How does informed trader trade the “reverse risk”? --OTM Option volume as risk factor for stock prediction

Jing Guo\textsuperscript{a}, Jie Xiao\textsuperscript{b}, Zhongxin Ni\textsuperscript{*}

School of Economics, Shanghai University, Shanghai 200444, China

\textsuperscript{*}Corresponding author e-mail: zhongxinni@i.shu.edu.cn, winejuiceraisins@163.com, 18717865113@163.com,

Abstract. This article researches the implied risks in the trading volume data of the SSE (Shanghai Security Exchange) 50 ETF options, and explores the power of these risks to predict the return of the underlying stocks. By analyzing the trading behavior of informed option traders, we believe that the trading volume of the OTM (out-of-money) options can capture these traders’ expectations of the “reverse risk”. We divide “reverse risk” into two types of risk information: the first one is the option selection process: the trader decides whether to choose a OTM option or a non-OTM option to hedge their private information, which represents whether they expect the future stock to reverse; the second risk information is measured by the change amount of the current OTM option trading volume, which depicts the new information shock brought by these informed traders’ behavior. The empirical result shows that the option selection process can significantly predict the future returns of the SSE 50 index, while the information shock caused by the change in OTM options trading volume has solid pricing power among individual stocks. We then use Granger causality test to explore the mechanism of risk transmission for the “reverse risk” trading process, and conclude as: the risk information held by informed traders(those who believe the future market will “reverse”) first affected their decision-making process, and their decisions in the options market will transmit the risk to the SSE 50 index trading market, and then cause the current information shocks, and these shocks will eventually affect the constituent stocks of the SSE 50.

1. Introduction

Financial derivatives have always been the focus of academic research, in which stock option is an important variety. Many studies have also proved that, as a hedging tool of the underlying stock, option implied information has an important effect on the price discovery process of the underlying stock. More importantly, many scholars have shown (Bates (1991), Garleanu et al. (2007), Lin et al. (2016)) that option data contain vital information and characteristics of future distribution of stock prices, and that these characteristics cannot be obtained using historical stock data. Those traders who possess this kind of information are called informed traders, Lin et al. (2016) mentions that options traders often have private information that is not available to stock traders, and that such information is valuable in guiding trading operations. Informed trading is more common in the derivatives market, where derivatives traders typically have profound research about underlying assets. If we assume that some option traders
have such private information about the underlying assets, they are bound to hedge that information, and the options market will be their first choice at this time (Black (1975), Faff and Hiller (2005)). Therefore, the information and preference of option traders is often reflected in changes in data such as the price and spread of the contract or the trading volume, and so on.

In fact, scholars have long studied the implied risk information in option trading behavior. Roll et al. (2010) has built the “option volume to stock volume ratio”, denoted as O/S ratio, their research shows that the O/S ratio is driven by informed trading. Roll et al. (2010) further evidence that changes in O/S are highly correlated with traders’ informed trading phenomena. Pan and Poteshman (2006), Roll (2010), Johnson and So (2012), Jin et al. (2012) and Lin et al. (2016) have all used this O/S measure and demonstrate this ratio has significant predictive power on the underlying asset. Johnson and So (2012) and Lin et al. (2016) both illustrate the mechanism behind O/S prediction ability from the perspective of informed traders' behavior. The basic assumption of both articles is that when a trader has private information and expects to benefit from it or hedge the relative risk, he or she makes a choice between the stock market and the option market, and the O/S ratio reflects this selective behavior of the trader. Often when a trader has negative information about a stock, i.e. knowing that the underlying asset is going to fall in the future, he or she chooses to hedge this information in the options market, at which point the ratio of option volume to volume of the stock will rise accordingly. This is usually because the short selling cost of the stock market is higher than the option market, and also the high leverage of options makes hedging more convenient. And within this research we focus on China's financial market, as we know that there is no direct short-selling mechanism in the Chinese stock market, so that makes these informed traders have no choice but to hedge in option market. Hence in this paper we directly use option volume data to capture the private information. In addition, we do not consider the volume of the overall option contracts, but rather the volume of the OTM (out-of-the-money) option. OTM options, as the name implies, are currently of no intrinsic value, and we believe that option traders suddenly conduct a high-volume OTM options trading reflects their expectation of underlying stock to “reverse” in the future.

Gkionis et al. (2018) suggests that the price of the OTM option implies a positive information of the underlying stock, and the article suggests that the price of the OTM option determines the RNS (risk neutral distribution skewness) of the stock's return. Gkionis et al. (2018) then further points out that the change of RNS brought by the fluctuation of the market demand for OTM options is in fact highly related to the current tail risk of the underlying stock. But it is inconvenient to extract the tail risk by directly using the stock’s historical data, since the current asset distribution can only be obtained using statistical methods after the asset price has been realized for a period of time to come, therefore we expect that the use of OTM option volume data provides a forward-looking picture of the underlying stock's RNS or tail risk, and also implies a positive information of the underlying stock. Furthermore, we take over the trader behavior analysis framework of Johnson and So (2012) and Lin et al. (2016) to examine the mechanisms of how OTM option volumes may affect future stock movements. First, as we illustrated above, we need to figure out what type of options these traders will choose to trade based on their private information. Obviously, if you expect the market to be stable in the future or the variation trend remains the same, ATM (at-the-money) or ITM (in-the-money) options are your best options, and we know that these types of trading have little to do with so-called tail risk. But what if you expect a big reversal in the future? It can be imagined that at this time the volume of OTM options suddenly increased, and this phenomenon is also revealing the increase in the tail risk of the underlying stock, so the future of the market will inevitably compensate for this risk in the future (Bollerslev et al. (2015), Bollerslev and Todorov (2011)). We expect that when traders choose between OTM and non-OTM (ATM or ITM) options, their sudden preference for OTM options will contribute to future stock market excess returns. But this analysis is vague, when a trader chooses OTM options between two types of options, how much will he trade to hedge his risk information? This depends on the extent to which he or she thinks the "future reversal of the market" is, so we expect that the change in the current period of the OTM option volume will be able to more specifically describe the impact of this "reverse risk" expectation on the stock market. Therefore, taking China's financial market as the background, we use the OTM option
contract volume of the SSE 50 ETF option as risk information to describe the strength of the informed trader's expectation that the market will reverse in the future, and then explore the ability of such risk information to predict the underlying stock. In our empirical experiments we find that the choice making between OTM and non-OTM options by informed traders directly affect the future trend at the macro level (SSE 50 Index), and the new risk shocks caused by informed traders in the current period after selecting the OTM option has more microscopic effect and a significant pricing power among the constituent individual stocks of the SSE 50.

