steady-state probability and state transition probability of basis sequences for stock index futures.

Stock index futures	Probability			
	π_1	π_2	P	Q
CSI 300	0.766	0.234	0.827	0.435
SSE 50	0.835	0.165	0.907	0.532
CSI 500	0.765	0.235	0.788	0.307

zero, illustrating that the original hypothesis is rejected and the sequences have no unit root. Therefore, we can deduce that the three basis sequences are stationary. Table 3 also gives the ARCH-LM test results of the basis sequences, and the P-value is zero, which means that there are significant heteroscedasticity effects in basis sequences, so the GARCH model can be used to measure the variance of basis sequences.

3.3. The Estimation of Parametric and Semiparametric RS-GARCH Models Based on Two-State Markov Process. In this section, we divide the basis sequences into two states and establish a two-state Markov chain. Table 4 shows the transition times between various states of basis sequences during the sample period.

Calculating the number of four state transitions in Table 4, we divide each transition time by the total transition times to obtain transition probability P and Q. Further, the steady-state probability π_1 and π_2 can be derived according to Eq. (9). The results are shown in Table 5. It can be seen that π_1 are all greater than π_2, which indicates that the basis is likely to be at a low volatility state. In addition, as the study in Section 2.3.1, P are all greater than Q for the three basis sequences in Table 5; it indicates that the basis at the low volatility state remains unchangeable with the characteristic of stronger continuity. This is also in line with the fact that the bull market in Chinese securities market has a short duration with high volatility and the bear market lasts for a long time with small price fluctuation. In addition, the steady-state probability and state transition probability shown in Table 5 will be taken as the initial parameter values for solving the RS-GARCH model later.

Further, combined with the estimation methods of Steps 1–10 in Section 2.3.2, we program with Matlab to derive the parameters of RS-GARCH model (detailed code in Appendix), and the results are shown in Table 6.

In Table 6, x_6, x_7, x_9, x_{10} are all greater than zero, $x_6 + x_7 < 1, x_9 + x_{10} < 1$, which ensures the stationarity of the conditional variance (Bollerslev, 1986) [12]. x_9 and x_{10} represent the high volatility state; x_6 and x_7 indicate the low volatility state. $x_9+x_{10} > x_6+x_7$ shows that the high volatility state is more susceptible to the previous state than the low volatility state (Bollerslev, 1986) [12]. According to Table 6, we estimate the conditional variance of the basis by using Steps 1–4 of the semiparametric RS-GARCH model in Section 2.3.3 (2). Due to space limitations, we only give the descriptive statistical results in Table 7. It can be seen that the conditional variance of SSE 50 is smaller, while that of CSI 300 is larger.

Furthermore, we adopt three error measure methods, which are mean square error (MSE), mean absolute error (MAD), and error of Gaussian quasi-maximum likelihood loss function (QLIKE), to assess the fitting effect of models. And we introduce the GARCH model for comparison, where $GARCH = c(1) + c(2)RESIDE(-1)^2 + c(3)GARCH(-1)$. It is found that $c(2)$ and $c(3)$ are significant, and $c(2)$ +

Table 6: The estimation of parameters of RS-GARCH model.

The parameters of RS-GARCH model	CSI 300 stock index futures	SSE 50 stock index futures	CSI 500 stock index futures
x_1	0.049 ∗ ∗∗ (10.005)	0.047 ∗ ∗∗ (4.003)	0.052 ∗ ∗∗ (2.996)
x_2	0.393 ∗ ∗∗ (40.065)	0.807 ∗ ∗∗ (23.966)	0.871 ∗ ∗∗ (29.950)
x_3	0.060 ∗ ∗∗ (18.909)	0.074 ∗ ∗∗ (6.739)	0.111 ∗ ∗∗ (4.756)
x_4	0.901 ∗ ∗∗ (28.0127)	0.920 ∗ ∗∗ (12.074)	0.973 ∗ ∗∗ (15.209)
x_5	−0.019 ∗ ∗∗ (−23.698)	−0.006∗ (−1.4339)	−0.044 ∗ ∗∗ (−4.986)
x_6	0.396 ∗ ∗∗ (61.672)	0.391 ∗ ∗∗ (9.330)	0.352 ∗ ∗∗ (3.960)
x_7	0.067 ∗ ∗∗ (76.465)	0.173 ∗ ∗∗ (4.115)	0.168 ∗ ∗∗ (4.089)
x_8	9.999 ∗ ∗∗ (36.156)	1.986 ∗ ∗∗ (8.311)	1.971 ∗ ∗∗ (12.821)
x_9	0.400 ∗ ∗∗ (28.824)	0.493 ∗ ∗∗ (12.645)	0.508 ∗ ∗∗ (16.241)
x_{10}	0.299 ∗ ∗∗ (19.059)	0.295 ∗ ∗∗ (12.343)	0.299 ∗ ∗∗ (13.307)

