Measuring Information Capacity of Volatility Risk Premium Using Quantile Regression

Yifan Ji
School of Economics, Shanghai University, Shanghai, China

*Corresponding author e-mail: jyf0327@shu.edu.cn

Abstract. Volatility is a very important variable in financial research, and so is the volatility risk premium. This article constructs upward, downward and overall volatility risk premiums based on the daily data of Shanghai 50 ETF options and the 5-minute high-frequency data of the Shanghai 50 Index from February 9, 2015 to February 28, 2020. The quantile regression method is used to predict the return of the Shanghai 50 Index in the next 14 days, 30 days, 60 days, 90 days, 180 days, 270 days, and 360 days. Our study finds that the 0.05 quantile has the best prediction effect, and the overall volatility risk premium is better than the upward and downward volatility risk premiums in most of the time, which shows overall risk premium has more information than others.

Keywords: Volatility Risk Premium, Quantile Regression, Index Return Prediction, High-Frequency Data

1. Introduction
Volatility is a very important variable in financial research. Asset pricing and risk management are inseparable from volatility indicators. At the same time, with the continuous development of the option market, the volatility of options has become the focus of scholars’ research. Volatility can be used as a quantitative value of market views, reflecting investors' expectations of future volatility, and also reflecting investors' expectations of future returns.

Giot and Laurent (2003) points out that volatility indicators have a specific relationship with future market index returns, indicating that the relationship between volatility and market returns will be one of the important information connotations of volatility indicators [1]. Many subsequent scholars also prove the above viewpoint. At the same time, volatility can be divided into implied volatility (IV) and realized volatility (RV). Implied volatility is generally calculated by financial models while the realized volatility is calculated from the actual data of the underlying asset. The difference between the realized volatility and the implied volatility is called the volatility risk premium (VRP). Recent studies on option volatility also divide volatility into upward volatility (good volatility) and downward volatility (bad volatility), and study the predictive ability of good and bad volatility respectively (Barndorff-Nielsen et al. (2010), Sega et al. (2015), Barunik et al. (2016), Chen (2019)) [2-5]. Why is the above division necessary? Because volatility is not necessarily bad, upward volatility brings gains, while downward volatility brings losses. Everyone hates downward volatility but likes upward
volatility. We also use the same method to obtain good and bad volatility risk premiums, and use volatility risk premiums to predict future returns. The risk premium of volatility combines the information of implied volatility and realized volatility, contains more information, and logically has a better prediction effect. Feunou et al. (2018) construct the upward volatility risk premium, the downward volatility risk premium, and the overall volatility risk premium, and use them to predict the S&P500 data and find the downward volatility risk premium prediction ability is better [6]. Kilic et al. (2019) also construct upward and downward volatility risk premiums and overall volatility risk premiums, and find that upward and downward volatility risk premiums can jointly explain the part of market returns that exceed 20% [7].

China’s earliest 50ETF options were listed for trading on February 9, 2015. The Shanghai and Shenzhen 300ETF options listed at the end of last year have further improved Chinese options market. Therefore, studying option volatility plays a very important role in Chinese market. This paper mainly studies the information capacity of the volatility risk premium of 50ETF options by predicting the return of the Shanghai 50 Index using the quantile regression.

2. Indicators and Models

2.1. Implied Volatility

There are many calculation methods for implied volatility indicators, but the implied volatility without a model is more commonly used. Refer to the article by Bskshi et al. (2003) [8], the implied volatility of option can be expressed as a linear combination of a series of OTM call options and put options. Therefore, the implied volatility can be expressed as:

\[
iV_t = \int_{K}^{\infty} \frac{2(1-\log \left(\frac{K}{S_t}\right))}{K^2} C(t,t+i,K)dK + \int_{0}^{\infty} \frac{2(1-\log \left(\frac{K}{S_t}\right))}{K^2} P(t,t+i,K)dK
\]

(1)

where \(C(t,t+i,K)\) and \(P(t,t+i,K)\) represents the price of OTM call options and put options with a strike price of \(K\) at time \(t\) and expiration date \(t+i\). \(S_t\) is the price of the underlying asset. At the same time, considering that the number of options in the Chinese options market is not very large, we choose options with expiration dates within a certain range in actual operation. The expiration date chosen in this article is 10-60 days.