The second part of this paper focuses on data sources, data characteristics and the construction method of all risk factors related to the volume of OTM options used in the follow-up empirical research, and then part three is prediction analysis of the future returns of the SSE 50 index and the fourth part mainly studies the pricing power of the OTM option volume risk among individual stocks. In Part 5, we use the Granger test to explore the transmission mechanism of implied risk in the OTM option volume, then prat 6 presents our conclusion.

2. Measures of the OTM option volume risk

2.1. Data
The option and stock data for this research is downloaded from Wind terminal, and the option data contains the daily average price of 1874 option contracts from February 9, 2015 to May 30, 2019, also these contacts’ execution price, volume and option type (put or call), in which there are 937 put options and 937 call contracts. Based on the findings of Johnson and So (2012), we know that the option trading data near expiration is best suited to extract private information, so we use the remaining maturity τ and option moneyness \( k \) to filter the sample at each point in time (further, due to the limitations of the SSE 50 ETF option data itself, we cannot consider the call and the call option separately, otherwise there will be a large number of null values in the dataset). Learn from the results of Pan (2000), Foresi and Wu (2005) and Xing et al. (2010), on day \( t \), we select the option contracts of the \( \tau \) in 8 days and 45 days, and \( k < 0.05 \), where moneyness \( k \) is defined as

\[
k = \begin{cases} 
\log(S/K) & \text{for call} \\
\log(K/S) & \text{for put.}
\end{cases}
\]

After the sample screening, use the average of the log volume (the original volume data unit is 10,000) of all the contracts as the daily OTM volume.

Table 1. Descriptive Statistics on Sample Characteristics

|       | Min   | Median | Mean  | Maximum |
|-------|-------|--------|-------|---------|
| \( n_t \) | 1.00  | 10.00  | 11.90 | 36.00   |
| \( \bar{\tau}_t \) | 10.80 | 27.00  | 26.60 | 45.00   |
| \( \bar{k}_t \) | 0.05  | 0.10   | 0.11  | 0.34    |
| Volume | 4.68  | 7.96   | 7.98  | 11.62   |

Table 1 shows the basic characteristics of the sample: the number of option contracts \( n_t, t = 1, \cdots, N \); the average remaining maturity period \( \bar{\tau}_t, t = 1, \cdots, N(\bar{\tau}_t \text{ is the average remaining maturity period of all } n_t \text{ option contracts}) \), average value status \( \bar{k}_t, t = 1, \cdots, N(\bar{k}_t \text{ is the average value status of } n_t \text{ option contracts}) \), and the OTM option trading volume (referred to as “volume” in the table) has been presented in statistics. As shown in Table 1, the average moneyness of the sample option contract is \( k = 0.11 \); the average remaining expiration maturity of the sample option contract is 26.6 days; the maximum value of the volume is 11.62, the minimum value is 4.68 and average 7.98.

The time series of the daily OTM option volume is shown in Figure 1, and we can see that the series presents a clear trend and unit root instability. Obviously, this is because China’s options market is in
the early stages of development, and the market capacity has been expanding since the listing of SSE 50ETF options. However, due to the significant non-stationarity of the unit root of the series, we first perform detrending before using the series as a risk factor. This article mainly uses two methods to deal with the trends of the OTM options volume series. First, we consider the ratio of the volume of OTM options and the volume of non-OTM options. Because the trends and dimensions of the two series are similar, using the ratio as a risk measure will avoid non-stationary phenomena, and we denote this ratio as $OTM_{VR}$. According to the analysis above, the first step of our characterization of the behavior of informed traders is a trader’s choice between two types of options, and the second step is the number of transactions performed by the trader, so $OTM_{VR}$ risk is our first stage measure of the risk contained in the informed trading process. Later, because we want to characterize the changes in the volume of the OTM options during the process, we use statistical methods to remove the trend terms in the series, leaving the innovation part of the series to characterize the current increased “market reverse risk”, denoted as $OTM_{V}$. For the empirical research in this article, we will use these two measurement methods as our

\[ \text{Figure 1. SSE 50 ETF Options Inflated Options Day Volume, the image above is a sequence of time series diagram, and for the below ones, left and right are ACF test and PACF test results respectively} \]

measure of “reverse risk information implied in the behavior of trading in OTM options”, and more specifically $OTM_{VR}$ measures whether the underlying stock will reverse or not, and $OTM_{V}$ measures the magnitude that expected by the informed traders. They will collectively be referred to as $OTM_{V}$ risks through the whole document for convenience.

2.2. Construction of OTM option volume risk measurement
Based on the theoretical analysis, we expect informed traders to first make a decision between OTM and non-OTM options. Our basic assumption is that if most options traders expect the trend of the SSE
50 Index to reverse in the future, the current $OTM_{VR}$ risk increases. Regarding the calculation method of $OTM_{VR}$ risk, specifically, on day $t$, we use the criteria in 2.1 to screen for OTM options. For non-OTM options, we also choose $\tau$ within 8 days and 45 days, and $k \geq 0.05$. The daily ratio of the average trading volume is denoted as $OTM_{VR,t}$. As can be seen from Figure 2, compared to the OTM option volume series in Figure 1, $OTM_{VR}$ series avoids the problem of strong sequence trend and can be used directly for further research.