Note: ∗, ∗∗, and ∗ ∗ ∗ mean that the parameters are significant at significance levels of 10%, 5%, and 1%, respectively, and it is t-statistic in parentheses.

Table 7: The estimation of conditional variance for parametric and semiparametric RS-GARCH models.

Stock index futures	Model	Mean value	Standard deviation	Maximum	Minimum
SSE 50	Parametric RS-GARCH	0.556	0.841	9.970	0.235
	Semi-parametric RS-GARCH	0.522	0.668	7.353	0.279
CSI 500	Parametric RS-GARCH	1.023	1.715	28.663	0.488
	Semi-parametric RS-GARCH	0.9705	1.001	18.758	0.729
CSI 300	Parametric RS-GARCH	2.461	1.419	26.839	0.934
	Semi-parametric RS-GARCH	2.409	1.225	27.455	0.933

Table 8: The parameter estimation results of GARCH model.

Stock index futures	GARCH parameters $c(1)$	$c(2)$	$c(3)$
SSE 50	0.062(0.000)	0.404(0.000)	0.556(0.000)
CSI 500	0.687(0.018)	0.600(0.000)	0.347(0.000)
CSI 300	0.303(0.000)	0.542(0.000)	0.388(0.000)

$c(3) < 1$ indicates that the conditional variance sequences are stationary and predictable, so GARCH (1,1) can be used for regression (Bollerslev, 1986) [11]. The parameter estimation results are shown in Table 8. Where MSE $= (1/n)\sum_{t=1}^{n}(\sigma_t - \widehat{\sigma}_t)^2$, MAD $= (1/n)\sum_{t=1}^{n}|\sigma_t - \widehat{\sigma}_t|$, QLIKE $= (1/n)\sum_{t=1}^{n}(\ln \widehat{\sigma}_t^2 + \sigma_t^2/\widehat{\sigma}_t^2)$, σ_t^2 is the square value of the difference between the basis on day t and the mean value of basis from day 1 to day t, and $\widehat{\sigma}_t^2$ is the estimated value of the conditional variance by using the method in Section 2.3.2.

In addition, we use the parametric and semiparametric RS-GARCH models to estimate the conditional variance of basis sequences for the three futures, calculate the fitting error, and use the GARCH model for comparison, which are shown in Figures 5–7. In the figures, the ordinate is the result of the fitting error, and the abscissa is the fitting error method. For the convenience of observation, the errors of different models are connected by different curves, and the three curves are the fitting errors of GARCH model and parametric and semiparametric RS-GARCH, respectively. It can be seen that the semiparametric model has the best fitting effect, followed by the parametric model, while GARCH model has the worst fitting effect. For the three futures, the fitting errors

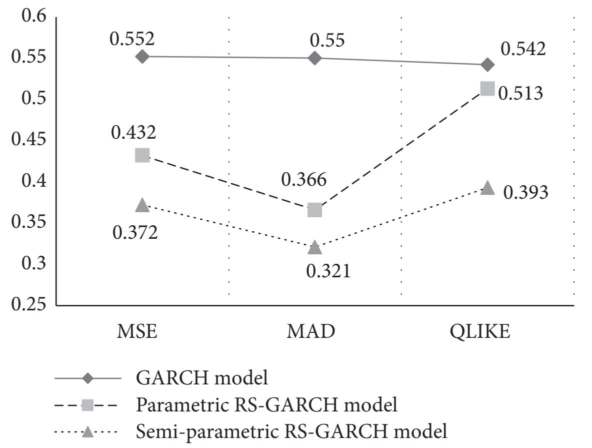

FIGURE 5: The model fitting error results in sample period (SSE 50).

