

## **Superstitious Numbers and Price Clustering : A Study Based on CSI 300 Stock Index Futures**

**Liu Fengyuan**

Professor, Department of Law and Finance  
East China University of Political Science and Law, Shanghai  
China.

**Liu Changqing**

Deputy Professor, Department of Education Economy and Management  
East China University of Political Science and Law, Shanghai  
China.

### **Abstract**

*This paper studies the price clustering phenomenon of CSI 300 stock index futures. The empirical results indicate that it is notable the ending number of the transaction price clusters at '0' and avoids '4', which is in line with the attractiveness hypothesis and the view that Chinese have a certain degree of superstitious number. Further regression analysis shows that the greater the price fluctuation, the more obvious the clustering is; the higher the trading volume, the less obvious the clustering is. However, the opening prices, closing prices and open interest have no significant impact on the price clustering.*

**Keywords:** price clustering, financial futures, CSI 300 stock index futures

### **Can Superstitious Numbers Be Believed In The Stock Market?**

#### **1. Introduction**

Traditional financial theory holds that in a market where there is no friction or bias, the transaction price should be evenly distributed among all possible prices (Niederhoffer, 1965), that is, the ending numbers of the price should be evenly distributed. While in the practice of financial market, the transaction price of financial assets tends to end in certain numbers, in what is known as price clustering. Osborne (1962) is the first scholar who discovered the phenomenon of price clustering in the stock market. Many scholars have studied the phenomenon of price clustering, and it has become a well-known phenomenon in the stock and foreign exchange markets that the ending numbers of transaction price clusters at 0 or 5 (such as Sopranzetti and Datar, 2002; Harris, 1991, etc.). Similar results have been obtained from studies on futures market. For example, the research on stock index futures and options market conducted by AP Gwilym and Alibo (2003), Schwartz et al. (2004), Chung and Jiang (2006) studied , and the research on Taiwan stock index futures by Kuo, Lin, and Zhao (2015) and Bhattacharya, Kuo, Lin, and Zhao (2017).

Goodhart and Curcio (1991) proposed the attraction hypothesis to explain the price clustering. They believe that the reason why asset prices tend to be round numbers is that they are more attractive to investors. For example, the attractiveness of the price ending in 0 is greater than that ending in 5, while the attractiveness of price ending in 5 is greater than that ending in other numbers. Harris (1991) proposed the negotiation hypothesis to explain the phenomenon of price clustering. In his opinion, a smaller fluctuation unit would increase the bargaining cost of the trader. In order to reduce the bargaining cost, the trader chooses to trade only at the round price, so that the order could be closed as soon as possible, thus forming the price clustering. Christie et al. (1994) put forward the collusion hypothesis, which holds that in order to make more profits, NASDAQ market makers collude with each other to maintain a high bid-ask spread, resulting in price clustering.

In Chinese culture, '4' is considered as an unlucky number, while '8' is considered as a lucky number, which has been studied by some scholars. For example, brown, Chua, and Mitchell (2002) studied the daily closing endings of six stock markets in the Asia-Pacific region, and the prices ending in 0 or 5 showed obvious clustering phenomenon; meanwhile, the probability of prices ending in '8' in Hong Kong market is significantly higher than that ending in '4'. Ke et al. (2017) studied the closing prices of Taiwan stock exchange from 2001 to 2013, and found that the frequency of prices ending in '8' was significantly higher than prices ending in '2', while the frequency of prices ending in '6' was significantly higher than prices ending in '4'. Researchers believe that bias or superstitious numbers in Chinese culture is the main reason why traders prefer '8' and avoid '4'.

For the phenomenon of price clustering in the domestic financial market, the past literature emphasizes on the securities market. This paper will study that domestic investors may avoid or prefer certain numbers due to their concerns about transaction costs or cultural reasons when they conduct financial futures trading, which result in the different cluster states of transaction prices, and will analyze the factors that affect price clustering through the model. At the same time, the price clustering phenomenon may be different because of transaction costs, prices and other factors in a bear market or bull market. In order to compare the price clustering phenomenon in different market environments of bull and bear markets, this paper divides the research period into bull market period and bear market period.

## 2. Data Sources and Clustering Phenomenon

This paper takes the financial futures with the largest trading volume in China - CSI 300 index futures as the research object. For comparative study, the samples are taken respectively during the bull market period from 20170101 to 20171231 and the bear market period from 20180101 to 20181231. In the test of the price clustering phenomenon, the transaction price per minute per day is adopted. There are 487 trading days and 117499 transaction prices in 2 sample intervals. In the study of factors affecting the degree of price cluster, daily opening price, closing price, trading volume, open interest and other daily data are used. All the data comes from The Wind Financial Terminal.

