Sign realized jump risk and the cross-section of stock returns: Evidence from China’s stock market

Youcong Chao¹, Xiaoqun Liu²*, Shijun Guo³

1 Business School, Central South University, Changsha, Hunan, P. R. China, 2 Tourism School, Hainan University, Haikou, Hainan, P. R. China, 3 Business School, The University of Queensland, St. Lucia, Brisbane, Australia

* xiaoquliniu22@163.com

Abstract

Using 5-minute high frequency data from the Chinese stock market, we employ a non-parametric method to estimate Fama-French portfolio realized jumps and investigate whether the estimated positive, negative and sign realized jumps could forecast or explain the cross-sectional stock returns. The Fama-MacBeth regression results show that not only have the realized jump components and the continuous volatility been compensated with risk premium, but also that the negative jump risk, the positive jump risk and the sign jump risk, to some extent, could explain the return of the stock portfolios. Therefore, we should pay high attention to the downside tail risk and the upside tail risk.

Introduction

With the development of financial econometrics, the measurement of financial risk variables has made a major breakthrough, which makes the study on time series and cross-sectional stock return-risk relations revealed by the Capital Asset Pricing Model (CAPM) model and the three-factor model has achieved great success. Whether risk measures, such as conditional skewness, tail risk, jumps and rare disaster risk, have a certain risk premium and the similarities and differences among them have become a main focus within the academic circle.

In the frame of the GARCH-class model based on low-frequency data, the short-run and long-run components of market risk have different contributions for the cross-sectional pricing of the volatility risk, and the short-run components capture market skewness risk, which is interpreted as a measure of the tightness of financial constraints¹. Chang estimates a conditional and forward-looking implied market skewness of market returns from the daily Standard & Poor’s 500 index option data, and the results show that the market skewness risk premium is statistically and economically significant and cannot be explained by other common risk factors, such as the market excess return or the size, book-to-market, momentum, and market volatility factors, or by firm characteristics². Chan and Feng find that time-varying jump risk premia are significantly associated with extreme news events or jumps in four stock market index futures returns³. Furthermore, there is a new perspective to investigate...
the return-(high order) risk trade-off relationship under the macro-economic framework with time-varying uncertainties. Some researchers deem that the extended long-run risk model (Extended LRR) that embeds time-varying jumps in the model in various ways (learning, model uncertainty, etc.) can measure the risk premium\cite{4,5}. Kelly and Jiang propose a bivariate prediction equation with the tail risk factors and macro-economic variables\cite{6}, which can be derived by the LRR model\cite{7} and the Disaster Risk model\cite{8}. It can be seen that jumps affect the equity premium in indirect ways; they obtained the estimated jump value through the options, or the jump was embedded in the conditional variance, conditional skewness, kurtosis and other conditions of higher moments of dynamic equations.

Under the framework of high-frequency data with a nonparametric method, the presence of jumps in stocks, portfolios and indexes has been confirmed\cite{9–11}. A number of studies directly estimate the jump intensity, jump mean, and jump volatility and use those realized jump measures as proxies for the entire realized jump risk, rather than as the proxy of skewness, to illustrate its explanatory power on the cross-sectional asset returns\cite{12–15}. Furthermore, Bollerslev and Todorov estimated the jump tail based on the new Extreme Value Theory and high frequency data and combined the jump tail with the investors’ expectation of their fear of extreme events that may occur, proving that the time-varying compensation needed for the expectation of a disaster can explain large parts of the average equity premium and variance risk premium\cite{9}. Another attraction for using a nonparametric method with high frequency data to identify the realized jump variables is that this method can calculate the positive and negative jumps and their distributions in a simple yet accurate way. In regard to stock price movements, large negative price changes, which occur to a certain, large enough extent, would cause much greater future volatility compared to a positive price change of the same magnitude\cite{16}. Further, Guo et al. considered the difference between the upside and downside risks, named the sign jump volatility, to investigate the possible influence on the index return; the empirical results showed that sign jump risk is an important determinant of equity premium, and this risk factor needs to be priced\cite{17}.

To gain insights into the existing studies, we find that all the empirical analysis only use either the index or individual stocks for a time series regression to probe into whether skewness, tail risk, rare disaster risk and realized jump could predict future returns. However, skewness, rare disaster and jump are tails in the theory of statistics, so it is important to distinguish positive and negative tails. Therefore, from the perspective of the portfolio level, the paper, on the cross-sectional level, further investigates whether positive, negative realized jumps and sign realized jumps could predict the excess portfolio return with the Fama-MacBeth two-step regression test. According to the theoretical relations between extreme positive, negative returns and the asymmetric character of volatility, the non-parametric method with high frequency data could be used to make accurate estimation to positive and negative realized jump variables representing upside and downside risk, which will shed new light on our understanding and grasp of financial market rules.

The remainder of the paper is organized as follows. Section 2 describes the identification method for the high frequency positive, negative and sign realized jumps and introduces the jump risk factor model that we employ, and Section 3 discusses the high-frequency data that we study and presents descriptive statistics. Empirical results are presented in Section 4, and we conclude in Section 5.

**Methods and materials**

**Estimation of realized jumps**

While jumps are known to be very crucial in asset pricing, the estimation of jump components by the parametric model has been questioned for stability over different sample time periods.
Recently, realized variance utilizing the information in the intraday data have been used to measure and forecast volatilities\cite{18,19}. More recently, Barndorff-Nielsen and Shephard proposed a series in their seminal work on bi-power variation measure used to divide the realized variance into continuous diffusion volatility and jump volatility, which make different contributions to capturing the dynamic characteristics of the volatility model\cite{20,21}.