Combined with the indicator construction method of Kilic et al. (2019) [7], we express the upward and downward implied volatility as:

\[
iV^u_t = \int_{S_t}^{\infty} \frac{2(1-\log \left(\frac{K}{S_t}\right))}{K^2} C(t,t+i,K)dK
\]

(2)

\[
iV^d_t = \int_{0}^{S_t} \frac{2(1-\log \left(\frac{S_t}{K}\right))}{K^2} P(t,t+i,K)dK
\]

(3)

At the same time, we consider that China’s stock market cannot be short-sold, so some of the information is actually lost. Therefore, it is best to use futures prices when pricing the underlying asset [8]. In addition, China’s 50ETF options also have a dividend-sharing mechanism, and some adjustments are also needed. According to Merton (1973) [9] ideas and the PCP formula, we can get:

\[
C_t + KB_t = P_t + S_t = P_t + F_t^2 B_t
\]

(4)

Where \(B_t\) is the price of a zero coupon bond with the same maturity date of the option. Therefore, the above implied volatility formula becomes:

\[
iV_t = \int_{S_t}^{\infty} \frac{2(1-\log \left(\frac{K}{P_t}\right))}{K^2} C(t,t+i,K)dK + \int_{0}^{\infty} \frac{2(1-\log \left(\frac{K}{P_t}\right))}{K^2} P(t,t+i,K)dK
\]

(5)

\[
iV^u_t = \int_{S_t}^{\infty} \frac{2(1-\log \left(\frac{K}{P_t}\right))}{K^2} C(t,t+i,K)dK
\]

(6)
2.2. Realized Volatility

The daily realized volatility is defined by the square root of the daily realized variance (Barndorff-Nielsen and Neil, 2004) \[10\].

\[
i_{t} = \sum_{j=1}^{H_{t}} r_{t,j}^{2} \tag{7}
\]

Where \(H_{t}\) represents the number of high-frequency transactions in a day, we use five minutes of data, so \(H_{t}=48\).

At the same time, the realized volatility is divided into upward realized volatility and downward realized volatility, expressed as:

\[
r_{t}^{u} = \sum_{j=1}^{H_{t}} r_{t,j}^{2} \ 1(r_{t,j} > 0) \tag{8}
\]

\[
r_{t}^{d} = \sum_{j=1}^{H_{t}} r_{t,j}^{2} \ 1(r_{t,j} < 0) \tag{9}
\]

2.3. Volatility Risk Premium

The volatility risk premium is defined as the difference between the volatility under the P and Q measures, that is, the difference between the realized volatility and the implied volatility:

\[
vr_{t} = r_{t} - i_{t} \tag{10}
\]

The upward and downward volatility risk premium can be expressed as:

\[
vr_{t}^{u} = r_{t}^{u} - i_{t}^{u} \tag{11}
\]

\[
vr_{t}^{d} = r_{t}^{d} - i_{t}^{d} \tag{12}
\]

2.4. Quantile Model

The linear quantile regression model can be expressed as:

\[
y_{i} = \beta_{0}(\tau) + \beta_{1}(\tau)x_{i} + \epsilon_{i}, \quad (i = 1, \ldots, n) \tag{13}
\]

Where \(\tau \in (0,1)\), \(Q_{\tau}(r|x_{i}) = \beta_{0}(\tau) + \beta_{1}(\tau)x_{i} + \epsilon_{i}, \quad (i = 1, \ldots, n)\) is defined as the sample conditional quantile function at the quantile of \(\tau\), it is assumed that the value of the parameter to be estimated varies with the quantile. In our model, \(y\) represents the return of the Shanghai 50 Index, and \(x\) represents different volatility risk premiums.

