Nonlinear Dynamic Characteristic Analysis of Conditional Jump Duration

Jiayi Xu ^{a,*}, Lianqian Yin ^b

International Business School of Jinan University, Zhuhai, Guangdong, China

^{a,*} 1240413833@qq.com, ^b lianqian@jnu.edu.cn

Abstract. Based on the five-minute tick market data of the Shanghai Composite Index from 2015 to 2019, the time duration of the jumps is realized by extracting the dynamic behavior model of the asset price to achieve the jumps, the dynamic process of the jump changes is described, and the broad ACD (1,1) model is established for the time duration sequence, which achieves a good fit for the dynamic behavior of the jump time duration. The results show that there is also aggregation effect in the time duration of the jumps, which shows that the intermittent phenomenon of the jump in asset prices opens up new ideas for further studying the characteristics and movement laws of asset price jump behavior from the perspective of jumping intensity.

Keywords: Realized Jump; Time Duration; Autoregressive Conditional Duration Model (ACD).

1. Introduction

In order to obtain more information included in market data, with the rapid development of computer technology and the reduction of the cost of data sample collection, recording, storage and analysis, more and more scholars begin to study the microstructure of financial markets more deeply based on high frequency data.

At present, the research based on high frequency data to measure the level of market volatility has achieved some fruitful results. Merton(1980) points out that if the frequency of financial asset prices observed in a certain period of time tends to infinity, the square of the rate of return in each observation interval will converge to the secondary variation of the observation interval according to the probability. Based on this idea, Andersen, Bollerslev and Huang (2007) further develop the theory of realizing volatility RV and double power variance BV, and estimate the jump of financial asset price by making difference between RV and BV, that is, discontinuous jump partial variation JV.

On the other hand, The definition of high frequency data is not limited to market information with higher sampling frequency, data of real-time sampling by transaction record which compared with traditional econometrics, can reflect the nature of financial market transactions at unequal intervals. The interval is called duration, further by drawing on concepts similar to ARCH models in volatility research methods, Engle and Russell(1998) proposed autoregressive conditional duration (ACD) model to describe the evolution of stock time duration in high frequency trading. With the elimination of the intraday effect, good fitting of the ACD model shows that there is clustering in the duration. Zhang, Rusell and Tsay(2001) extend the ACD model to account for nonlinearities and structural mutations in the data, and propose a ACD model with a generalized gamma distribution, that is, the GACD model.

Based on the above ideas, in order to explore whether there is clustering phenomenon in the duration of realized jump, this paper first extracts the time duration from the realized jump and describes the process of the dynamic change of realized jump. Then the ACD model with generalized random distribution is introduced to model and analyze the dynamic behavior characteristics of jump duration. It is worth noting that this kind of estimation based on high frequency data is often disturbed by market noise in empirical research. Dan-Ma, Yuping-Yin (2012) discussed the effects of these noises and proposed to pre-filter the data to minimize noise interference. However, the jump information would also be filtered away as noise at the same time. Therefore, this paper introduces the sparse sampling method proposed by Bandi and Russel(2006) to reduce the noise of the jump sequence. This method observes the asset price income at equal intervals. Although some extremely

high frequency information is lost, the original structure of the data is retained, letting it more suitable for jump estimation.

2. Theoretical Analysis

2.1 Jump Variation, Jump Text, and Realized Jumps

The average time of continuous trading on the t day is divided into m intervals, and the length of each interval is expressed as Δ . Then the last price in the i^{th} interval of the t^{th} trading day can be defined as $P_{t,i}$, the highest price as $H_{t,i}$, the lowest price as $L_{t,i}$, $i=1,2, \dots, m$. Therefore, the closing price on the t^{th} day is $P_{t,m}$, the i^{th} return on the t^{th} day is: $r_{t,i} = \ln P_{t,i} - \ln P_{t,i-1}$, $i \geq 2$. we can further define the realization of volatility RV as the sum of square of the intra-day rate of return on financial assets. The definition is as follows:

$$RV_t = \sum_{i=1}^m r_{t,i}^2 \quad (1)$$

The double power variation BV as the product of the absolute value of the two adjacent returns before and after the financial asset day. The definition is as follows:

$$BV_t = \frac{\pi}{2} \frac{m}{m-1} \sum_{i=2}^m |r_{t,i}| |r_{t,i-1}| \quad (2)$$

Refer to Tauchen and Zhou (2011) [7], we test whether significant jump behavior occurs within the t^{th} day. The test statistics of significant jump behavior is as follows:

$$ZJ_t = \frac{(RV_t - BV_t) / RV_t}{\sqrt{[(\frac{\pi}{2})^2 + \pi - 5] \frac{1}{m} \max(1, \frac{TQ_t}{BV_t^2})}} \xrightarrow{\Delta \rightarrow 0, d} N(0,1) \quad (3)$$