Under the reasonable presumption that jumps in financial markets are usually rare and large, we assume that there is at most one jump per day and that the jump size dominates the contributions to capturing the dynamic characteristics of the volatility model\cite{20,21}. A series in their seminal work on bi-power variation measure used to divide the realized variance into continuous diffusion volatility and jump volatility, which make different contributions to capturing the dynamic characteristics of the volatility model\cite{20,21}.

Let $p_t = \log(P_t)$ denote the time $t$ logarithmic price of the asset, and assume that it evolves in continuous time as a jump–diffusion process:

$$dp_t = \mu_t dt + \sigma_t dW_t + J_t dq_t$$

(1)

Where $dq_t$ is a Poisson jump process with intensity $\lambda_p$ and $I_s$ is the corresponding log jump size distribution following a normal distribution $N(\mu_s, \sigma_s)$. Consider

$$RV_t = \sum_{j=1}^{M} \sigma_t^2 \int_{t_{j-1}}^{t_j} ds + \sum_{j=1}^{M} \int_{t_{j-1}}^{t_j} \sigma_t^2 dq_t$$

(2)

where $M$ is the total number of trades between time $t$ and $t + 1$, $j$ is the indication for each trade, $RV_t$ is the realized volatility at time $t$ and $r_{ij} = \log(P_{ij}P_{ij-1})$. In the realistic financial markets, the price volatility of financial asset is not continuous but contains jumps due to the influence aroused by information shocks on the market. To measure discrete jump variation ($RV_t$), a method called realized bipower variation ($BV_t$) is proposed\cite{21,22}:

$$BV_t = \frac{\pi}{2} \frac{1}{M-1} \sum_{j=2}^{M} |r_{ij}| |r_{ij-1}| \to \int_{0}^{t} \sigma_t^2 ds$$

(3)

Where $\sqrt{\pi/2} = E(\Pi_t)$. $\Pi_t$ is a standardized normal distribution random variable, and $M/(M-1)$ is the amendment to sample size. According to Barndorff-Nielsen and Shephard, the difference between $RV_t$ and $BV_t$ is just the consistent estimator of the discrete jump variation when $M \to \infty$; that is,

$$RV_t - BV_t = RJV_t$$

(4)

To enhance the accuracy of discrete jump variation, the $Z_t$ statistic based on the bipower variation theory is used to identify the discrete jump variation\cite{21,23,24}. $Z_t$ is given as follows:

$$Z_t = \frac{RV_t - BV_t}{RV_t} \sqrt{\left(\frac{2}{\pi} + \pi - 5\right) \frac{\pi}{M} \max\left(1, \frac{TP_t}{\pi^3}\right)}$$

(5)

Where,

$$TP_t = \frac{M}{M-2} \cdot \frac{M}{4[\Gamma(7/6)/\Gamma(1/2)]^3} \cdot \sum_{j=2}^{M} |r_{ij}|^{1/2} |r_{ij-1}|^{1/2} |r_{ij-2}|^{1/2}$$

(6)

Under the significance level, $1 - \alpha$, we obtain the estimation of the discrete jump variation:

$$RJV_t = I(Z_t > \Phi^{-1}_\alpha) \sqrt{[RV_t - BV_t]}$$

(7)

where $\Phi$ is the cumulative distribution function of a standard normal, and $I_{t,x} = I(Z_t \geq \Phi^{-1}_\alpha)$ is the resulting indicator function for whether there is a jump during the day; $I_{t,x} = I(Z_t \geq \Phi^{-1}_\alpha)$
equals 1 when a jump is detected on day \( t \) and equals 0 otherwise. In the process of actual operation, we need to choose an appropriate \( \alpha \); previous studies revealed that when jump contributions are 10% and 80%, the significance level should be 0.99 and 0.999, respectively [12].

With the above test of the \( Z_t \) statistic and the related bipower variation theory, we obtain the estimators of \( BV_t \) and \( RJV_t \) and then calculate the monthly jump size (\( Size_{RJV} \)), monthly jump size mean (\( Mean_{RJV} \)), monthly jump arrival rate (\( Arr_{RJV} \)), and monthly jump size standard deviation (\( Std_{RJV} \)). The jump components are defined as follows:

\[
Size_{RJV_{\text{month}}} = \sum Size_{RJV_{\text{day}}} \tag{8}
\]

\[
Mean_{RJV_{\text{month}}} = \frac{Size_{RJV_{\text{month}}}}{N_{RJV_{\text{days}}}} \tag{9}
\]

\[
Arr_{RJV_{\text{month}}} = \frac{N_{RJV_{\text{days}}}}{\text{days}} \tag{10}
\]

\[
Std_{RJV_{\text{month}}} = \left( \frac{1}{N_{RJV_{\text{days}}}} \sum (Size_{RJV_{\text{day}}} - Mean_{RJV_{\text{day}}})^2 \right)^{1/2} \tag{11}
\]

Where \( N_{RJV_{\text{days}}} \) is the total number of days when a realized jump occurs, and “days” denotes the trading days in a month. We follow previous studies and set the confidence level \( \alpha \) at 0.95 in this paper [12, 15].

Construction of empirical model

The realized negative (positive) semi-variance can distinguish downside volatility risk from upside volatility risk. Based on this method, the signed jump variation (difference between the upside risk and the downside risk) is further defined as containing continuous volatility and positive or negative jump variation [22, 25, 26], but our sign jump variation is the difference between the positive jump volatility and the negative jump volatility.