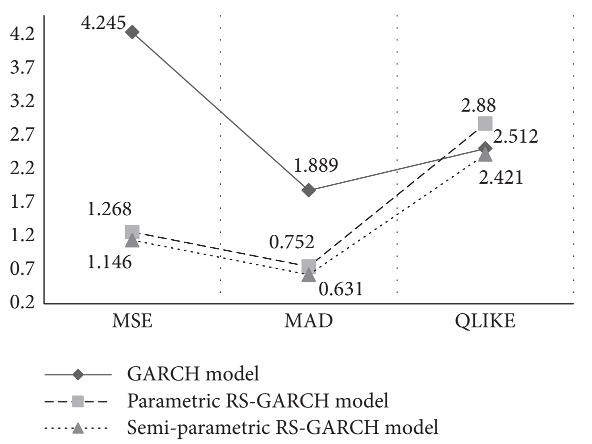

FIGURE 6: The model fitting error results in sample period (CSI 500).

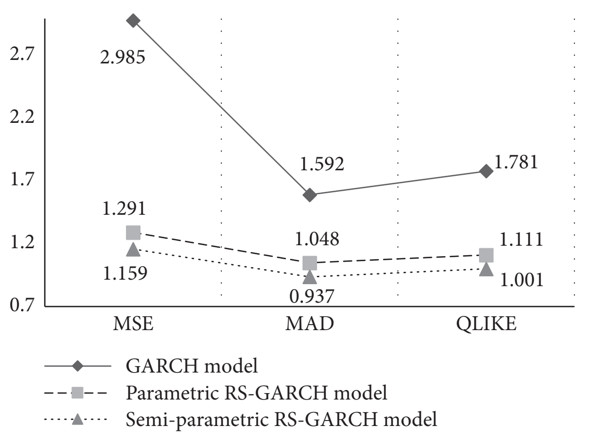

FIGURE 7: The model fitting error results in sample period (CSI 300).

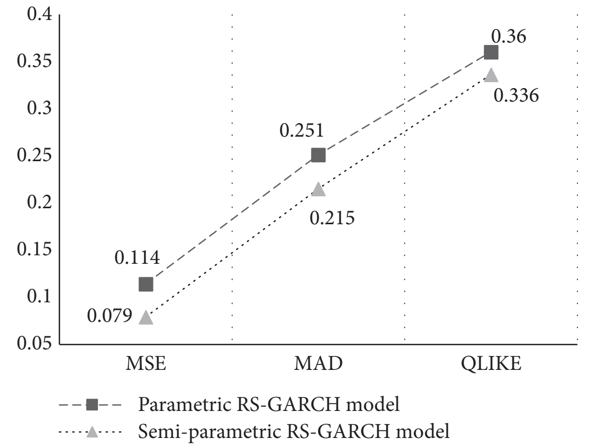

FIGURE 8: The model fitting error results out of sample period (SSE 50).

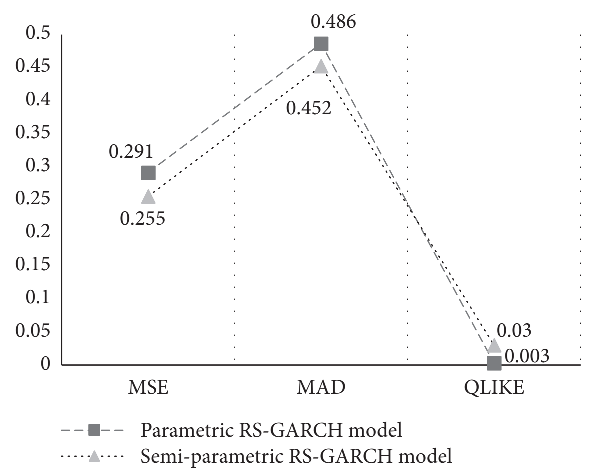

FIGURE 9: The model fitting error results out of sample period (CSI 500).

(MSE, MAD, and QLIKE) of the semiparametric RS-GARCH model are lower than the parametric method. Further, we select the transaction data from June 1, 2018, to September 28, 2018, as out-of-sample data and compare the fitting effect of the two RS-GARCH models (Figures 8–10); the results show that the semiparametric model is still better. In summary, it can be found that the semiparametric RS-GARCH model constructed in this paper improves the estimation accuracy of conditional variance of the basis, so this model can better estimate the daily volatility of stock index futures in China.