The above solution is obtained when the expectation of the weighted absolute distance reaches the minimum value, namely

\[
\min_{\beta_{0}(\tau), \beta_{1}(\tau)} \sum_{i} 1(y_{i} > x_{i}) \tau|y_{i} - \beta_{0}(\tau) - \beta_{1}(\tau)x_{i}| + \sum_{i} 1(y_{i} < x_{i}) (1-\tau)|y_{i} - \beta_{0}(\tau) - \beta_{1}(\tau)x_{i}| \tag{14}
\]

where \(X_{i} = \beta_{0}(\tau) + \beta_{1}(\tau)x_{i}\).

Quantile regression can obtain different regression lines by setting different \(\tau\) values, so that the data information of the local area can be obtained through a specific regression line, and the variable relationship of the overall data can be obtained by combining all the regression lines.

We use the volatility risk premium, the upward volatility risk premium and the downward volatility risk premium to predict the return of the Shanghai 50 Index in the next 14 days, 30 days, 60 days, 90 days, 180 days, 270 days and 360 days, respectively. Look at the effect of its prediction and the strength of its interpretation. In addition, the average return and cumulative return can be used when calculating future returns, the results reported in this article are the results of cumulative returns.

3. Data and Data Analysis
3.1. Data
Our data is obtained from JoinQuant (JQ) database which contains all high-frequency securities, option data, and futures data. In this study, we sample SSE 50ETF option data at the daily frequency and choose the period from February 9 2015 to February 28 2020, for a total of 2216 observations.

The bond price is obtained from the ChinaBond database which contains bond data and research analysis. Shanghai Inter-bank Offered Rate (“Shibor”), which is called the “Libor” in China, has significance for the marker-based interest rate reform in China. Thus, we use it as the risk-free interest rate.

3.2. Data Analysis
We carry out simple descriptive statistics on implied volatility, realized volatility and volatility risk premium, and get the following Table 1.

From Table 1 we can see that the mean and variance of realized volatility are relatively small, but the mean variance of implied volatility and volatility risk premium are relatively large. At the same time, we can see that the mean and variance of all downward volatility are the largest, which means that the volatility of losses is greater than that of gains.

Table 1. Summary Statistics

|     | $ru^i$ | $ru^d$ | $rv$ | $v_{P^2}$ | $v_{P^d}$ | $v_{P^u}$ | $\text{iv}^u$ | $\text{iv}^d$ | $\text{iv}_{P^u}$ | $\text{iv}_{P^d}$ |
|-----|--------|--------|------|-----------|-----------|-----------|-------------|-------------|----------------|----------------|
| mean| 0.01   | 0.01   | 0.02 | 0.29      | 0.31      | 0.60      | 0.28        | 0.30        | 0.58           | 0.58           |
| std | 0.03   | 0.04   | 0.06 | 0.38      | 0.51      | 0.69      | 0.37        | 0.50        | 0.67           | 0.67           |
| min | 0.00   | 0.00   | 0.00 | 0.01      | 0.02      | 0.04      | 0.01        | 0.31        | 0.04           | 0.04           |
| 25% | 0.00   | 0.00   | 0.00 | 0.09      | 0.11      | 0.21      | 0.09        | 0.10        | 0.20           | 0.20           |
| 50% | 0.00   | 0.00   | 0.01 | 0.18      | 0.20      | 0.40      | 0.17        | 0.19        | 0.39           | 0.39           |
| 75% | 0.01   | 0.01   | 0.02 | 0.32      | 0.35      | 0.72      | 0.31        | 0.35        | 0.70           | 0.70           |
| max | 0.49   | 0.85   | 1.03 | 2.78      | 7.00      | 7.15      | 2.78        | 7.00        | 7.14           | 7.14           |

Note: all the values are percentage.