When estimating the realized jump volatility and extracting its characteristic components, we assume that there is at most one jump in a trading day, which has a dominant effect on the stock return once it occurs, and this means that extreme changes in the stock returns are accompanied by the presence of jumps. The non-parametric method can not only identify jump volatility and estimate distribution parameters but can also calculate monthly jump factors accurately and easily, including the sum of monthly positive and negative jumps, monthly jump intensity, monthly jump mean, monthly jump arrival rate and monthly jump standard deviation, which help depict extreme changes in the returns and verify the asymmetric characteristic of volatility more effectively. At this time,

\[
P_r \cdot Size_{RJV} = I\{r \geq 0\} \cdot Size_{RJV} \tag{12}
\]

\[
N_r \cdot Size_{RJV} = I\{r \leq 0\} \cdot Size_{RJV} \tag{13}
\]

\[
Sign \cdot Size_{RJV} = P_r \cdot Size_{RJV} - N_r \cdot Size_{RJV} \tag{14}
\]

Correspondingly, we are able to get the positive, negative and sign jump mean, arrival rate, and the standard deviation, expressed as, \( P_r \cdot Mean, P_r \cdot Arr, P_r \cdot Std, N_r \cdot Mean, N_r \cdot Arr, N_r \cdot Std, Sign \cdot Mean, Sign \cdot Arr, Sign \cdot Std \), respectively.

Meanwhile, based on the simple linear relations between jump components [12, 13], we use the lagged one month realized jump variation components as the predictive value of the conditional variance [27], that is, the regression
model is constructed as shown below:
\[ r_{i,t+1} - r_{i,t} = \gamma_0 + \gamma_1 \text{Mean}_RJV_{i,t} + \gamma_2 \text{Arr}_RJV_{i,t} + \gamma_3 \text{Std}_RJV_{i,t} + \gamma_4 BV_{i,t} + \epsilon_{i,t+1} \] (15)

Based on this model, we consider the negative and positive jump volatility components and sign jump volatility. Thus, the following three equations are constructed as follows, while equation Eq (16) and Eq (17) consider only the negative and positive jump components, respectively, and equation Eq (18) includes only the factor of sign jump components, Eq (19) and Eq (20) expand Eq (16), Eq (17) and Eq (18) to portray the significance of the sign jump factors with the control of the negative or positive jump factors.
\[ r_{i,t} - r_{i,t-1} = \gamma_0 + \gamma_1 N \text{Mean}_{i,t-1} + \gamma_2 N \text{Arr}_{i,t-1} + \gamma_3 N \text{Std}_{i,t-1} + \gamma_4 BV_{i,t-1} + \epsilon_{i,t} \] (16)
\[ r_{i,t} - r_{i,t-1} = \gamma_0 + \gamma_1 P \text{Mean}_{i,t-1} + \gamma_2 P \text{Arr}_{i,t-1} + \gamma_3 P \text{Std}_{i,t-1} + \gamma_4 BV_{i,t-1} + \epsilon_{i,t} \] (17)
\[ r_{i,t} - r_{i,t-1} = \gamma_0 + \gamma_1 \text{Sign}_\text{Mean}_{i,t-1} + \gamma_2 \text{Sign}_\text{Arr}_{i,t-1} + \gamma_3 \text{Sign}_\text{Std}_{i,t-1} + \gamma_4 BV_{i,t-1} + \epsilon_{i,t} \] (18)
\[ r_{i,t} - r_{i,t-1} = \gamma_0 + \gamma_1 N \text{Mean}_{i,t-1} + \gamma_2 N \text{Arr}_{i,t-1} + \gamma_3 N \text{Std}_{i,t-1} + \gamma_4 \text{Sign}_\text{Mean}_{i,t-1} \\
+ \gamma_5 \text{Sign}_\text{Arr}_{i,t-1} + \gamma_6 \text{Sign}_\text{Std}_{i,t-1} + \gamma_7 BV_{i,t-1} + \epsilon_{i,t} \] (19)
\[ r_{i,t} - r_{i,t-1} = \gamma_0 + \gamma_1 P \text{Mean}_{i,t-1} + \gamma_2 P \text{Arr}_{i,t-1} + \gamma_3 P \text{Std}_{i,t-1} + \gamma_4 \text{Sign}_\text{Mean}_{i,t-1} \\
+ \gamma_5 \text{Sign}_\text{Arr}_{i,t-1} + \gamma_6 \text{Sign}_\text{Std}_{i,t-1} + \gamma_7 BV_{i,t-1} + \epsilon_{i,t} \] (20)

It is noteworthy that explanatory variables of Eq (15) represent the following four different combinations: Mean_RJV, Arr_RJV, Std_RJV or Size_RJV, Std_RJV or Size_RJV, Mean_RJV, Std_RJV or Size_RJV, Mean_RJV, Std_RJV; during the actual operation of 84 successive months with the OLS regressions, these four combinations are considered separately, and by comparing the adjusted goodness of fit, squared residuals and other indicators, and conducting the overall simulation of 84 months, we determine which of these combinations are more reasonable.

Data set
Different sampling frequencies of intraday data can affect the research results. On the one hand, low sampling frequency may fail to depict the actual volatility information on that day. On the other hand, high sampling frequency may lead to the problem of micro-noise, which may affect the results. As suggested by previous studies[21,24], five-minute high-frequency data are used for return calculation. The data was collected from the CSMAR high-frequency database from January 4, 2007 to December 31, 2013 for all A-share stocks traded in China’s stock market. The trading hours of China’s stock market span two sessions on a trading day: a morning session from 9:30am to 11:30am and an afternoon session from 1:00pm to 3:00pm. At most, 49 transactions can be collected per day (including one overnight transactions and 48 intraday transactions), and there are 84 months' data in total. The value-weighted monthly returns of 25 portfolios were obtained directly from the RESSET database. We calculate all monthly jump components with high-frequency transactions and run the regression on monthly return over monthly jump components.

