To appear in the 4th VSBF

The Gap-filling Phenomena in the Stock Markets

XIAOHANG LIU†§, YAN WANG‡, ZIYAN JIANG§, QINGHUA CHEN*† and HONGGANG LI†

†School of Systems Science, Beijing Normal University,Beijing 100875, China
‡Department of Mathematics, University of California, Los Angeles 90095, USA
§School of Government, Beijing Normal University,Beijing 100875, China

There is a series of stylized facts in the stock markets, such as the fat-tailed distribution, volatility clustering of returns, asymmetric leverage effect, and they even have become the standard criteria for the verification of theoretical models. Beyond those facts, there are some superstitions that have not yet been verified by scholars, but spreading among stock traders. For example, to Chartists, there is an axiom about the stock gap which is "gaps always get filled". Chartists would like to take advantage of the regular pattern about the gap behavior to decide how they trade since they believe that the gaps always exist universally and the gaps should be closed quickly. However, there is no scientific demonstration of the existence of these statements about the gap since there is a lack of sufficient academic literature to discuss this kind of issue. This paper discusses some characteristics of the gap by collecting empirical data and reveals some similarities and differences of gaps between Chinese and American stock markets. The number of gaps in the United States is more than that in China and the time to fill the gap is slightly shorter. These indicate that the price movement in the US stock market is indeed more active than that in the Chinese market. By applying a random exchange process on the original data, there are significant changes among the statistical results, which means the real data series has some inherent structure behind the price variation. Further, this paper counts the no-trend data and its random shuffling series. Some differences suggest that the overall trend hinders the gaps' generation and slows down the gaps' refilling process to a certain extent.

Keywords: Price gap; Gap-filling; Stock market; Random exchange process; Detrending;

JEL Classification: E44; G410

1. Introduction

Statistical analysis on actual data of financial markets has discovered some stylized facts such as the fat-tailed distribution of returns, volatility clustering of returns, and slow decay of autocorrelation in absolute returns etc (Cont 2001, Bouchaud *et al.* 2002). These widespread phenomena could be found in various financial markets and have attracted the attentions of a large number of scholars. They explore various mechanism models to reproduce these statistical characteristics, including nonlinear adaptive systems (Hommes 2002), evolutionary percolation model (Wang *et al.* 2005), dynamic heterogeneous agent models (Hommes 2006, Schmitt and Westerhoff 2017), even Ising model, a kind of pure physical processing (Lima 2017). It is not an overstatement that these stylized factors are recognized as the touchstones for judging a model if it is good enough. The scholars adherence that the more typical facts that can be reproduced, the more successful the model is (Lehoczky and Schervish 2018).

*Corresponding author. Email: qinghuachen@bnu.edu.cn

However, more possible facts in financial markets are being discovered and tested still in processing. For instance, the power law relation between skewness and kurtosis (Cristelli 2014), the universal and stationary price formation mechanism (Sirignano and Cont 2018), and spurious trend switching (Filimonov and Sornette 2012). In contrast, the existence of the stock price gaps, and how the gaps are closed have not grab enough investigation even some scholars have long time noticed that evidences as "the market abhors a vacuum and all gaps will be filled" (Peacock 1997). On contrary, focus on the gap is mainly popular with stock traders in whole world. In the actual scene, Chartists conclude some rules to guide them in their practical activities based on graphic analysis. They believe if a gap is not closed by the next minor reaction, there is a higher chance that it will be shouted by the very near coming trade day. Only when a gap in the stock price be replenished in short time, and then the stock price will continue to move forward (Guo 2004). Is that true? There are no reliable academic answers to this basic question yet. Until now, there are only a few academic literatures on the gap which mainly focus on the similarity and difference of statistics of the trade volume in transactions period such as before, during and after the gap (Dahlquist and Bauer 2012, Caporale and Plastun 2017). The related discussions are still very insufficient.

This paper focuses on the statistics of gaps and gap-filling phenomena. we try to reveal how often gaps occur and how quick the gap can be filled from empirical data through comparing the reality in empirical statistics and results from a random exchange process. The sections of the article will be arranged as follows. In section 2, a brief introduction to the gap phenomenon and our data sources will be present. In section 3 and section 4, a stochastic exchange process will be presented, and the results will be compared with the empirical ones. In section 5, the statistical results of no-trend data will be shown that the influence of overall trend on the gap-filling phenomenon is enclosed. In last section, we conclude and discuss.

2. Background and Data Source

2.1. *Gap and Gap-filling Phenomena*

The candlestick-graph is a style of financial chart used to describe price movements. Every "candlestick" typically shows one day and "candlesticks" express the whole situation in a continual trading session (Morris 2006). As shown as the legend located at left upper part in figure 1, each candlestick is like a combination of line-chart and a bar-chart that represents all four important pieces of information for that day: the open, the close, the high and the low which means opening price, closing price, maximum price and minimum price in one day respectively.

Candlestick graph is one of the most convenient tools to find the trend and the gap of the stock. The figure 1 exhibits the prices of a certain trading sessions for Shahe Industrial Co., Ltd. (stock code 000014) that list in Shenzhen stock exchange market in China. In this period, from 25th April to 15th June, 2018. On the price chart, it is obvious to find 4 vacancies or spaces appears between the bars that indicating the gaps or windows on 4th May, 7th May, 14th May and 29th May respectively. Precisely, gaps happen when the low price is higher than the next day's high price, or high price is lower than the next day's low price that no shares were traded within a particular price range.