Since the minimum change price of CSI 300 stock index futures is 0.2, so prices can only end in 0, 2, 4, 6, and 8, while prices will not end in 1, 3, 5, 5, 7 and 9. We examine whether the occurrence frequency of the endings of transaction price follows the uniform distribution. If it does not follow the uniform distribution, it means that the probability of occurrence of some Numbers is significantly higher, that is, there is a “price clustering” phenomenon of preference for numbers.

Data in Table 1 show that the hypothesis that the endings of transaction price follow the average distribution is rejected at the significance level of 1%. Also, the proportion of prices ending in 0 is the highest in both bull markets in 2017 and bear markets in 2018, which is consistent with the attractiveness hypothesis put forward by scholars such as Goodhart and Curcio (1991).

Table 1. Intra-day price clustering phenomenon of CSI 300 index futures

Endings of Price	Bull market in 2017		Bear market in 2018	
	Frequency	Percentage	Frequency	Percentage
0	15062	25.59084	14731	25.12022
2	11127	18.90514	11152	19.01709
4	9893	16.80854	10487	17.88309
6	11496	19.53209	11107	18.94035
8	11279	19.1634	11165	19.03926
Total	58857	100%	58642	100%
Chi-Square Test	$\chi^2_{4,1\%} = 1281.92$		$\chi^2_{4,1\%} = 988.41$	

At the same time, mantissa “4” appeared in the lowest proportion in both study samples. It is consistent with the Chinese digital culture or superstitious number theory proposed by Brown, Chua, and Mitchell (2002). At the same time, the ending number “4” appeared in the lowest proportion in both study samples. It is consistent with the theory of Chinese number culture or superstitious number put forward by Brown, Chua, and Mitchell (2002) and other scholars. Their researches show that there is a phenomenon of price clustering in the Chinese securities market, and it is obvious that market participants avoid the number ‘4’. In China, it is common to avoid the number “4” as much as possible in daily life. For example, some high-end real estates use letters to replace the floor numbers consisting of number 4. The News of removing number “4” from the license plate is often seen in the press. The stock exchange, as a self-regulatory organization, also tries to avoid number “4” and use the lucky number “8” instead. The domestic listing code consists of six figures, as of November 25, 2019, there were a total of 3733 listed companies, of which only 204 companies were listed with the listing code ending in ‘4’, accounting for 5.46% of all listed companies. In the meantime, all the codes of the 56 companies in SCI tech Innovation Board which started trading on July 22, 2019 are prefixed with the lucky number ‘688’. There is neither number “4” in the codes endings nor number “4” in the last three digits of these codes.

In addition, there was no significant difference in the phenomenon of price clustering during the bull market in 2017 and the bear market in 2018, which was believed to be related to the small difference in the additional costs caused by preference for number during the two research periods. The data show that the average closing price of CSI 300 index in 2017 was 3662.20, and 3595.80 in 2018, while the minimum change price of CSI 300 index futures contract is 0.2. Now it is assumed that the latest transaction price of the contract is 3662.20, and a bullish investor wants to close the deal in a hurry, he can quote 3662.40 or 3662.60, but he does not like the number ‘4’, so he directly quotes 3662.60. At this time, the increased cost is 0.20 points, and the increased cost only accounts for  $100 * (0.2 / 3662.60) = 0.55\%$  of the total contract value. Similarly, in the bear market in 2018, he could directly quote 3596, and the increased cost due to preference for number only accounted for  $100 * (0.2 / 3696.0) = 0.56\%$  of the total contract value. This is consistent with the results of Liu Fengyuan's (2008) study on the price clustering of the stock market. He believes that when quoting, traders are influenced by both investment rationality and superstitious number. When the cost of preference for number is low, they tend to quote the price they like; on the contrary, when the cost of preference for number is too high, investors are relatively rational and they will temporarily abandon the preference for superstitious number for saving costs.

### 3. Regression Analysis

After the clustering analysis of the transaction price per minute, we then use the daytime data to study the factors that affect the price clustering. First of all, each variable is defined:

CL (clustering): the degree of price clustering, by the daily trading volume ending in 0 divided by the total volume of the day;

OI (open interest): open interest of the day;

OP (open price): daily opening price;

CP (close price): daily closing price;

VOL (volatility): price volatility;

VOLUME: daily trading volume

The skewness of the data in Table 2 shows that all daytime data variables are skewed to the right. At the same time, the Jarque-Bera value indicates that all variables reject the assumption that they conform to the normal distribution at the significant level of 1%.