![Figure 2. $OTM_{VR}$ series](image)

Taking into account the differences caused by different processing methods, we choose different algorithms to remove the trend of the OTM options trading volume to ensure the robustness. Our first measurement is the difference of OTM option volume series, we denote it as $OTM_{V1}$, simply use the difference between the value of the current OTM option volume and the value of the previous period. Secondly, we use the ARIMA (autoregressive integrated moving average) model to extract the innovation part of OTM option volume series, and the measurement is denoted as $OTM_{V2}$. For the ARIMA method, first, we train the model using OTM option volume data to decide the ARIMA orders, and the results of the model estimation results are shown in Table 2. We can see that the standard errors of the AR terms and the MA terms are very small relative to the coefficient estimates of the terms. We then use the residual series of the model as $OTM_{V2}$. Finally, we use the Empirical Mode Decomposition (EMD) method. The EMD method is a signal processing method that can be used to process serial data with tendency or periodicity. The EMD method decomposes a complex signal into a finite number of intrinsic mode functions (IMFs), and each IMF component decomposed contains certain characteristic information of the original signal. The final IMF obtained by EMD decomposition is to deal with the sequence trend after noise. This IMF can more clearly characterize the long-term trend and periodicity of the original series. Therefore, this paper uses the EMD method to decompose the original series, and then uses the sum of the first two IMFs with no trend and periodicity to obtain the third measurement as $OTM_{V3}$.

In summary, this paper uses three different calculation methods to decompose the original OTM option volume series: $OTM_{V1}$, $OTM_{V2}$, and $OTM_{V3}$. It can be seen from Figure 3 that after the decomposition, the new measure of risks no longer has significant unit root non-stationarity.

| parameter | ARIMA (3,1,1) estimation | Sd |
|-----------|--------------------------|----|
| AR1       | 0.073                    | 0.031 |
| AR2       | 0.130                    | 0.037 |
| AR3       | 0.059                    | 0.031 |
| MA1       | -0.987                   | 0.005 |
| AIC       | 1554.49                  |     |
| BIC       | 1579.26                  |     |
| Log Likelihood | -772.25             |     |
Figure 3. $OTM_{V1}$, $OTM_{V2}$ and $OTM_{V3}$ series, where a1, a2, b1, b2, c1, and c2 are the timing diagram and ACF test results of the $OTM_{V1}$ variable, the timing diagram of the $OTM_{V2}$ variable, and the ACF test result of timing diagram of $OTM_{V3}$ variable and ACF test results.

3. Forecast analysis of $OTMV$ factors on returns of SSE 50

As we expect the impact of $OTMV$ risks on stock returns and the lack of immediate risk hedging channels in Chinese financial market, and the option trading operation itself takes time and cost, we believe that the forecast analysis in this section better uses weekly or even lower frequency data (Bollerslev et al. (2015), Lin et al. (2016)), so that risk measurement of the OTM option volume can capture the risk information properly. However, due to the lack of SSE 50 ETF option data, we use weekly data for predictive analysis.
In order to investigate the robustness of the risk factor of $OTMV$, we also consider the univariate regression model and multivariate regression model to analyze the future returns of the SSE 50 Index. For multivariate models we consider control variables: log (P/D) (logarithmic ratio of price and dividend), historical returns: last month returns (denoted as 1 Month L. R.) and last three month returns (denoted as 3 Month L. R.), and use the current risk $OTMV_t$ to perform the following univariate regression firstly on the daily average excess return ($IndRe_{t,T}$) of the SSE 50 within a period of time $T$ in the future:

$$IndRe_{t,T} = \alpha_{t,T} + \beta_{OTMV,t} \cdot OTMV_t + \epsilon_{t,T},$$  \hspace{1cm} (2)

and for multivariate regression we use the following model:

$$IndRe_{t,T} = \alpha_{t,T} + \beta_{OTMV,t} \cdot OTMV_t + \beta_{Contrs,t} \cdot Contrs_t + \epsilon_{t,T},$$  \hspace{1cm} (3)

where $\beta_{OTMV,t}$ is the regression coefficient of the risk of imaginary options volume risk to future stock index excess returns, the vector $\beta_{Contrs,t}$ is the coefficients vector of all control variables, $IndRe_{t,T}$ uses cumulative excess returns within the time period between $t$ to $t + T$, and we choose different value of $T$ also for the concern of the robustness. Descriptive statistics of all explanatory variables and correlation coefficients between the variables are shown in Table 3. Panel A shows descriptive statistics of each explanatory variable. It can be seen that the average values of the risk factors $OTM_{V1}$, $OTM_{V2}$, and $OTM_{V3}$ in the three calculation methods to be considered in this paper are around 0, and the expected values are not much different. The connotations represented by the following risk factors are consistent; Panel B shows the correlation coefficient between the variables. The correlation coefficient between $OTM_{V1}$ and $OTM_{V2}$ is higher, which is 0.874, which is consistent with the calculation method of the factors; among the other explanatory variables, correlation coefficients are at a low level, so and there is no multicollinearity problem in the regression model.

### Table 3. Descriptive Statistics and Correlation Coefficients

| Panel A | Descriptive Statistics |
|---------|------------------------|
| $OTM_{V1}$ | $OTM_{V2}$ | $OTM_{V3}$ | $OTM_{VR}$ | Log(P/D) | 1 Month L.R. | 3 Month L.R. |
| Mean | 0.00 | 0.02 | 0.00 | 0.56 | 1.08 | 0.002 | 0.01 |
| Median | 0.00 | 0.02 | 0.01 | 0.51 | 1.07 | 0.006 | 0.02 |
| Min | -0.62 | -0.54 | -0.71 | 0.05 | 0.42 | -0.218 | -0.35 |
| Max | 0.49 | 0.59 | 0.69 | 1.45 | 1.36 | 0.207 | 0.39 |

| Panel B | Correlation Coefficients |
|---------|--------------------------|
| $OTM_{V1}$ | 1.000 | 0.874 | 0.257 | 0.030 | 0.004 | 0.039 | 0.018 |
| $OTM_{V2}$ | 1.000 | 0.215 | 0.210 | -0.003 | -0.078 | -0.072 |
| $OTM_{V3}$ | 1.000 | -0.068 | 0.056 | 0.033 | -0.007 |
| $OTM_{VR}$ | 1.000 | 0.031 | -0.319 | -0.148 |
| Log(P/D) | 1.000 | -0.235 | -0.494 |
| 1 Month L.R. | 1.000 | 0.486 |
| 3 Month L.R. | 1.000 |

The regression results are shown in Table 4. Panel A and Panel B correspond to $T$ values of 6 months and 12 months, respectively. The first four columns in each Panel are
univariate regression models using \textit{OTMV} risk as explanatory variable; the multivariate four columns are using \textit{OTMV} risk plus all control variables. In the table we show the estimated parameters with their $t$-statistics in the parentheses below (these in bold are the cases that the test results are significant under critical value of 10\%).