Construction of the investment portfolios
In the past, researchers mainly used the composite index or individual stocks with trading activity, larger market capitalization to study the dynamic characteristics of the realized jump
volatility, and its model-predictive\cite{9,10}; one possible reason is that the calculation of the jump process for the dynamic portfolios is complex and cumbersome based on the index and individual stocks. However, for the return-risk trade-off, on the one hand, although the index has been largely used as the sample, it is hard to arrive at a robust conclusion for the individual stocks, which is sensitive to the presence of micro-noise; on the other hand, more importantly, in addition to the time series dimensions, cross-sectional data are indispensable for complete research. In fact, 25 portfolios constructed in the Fama-French model laid a theoretical and empirical economic foundation for the future portfolio-based empirical analysis. It is of great economic application significance to explore whether the realized jump risk may predict the excess portfolios return, which is a salutary supplement of the index from the market level and individual stocks from a firm level.

In this paper, based on the Fama-French three-factor asset pricing model, we use the 25 portfolios that are determined by the size and book-to-market ratio and their constituent stocks as a testing asset. The group forming strategy is as follows: the size is market value at the end of June of year T, the book-to-market is the book value of equity in fiscal year T-1 divided by market value at the end of December, T-1 year; we take the respective quintiles of the market value and the book-to-market, then we obtain 25 groups at the intersection of the two quintiles and finally, according to these portfolios, we calculate their monthly market capitalization-weighted average returns from July, T year to June, T+1 year.

Portfolios with constituent stocks are time-varying every year from 2007 to 2013, that is, in fact, there are \(7 \times 25 = 175\) sets of data. The daily realized jump intensity, mean, standard deviation and arrival rate for each stock are calculated based on five-minute high frequency data of each constituent stock in each group of 25 portfolios per year. Then, all 12 of the monthly jump data for each stock are simply additive combinations of the corresponding daily jump data; finally, the monthly jump components for this portfolio are the average of all constituent stocks. Synthesizing all the data of 7 years, every jump component of each portfolio has 12 monthly jump data every year; then, we obtain 84 sets of time series data of the four jump components for each portfolio. It is worth noting that time varying constituent stocks of the 25 portfolios are not what matter; it is the fact that \(s1b1 \ldots s1b5, s2b1 \ldots s2b5, \ldots, s5b1 \ldots s5b5\) as a whole achieve their significance. Therefore, the constituent stocks that calculate the jump values of all 25 portfolios in 2007 are different from those in 2008 as well as in the seventh year, 2013, as well. By analogy, although different stocks constitute the specified portfolios every year, they can still reflect the time-varying realized jump volatility of the 25 portfolios in 7 years.

**Results and discussions**

**Summary statistics of portfolios and jumps**

Panel A of Table 1 reports the mean, variance, and Sharpe ratio of the market portfolio in the sample interval; group B shows the mean, variance, and Sharpe ratio of the 25 portfolios, which were selected based on their market value and book value; \(s1 \ldots s5\) represent the market values from low to high, and \(b1 \ldots b5\) are the book-to-market ratios from low to high, i.e., \(s5b5\) represents the equity portfolio with the largest market value and the highest book-to-market ratio. As shown, the mean returns of all 25 portfolios are higher than the market portfolio, and 22 portfolios’ Sharpe ratio is larger than that of the market portfolio.

Panel B of Table 1 and Figs 1 and 2 show the average returns of China’s A-share market portfolios for seven years is consistent with the results of Fama and French\cite{28}. It indicates that the average return of stock portfolio has a reverse correlation with the firm size and has a positive correlation with the book-to-market ratio.
As shown in Fig 1 and panel B of Table 1, under control of the smallest B/M value b1, the return of the smallest size portfolio s1b1 is 0.0320 higher than other 4 portfolios; in detail, the return of s2b1, s3b1, s4b1 and s5b1 are 0.0249, 0.0228, 0.0108 and 0.0115, respectively. Likewise, when b2, b3, b4 and b5 are respectively controlled, the returns of portfolios are decreasing with the increasing size. Similarly, see from Fig 2 and Panel B of Table 1, when size is under control, the return of the highest BM portfolio s1b5 is higher than the lowest BM portfolio s1b1, when s1 is under control. The same results are found when s2, s4 or s5 is controlled. Note that the positive relationship between BM and returns does not appear when s3 is under control, and the effect of BM on portfolios’ return is not strictly increased with the increasing BM value, which is different from the size effect.

![Fig 1. Return of Size effect with Fixed B/M bi.](https://doi.org/10.1371/journal.pone.0181990.g001)
Table 2 reports the portfolios’ average positive and negative realized jump components in the sample interval, including the jump size (\(P(N)_{\text{Size}}\)), mean (\(P(N)_{\text{Mean}}\)), arrival rate (\(P(N)_{\text{Arr}}\)) and standard deviation mean (\(P(N)_{\text{Std}}\)) (Only results of s1b1-s1b5 are presented here; see other results in Tables A-D in S1 Text). It is obvious that the positive and negative realized jump components series have significant characteristics of "leptokurtosis and heavy tails". Meanwhile, the number of positive jumps (\(NP\)) is larger than the number of negative ones (\(NN\)) for each portfolio. In this case, although the impact of negative shocks is higher than the positive shocks, the occurrence of the former has a lower probability. Consequently, the positive jump intensity and other positive jump components are higher than those of the negative jumps. We also find that the means of the sign realized jump intensities for all portfolios are positive and appear as the property "high peak and fat tail" (see S1 Table).

![Fig 2. Return of BM effect with Fixed Size si.](https://doi.org/10.1371/journal.pone.0181990.g002)

**Fig 2.** Return of BM effect with Fixed Size \(si\). When \(s1, s2, s4\) and \(b5\) are respectively controlled, the return of the highest BM portfolio is higher than the lowest BM portfolio. Note that the positive relationship between BM and returns does not appear when \(s3\) is under control.