Table 2 : Descriptive statistics of daytime data

	CL	OI	OP	CP	VOL	VOLUME
Average	24.78	28114.03	3630.70	3629.09	51.98	18271.91
Maximum	43.24	54220	4412.8	4404.6	259.8	59022
Minimum	14.19	6100	3002.6	2995.8	13	4154
Standard deviation	3.86	8546.26	323.75	324.23	29.07	8739.23
Skewness	0.440699	0.71062	0.240969	0.229975	1.871363	1.618116
Jarque-Bera	32.47	65.71	23.33	22.10	1181.61	380.33
Number of samples	487					

Generally speaking, the time series of stock prices are unstable due to its long-term trend, which may lead to the existence of spurious regression when adopting traditional regression methods. Therefore, unit root tests must be carried out on the original series before conducting further research. In this paper, ADF unit root test is taken to test the above six variables.

Table 3 : Unit root test of daytime raw data

Variables	With intercept term		With intercept term and trend term	
	LAG	ADF	LAG	ADF
CL	3	-20.08934	3	-20.28277
OI	1	-4.661178	1	-5.305528
OP	1	-0.917722	1	-1.242621
CP	1	-0.968405	1	-1.288269
VOL	2	-5.600320	2	-9.459626
VOLUME	5	0.202258	5	-1.545634

It can be seen from Table 3 that the three variables, opening price, closing price and trading volume, cannot reject the null hypothesis that the unit root exists. Considering that the economic meaning of these three variables is difficult to explain after differential treatment, also the main purpose of this paper is to study the impact of price level and transaction volume on price clustering, and the price volatility variable can measure the impact of price change on price clustering. Therefore, these three variables are sorted in ascending order separately, and assign corresponding values 1, 2, 3... 487 respectively, then taking logarithm lg (1), lg (2)... lg (487). With this process, LG (rankOP), LG (rankCP) and LG (rankVOLUME) are used to replace the original opening price (OP), closing price (CP) and trading volume (VOLUME). A unit root test on adjusted variables shows that all variables reject the hypothesis that a unit root exists, as in Table 4 and Table 5.

Table 4 : Unit root test of daytime raw data (first-order difference)

Variables	With intercept term		With intercept term and trend term	
	LAG	ADF	LAG	ADF
ΔCL	3	-14.37087	3	-14.35954
ΔOI	2	-22.14202	2	-22.12766
ΔOP	1	-21.77851	1	-21.90305
ΔCP	1	-23.44577	1	-23.56408
ΔVOL	4	-15.25708	4	-15.24365
ΔVOLUME	4	-15.28291	4	-15.35018

Table 5 : Unit root test of adjusted data

Variables	With intercept term		With intercept term and trend term	
	LAG	ADF	LAG	ADF
CL	3	-20.08934	3	-20.28277
OI	1	-4.661178	1	-5.305528
LG (rankOP)	1	-3.067036	1	-3.204823
LG (rankCP)	1	-3.240814	1	-3.307428
VOL	2	-5.600320	2	-9.459626
LG (rankVOLUME)	3	-4.471703	1	-7.756223

After conducting the unit root test, we learn from the method proposed by Harris (1991) to take the ordinary least square method (OLS) to study the influence of LG (rankOP), LG (rankCP), open interest (OI), price volatility (vol), LG (rankVOLUME) on the price clustering. Because the correlation coefficient of LG (rankop) and LG (rankcp) is up to 0.973, multicollinearity exists between the two variables. Accordingly, one of the two variables is selected for regression analysis:

$$CL_t = C_0 + C_1LG(OI_t) + C_2LG(rankOP_t) + C_4VOL_t + C_5LG(rankVOLUME_t) + \varepsilon_t$$

The results of parameter estimation in Table 6 show that the coefficients of price volatility (VOL) and trading volume (LG (rank volume)) are 0.043415 and -1.06952 respectively, indicating that price volatility and trading volume (LG (rank volume)) have a significant impact on price clustering.

Table 6 : OSL parameter estimation of influencing factors of price clustering

Constant term	Volatility (VOL)	Opening price (LG (rankOP))	Trading volume (LG (rankVOLUME))	Open interest (LG (OI))	R Square	Adjusted R Square	F statistic
13.66538	0.043415	0.678003	-1.06952	2.199583	0.46702	0.45831	94.03570
(2.092479)**	(6.664106)**	(1.513295)**	(-2.02326)**	(1.548274)**			

Note: t statistic is in brackets, \*\* indicates 5% significance level

It shows that the higher the price volatility is, the more obvious the price clustering is, while the higher the trading volume is, the less obvious the price clustering is. In addition, the opening price (LG (rankOP)) and open interest have no significant impact on the price clustering.