Clearly that the \textit{OTM}_{VR} risk factor has a significant ability to explain the future SSE 50 return. The estimated coefficient of \textit{OTM}_{VR} for the next 6 months is about 0.22, and the return coefficient for the next 12 months is about 0.26. From the descriptive statistics of \textit{OTM}_{VR}, when \textit{OTM}_{VR} increases from a minimum value of 0.05 to a maximum value of 1.45, the future The cumulative excess return in 6 months will therefore increase by 30.8\%, and the cumulative excess return in the next 12 months will therefore increase by 36.4\%. Comparing the observations of $\tau$ for 6 months and 12 months, we can find that the coefficient of \textit{OTM}_{VR} has shown signs of decay, indicating that the impact of the OTM option volume risk on the future returns of the SSE 50 began to weaken after 6 months. The above regression results show that when the explanatory variable contains \textit{OTM}_{VR}, the R-squared (goodness of fit) of the model is significantly improved. For example, the goodness of fit of the model that includes \textit{OTM}_{VR}’s

### Table 4. Prediction on SSE 50 index using \textit{OTMV} risks

#### Panel A

| Univariate | Multivariate |
|------------|--------------|
| Intercept  | -0.11 (6.99) | -0.00 (6.99) |
| \textit{OTM}_{VR} | 0.20 (7.95) | 0.22 (9.95) |
| \textit{OTM}_{V1} | 0.00 (0.03) | 0.01 (0.12) |
| \textit{OTM}_{V2} | 0.00 (0.08) | 0.01 (0.15) |
| \textit{OTM}_{V3} | 0.00 (0.05) | 0.01 (0.16) |
| Log(P/D)   | 0.35 (4.77) | 0.52 (5.71) |
| 1 Month L.R. | -0.08 (1.15) | 0.01 (0.16) |
| 3 Month L.R. | -0.03 (0.45) | -0.03 (0.45) |
| $R^2$      | 24.88 | 45.56 |

#### Panel B

| Univariate | Multivariate |
|------------|--------------|
| Intercept  | -0.15 (6.68) | -0.03 (2.03) |
| \textit{OTM}_{VR} | 0.23 (6.48) | 0.26 (9.36) |
| \textit{OTM}_{V1} | -0.05 (0.34) | -0.07 (0.56) |
| \textit{OTM}_{V2} | -0.00 (0.05) | -0.01 (0.16) |
| \textit{OTM}_{V3} | -0.03 (0.43) | 0.01 (0.15) |
| Log(P/D)   | -0.72 (8.14) | -0.82 (7.41) |
| 1 Month L.R. | 0.42 (2.61) | -0.01 (0.06) |
| 3 Month L.R. | 0.18 (2.03) | 0.16 (1.41) |
| $R^2$      | 42.02 | 57.64 |
multivariate regression for the next 12 months is 57.64%. This result proves that the $OTM_V$ factor has a strong ability to explain the future cumulative excess return of the SSE 50 Index. In addition, we found that although $OTM_V$’s future returns on the SSE 50 Index are very significant and robust, the three factors $OTM_{V1}$, $OTM_{V2}$, and $OTM_{V3}$ have no predictive power for SSE 50 index returns.

The regression results clearly tell us that when the risk of $OTM_V$ is too high, the SSE 50 Index will compensate for this risk in the future, so there will be positive returns to come; but the three factors of $OTM_V$ do not have similar predictive power. This illustrate that the option trader’s choice between OTM options and non-OTM options has a significant impact on the macro-level trend of the stock market (although we can argue that the SSE 50 index does not represent the macro state of the Chinese stock market, but comparing with the individual stocks, the index itself is relatively macroscopic, representing a signal which is more relevant to the systematic risk of the stock market). And the current increased shocks caused by $OTMV$’s innovation risk has no significant predictive power at this level. This confirms the above analysis of the prediction mechanism of informed trading behavior. We believe that informed traders first judge the market based on their private information, and then choose between OTM options and non-OTM options. This selection behavior will be based on their private information about the underlying assets (SSE 50 Index), and it has a strong predictive effect; after a trader chooses the OTM option as a hedging tool, he must further determine the transaction volume, which will subsequently bring a shock to the current OTM option transaction volume, we expect this type of innovation risk will have a deeper impact on the micro level, so in the following empirical section, we will explore the pricing power cross-sectionally among individual stocks.

4. Research on the cross-sectional pricing ability of $OTMV$

For robustness concern, we also include other classic stock excess return pricing factors when analyzing in this section, which are $MktRe$: excess market return (here we use bank’s one-year deposit interest rates from as risk-free return), and the Fama and French pricing factors: $SMB$ and $HML$, and the momentum factor $UMD$, and all those factor data downloaded from Guotai Junan terminal. When it comes to pricing power of the risk factor in cross-section (Chang et al. (2013)), the basic assumption is that when a stock has an exposure on this risk factor, the risk exposure will bring a future risk compensation, that is, the risk premium. Therefore, in order to study the pricing power of the OTM option volume risk factor in the cross-section of individual stocks, first use all constituent stocks of the SSE 50 index to calculate the exposure of the individual stock returns on the factor, and the model is regressed as follows ($t$ in $[1, T]$)

$$R_{i,t} = \beta_0^i + \beta_m^i \cdot (R_{m,f} - r_{f,t}) + \beta_{OTMV}^i \cdot OTMV_t + \epsilon_{i,t},$$

(4)

where $R_{i,t}$ is the excess return of the $i$-th stock (minus the current risk-free return $r_{f,t}$, and $R_{m,f}$ is the market portfolio return). Then $\beta_{OTMV}^i$ is the $OTMV$ risk exposure of $i$-th stock. If the $OTMV$ risk does have pricing power across individual stocks, at this time, part of the future excess returns of $i$-th stock should be brought by the stock’s $OTMV$ risk exposure. After the exposure estimation of all stocks ($i = 1, \cdots , N$), we use the exposure to the $OTMV$ risk to calculate the premium contribution of $OTMV$ risk factors (premium or price of risk) $i = 1, \cdots , N$,

$$E[R_i] = \lambda_0 + \lambda_m \cdot \beta_m^i + \lambda_{OTMV} \cdot \beta_{OTMV}^i.$$  