Table 2. Descriptive statistics of the jump components for s1b1-s1b5.

| Portfolios | NN   | NP   | N_Size | N_Mean  | N_Arr  | N_Std  | BV     |
|------------|------|------|--------|---------|--------|--------|--------|
| s1b1       | 1.1001*** | 2.9642*** | 0.0233*** | 0.0136*** | 0.0557*** | 0.0113*** | 0.0221*** |
| s1b2       | 1.1924*** | 2.9022*** | 0.0259*** | 0.0146*** | 0.0603*** | 0.0120*** | 0.0238*** |
| s1b3       | 1.2059*** | 2.8771*** | 0.0292*** | 0.0162*** | 0.0613*** | 0.0132*** | 0.0260*** |
| s1b4       | 1.2401*** | 2.9857*** | 0.0307*** | 0.0171*** | 0.0630*** | 0.0141*** | 0.0271*** |
| s1b5       | 1.2413*** | 3.2091*** | 0.0300*** | 0.0169*** | 0.0629*** | 0.0139*** | 0.0282*** |

Panel A Descriptive Statistics of the Positive Jump Components

| Portfolios | NN   | NP   | P_Size | P_Mean  | P_Arr  | P_Std  | BV     |
|------------|------|------|--------|---------|--------|--------|--------|
| s1b1       | 1.1001*** | 2.9642*** | 0.0595*** | 0.0195*** | 0.1488*** | 0.0028*** | 0.0221*** |
| s1b2       | 1.1924*** | 2.9022*** | 0.0628*** | 0.0203*** | 0.1463*** | 0.0032*** | 0.0238*** |
| s1b3       | 1.2059*** | 2.8771*** | 0.0689*** | 0.0223*** | 0.1463*** | 0.0038*** | 0.0260*** |
| s1b4       | 1.2401*** | 2.9857*** | 0.0729*** | 0.0226*** | 0.1526*** | 0.0037*** | 0.0271*** |
| s1b5       | 1.2413*** | 3.2091*** | 0.0837*** | 0.0239*** | 0.1647*** | 0.0039*** | 0.0282*** |

Panel B Descriptive Statistics of the Negative Jump Components

***, **, * represent the 1%, 5%, and 10% significance levels of the Jarque-Bera Test, respectively.

https://doi.org/10.1371/journal.pone.0181990.t002
Fama-MacBeth cross-sectional regressions

For each portfolio, we make its excess return the dependent variable and its jump volatility factors and the continuous volatility the explanatory variables to do a time series regression analysis with the 84 sets of time series data, and we obtain the corresponding coefficients of these risks premium factors. Then, we calculate the means and standard deviations of the above sets of coefficients, namely, \( r_{kt}(k = 0,1,2,3,4) \), which are the explanatory variables in the second step of the Fama-MacBeth cross-sectional regression, and all 84 monthly average excess returns are the dependent variables. T-statistics are used to test whether the jump volatility components can significantly affect the cross-sectional returns of the portfolios.

First, for all 84 months, we conduct regression analysis with Eq (15) and Eq (16) to Eq (20) six times each. Only the 12-month results indicate that the jump volatility components do not significantly affect the cross-sectional returns of the portfolios for only 12 months, while the results of the remaining 72 months are significant, which confirms that jump components can explain and predict the cross-sectional returns of the 25 portfolios with different combinations of negative, positive, and sign jump volatility. Table 3 provides the summary of the number of significant months for the 25 portfolios.

Next, as shown in Table 4, for Eq (15), out of the 84 months, there are 29 months whose jump components explain the excess return in the form of a linear combination of Mean, Arr and Std, and there are 22 months whose explanatory variables are the linear combination of Size, Mean and Std; additionally, there are 21 months in the form of a linear combination of Size, Arr and Std. More generally, the majority of the jump components’ coefficients are positive, which is in accord with the economic theory of high risk with high risk premium. There are also 8 months whose continuous volatility (BV) is significant; in general, the negative coefficient of BV is always accompanied by the negative coefficients of the Arr and Std, and there are also positive coefficients of the Mean and Std. For those negative jump and continuous volatility values, it is possible to interpret them from the perspective of the negative relationship between idiosyncratic volatility and return.

Third, as for Eq (16) and Eq (17), the linear combinations of \( P_{\text{Mean}}, P_{\text{Arr}} \) and \( P_{\text{Std}} \) and \( N_{\text{Mean}}, N_{\text{Arr}} \) and \( N_{\text{Std}} \) have the largest number of significant months, 26 and 18, respectively. There are both negative and positive coefficients of the negative or the positive jump components, and the numbers of negative coefficients of the negative jump mean and arrival rate for the negative jump components are larger than for the corresponding positive coefficients; the ratios are 8/4 and 8/3, respectively. However, the numbers of positive coefficients of the positive jump mean, standard deviation and arrival rate are larger than for the corresponding negative coefficients, which suggest that part of the increased current volatility comes from the negative or positive jumps of the past month. Under the four linear combinations of the negative and positive jump components, the numbers of months with significant continuous volatility (BV) are 14 and 11, respectively, and the coefficients are half positive and half negative.

| Regression Equations                                      | Significant Months |
|-----------------------------------------------------------|--------------------|
| Significant Months for Jump (Eq 15)                       | 40 Months in Total |
| Significant Months for Negative Jump (Eq 16)              | 34 Months in Total |
| Significant Months for Positive Jump (Eq 17)              | 43 Months in Total |
| Significant Months for Sign Jump (Eq 18)                 | 41 Months in Total |
| Significant Months for Sign and Negative Jump (Eq 19)    | 55 Months in Total |
| Significant Months for Sign and Positive Jump (Eq 20)    | 42 Months in Total |

https://doi.org/10.1371/journal.pone.0181990.t003
negative. The jump volatility and the continuous volatility are both significant in the regression model, which indicates that the two types of volatility make different contributions to explaining the excess return of the portfolios. Meanwhile, for some portfolios, their jump component factors are significant and yet their continuous volatility is non-significant, which means that it is necessary to add the extreme jump volatility risk to the equilibrium analysis due to different cross-sectional characteristics when we discuss the relationship between return and risk.