The coefficient of price volatility (VOL) is significantly positive, indicating that in the face of high price volatility, investors will simplify the transaction process and offer prices at preferred prices in order to close the deal as soon as possible and reduce the decision-making cost. Therefore, the higher the volatility, the more obvious the price clustering is, which is consistent with the research results of Harris, L. (1991) and Hornick, Cherianand, and Zakay (1994) and other scholars.

The coefficient of trading volume (LG (rankVOLUME)) is significantly negative, indicating that the relevant information of the contract has been known to the market through a large number of transactions traded by participants, so the phenomenon of clustering has diminished, which is consistent with the research results Harris, L. (1991), ap Gwilym et al. (1998), Schwartz, A. L. (2004) and other scholars. These scholars believe that trading volume is the contrary indicator of expected price discrimination.

#### **4. Conclusion**

Existing empirical studies show that the phenomenon of price clustering is widespread in such markets as stock, gold, foreign exchange, interest rates, but most of the research focuses on the stock market. There is little literature on the financial futures, especially on the price clustering in Chinese financial futures market. Thus, this paper makes an empirical study on the price clustering of CSI 300 Index, the financial futures with the largest trading volume in China to date.

The empirical results show that whether it is in a bull market or in a bear market, the CSI 300 futures index shows a phenomenon of price clustering: the quotation obviously tends to end in '0' and avoids ending in "4". The quotation tending to end in '0' is consistent with the attractiveness hypothesis, and avoiding ending in "4" is related to the Chinese superstitious number.

#### **References**

- Adam L. et al. (2004) , Clustering in the futures market: evidence from S&P 500 futures contracts. *Journal of Futures Markets* 24(5):413 - 428 · May 2004
- ap Gwilym, O. and Alibo, E. (2003). Decreased price clustering in FTSE100 futures contracts following the transfer from floor to electronic trading. *The Journal of Futures Markets*, Vol.23, pp.647-659.
- Brown, P., A. Chua and J. Mitchell (2002), "The influence of cultural factors on price clustering: Evidence from Asia-Pacific stock markets," *Pacific-Basin Finance Journal*, 10:3, 307-332. Doi: 10.1016/S0927-538X(02)00049-5
- Bhattacharya, U., W.Y. Kuo, T.C. Lin, and J. Zhao (2017), "Do superstitious traders lose money?" *Management Science*, Vol 64 Issue 8, 2018 3772-3791 . Doi: 10.1287/mnsc.2016.2701
- Chung, C. and Chiang, S. (2004). Price clustering in E-mini and floor-traded index futures. 2004 Financial Management Association/Europe Conference, Switzerland.
- Christie, W. G. and Schultz, P. H. (1994). Why do NASDAQ market makers avoid odd-eighth quotes? *Journal of Finance*, Vol.49, pp.1813-1840.
- Harris, L. (1991). Stock price clustering and discreteness. *Review of Financial Studies*, Vol.4, pp.389-415.
- Ke, W.C., H.L. Chen, H.W. Lin, and Y.C. Liu (2017), "The impact of numerical superstition on the final digit of stock price," *North American Journal of Economics and Finance*, 39,145–157. Doi: 10.1016/j.najef.2016.10.003
- Kuo, W.Y., T.C. Lin, and J. Zhao (2015), "Cognitive limitation and investment performance: evidence from limit order clustering," *Review of Financial Studies*, 28: 3, 839-875. Doi:10.1093/rfs/hhu044
- Niederhoffer, V. (1965). Clustering in Stock Prices, *Operations Research*, Vol.13, pp. 258-265.
- Schwartz, A. L., Van Ness B. F., and Van Ness R.A. (2004). Clustering in the futures market: Evidence from S&P 500 futures Contracts, *The Journal of Futures Markets*, Vol.24, pp.413-428.
- Sopranzetti, B. J. and Datar V. (2002). Price clustering in foreign exchange spot markets. *Journal of Financial Markets*, Vol.5, pp.411-417.
- Osborne, M. F. M. (1962). Periodic structure in the Brownian motion of stock prices. *Operations Research*, Vol.10, pp.345-379.
- Liu Fengyuan. Social culture, number preference and stock quotation: A Study on price clustering in China's stock market *China soft science*, 2008, (6).