As can be seen from the Figure 1. below, whether it is implied volatility, realized volatility, or volatility risk premium, there have been two major fluctuations, namely the stock market crash in 2015 and the COVID-19 in early 2020. At the same time, we can see that the changes in the implied volatility brought about by COVID-19 is higher than the changes brought about by the stock market crash in 2015. However, the realized volatility during the 2015 stock market crash was greater than that of the 2020 COVID-19. This is because many fluctuations have not yet been fully revealed. However, IV contains future information and it shows that the fluctuation of the COVID-19 is much greater than that of 2015. Therefore, we can say that there is a high possibility that the fluctuation of realized volatility will exceed 2015.

(a) the time serious of implied volatility
Figure 1. The time series of implied volatility, realized volatility and volatility risk premium

4. Analysis of Quantile Model Results

We use volatility risk premium, upward volatility risk premium, and downward volatility premium to predict the next 14 days, 30 days, 60 days, 90 days, 180 days, 270 days and 360 days of the SSE 50 Index's returns. At the same time, we chose different quantiles. The selected quantiles are 0.05, 0.25, 0.45, 0.65, 0.85. At the same time, we also reported the results of simple OLS regression and compared the above results. And the forecast results of the Shanghai 50 Index are shown in Table 2-4.

| h   | $\beta_0$ | $\beta_1$ | $R^2$ | $\beta_0$ | $\beta_1$ | $R^2$ | $\beta_0$ | $\beta_1$ | $R^2$ | $\beta_0$ | $\beta_1$ | $R^2$ | $\beta_0$ | $\beta_1$ | $R^2$ | $\beta_0$ | $\beta_1$ | $R^2$ | $\beta_0$ | $\beta_1$ | $R^2$ | $\beta_0$ | $\beta_1$ | $R^2$ |
|-----|-----------|-----------|-------|-----------|-----------|-------|-----------|-----------|-------|-----------|-----------|-------|-----------|-----------|-------|-----------|-----------|-------|-----------|-----------|-------|-----------|-----------|-------|
| 14  | -0.03     | 22.29     | 0.13  | -0.03     | 30.46     | 0.25  | -0.03     | 51.00     | 0.25  | -0.06     | 58.82     | 0.27  | -0.12     | 58.17     | 0.25  | -0.14     | 49.81     | 0.17  | -0.09     | 45.90     | 0.18  |
| 30  | -0.01     | 7.33      | 0.05  | 0.02      | 19.14     | 0.12  | 0.03      | 33.88     | 0.09  | 0.04      | 40.45     | 0.13  | 0.13      | 38.08     | 0.09  | 0.21      | 38.68     | 0.04  | 0.06      | 26.08     | 0.08  |
| 60  | 0.01      | 3.40      | 0.03  | 0.02      | 6.61      | 0.04  | 0.05      | 23.00     | 0.12  | 0.07      | 26.80     | 0.12  | 0.13      | 37.16     | 0.11  | 0.17      | 30.10     | 0.07  | 0.06      | 17.99     | 0.06  |
| 90  | 0.03      | 2.07      | 0.01  | 0.04      | 3.88      | 0.02  | 0.07      | 17.99     | 0.06  | 0.09      | 17.58     | 0.08  | 0.17      | 29.05     | 0.10  | 0.20      | 26.12     | 0.05  | 0.06      | 15.63     | 0.05  |
| 180 | 0.05      | 0.00      | 0.01  | 0.01      | 1.36      | 0.01  | 0.07      | 13.6     | 0.04  | 0.09      | 11.91     | 0.01  | 0.11      | 29.5     | 0.01  | 0.21      | 15.17     | 0.04  | 0.08      | 7.68      | 0.01  |
| 270 | 0.02      | 0.00      | 0.01  | 0.02      | 1.36      | 0.01  | 0.07      | 13.6     | 0.04  | 0.09      | 11.91     | 0.01  | 0.11      | 29.5     | 0.01  | 0.21      | 15.17     | 0.04  | 0.08      | 7.68      | 0.01  |
| 360 | 0.06      | 0.01      | 0.01  | 0.06      | 1.36      | 0.01  | 0.07      | 13.6     | 0.04  | 0.09      | 11.91     | 0.01  | 0.11      | 29.5     | 0.01  | 0.21      | 15.17     | 0.04  | 0.08      | 7.68      | 0.01  |
At the same time, we can see that all the prediction coefficients are positive. It can be seen that as the volatility risk premium increases, the return of the stock market will rise in the future. This is not difficult to understand. Traders in the options market generally have special information that stock market traders don’t have. The part where the realized volatility is higher than the implied volatility can be expressed as an overreaction of stock market traders to risks, and the risks of this part have exceeded the risks expected by option market traders. This part of the risk of overreaction must be compensated, and this part of the compensation is the future stock market rise, so all regression coefficients are positive. In addition, it can be seen that as the number of forecast days increases, the predicted coefficients also show an inverted U shape, that is, the predicted coefficients change from low to high and back to low. It can be seen that each unit of change in the volatility risk premium will have an effect of increasing and then decreasing in the future stock market returns. The best prediction effect is in the mid-term (about 90 days). Besides, we can see that the overall volatility risk premium has the higher $R^2$ than upward and downward volatility risk premium, which shows overall volatility risk premium has more information than others.