As is seen from the results, different types of linear combinations of the jump components should be considered when we study the effect of jump volatility on the excess return of the portfolio. Through the analysis and comparison of the different jump components’ combinations, the optimal linear portfolio regression with the highest value of goodness-of fit and the least MSE could be found.

Furthermore, according to the studies concerning sign realized jump and positive and negative realized jump means[13,17], in this paper, the sign jump components in terms of a certain linear combination are directly treated as the explanatory variables using the same regression method as Eq (15), Eq (16) and Eq (17) to explore whether these factors contain some additional information for predicting the stock return. Eq (18) considers only the sign jump variables, Eq (19) and Eq (20) test the robustness of the sign jump factor under the

| Linear Combinations | Size | Arr | Std | BV | Number of Significant Months |
|----------------------|------|-----|-----|----|-----------------------------|
| Size / Arr / Std     | 4+, 4-| 6+, 5-| 7+, 5-| 3+, 2-| 21 |
| P_Size / Arr / Std   | 7+, 2-| 4+, 6-| 3+, 8-| 1+, 2-| 19 |
| N_Size / Arr / Std   | 3+, 4-| 3+, 4-| 3+, 3-| 1+, 2-| 14 |
| Sign_Size / Arr / Std| 8+, 4-| 8+, 7-| 3+, 8-| 5+, 4-| 24 |
| P_Sign_Size / Arr / Std | — | — | — | — | 22 |
| N_Sign_Size / Arr / Std | — | — | — | — | 31 |
| Size / Mean / Std    | 5+, 3-| 3+, 6-| 7+, 4-| 1+, 2-| 22 |
| P_Size / Mean / Std  | 11+, 3-| 7+, 4-| 2+, 5-| 2+, 2-| 23 |
| N_Size / Mean / Std  | 3+, 8-| 8+, 2-| 2+, 7-| 1+, 1-| 13 |
| Sign_Size / Mean / Std | 6+, 2-| 3+, 6-| 6+, 4-| 2+, 1-| 20 |
| P_Sign_Size / Mean / Std | — | — | — | — | 13 |
| N_Sign_Size / Mean / Std | — | — | — | — | 24 |
| Size / Arr / Mean    | 4+, 2-| 4+, 6-| 5+, 5-| 1+, 3-| 18 |
| P_Size / Arr / Mean  | 7+, 0-| 2+, 6-| 2+, 6-| 1+, 2-| 15 |
| N_Size / Arr / Mean  | 2+, 4-| 2+, 4-| 4+, 3-| 1+, 1-| 13 |
| Sign_Size / Arr / Mean | 7+, 7-| 7+, 8-| 4+, 5-| 4+, 4-| 21 |
| P_Sign_Size / Arr / Mean | — | — | — | — | 19 |
| N_Sign_Size / Arr / Mean | — | — | — | — | 20 |
| Mean / Arr / Std     | 8+, 5-| 9+, 7-| 7+, 6-| 1+, 3-| 29 |
| P_Mean / Arr / Std   | 7+, 6-| 9+, 9-| 2+, 6-| 3+, 1-| 26 |
| N_Mean / Arr / Std   | 4+, 8-| 3+, 8-| 8+, 3-| 2+, 2-| 18 |
| Sign_Mean / Arr / Std | 1+, 3-| 6+, 4-| 4+, 5-| 2+, 4-| 25 |
| P_Sign_Mean / Arr / Std | — | — | — | — | 22 |
| N_Sign_Mean / Arr / Std | — | — | — | — | 31 |

The symbols “+” and “-” represent the positive and negative coefficients of the variables, respectively.

https://doi.org/10.1371/journal.pone.0181990.t004
control of the positive and negative jump components. In terms of the economic theory, not only the asymmetric characteristic of the volatility or the theory of leverage but also the loss-aversion of prospect theory can explain the asymmetry: the reduction in the subjective utility of a unit loss is greater than the increase in the utility of the same amount of gain. The explanation is verified by empirical analysis, which finds that the mean of the positive jump components is larger than the mean of the negative ones [26].

The empirical results show that there are 43 and 34 months for which positive and negative jumps are significant, respectively, and 41 months for which the sign jumps are significant, which is almost the same as the situation when the linear jump combination is only considered, for which the number of significant months is 42. In addition, the number of significant months is 42 when the positive jump and the sign volatility are both considered, while the number of significant months reaches 55 when the negative jump and the sign volatility are added into the regression. The results demonstrate that both the negative and positive jumps are the common factors that can explain and predict the excess returns; in particular, the negative jumps have greater explanatory power when the sign jump factors are under control. There is no significant law for the positive or negative value of the coefficients of the significant jump variables, but it might be a problem deserving of systematic search that the sign of the same variable is different under different linear combinations.

Finally, all the above 84 coefficients of the jump components and continuous volatility are used as explanatory variables for the second step of the Fama-MacBeth cross-sectional regression, and the average excess returns of all months through seven years are treated as explained variables. The significance of the coefficients of these variables is tested by T-statistics or P-Values.