3.4. Estimation and Test of VaR of the Basis for Stock Futures Hedging. According to the conditional variance of basis sequences obtained in Section 3.2, we use Eq. (5) and Eq. (6) in Section 2.2 to obtain the VaR sequences for long and short hedges on day *t*. Table 9 gives the descriptive statistical results of VaR sequences for stock index futures in China. It can be seen that VaR calculated by the semiparametric method is smaller than that of the parametric method, which indicates that the semiparametric method is more conservative in risk estimation. Moreover, as can be seen from Table 9, the standard deviation of VaR calculated by semiparametric method is smaller than that of parametric model as a whole, which also shows that the result by adopting semiparametric method is more stable.

Table 9: The estimation results of VaR of the basis for stock index futures.

Stock index futures	Long/Short	Model	Confidence level	Mean value	Standard deviation	Maximum	Minimum
SSE 50	VaR_L	RS-GARCH	95%	1.537	1.179	10.908	0.039
			97.5%	1.747	1.265	11.759	0.243
			99%	1.991	1.368	12.748	0.479
		Semi-parametric RS-GARCH	95%	1.533	1.107	11.825	0.082
			97.5%	1.742	1.177	12.680	0.091
			99%	1.985	1.262	13.673	0.332
	VaR_S	RS-GARCH	95%	0.654	0.735	5.303	0.003
			97.5%	0.863	0.766	6.017	0.021
			99%	1.107	0.817	6.846	0.023
		Semi-parametric RS-GARCH	95%	0.651	0.729	5.760	0.008
			97.5%	0.860	0.747	6.525	0.012
			99%	1.103	0.780	7.413	0.015
CSI 500	VaR_L	RS-GARCH	95%	2.889	2.012	18.163	0.075
			97.5%	3.451	2.117	19.850	0.248
			99%	3.535	2.244	21.809	0.662
		Semi-parametric RS-GARCH	95%	2.830	1.791	15.028	0.121
			97.5%	3.188	1.846	15.684	0.405
			99%	3.117	1.913	16.446	0.066
	VaR_S	RS-GARCH	95%	0.225	1.397	7.901	0.005
			97.5%	0.523	1.404	9.015	0.009
			99%	0.871	1.415	10.310	0.024
		Semi-parametric RS-GARCH	95%	0.166	1.343	5.117	0.004
			97.5%	0.453	1.345	5.653	0.005
			99%	0.787	1.365	6.564	0.032
CSI 300	VaR_L	RS-GARCH	95%	2.693	0.949	15.107	1.803
			97.5%	3.179	1.031	16.739	2.314
			99%	3.742	1.129	18.635	2.668
		Semi-parametric RS-GARCH	95%	2.677	0.902	15.536	1.509
			97.5%	3.159	0.973	17.187	2.150
			99%	3.720	1.0576	19.105	2.667
	VaR_S	RS-GARCH	95%	2.374	0.503	13.633	1.169
			97.5%	2.859	0.540	15.210	1.474
			99%	3.423	0.599	17.041	1.828
		Semi-parametric RS-GARCH	95%	2.358	0.4596	7.375	0.293
			97.5%	2.840	0.4783	8.306	1.431
			99%	3.401	0.5153	9.387	1.827

Note: VaR_L denotes the VaR of the basis for the long hedge, and VaR_S is the VaR of the basis for the short hedge.

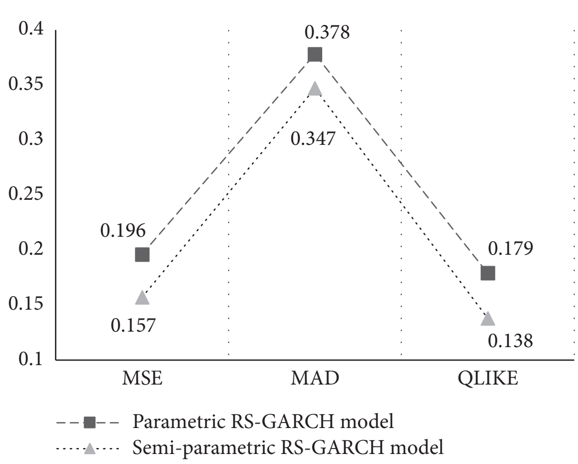

Figure 10: The model fitting error results out of sample period (CSI 300).