(5)

In the research of $OTMV$’s pricing power of individual stocks, $\lambda_{OTMV}$ is called the price of risk factor, and the pricing power of the factor is reflected in the significance and value of the estimation of $\lambda_{OTMV}$. Based on the assumptions of formula (1) and (2), we use $OTMV$ beta-based portfolios sorting method to study the pricing power of $OTMV$ risk in the pricing of individual stocks (Chang et al. (2013)).
4.1. OTMV beta-based portfolio sorting method

In order to verify the above theoretical hypothesis, we first use a beta-based portfolio sorting method, firstly we estimate the exposure of individual stocks on OTMV, and then use the risk exposure to rank individual stocks ascendingly and have 50 stocks divided into 5 portfolios: $Q_1, Q_2, Q_3, Q_4$ and $Q_5$. These 5 portfolios contain 10 stocks each and $Q_1$ is the portfolio with the lowest OTMV risk exposure and $Q_5$ the highest. For the first step, assuming the current time period is $t$, using the historical data from period $t - k$ to $t - 1$ (we use the data for last 52 weeks, therefore $k$ equals to 52 weeks, roughly 1 year), fit the regression model (1), and obtain the exposure of all stocks $\beta^i_{OTMV}, i = 1, \cdots, N$. We divide the stock into 5 asset portfolios based on the size of the $\beta^i_{OTMV}$, and record the excess returns (Ex. Re. for short) of next time period (time from $t$ to $t + \tau$) after every portfolio formation process, and at the same time record the CAPM model’s alpha (“C-alpha” for short), Fama-French three factor model’s alpha (“3F-alpha” for short) and Carhart four factor model’s alpha (4F-alpha for short) and five-factor model alpha (5F-alpha for short) of each portfolio. Finally, we use a two-sample t-test to examine whether there is a significant difference between the excess return (or alphas) of the portfolio $Q_5$ and that of portfolio $Q_1$.

Table 5 and Table 6 reports the results of portfolio returns based on OTMV$_{V1}$, OTMV$_{V2}$, OTMV$_{VR}$ and OTMV$_{V3}$ factors; In each table, Panel A, Panel B, and Panel C use $\tau$ equals to 2 months, 3 months, and 6 months (since we find that the results of each factor are not significant while using $\tau = 1$ month, so these cases are not presented here.), respectively. Among them, columns $Q_1$ to $Q_5$ present 5 portfolios average excess returns and alphas; columns $Q_5 - Q_1$ present the difference between the excess return or alphas of the portfolio $Q_5$ and that of portfolio $Q_1$ and also the t-value of the two-sample test in the parenthesis below (these in bold are the cases that the test results are significant).

### Table 5. OTMV(OTMV$_{V1}$ and OTMV$_{V2}$) beta-based Portfolio Returns

| Panel A | Prediction for next 2-month return | OTMV$_{V1}$ | OTMV$_{V2}$ |
|---------|-----------------------------------|-------------|-------------|
|         | $Q_1$ | $Q_2$ | $Q_3$ | $Q_4$ | $Q_5$ | $Q_5-Q_1$ | $Q_1$ | $Q_2$ | $Q_3$ | $Q_4$ | $Q_5$ | $Q_5-Q_1$ |
| Ex. Re. | 0.038 | 0.094 | 0.090 | 0.122 | 0.147 | 0.109 | (1.81) | 0.065 | 0.095 | 0.025 | 0.129 | 0.168 | 0.103 | (1.73) |
|         | C-alpha | 0.036 | 0.094 | 0.098 | 0.126 | 0.136 | 0.100 | (1.58) | 0.055 | 0.104 | 0.0285 | 0.134 | 0.159 | 0.104 | (1.66) |
|         | 3F-alpha | 0.013 | 0.075 | 0.086 | 0.092 | 0.161 | 0.148 | (2.22) | 0.039 | 0.084 | 0.010 | 0.098 | 0.186 | 0.147 | (2.27) |
|         | 4F-alpha | 0.018 | 0.071 | 0.086 | 0.103 | 0.143 | 0.161 | (2.42) | 0.011 | 0.083 | 0.009 | 0.100 | 0.175 | 0.163 | (2.54) |

| Panel B | Prediction for next 3-months return | OTMV$_{V1}$ | OTMV$_{V2}$ |
|---------|-----------------------------------|-------------|-------------|
|         | $Q_1$ | $Q_2$ | $Q_3$ | $Q_4$ | $Q_5$ | $Q_5-Q_1$ | $Q_1$ | $Q_2$ | $Q_3$ | $Q_4$ | $Q_5$ | $Q_5-Q_1$ |
| Ex. Re. | 0.016 | 0.079 | 0.077 | 0.111 | 0.146 | 0.130 | (2.66) | 0.062 | 0.069 | 0.017 | 0.106 | 0.164 | 0.101 | (2.11) |
|         | C-alpha | 0.004 | 0.068 | 0.070 | 0.102 | 0.130 | 0.126 | (2.48) | 0.042 | 0.065 | 0.008 | 0.098 | 0.152 | 0.109 | (2.20) |
|         | 3F-alpha | -0.01 | 0.054 | 0.054 | 0.075 | 0.148 | 0.162 | (3.08) | 0.035 | 0.041 | -0.01 | 0.070 | 0.173 | 0.138 | (2.71) |
|         | 4F-alpha | -0.04 | 0.055 | 0.053 | 0.088 | 0.131 | 0.168 | (3.11) | 0.015 | 0.041 | -0.01 | 0.075 | 0.159 | 0.144 | (2.75) |
significant under critical value of 10%); all returns or alphas in the portfolio return table are annualized (the values shown in the table are in units of percentages.). Based on the results in table 5 and 6, although the test for \( \tau = 2 \) months is not very consistent, but for \( \tau = 6 \) months and 12 months, \( OTM_{V1} \), \( OTM_{V2} \) and \( OTM_{V3} \) factors’ portfolio returns are consistently significant almost under different prediction intervals and different model assumptions, so the overall conclusion can be claimed as the \( OTM_{V} \) risk has strong and consistently significant pricing power for the risk premium of individual stocks, the price of risk factors is significantly positive and it is stable over a long periods of