At the same time, in order to compare the upward volatility risk premium, downward volatility risk premium and overall volatility risk premium more clearly, the result of adjusted $R^2$ of Quantile regression is presented in Figure 2. It can be clearly seen that the adjusted $R^2$ of quantile regression at 5% is the best, and all quantile regressions are greater than the results obtained by OLS regression.
5. Conclusion
We construct upward, downward and overall volatility risk premiums based on the daily data of SSE 50 ETF options and the 5-minute high-frequency data of the SSE 50 Index from February 9, 2015 to February 28, 2020. Using quantile regression method to predict the return of the Shanghai 50 Index in the next 14 days, 30 days, 60 days, 90 days, 180 days, 270 days, and 360 days. Through the research above, our study finds that the 0.05 quantile has the best prediction effect, and the overall volatility risk premium is better than the upward and downward volatility risk premiums in most of the time. It indicates that overall volatility risk premium has more information than others and 0.05 quantile is the best quantile to measure their information capacity.

References
[1] P. Giot and S. Laurent, “Value-at-Risk for Long and Short Trading Positions,” J. Appl. Econom., vol. 18, no. 6, pp. 641–663, 2003.
[2] O. E. Barndorff-Nielsen, S. Kinnebrock, and N. Shephard, “Measuring downside risk-realised semivariance,” Econ. Pap., 2010.
[3] G. Segal, I. Shaliastovich, and A. Yaron, “Good and bad uncertainty: Macroeconomic and financial market implications,” J. financ. econ., vol. 117, no. 2, pp. 369–397, 2015.
[4] Barunik, Vacha, Lukas, Jozef, Kocenda, and Evzen, “Asymmetric connectedness on the US stock market: Bad and good volatility spillovers,” J. Financ. Mark., 2016.
[5] Y. Chen, W. Li, and F. Qu, “Dynamic asymmetric spillovers and volatility interdependence on China’s stock market,” Phys. A Stat. Mech. its Appl., vol. 523, pp. 825–838, 2019
[6] B. Feunou, M. R. Jahan-Parvar, and C. Okou, “Downside Variance Risk Premium,” J. Financ. Econom., vol. 16, no. 3, pp. 341–383, 2018.
[7] M. Kilic and I. Shaliastovich, “Good and Bad Variance Premia and Expected Returns,” Soc. Sci. Electron. Publ., vol. 65, no. 06, pp. 2522–2544, 2019.
[8] G. Bakshi, “Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options,” Rev. Financ. Stud., 2003.
[9] R. C. Merton, “The Relationship between Put and Call Option Prices: Comment,” J. Finance, vol. 28, no. 1, pp. 183–184, 1973.
[10] O. E. Barndorff-Nielsen and S. Neil, “Power and Bipower Variation with Stochastic Volatility and Jumps,” J. financ. econ., vol. 2, no. 1, pp. 1–37, 2004.