Furthermore, we adopt the Kupiec backtesting method to estimate the VaR of the basis for the three stock index futures, and the results are shown in Table 10. It can be found that, in the acceptance domain, the lower the number of failures, the better the prediction effect of the model. However, if the failure rate is excessively low, it means that the model is too conservative and investors may miss investment opportunities. In Table 10, the acceptance domain of LR statistic is (0.00016,6.635) at the significance level of 1%, (0.001,5.024) at the significance level of 2.5%, and (0.004,3.841) at the significance level of 5%. It can be seen that, at 95% confidence level, a few LR statistics fall into the rejection domain, indicating that both the parametric and semiparametric RS-GARCH models can reasonably estimate the VaR of the basis. Further, Table 10 shows that the higher the confidence level, the fewer the times of failing the test. At three different confidence levels, the failure times of the

Table 10: The Kupiec backtesting of VaR of the basis for stock index futures.

Stock index futures	Long/Short	Model	Confidence level	Failure days	Failure rate	LR statistics
SSE 50	VaR_L	RS-GARCH	95%	22	2.88%	8.480∗
			97.5%	14	1.83%	1.537
			99%	9	1.18%	0.231
		semi-parametric RS-GARCH	95%	20	2.62%	10.969∗
			97.5%	13	1.70%	2.246
			99%	5	0.65%	1.050
	VaR_S	RS-GARCH	95%	13	1.70%	23.24∗
			97.5%	9	1.18%	6.792∗
			99%	7	0.92%	0.056
		semi-parametric RS-GARCH	95%	12	1.57 %	25.544∗
			97.5%	6	0.79%	12.533∗
			99%	3	0.39%	3.6997
CSI 500	VaR_L	RS-GARCH	95%	38	4.97%	0.001
			97.5%	27	3.53%	2.976
			99%	15	1.96%	5.591
		semi-parametric RS-GARCH	95%	36	4.71%	0.136
			97.5%	26	3.40%	2.301
			99%	15	1.96%	5.591
	VaR_S	RS-GARCH	95%	18	2.36%	13.868∗
			97.5%	12	1.57%	3.112
			99%	7	0.92%	0.055
		semi-parametric RS-GARCH	95%	14	1.83%	21.093∗
			97.5%	8	1.05%	8.441∗
			99%	5	0.65%	1.049
CSI 300	VaR_L	RS-GARCH	95%	89	4.51%	1.0466
			97.5%	44	2.23%	0.622
			99%	24	1.22%	0.865
		semi-parametric RS-GARCH	95%	85	4.30%	2.111
			97.5%	42	2.13%	1.1891
			99%	22	1.11%	0.249
	VaR_S	RS-GARCH	95%	77	3.9%	5.434∗
			97.5%	36	1.82%	4.096
			99%	11	0.56%	4.633
		semi-parametric RS-GARCH	95%	74	3.75%	7.123∗
			97.5%	35	1.77%	4.770
			99%	11	0.56%	4.633

Note: ∗ indicates that the LR statistic falls within the rejection domain.

semiparametric RS-GARCH model are less than or equal to that of the parametric RS-GARCH model, which shows that the former model can effectively improve the estimation accuracy of VaR of the basis for stock index futures in China. In addition, at 95% confidence level, some LR statistics are larger, falling into the rejection domain, and the failure rate is far lower than 5%, which indicates that the model at this confidence level overestimates the risk to a certain extent. Moreover, we find that the VaR of the basis for Chinese stock index futures can better characterize the leptokurtosis and fat-tail of the basis sequences. Therefore, it is recommended to use the confidence level of 97.5% and 99% for VaR estimation. It can be further deduced that the semiparametric RS-GARCH model can effectively improve the measurement accuracy of the basis risk for Chinese stock index futures.

4. Conclusion

This paper aims to investigate the VaR measurement of the basis for CSI 300, SSE 50, and CSI 500 stock index futures in China. Since the RS-GARCH model can effectively deal with the structural transition of variance and the two-state Markov process can transform the original 2^N state transition paths into two paths, thus effectively reducing the computational dimension and complexity, this paper firstly constructs a parametric RS-GARCH model based on two-state Markov process to obtain the conditional variance of the basis. If there are more explanatory variables, the convergence rate of kernel estimation and local linear estimation will decrease sharply when using the multivariate nonparametric regression method. In view of this, this paper constructs the semiparametric RS-GARCH model based on two-state Markov process to estimate the conditional variance of the basis and adopts the log-likelihood function and kernel estimation with invariable bandwidth method for model solution. On this basis, this paper calculates the VaR of long and short hedges for Chinese stock index futures and carries out the Kupiec backtesting.