| Panel C | Prediction for next 6-months return |
|---------|-------------------------------------|
|         | \( OTM_{V1} \) | \( OTM_{V2} \) | \( OTM_{V3} \) |
|         | \( Q_1 \) | \( Q_2 \) | \( Q_3 \) | \( Q_4 \) | \( Q_5 \) | \( Q_2-Q_1 \) | \( Q_1 \) | \( Q_2 \) | \( Q_3 \) | \( Q_4 \) | \( Q_5 \) | \( Q_2-Q_1 \) |
| Ex. Re. | 0.017 | 0.058 | 0.074 | 0.107 | 0.167 | 0.150 (3.67) | 0.054 | 0.037 | 0.037 | 0.100 | 0.099 | 0.124 (3.05) |
| C-alpha | 0.005 | 0.049 | 0.065 | 0.098 | 0.155 | 0.150 (3.45) | 0.036 | 0.033 | 0.038 | 0.091 | 0.167 | 0.131 (3.07) |
| 3F- alpha | -0.01 | 0.025 | 0.036 | 0.075 | 0.149 | 0.159 (3.71) | 0.023 | 0.007 | 0.012 | 0.058 | 0.167 | 0.144 (3.47) |
| 4F-alpha | -0.02 | 0.018 | 0.036 | 0.076 | 0.136 | 0.151 (3.50) | 0.015 | 0.005 | 0.012 | 0.058 | 0.153 | 0.138 (3.26) |

| Table 6. \( OTMV \) ( \( OTM_{V1} \) and \( OTM_{V2} \)) beta-based Portfolio Returns |
|-----------------|-----------------|-----------------|
| Panel A         | Prediction for next 2-month return |
| \( OTM_{VR} \)  | \( OTM_{V2} \)  | \( OTM_{V3} \)  |
| \( Q_1 \) | \( Q_2 \) | \( Q_3 \) | \( Q_4 \) | \( Q_5 \) | \( Q_2-Q_1 \) | \( Q_1 \) | \( Q_2 \) | \( Q_3 \) | \( Q_4 \) | \( Q_5 \) | \( Q_2-Q_1 \) |
| Ex. Re. | 0.162 | 0.055 | 0.035 | 0.071 | 0.158 | -0.003 (-0.06) | 0.087 | 0.023 | 0.069 | 0.109 | 0.196 | 0.108 (1.81) |
| C-alpha | 0.150 | 0.053 | 0.037 | 0.079 | 0.159 | 0.010 (0.15) | 0.088 | 0.023 | 0.070 | 0.108 | 0.194 | 0.105 (1.69) |
| 3F- alpha | 0.155 | 0.054 | 0.006 | 0.046 | 0.158 | 0.003 (0.04) | 0.089 | -0.01 | 0.055 | 0.088 | 0.193 | 0.104 (1.61) |
| 4F-alpha | 0.146 | 0.036 | -0.01 | 0.043 | 0.160 | 0.014 (0.23) | 0.084 | -0.00 | 0.043 | 0.069 | 0.182 | 0.097 (1.51) |

| Panel B         | Prediction for next 3-months return |
| \( OTM_{VR} \)  | \( OTM_{V2} \)  | \( OTM_{V3} \)  |
| \( Q_1 \) | \( Q_2 \) | \( Q_3 \) | \( Q_4 \) | \( Q_5 \) | \( Q_2-Q_1 \) | \( Q_1 \) | \( Q_2 \) | \( Q_3 \) | \( Q_4 \) | \( Q_5 \) | \( Q_2-Q_1 \) |
| Ex. Re. | 0.140 | 0.042 | 0.032 | 0.062 | 0.141 | 0.001 (0.03) | 0.085 | 0.022 | 0.049 | 0.092 | 0.175 | 0.090 (1.86) |
| C-alpha | 0.116 | 0.028 | 0.019 | 0.060 | 0.141 | 0.025 (0.52) | 0.074 | 0.007 | 0.040 | 0.086 | 0.164 | 0.090 (1.86) |
| 3F- alpha | 0.116 | 0.026 | -0.00 | 0.031 | 0.137 | 0.021 (0.42) | 0.076 | -0.02 | 0.026 | 0.063 | 0.168 | 0.091 (1.78) |
| 4F-alpha | 0.111 | 0.019 | -0.01 | 0.022 | 0.137 | 0.026 (0.51) | 0.073 | -0.02 | 0.019 | 0.051 | 0.155 | 0.081 (1.57) |

| Panel C         | Prediction for next 6-months return |
| \( OTM_{VR} \)  | \( OTM_{V2} \)  | \( OTM_{V3} \)  |
| \( Q_1 \) | \( Q_2 \) | \( Q_3 \) | \( Q_4 \) | \( Q_5 \) | \( Q_2-Q_1 \) | \( Q_1 \) | \( Q_2 \) | \( Q_3 \) | \( Q_4 \) | \( Q_5 \) | \( Q_2-Q_1 \) |
| Ex. Re. | 0.126 | 0.060 | 0.057 | 0.065 | 0.113 | -0.01 (-0.3) | 0.098 | 0.045 | 0.043 | 0.059 | 0.170 | 0.071 (1.77) |
| C-alpha | 0.107 | 0.046 | 0.045 | 0.063 | 0.108 | 0.00 (0.02) | 0.086 | 0.033 | 0.033 | 0.051 | 0.161 | 0.075 (1.79) |
| 3F- alpha | 0.105 | 0.026 | 0.018 | 0.028 | 0.094 | -0.01 (-0.28) | 0.078 | 0.000 | 0.007 | 0.023 | 0.161 | 0.082 (2.04) |
| 4F-alpha | 0.100 | 0.022 | 0.012 | 0.021 | 0.094 | 0.01 (-0.24) | 0.078 | -0.00 | 0.004 | 0.004 | 0.148 | 0.070 (1.69) |
time. However, the $OTM_{VR}$ risk factor’s portfolio returns are not significant at all, it means that $OTM_{VR}$ does not contribute to the risk premium of individual stocks significantly.