The results are shown in Table 5: For Eq (19), whose explanatory variables are the linear combination of \( N_{\text{Sign Size}} , N_{\text{Sign Arr}} , N_{\text{Sign Std}} \) and \( BV \), the sign jump arrival rate passes the t-statistic test, which refutes the null hypothesis of a zero coefficient at a significance level of 10%, and the coefficient is significantly negative. In addition, as for Eq (18), the linear combination of \( \text{Sign Size} , \text{Sign Arr} , \text{Sign Std} \) and \( BV \) is used to conduct OLS regression, and the coefficients of the sign jump intensity and the realized continuous volatility are significantly positive at the 5% significance level, which indicates that both the sign jump risk and the continuous volatility risk obtain risk compensation. According to Eq (15), the regression results demonstrate that the realized standard deviation has a negative risk price, which can be

| Table 5. Cross-sectional regression results. |
|-----------------------------------------------------|
| **Panel A. Results of Eq (19)**                      | **Panel B. Results of Eq (18)**                      | **Panel C. Results of Eq (15)**                      |
| **Variables**                                       | **Coefficients**                                    | **Variables**                                       | **Coefficients**                                    | **Variables**                                       | **Coefficients**                                    |
| Intercept                                           | 0.0155 (1.1345)                                     | Intercept                                           | 0.0125 (0.9667)                                     | Intercept                                           | 0.0195 (1.5399)                                     |
| \( N_{\text{Size}} \)                               | -0.0041 (-0.7276)                                   | \( \text{Sign Size} \)                              | 0.0330** (2.0146)                                   | Mean                                                | 0.0033 (0.9471)                                     |
| \( N_{\text{Arr}} \)                                | -0.0222* (-1.7815)                                  | \( \text{Sign Arr} \)                               | -0.0043 (-1.543)                                    | \( \text{Arr} \)                                    | -0.0528 (-1.5944)                                   |
| \( N_{\text{Std}} \)                                | -0.0032 (-0.8817)                                   | \( \text{Sign Std} \)                               | 0.0062 (1.4514)                                     | \( \text{Std} \)                                    | -0.0022** (-2.0074)                                 |
| \( \text{Sign Size} \)                              | 0.0109 (0.7663)                                     | \( \text{BV} \)                                     | 0.0084** (1.9053)                                   | \( \text{BV} \)                                     | 0.0054 (1.3892)                                     |
| \( \text{Sign Arr} \)                               | -0.0145 (-0.6161)                                   |                                                     |                                                     |                                                     |                                                     |
| \( \text{Sign Std} \)                               | 0.0023 (0.6011)                                     |                                                     |                                                     |                                                     |                                                     |
| \( \text{BV} \)                                     | 0.0046 (1.1958)                                     |                                                     |                                                     |                                                     |                                                     |
| Adj. \( R^2 \)                                     | 0.0284                                              | Adj. \( R^2 \)                                     | 0.0339                                              | Adj. \( R^2 \)                                     | 0.0613                                              |

This table reports the results of cross-sectional regressions of monthly portfolio returns in each month. T-statistics are shown in parentheses. ***, **, * represent the 1%, 5%, and 10% significance levels, respectively.

https://doi.org/10.1371/journal.pone.0181990.t005
interpreted by the highly positive correlation between the idiosyncratic volatility risk and the jump risk, and the idiosyncratic volatility has a negative risk price. In Eq (18), the coefficient for \( \text{Sign\_Size} \) has a statistically significant mean of 0.0330 (\( t = -2.0146 \)). The results strongly support that sign jump components are priced and indicate that the cross-sectional returns are monotonically increasing with sign jump size.

The above results shed light on the complex way in which jump risks perform; other possible situations should be considered carefully, which this paper has tried to discuss. Considering all these 24 types of cases together, the linear form of the sign realized jump intensity, the negative realized jump arrival rate and the realized jump standard deviation are able to explain some part of the observable cross-sectional returns.

**Conclusion**

For risk-averse investors, jump risk, which leads to return distribution with a fat tail, can greatly affect investors’ perception about how risky a company is or how it captures the risk information. This paper applies the realized jumps to be a measure of tail risk, where daily realized jump volatility is estimated from 5-minute high-frequency trading data of Chinese A-share stocks and the stocks are divided into 25 portfolios according to their size and book-to-market ratio. Then, a non-parametric method is applied to simply and accurately calculate the positive and negative jumps and their distribution variables based on the theory of the asymmetric characteristics, including jump intensity, jump mean, jump standard deviation and jump arrival rate.

Then, we argue that the financial constraints make investors fail to effectively diversify non-systematic risk away; hence, the tail risk should be priced, and jumps should be a good supplemental factor to explain the excess portfolio returns.

The Fama-MacBeth two-step regression test shows that both the continuous realized volatility component and the jump component have risk premia, and the jump tail risk, including the sign, negative and positive realized jump components can explain the portfolios’ excess returns as well; this sheds new light on stock return dynamics and helps us better understand and grasp the financial market rules, and such research is rare. Combining the concept of jumps with investor’s economic intuitions, such as fear of loss, hedging of the expected losses or higher risk compensation for expected losses, and the frightening expectation of catastrophic risks in the uncertain macroeconomic environment, it is beneficial for investors’ hedging risk control strategy to consider jumps. The results of this paper also tell us that we need to pay more attention to the role of jumps, particularly large negative jumps.

**Supporting information**

S1 Text. Descriptive statistics of the jump components for the other 20 portfolios. \( ***, **, * \) represent the 1%, 5%, and 10% significance levels of the Jarque-Bera Test, respectively. (DOCX)

S1 Table. Descriptive statistics of the sign jump intensity (sign_size) for Fama-French portfolios. \( ***, **, * \) represent the 1%, 5%, and 10% significance levels of the Jarque-Bera Test, respectively. (DOCX)

**Acknowledgments**

We thank the constructive comments and suggestions provided by the referees, which are all valuable and helpful for improving our paper.
Author Contributions
Data curation: YC XL.
Formal analysis: YC XL.
Investigation: YC XL SG.
Methodology: YC XL SG.
Resources: YC XL SG.
Writing – original draft: YC XL.