Through empirical research, the conclusions are as follows.

First, the basis volatility of hedging has obvious agglomeration for stock index futures in China, especially in the early stage of futures listing and before June 2015, while the basis is basically in the "normal" market after June 2015 (the basis is greater than zero).

Second, this paper divides the basis of stock index futures hedging into two states of high and low volatility and establishes a two-state Markov process, which finds that the duration of the basis sequence in a low volatility state is longer than that of the high volatility state in China.

Third, the parametric and semiparametric RS-GARCH models based on two-state Markov process can better describe the characteristics of leptokurtosis and fat-tail in Chinese stock index futures market; comparing the three fitting errors (MSE, MAD, and QLIKE) of the two models, we can find that the semiparametric RS-GARCH model is better.

Fourth, the mean value and standard deviation of VaR calculated by semiparametric RS-GARCH model are smaller than that of the parametric model, which indicates that the former is more conservative in risk estimation.

Fifth, the Kupiec backtesting shows that the number of failure days of VaR based on semiparametric RS-GARCH model is less than or equal to that of the parametric model, which demonstrates that the semiparametric RS-GARCH model is better, and it can avoid the incorrect setting of the model, thus making the estimation result of VaR closer to the actual loss value.

Appendix

```
%%%%%%rs-garch parametric estimation %%%%
function x= test1 (b,P,Q)
h1(1)=18969.3325;
h2(1)=20765.5881;
u1(1)=173.8939;
u2(1)=105.0945;
p(1)=0.5896;
h(1)=p(1) * (u1(1)^2 + h1(1)) + (1-p(1)) * (u2(1)^2 + h2(1)) - (p(1) * u1(1) + (1-p(1)) * u2(1))^2;
e(1)=b(1) - (p(1) * u1(1) + (1-p(1)) * u2(1));
f1(1)=exp((-(b(1) - u1(1))^2)/(2*h1(1)))/sqrt(2*pi*h1(1));
f2(1)=exp((-(b(1) - u2(1))^2)/(2*h2(1)))/sqrt(2*pi*h2(1));
s = @(x) 0;
for i=2:length(b)
    p(i)=(1-Q) * f2(i-1) * (1-p(i-1))/(f1(i-1) * p(i-1) + f2(i-1) * (1-p(i-1))) + P * f1(i-1) * p(i-1)/(f1(i-1) * p(i-1) + f2(i-1) * (1-p(i-1)));
    h1i=@(x) x(1) + x(2) * e(i-1)^2 + x(3) * h(i-1);
    h2i=@(x) x(4) + x(5) * e(i-1)^2 + x(6) * h(i-1);
    u1i=@(x) x(7) + x(8) * b(i-1);
    u2i=@(x) x(9) + x(10) * b(i-1);
    f1i=@(x) exp ((-(b(i) - u1i(x))^2)/(2*h1i(x)))/sqrt(2*pi*h1i(x));
    f2i=@(x) exp ((-(b(i) - u2i(x))^2)/(2*h2i(x)))/sqrt(2*pi*h2i(x));
    s=@(x) s(x) + log(f1i(x) * p(i) + f2i(x) * (1-p(i)));
    fun= @(x) -s(x);
    x0 = [28.3240 0.1874 0.1023 168.5506 0.2418 1.2899 -4.4329 0.9775 47.9077 0.6006];
    options = optimset('MaxFunEvals',4000);
    x = fminunc (fun,x0,options);