4.2. Performance of OTMV beta-based portfolio

In order to more intuitively see the pricing ability of different factors, we construct a $OTMV$ beta-based strategy using all four $OTMV$ risk factors. The specific method is similar to the beta-based portfolio formation method. First, the stocks’ exposure to $OTMV$ is estimated, and then the stocks are divided into portfolios based on the the risk exposures of the stocks. Suppose we now have the historical data from period $t - k$ to $t - 1$ at week $t$, first we fit the regression model (1) to obtain $\beta_{i}^{OOGV}$, for $i = 1, \cdots, 50$. Assuming that the price of the risk factor on future returns is positive (negative), we hold 10 stocks with the highest (low) $\beta_{i}^{OOGV}$ for the purpose of profiting (given that the Chinese stock market has no short mechanism, so we don’t consider shorting strategy).

Repeat the above operation continuously every week, then we get our $OTMV$ bate-based strategy, based on the position calculation and the daily market value of the portfolio’s closing price on the day, then calculate the simple cumulative return $r_{k} = (value_{k} - value_{1})/value_{1}$, where $value_{k}$ is the value of the portfolio on the $k$-th day, $value_{1}$ is the initial value of the investment portfolio. Figure 4 represents the series of cumulative returns of a single factor’s beta-based strategy. Because the results of $OTM_{V1}$, $OTM_{V2}$ and $OTM_{V3}$ obtained by different processing methods are similar, we only show one of the factors which has the highest cumulative return along with $OTM_{VR}$ factor. The cumulative returns of simply holding the SSE50 index at the same time are also plotted for comparison. Figure shows that from February 22, 2016 to February 2018, the cumulative return of the single-factor investment strategy based on the $OTM_{V}$ risk factor reached nearly 200%, which far outperformed the cumulative return of the SSE 50 Index at the same period (50%); as of February 2019, the cumulative return of $OTM_{VR}$ beta-based strategy reached nearly 150%, while the cumulative return of the SSE 50 Index during the same period was only 40%. This result is very consistent with the portfolio partition result based on the volatility option volume risk beta.

Therefore, we use two different methods to explore the pricing power of four risk factors on individual stocks, and they all state the same conclusion: the $OTM_{V}$ risk have significant pricing power in individual stocks, and the risk price is robustly positive across different experiment cases; $OTM_{VR}$ has been proven to have no significant individual stock pricing power under various assumptions. This also confirms our previous expectation: $OTM_{V}$ risks characterize the new risk shocks caused by informed traders in the current period after traders selecting the OTM option, which is microscopic, so that exposed individual stocks will gain a risk premium in the future. Combining the result in section 3, according to our assumptions: some options traders have private expectation for future “reverse risk” and they are make a choice in the options market according to that expectation, and this choice is believed to directly affect the future trend at the macro level (SSE 50 index), and then the shock impact
of their actions (depicted by the $OTM_V$ risk) is further transmitted to the more micro-level risks, which in turn affects some $OTMV$-sensitive stocks’ excess returns in the future. In order to verify our risk transmission mechanism, in the next section we use a strict statistical method to look at the causal relationship between the four $OTMV$ risk variables and the sentiment indicators of SSE 50 index.

5. Risk transmission mechanism of “reverse risk” trading

Through the previous research, we found that the ratio of OTM and non-OTM option trading volume has a significant ability to predict the index return, and there is no significant pricing power for individual stocks; while the innovation risk has no significant effect on the index return, but possess significant pricing power in cross section. In order to analyze the risk transmission and risk premium compensation mechanism of $OTMV$ risk, we use the proxy variables (including turnover rate (shorted as Turn. Rate), logarithmic trading volume (shorted as Volume) and last 60-days-historical volatility (shorted as His. Vol.)) of the trading sentiment of the SSE 50 index to conduct the Granger causality test with all four $OTMV$ risk factors. The Granger test results for two types of variables ($OTMV$ risk and trading sentiment indicators of the SSE 50 index) are presented in Table 7, where “$\rightarrow$” and “$\leftarrow$” indicate cause and effect directions, for instance, in row “Turn. Rate” and column “$OTM_V$”, “$\rightarrow$” indicate the test result for “Turn. Rate as the cause of $OTM_V$”, and “$\leftarrow$” indicates the test result for “$OTM_V$ as the cause of Turn. Rate”, also the $p$-values of the test are shown in parentheses below. Overall, we can see that the $OTM_V$ risk as cause of trading sentiment of SSE 50 index are consistently significant, but not vice versa; and the trading sentiment indicators of the SSE 50 index as the causes of $OTM_V$ risk is significant, but not vice versa. In conjunction with the former literature, next we explore the risk transmission mechanism of “reverse risk” trading.

Gkionis et al. (2018) proposed that the price of the OTM options implied positive information of the underlying stock. The article believed that the price determines the current RNS. Empirical research by Gkionis et al. (2018) proved that long-term holding of high RNS stocks will bring significant positive alpha. Gärleanu et al. (2009) also believe that the current higher RNS does reflect the additional demand for OTM options, and this demand is precisely because the trader caught the “stock underpricing”. We believe that the decisions of option traders in the options market reflect their private information about the underlying assets. In particular, the trading behaviors of these traders for OTM options reflect their expectations of the underlying stocks inevitably reversing in the future, to the raise of $OTMV$ risk will bring a positive compensation into the stocks’ future return.