Where the TQ is the third power variation on the t^{th} day, and the expression is:

$$TQ_t = \frac{1}{m\mu_{4/3}} \sum_{i=3}^m |r_{t,i-1}r_{t,i-2}|^{\frac{4}{3}} \quad (4)$$

$\mu_{4/3}$ is the adjustment factor, whose value is $2^{2/3} \frac{\Gamma(7/6)}{\Gamma(1/2)}$, $\Gamma(\cdot)$ represents the gamma function. And then we get the definition of the differential JV as follows:

$$JV_t = I_{\{ZJ_t > \Phi_{1-\alpha}\}} (RV_t - BV_t) \quad (5)$$

Where $I_{\{\cdot\}}$ refer to indicator function. If $ZJ_t > \Phi_{1-\alpha}$, then $I_{\{\cdot\}} = 1$, else $I_{\{\cdot\}} = 0$. Notice that $\Phi_{1-\alpha}$ stand for the $1-\alpha$ quantile of the standard normal distribution, α is the confidence level chosen by the estimation JV.

Ultimately, in order to obtain the value of the jump amplitude of the day, and consider that the significant difference between the realized volatility and the double power variation on day t may be caused by multiple jumps in the day, we assume that only one jump occurs in day t , and the amplitude of this jump is the cumulative sum of all jumps within the day. The definition of this jump on the day is as follows:

$$RJ_t = \text{sign}(\{r_{t,k} : |r_{t,k}| = \max_{i \in \{1,2,\dots,m\}} |r_{t,k}|\}) \sqrt{JV_t} \quad (6)$$

Where $\text{sign}(\{x_k : |x_k| = \max |x_i|\})$ represents the sign of x_k , x_k takes the maximum absolute among $i \in \{1,2,\dots,m\}$.

2.2 Autoregressive Conditional Duration Model

The jump duration model mainly considers the time interval between two adjacent jumps, noted as Δt_i . Because the sequence of jump duration does not present intra-days effect characteristic like the transaction duration sequence of financial assets does, there is no need to make the adjustment of eliminating intra-days effect for the sequence of jump duration.

Referring to the self-return conditional duration (ACD) model proposed by Engle and Russell(1998)[3], the idea of GARCH model is used to study the dynamic structure of time duration. Firstly, if setting $x_i = \Delta t_i$, we can define the conditional expectation of the duration between the $i-1$ jump and the i jump as $\varphi_i = E(x_i | F_{i-1})$. Where F_{i-1} is the information set which we can get from jump $i-1$. Thus, the ACD (r, s) model is defined as follows:

$$x_i = \varphi_i \varepsilon_i \quad (7)$$

where $\{\varepsilon_i\}$ is a sequence of non-negative random variables which distributes independently and uniformly, and satisfied the condition: $E(\varepsilon_i) = 1$. Further in the GACD (r, s) model extended by Zhang et al., ε_i obeys a standardized generalized gamma distribution and φ_i is assumed as the following forms:

$$\varphi_i = \omega + \sum_{j=1}^r \gamma_j x_{i-j} + \sum_{j=1}^s \omega_j \varphi_{i-j} \quad (8)$$

Like the GARCH model, the process $\eta_i = x_i - \varphi_i$ is a martingale difference sequence, and the ACD (r, s) model can be written as:

$$x_i = \omega + \sum_{j=1}^{\max(r,s)} (\gamma_j + \omega_j) x_{i-j} - \sum_{j=1}^s \omega_j \eta_{i-j} + \eta_i \quad (9)$$

The expression above can also be regarded as the form of ARMA processes without gaussian renewal, where for $j > r$, $\gamma_j = 0$; and for $j > s$, $\omega_j = 0$. Meanwhile, according to the expectation of both sides of the above expression, the basic conditions of weak stability of the ACD (r, s) model can be obtained:

$$E(x_i) = \frac{\omega}{1 - \sum_{j=1}^{\max(r,s)} (\gamma_j + \omega_j)} \quad (10)$$

Because the expectation of time duration is non-negative, the condition which requires the parameter of the model should be non-negative and bounded can be expressed as:

$$\omega > 0, 1 - \sum_j (\gamma_j + \omega_j) > 0 \quad (11)$$

For the ACD (r, s) model, suppose $i_0 = \max(r, s)$, $x_i = (x_1, \dots, x_i)'$, we can then defined the likelihood function of the duration as:

$$f(x_T | \theta) = \left[\prod_{i=i_0+1}^T f(x_i | F_{i-1}, \theta) \right] \times f(x_{i_0} | \theta) \quad (12)$$

Where θ represents the parameter vector of the model and T stands for the sample size. Since for the generalized ACD model, the impact of the likelihood function will decrease with the increase of the sample size T , which makes the parameter estimation of the ACD model very complicated by using the likelihood function. Therefore, the conditional likelihood method can be used to estimate the parameters of the model by edge probability density function $f(x_{i_0} | \theta)$.