    h1(i)=x(1) + x(2) * e(i-1)^2 + x(3) * h(i-1);
    h2(i)=x(4) + x(5) * e(i-1)^2 + x(6) * h(i-1);
    u1(i)=x(7) + x(8) * b(i-1);
    u2(i)=x(9) + x(10) * b(i-1);
    h(i)=p(i) * (u1(i)^2 + h1(i)) + (1-p(i))*(u2(i)^2 + h2(i))-(p(i) * u1(i) + (1-p(i)) * u2(i))^2;
    e(i)=b(i) - (p(i) * u1(i) + (1-p(i)) * u2(i));
    f1(i)=exp((-(b(i) - u1(i))^2)/(2*h1(i)))/sqrt(2*pi*h1(i));
    f2(i)=exp((-(b(i) - u2(i))^2)/(2*h2(i)))/sqrt(2*pi*h2(i));
end
%%%%%%%%%%%%%%%%%%%%%%%%
function [h1,h2,h,e,u1]= testrs (b,P,Q,x)
h1(1)=1.536164415241;
h2(1)=1.336363744144;
u1(1)=0.940125;
u2(1)=0.589625;
p(1)=0.5088;
h(1)=p(1) * (u1(1)^2 + h1(1)) + (1-p(1)) * (u2(1)^2 + h2(1)) - (p(1) * u1(1) + (1-p(1)) * u2(1))^2;
e(1)=b(1) - (p(1) * u1(1) + (1-p(1)) * u2(1));
f1(1)=exp((-(b(1) - u1(1))^2)/(2*h1(1)))/sqrt(2*pi*h1(1));
f2(1)=exp((-(b(1) - u2(1))^2)/(2*h2(1)))/sqrt(2*pi*h2(1));
```

```
for i=2:length(b)
    p(i)=(1-Q) * f2(i-1) * (1-p(i-1))/(f1(i-1) * p(i-1) + f2(i-1) * (1-p(i-1))) + P * f1(i-1) * p(i-1)/(f1(i-1) * p(i-1) + f2(i-1) * (1-p(i-1)));
    h1(i)=x(1) + x(2) * e(i-1)^2 + x(3) * h(i-1);
    h2(i)=x(4) + x(5) * e(i-1)^2 + x(6) * h(i-1);
    u1(i)=x(7) + x(8) * b(i-1);
    u2(i)=x(9) + x(10) * b(i-1);
    f1(i)=exp((-(b(i) - u1(i))^2)/(2*h1(i)))/sqrt(2*pi*h1(i));
    f2(i)=exp((-(b(i) - u2(i))^2)/(2*h2(i)))/sqrt(2*pi*h2(i));
    e(i)=b(i) - (p(i) * u1(i) + (1-p(i)) * u2(i));
    h(i)=p(i) * (u1(i)^2 + h1(i)) + (1-p(i))*(u2(i)^2 + h2(i))-(p(i) * u1(i) + (1-p(i)) * u2(i))^2;
end
h1
h2
h
e
%%%%%%%%%条件方差非参数估计%%%%%%%%%%%
function m = feicanshuguji(y,x,h)
n = length(y);
m = zeros(n,1);
p = ones(n,1);
for i =1:n
    w = zeros(n);
    for j =1:n
        e = ((x(j,1) - x(i,1))/h)^2 +((x(j,2)-x(i,2))/h)^2;
        w(i,j) = (2/pi)*(1-e)*(e<1);
    end
    m(i) = (p'*w*y)/(p'*w*p);
end
%%%%%%%%%%%%半参数RS-GARCH 计算条件方差%%%%%%%%%%%%%
function [h,e]= testrss (b,P,Q,x,h1,h2)
h1(1)=1.536164415241;
h2(1)=1.336363744144;
u1(1)=0.940125;
u2(1)=0.589625;
p(1)=0.5088;
h(1)=p(1) * (u1(1)^2 + h1(1)) + (1-p(1)) * (u2(1)^2 + h2(1)) - (p(1) * u1(1) + (1-p(1)) * u2(1))^2;
e(1)=b(1) - (p(1) * u1(1) + (1-p(1)) * u2(1));
f1(1)=exp((-(b(1) - u1(1))^2)/(2*h1(1)))/sqrt(2*pi*h1(1));
f2(1)=exp((-(b(1) - u2(1))^2)/(2*h2(1)))/sqrt(2*pi*h2(1));
for i=2:length(b)
    p(i)=(1-Q) * f2(i-1) * (1-p(i-1))/(f1(i-1) * p(i-1) + f2(i-1) * (1-p(i-1))) + P * f1(i-1) * p(i-1)/(f1(i-1) * p(i-1) + f2(i-1) * (1-p(i-1)));
    u1(i)=x(7) + x(8) * b(i-1);
    u2(i)=x(9) + x(10) * b(i-1);
    f1(i)=exp((-(b(i) - u1(i))^2)/(2*h1(i)))/sqrt(2*pi*h1(i));
    f2(i)=exp((-(b(i) - u2(i))^2)/(2*h2(i)))/sqrt(2*pi*h2(i));
    e(i)=b(i) - (p(i) * u1(i) + (1-p(i)) * u2(i));
    h(i)=p(i) * (u1(i)^2 + h1(i)) + (1-p(i))*(u2(i)^2 + h2(i))-(p(i) * u1(i) + (1-p(i)) * u2(i))^2;
end
h
e
```