References
1. Adrian T, Rosenberg J. Stock returns and volatility: pricing the short-run and long-run components of market risk. The Journal of Finance. 2008; 63(6): 2997–3030. https://doi.org/10.1111/j.1540-6261.2008.01419.x
2. Chang BY, Christoffersen P, Jacobs K. Market skewness risk and the cross section of stock returns. Journal of Financial Economics. 2013; 107(1):46–68. doi.org/10.1016/j.jfineco.2012.07.002.
3. Chan WH, Feng L. Time-varying jump risk premia in stock index futures returns. Journal of Futures Markets. 2012; 32(7): 639–659. https://doi.org/10.1002/fut.20540
4. Bansal R., Shaliastovich I. Learning and asset-price jumps. Review of Financial Studies. 2011; 24(8): 2738–2780. https://doi.org/10.1093/rfs/hhq023
5. Drechsler I, Yaron A. What’s vol got to do with it. Review of Financial Studies. 2011; 24(1):1–45. https://doi.org/10.1093/rfs/hhq085
6. Kelly B, Jiang H. Tail risk and asset prices. Review of Financial Studies. 2014; 27(10): 2841–2871. https://doi.org/10.1093/rfs/hhu039
7. Bansal R, Yaron A. Risks for the long run: a potential resolution of asset pricing puzzles. The Journal of Finance. 2004; 59(4):1481–1509. https://doi.org/10.3386/w8059
8. Barro RJ. Rare disasters and asset markets in the twentieth century. The Quarterly Journal of Economics. 2006; 121(3):823–866. https://doi.org/10.1162/qjec.121.3.823
9. Bollerslev T, Todorov V, Li SZ. Jump tails, extreme dependencies, and the distribution of stock returns. Journal of Econometrics. 2013; 172(2): 307–324. https://doi.org/10.1016/j.jeconom.2012.08.014
10. Bollerslev T, Law TH, Tauchen G. Risk, jumps, and diversification. Journal of Econometrics. 2008; 144 (1): 234–256. http://dx.doi.org/10.1016/j.jeconom.2008.01.006.
11. Todorov V, Bollerslev T. Jumps and betas: a new framework for disentangling and estimating systematic risks. Journal of Econometrics. 2010; 157(2): 220–235. doi.org/10.1016/j.jeconom.2009.11.010.
12. Tauchen G, Zhou H. Realized jumps on financial markets and predicting credit spreads. Journal of Econometrics. 2011; 160(1): 102–118. doi.org/10.1016/j.jeconom.2010.03.023.
13. Zhang BY, Zhou H, Zhu H. Explaining credit default swap spreads with the equity volatility and jump risks of individual firms. Review of Financial Studies. 2009; 22(12): 5099–5131. http://www.jstor.org/stable/40468339.
14. Zhou H, Zhu J Q. Jump risk and cross section of stock returns: evidence from China’s stock market. Journal of Economics and Finance. 2011; 35(3): 309–331. https://doi.org/10.1007/s12197-009-9097-z
15. Chen G, Liu X, Hsieh P, Zhao X. Realized jump risk and equity return in China. Discrete Dynamics in Nature and Society. 2014 June; https://doi.org/10.1155/2014/721635
16. Rabemananjara R, Zakoian JM. Threshold arch models and asymmetries in volatility. Journal of Applied Econometrics. 1993; 8(1): 31–49. https://doi.org/10.1012/jae.3950080104
17. Guo H, Wang K, Zhou H. Good jumps, bad jumps, and conditional equity premium. Bad Jumps, and Conditional Equity Premium. 2014 Oct 28. http://dx.doi.org/10.2139/ssrn.2516074.
18. Barndorff-Nielsen OE, Shephard N. Estimating quadratic variation using realized variance. Journal of Applied Econometrics. 2012; 17(5): 457–477. http://www.jstor.org/stable/4129267.
19. Andersen TG, Bollerslev T. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. International Economic Review. 1998; 39(4): 885–905.
20. Barndorff-Nielsen OE, Shephard N. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics. 2006; 4(1):1–30. https://doi.org/10.1093/jjfinec/nbi022
21. Barndorff-Nielsen OE, Shephard N. Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics. 2004; 2(1): 1–37. https://doi.org/10.1093/jjfinec/nbh001
22. Barndorff-Nielsen OE, Shephard N. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics. 2006; 4(1): 1–30. https://doi.org/10.1093/jjfinec/nbi022
23. Huang X, Tauchen G. The relative contribution of jumps to total price variance. Journal of Financial Econometrics. 2005; 3(4): 456–499. https://doi.org/10.1093/jjfinec/nbi025
24. Andersen TG, Bollerslev T, Diebold FX. Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility. Review of Economics and Statistics. 2007; 89(4): 701–720. https://doi.org/10.1162/rest.89.4.701
25. Barndorff-Nielsen OE. Shephard N. Measuring downside risk-realized semivariance. In: Tim Bollerslev JRRA, editor. Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle, Eds. New York: Oxford University Press. 2010. pp. 117–136.
26. Patton AJ, Sheppard K. Good volatility, bad volatility: Signed jumps and the persistence of volatility. Review of Economics & Statistics. 2015; 97(3): 683–697. https://doi.org/10.1162/rest_a_00503
27. Bali TG, Peng L. Is there a risk–return trade-off? Evidence from high-frequency data. Journal of Applied Econometrics. 2006; 21(8):1169–1198. https://doi.org/10.1002/jae.911
28. Fama EF, French KR. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics. 1993; 33(1): 3–56. https://doi.org/10.1016/0304-405X(93)90023-5