3. Data Specification

Refer to the sparse sampling method proposed by Bandi and Russell(2006)[6]. Using only a single sub-sample of the population with a total number of N that can be collected on day t at a frequency k , there are estimations with a number of N/k :

$$RV_{t,k}^{n^{(sparse)}} - BV_{t,k}^{n^{(sparse)}} = \sum_{i=1}^{n^{(sparse)}=[n/k]} r_{t,i}^2 - \sum_{i=1}^{n^{(sparse)}=[n/k]} |r_{t,i+1}||r_{t,i}| \xrightarrow{\Delta \rightarrow 0, p} JV_{t,k}^{n^{(sparse)}} \quad (13)$$

Based on the above ideas, this paper selects the market data of all trading days of the Shanghai Composite Index from January 5, 2015 to December 31, 2019 as sample data, and takes 5 min as a frequency interval to calculate the interval yield. All data can be obtained by Tushare API and the noise problem of high frequency data modeling is effectively solved.

Besides, when testing the significance of daily jump behavior, this paper sets the confidence level of 1%, that is, taking $\alpha=1\%$ to judge whether there is jump behavior within the day. The process of obtaining jump duration and using the generalized ACD modeling for analysis of jump duration is realized by python and Eviews software.

4. Empirical Results and Analysis

4.1 Variance Test

Considering the five-minute tick market data for the Shanghai 50 Index from January 5, 2015 to December 31, 2019. The jump amplitude is recorded by the z axis, the y axis records the time duration of the jump, and the x is the time axis. The dynamic process of realizing the jump changing during the sample period is showed by the three-dimensional scatter plot. The distribution of 3D scatter plot in the y axis section shows the calculated daily jump with 1218 observations at the confidence level of 1%, and the distribution of the z axis section shows the time duration sequence extracted from the realized jump sequence with 262 observations.

Fig 1. Realized Jump

Fig 2. The duration of the jump

It can be noted that both the jump and the jump duration present obvious "group fluctuation" phenomenon. For the jump time duration series, the autocorrelation and partial autocorrelation coefficients are calculated. Almost all the values jump up and down within the confidence interval, which indicates that the time duration series has weak autocorrelation indicating there is no need to introduce autocorrelation in conditional expectation model to satisfy the mean equation of ACD model. Further, the ARCH test of the time duration sequence shows that there is conditional heteroscedasticity at the confidence level of 5%(see Table 1).

4.2 Modeling Analysis of Jump Duration based on ACD (1,1)

Because of the sequence correlation in the jump duration, this paper fits the data into a generalized ACD (1,1) model, and the fitting model is as follows:

$$x_i = \varphi_i \varepsilon_i, \quad \varphi_i = 0.061 + 0.139x_{i-1} + 0.829\varphi_{i-1} \quad (14)$$

Where $\{\varepsilon_i\}$ is a random sequence which presents an independent identity distribution. The generalized random distribution with a model fitting parameter $\hat{\alpha} = 4.049$ has a standard error at 0.584. The distribution and modeling results of standardized residuals are shown in Table 2. The results also show that the standard deviation between the first order lag term and the Garch parameter in the fitting equation (14) is estimated to be 0.063 and 0.064, and the p values of the two estimates are less than 0.05, indicating that the estimate is significant at the level of 5%.

Table 1. Residual Distribution and ACD (1,1) Modeling Results

Statistics	Sample size	Mean	Median	Std.Error
Jump duration	262	0.974	0.836	0.398
	Skewness	Kurtosis	Maximum	Minimum
	1.288	4.876	2.717	0.357
Variables	Estimated coefficient	Std. error	z statistics	P value
CONST	0.061	0.061	0.997	0.319
RESID(-1)^2	0.139	0.063	2.197	0.028
GARCH(-1)	0.829	0.064	13.000	0.000
GED PARAMETER	4.049	0.584	6.931	0.000
GARCH TEST	F-test	probability	chi-square test	probability
	5.024	0.026	4.966	0.026

Finally, we analyze the standadized residuals of the established ACD (1,1) model by using Q-test with the statistics of $Q(10) = 9.8453$, $Q(20) = 13.133$. Obviously, there is no significant sequence correlation in the standardized renewal process $\{\varepsilon_i\}$. The autocorrelation and partial autocorrelation coefficients are shown in Fig 3 and 4, and the fitted model is suitable.