Based on the above theory, our $OTM_{VR}$ measure represents the decision-making process of option traders with “reverse risk” information, and $OTM_V$ risk represent the

| Table 7. Granger causality test on OTMV and Sentiment Indicator |
|---------------------------------------------------------------|
| $OTM_{VR}$ | $OTM_{V1}$ | $OTM_{V2}$ | $OTM_{V3}$ |
| Turn. Rate | $\rightarrow$ | $\rightarrow^{***}$ | $\rightarrow^{***}$ | $\rightarrow^{***}$ |
| Volume | $\rightarrow$ | $\rightarrow^{***}$ | $\rightarrow^{***}$ | $\rightarrow^{***}$ |
| His. Vol. | $\rightarrow$ | $\rightarrow$ | $\rightarrow^{**}$ | $\rightarrow$ |
| $p$-value | (0.53) | (0.002) | (-0.000) | (0.001) |
| $p$-value | (0.38) | (0.002) | (-0.000) | (0.001) |
| $p$-value | (0.07) | (0.836) | (0.021) | (0.658) |

| Turn. Rate | $\leftarrow^{***}$ | $\leftarrow$ | $\leftarrow$ | $\leftarrow$ |
| Volume | $\leftarrow^{***}$ | $\leftarrow$ | $\leftarrow$ | $\leftarrow$ |
| His. Vol. | $\leftarrow^{***}$ | $\leftarrow$ | $\leftarrow$ | $\leftarrow$ |
| $p$-value | (-0.000) | (0.221) | (0.488) | (0.238) |
| $p$-value | (-0.000) | (0.190) | (0.428) | (0.330) |
| $p$-value | (-0.000) | (0.599) | (0.855) | (0.988) |

Significance symbols: “$^{***}$”, “$^{**}$”, “$^{*}$” and “.” represent significant levels of 0.01%, 1%, 5% and 10%, respectively.
current risk shock brought by the decisions of these traders. From the perspective of risk transmission, Table 7 clearly tells us that the risk information held by these traders first affects their decision-making process and leads to their trading tendencies. Through these traders’ decisions in the options market, the risk will be transmitted to the SSE 50 index trading market, and then bring the innovated shocks of “reverse risk” into a more micro-level scoping, and eventually affect the constituent stocks of the SSE 50.

6. Conclusions
In this paper we firstly analyze the trading behavior of informed option traders. We believe that the trading volume of the OTM options can capture the private risk information of these traders, more precisely their expectations of the “reverse risk” of the stock market. We divide “reverse risk” into two types of risk information: the first one is the option selection process of the trader: the trader decides whether to choose a OTM option or a non-OTM option to hedge their private information, which represents whether they expects the future stock to reverse; the second risk information is measured by the change amount of the current OTM option trading volume, which depicts the new information shock brought by the trader's behavior choice. For empirical experiments, in order to verify the prediction ability of the two types of risks on the underlying stocks, we study the effect of the trading volume risk of OTM options on both SSE 50 index and index component stocks using different statistical models. The empirical result shows that the option selection process can significantly predict the future returns of the SSE 50 index, while the information shock caused by the change in OTM options trading volume has solid pricing power at the micro level of individual stocks. We then use Granger causality tests to confirm that the mechanism of risk transmission for the "reverse risk" trading process is: the risk information held by traders first affected their decision-making process leading to their trading tendency, and their decisions on the options market will transmit the risk to the SSE 50 index trading market, and then cause the current information shocks, and finally these shocks will eventually affect the constituent stocks of the SSE 50.

References
[1] Andersen, T. G. , Fusari, N. , & Todorov, V. . (2019). The pricing of tail risk and the equity premium: evidence from international option markets. Journal of Business and Economic Statistics, 1-27.
[2] Bates, D. S. (1991). The crash of ’87: was it expected? the evidence from options markets. The Journal of Finance[J], 46(3), 1009-1044.
[3] Chang, B. Y. , Christoffersen, P. , & Jacobs, K. (2013). Market skewness risk and the cross-section of stock returns. Social Science Electronic Publishing, 107(1), 46-68.
[4] Black, F., Scholes, M. S. (1973). The pricing of options and corporate liabilities. Journal of Political Economy[J], 81(3), 637-654.
[5] Bollerslev, T., Todorov, V., & Lai, X. (2015). Tail risk premia and return predictability. Journal of Financial Economics, 118(1), 113-134.
[6] Bollerslev, T., & Todorov, V., (2011). Tails, fears, and risk premia. Social Science Electronic Publishing, 66(6), 2165-2211.
[7] Easley, D., O’Hara, M. , & Srivivas, P. S. (1998). Option volume and stock prices: evidence on where informed traders trade[J]. The Journal of Finance, 53(2), 431-465.
[8] Fama, Eugene F., and J. D. Macbeth. "Risk, Return, and Equilibrium: Empirical Tests." Journal of Political Economy 81.3(2000):607-636.
[9] Garleanu, N., Pedersen, L. H., Poteshman, A. M. (2009). Demand-based option pricing[J]. Review of Financial Studies, 22(10), 4259-4299.
[10] Jin, W., Livnat, J., Zhang, Y. (2012). Option prices leading equity prices: do option traders have an information advantage[J]? Journal of Accounting Research, 50(2), 32.
[11] Johnson, T. L., So, E. C. (2012). The option to stock volume ratio and future returns[J]. Journal
of Financial Economics, 106(2), 262-286.

[12] Lin, T., Ge, L., Pearson, N. (2016). Why does the option to stock volume ratio predict stock returns[J]. Journal of Financial Economics, 120(3), 601-622.

[13] Mayhew, D. S. (2002). Risk-neutral skewness: evidence from stock options[J]. The Journal of Financial and Quantitative Analysis, 37(3), 471-493.

[14] Merton, R. C. (1976). Option prices when underlying stock returns are discontinuous[J]. Journal of Financial Economics, 3(1-2), 125-144.

[15] Pan, J. (2002). The jump-risk premia implicit in options: evidence from an integrated time-series study[J]. Journal of Financial Economics, 63(1), 3-50.

[16] Pan, J., Poteshman, A. M. (2004). The information in option volume for stock prices[J]. The Review of Financial Studies, 19(3), 871–908.

[17] Roll, R., Schwartz, E., Subrahmanyam, A. (2010). O/S: the relative trading activity in options and stock[J]. Journal of Financial Economics, 96(1), 1-17.

[18] Wu, L., Foresi, S. (2005). Crash-o-phobia: a domestic fear or a worldwide concern[J]? The Journal of Derivatives 13(2), 8-21.

[19] Xing, Y., Zhang, X., Zhao, R. (2010). What does the individual option volatility smirk tell us about future equity returns[J]? Journal of Financial and Quantitative Analysis, 45(03), 641-662.

[20] Yan, S. (2011). Jump risk, stock returns, and slope of implied volatility smile[J]. Journal of Financial Economics, 99(1), 216-233.

[21] Zhang, J., Xiang, Y. (2008). The implied volatility smirk[J]. Quantitative Finance, 8(3), 263-284.