Data Availability

The authors declare that the data will be available upon acceptance of this article. The data used to support the findings of this study are included within the article; it is true and reliable. If readers need related data, they can contact the authors.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant number 71171155; Xi'an Social Sciences Planning Fund Major Projects under Grant number 17J92; Special Scientific Research Project of Education Department of Shaanxi Province under Grant numbers 18JK0535 and 16JK1527; Xi'an Development and Reform Commission Regional Economic Issues under Grant number SXTY2018-08-15; and Humanities and Social Sciences Research Planning Fund of Ministry of Education under Grant number 19YJA630080.

References

[1] J. P. Morgan, "Measuring the risk in value at risk," *Financial Analysts Journal*, pp. 47–55, 1996.

[2] H. Working, "Futures trading and hedging," *American Economic Review*, vol. 43, no. 3, pp. 314–343, 1953.

[3] P. Fraser and A. J. McKaig, "UK metal and energy basis variation and a common source of risk: A note," *International Journal of Finance & Economics*, vol. 5, no. 2, pp. 155–164, 2000.

[4] A. Gulley and J. E. Tilton, "The relationship between spot and futures prices: An empirical analysis," *Resources Policy*, vol. 41, no. 1, pp. 109–112, 2014.

[5] Z. Zheng and W. Huang, "Macroeconomic influences, risk premiums and basis variation of futures markets," *Quarterly Journal of Finance*, vol. 8, no. 6, pp. 221–230, 2013.

[6] E. F. Fama and K. R. French, "Commodity futures prices: some evidence on forecast power, premiums, and the theory of storage," *The Journal of Business*, vol. 60, no. 1, pp. 55–73, 1987.

[7] U. Broll, P. Welzel, and K. P. Wong, "Futures hedging with basis risk and expectation dependence," *International Review of Economics*, vol. 62, no. 3, pp. 213–221, 2015.

[8] X. Zhuang T, Y. Li, L. Guo H, and Q. Li X, "Research on basis risk of steel futures market based on VaR," *Journal of Systems Management*, vol. 25, no. 4, pp. 669–676, 2016.

[9] 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.

[10] C. Francq and J.-M. Zaköan, "The L2-structures of standard and switching-regime GARCH models," *Stochastic Processes and Their Applications*, vol. 115, no. 9, pp. 1557–1582, 2005.

[11] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.

[12] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.

[13] C. G. Lamoureux and W. D. Lastrapes, "Persistence in variance, structural change, and the GARCH model," *Journal of Business and Economic Statistics*, vol. 8, no. 2, pp. 225–234, 1990.

[14] S. F. Gray, "Modeling the conditional distribution of interest rates as a regime-switching process," *Journal of Financial Economics*, vol. 42, no. 1, pp. 27–62, 1996.

[15] R. Sajjad, J. Coakley, and J. C. Nankervis, "Markov-switching GARCH modelling of value-at-risk," *Studies in Nonlinear Dynamics & Econometrics*, vol. 12, no. 3, article 1522, 2008.

[16] T. Elenjical, P. Mwangi, B. Panulo, and C.-S. Huang, "A comparative cross-regime analysis on the performance of GARCH-based value-at-risk models: evidence from the Johannesburg stock exchange," *Risk Management*, vol. 18, no. 2-3, pp. 89–110, 2016.

[17] J. Yang P and H. Zhang C, "Estimation of Shanghai and Shenzhen stock market based on Markov regime switching models," *Chinese Journal of Management Science*, vol. 21, no. 2, pp. 42–49, 2013.

[18] H. Peng F and Y. Chen, "The estimation of optimal hedging ratio of copper future market in China——based on Markov regime-switching garch model," *Chinese Journal of Management Science*, vol. 23, no. 5, pp. 14–22, 2015.

[19] P. Jorion, *Market Risk*, Wiley Press, New York, NY, USA, 2010.

[20] P. H. Kupiec, "Techniques for verifying the accuracy of risk measurement models," *The Journal of Derivatives*, vol. 3, no. 2, pp. 73–84, 1995.

[21] Z. Ye A, *Nonparametric Econometrics*, Nankai University Press, Nanjing, China, 2003.