Fig 3. Standardized renewal ACF

Fig 4. Standardized renewal PACF

Furthermore, it is worth noting that the expectation of jump time duration is $0.061 / (1 - 0.829 - 0.139) \approx 1.91$ days, the sample mean value is 1.86 days close to the jump duration, and the estimation

coefficient $\gamma_1 + \omega_1 \approx 0.968$ implies some continuity of the jump adjustment duration. Similarly, we use axis y to record the predicted jump time duration in the sample, the z axis records the amplitude of the jump at the end point of the time duration, and axis x is used as the end point jump times. The 3D scatter plot (Fig 5) and the time duration scatter distribution (Fig 6) of the jump-time duration fitting of the ACD model are obtained as follows:

Fig 5. Duration of realized jumps

Fig 6. Predicted duration

4.3 Robustness Test

Robustness test can be achieved by three methods. The first one is replacing dependent variables, main independent variables and relaxing variable conditions, the second is adding tool variables such as missing variables, various kinds of virtual variables, etc, and the third is updating data sources and changing sample period, sample size, or classification methods. To verify the robustness of the results of this study, this paper adopts the method of updating data sources, that is establishing generalized ACD models for CSI 300 index, B stock index and fund index respectively. The results show that the main conclusions of this study are robust. And the test results are shown in the table below.

Table 2. Robustness Test

Sample (Sample size)	Variables	Estimated coefficient	Std. error	z statistics	P value
CSI 300 index (541)	CONST	2.459	0.261	9.426	0.000
	RESID(-1)^2	-0.035	0.020	-1.749	0.080
	GARCH(-1)	-0.843	0.183	-4.598	0.000
	GED PARAMETER	5.340	0.554	9.638	0.000
B stock index (597)	CONST	2.786	0.170	16.420	0.000
	RESID(-1)^2	0.022	0.014	1.546	0.122
	GARCH(-1)	-0.957	0.061	-15.694	0.000
	GED PARAMETER	4.880	0.540	9.036	0.000
fund index (520)	CONST	0.474	0.268	1.767	0.077
	RESID(-1)^2	-0.059	0.026	-2.239	0.025
	GARCH(-1)	0.687	0.205	3.349	0.001
	GED PARAMETER	6.503	0.976	6.664	0.000

5. Conclusion

Based on the theoretical framework of the jump and time duration model, a generalized ACD model is established for the time duration of the jump in this paper. The results of the model estimation are well fitted to the dynamic behavior of the jump time duration. And it shows that the time duration of jump also has aggregation effect. The longer time duration reflects a period without new news, which indicates a steady trend of asset price. On the contrary, the shorter time duration indicates the sharp fluctuation of asset price.

By analyzing and predicting the dynamic behavior of jump time duration using ACD model, we can then open a way for the study of jump theory not only in jump amplitude, but also in the

perspective of jump intensity. This method not only reflects the fluctuation characteristics of high frequency market data through daytime jump, but also make the use of the existing model which can only be applied when analyzing low frequency data before with good fitting effect. More importantly, it is of great theoretical and practical significance to further study the characteristics and influence of asset price jump behavior, grasp the law of asset price movement, and promote the stability of financial market.

Acknowledgments

The work was supported by the National Natural Science Foundation of China Youth Project: arbitrary fuzzy complex data reconstruction and jump detection method research under Grant Nos.6507010083 and the teaching reform research project (international curriculum project) of Jinan University and other projects.

References

- [1] Merton RC. On estimating the expected return on the market: An exploratory investigation[J]. *Journal of Financial Economics*, 1980, 8(4): 323–361.
- [2] Andersen TG, Bollerslev T, Huang X. A Reduced Form Framework for Modeling Volatility of Speculative Prices Based on Realized Variation Measures[EB/OL]. [2007]. Working paper series in Department of Economics of University of Oklahoma.
- [3] Engle, R.F. and Russell, J.R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* 66: 1127-1162.
- [4] Zhang, M.Y., Russell, J.R., and Tsay, R.S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics* 104: 179-207.
- [5] Ma Dan, and Yin Youping. Analysis of noise, Jump and High Frequency Price Fluctuation--Realization of Fluctuation Based on Threshold Preaverage. *Financial research*. 04(2012):124-139.
- [6] Bandi F, Russell J. Separating Market Microstructure Noise from Volatility[J]. *Journal of Financial Economics*, 2006, 79(3).
- [7] Tauchen G, Zhou H. Realized Jumps on Financial Markets and Predicting Credit Spreads[J]. *Journal* 2011 (1).
- [8] Lianqian-Yin, Yutian-Zhou, Huamei-Wu. Research on China's Asset Price Jump Behavior--Analysis Based on High Frequency Dataset [J]; *Forum on Statistics and Information*, 2015,30(5): 57-62.
- [9] Kehong-Geng, Shiyang-Zhang. A ACD-GARCH-V Model of UHF Time Series in Financial Markets[J]. *Statistics and decision-making*, 2007, 232(4): 81-83.
- [10] Lina-Dai, Jing-Rui. A Nonparametric ACD Model and Its Empirical Study in China Stock Market[J]. *Forum on Statistics and Information*, 2006(3): 96-99.