Price gaps can be divided into two categories that up gaps and the down ones. To form an up gap, the lowest price after the market closed must higher than the highest price of the previous day, and a down gap is formed oppositely–the high price after the market closed must be lower than the lowest price of the previous day. There are 2 up gaps and 2 down gaps in figure 1. Up gaps occurred on 4th May and 7th May and down gaps would be found on 14th May and 29th May.

A gap is filled or closed means that the price movement usually retraces at space at a later time (few days to a few weeks, even longer). For an up gap, the lowest price on the day of filling the gap must lower or equal to the lower boundary of the gap (high price on the previous transaction

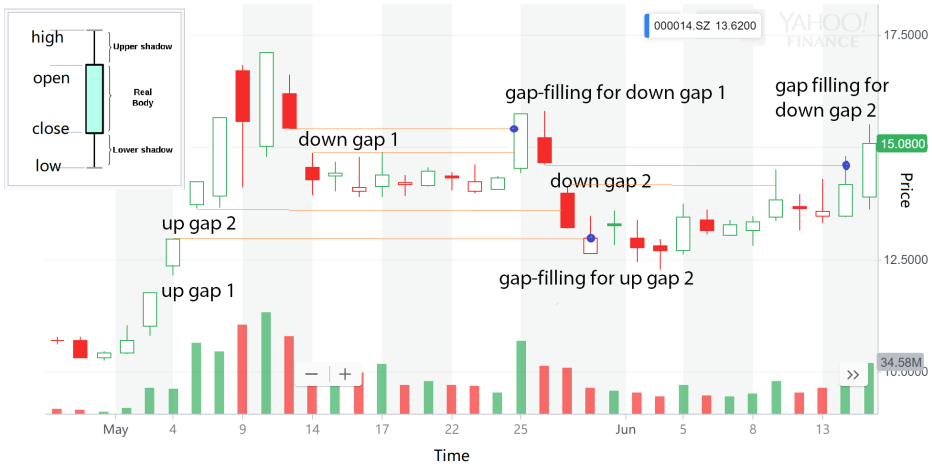

Figure 1. Phenomena of gap and gap-filling. This plot is got from https://sg.finance.yahoo.com. This candlestick plot is for Shahe Industrial Co., Ltd. (000014.SZ) in the Shenzhen stock exchange (SZSE).

date). For a down gap, The highest price on the day of filling the gap must higher or equal to the upper boundary of the gap (low price on the previous transaction date). In that case, the "up gap 2" was filled in 30th May; the "down gap 1" and "down gap 2" was filled in 25th May and 14th June respectively. Until the end of the period, the "up gap 1" has not closed yet.

Gaps are meaningful that can provide clues about the price movement. The breakup from price continuity implies that something important has happened to the fundamentals or the psychology of the crowd that has triggered for one stock even the whole market. By the inception of technical analysis, these holes have always been in the limelight of the Chartist.

2.2. *Data Source and Clearing*

We collected all the available stock data from Shanghai stock exchange (SHSE), the main board of Shenzhen stock exchange (SZSE-1), the second board of Shenzhen stock exchange (SZSE-2), New York stock exchange (NYSE) and NASDAQ stock exchange (NASDAQ) respectively. The data of China Shanghai exchange and Shenzhen exchange are from tushare, which is a free, open source Python financial data interface package. It can collect and store stocks and other financial data. (see more information at `http://tushare.org/`). The transaction data in the United States come from the Yahoo Finance website which provides historical reference data and foreign exchanges (see more details on `https://finance.yahoo.com`). The data that we used in our research are all daily price data, including the open, close, high and low prices. We use backward adjusted prices to get rid of effect of splits and dividends. These stock markets are representative markets. SHSE, and NYSE markets are like a pair of counterparts, and they list mainly industrial enterprises. SZSE-2 is similar to NASDAQ and they list mainly high-tech firms.

Table 1. Information of Empirical Data

Stock market	Stocks No.	Time span	Average trade days	Source
SHSE	1384	2008/1/2-2017/12/29	1605.37	tushare
SZSE-1	1361	2008/1/2-2017/12/29	1751.18	tushare
SZSE-2	733	2009/12/30-2018/10/19	1098.85	tushare
NYSE	2005	2008/1/2-2017/12/29	2508.02	Yahoo
NASDAQ	1838	2008/1/2-2017/12/29	1838.76	Yahoo

As shown in table 1, data spans from early 2008 to the end of 2017 while the SZSE-2 did not start until the end of 2009, so its data ranged from December 2009 to October 2018. The descending order of average trading time is NYSE, NASDAQ, SZSE-1, SHSE, SZSE-2. The average trading

time in China is significantly smaller than that in the US and the difference stems from the fact that Chinese stocks always suspend because of holidays and other issues.

We carried out a simple check and cleaning of the data. The Chinese stock data which are from tushare do not contain the data when the stock is suspended. However, the data from Yahoo will show NA when the stock is suspended and we delete some stocks with NA or other abnormal volumes.

3. Random shuffling

In this subsection, we will present a randomly shuffling process. To figure out whether there is an inherent mechanism of gap generation and complement, we use the random shuffling method to generate new sequences on the basis of original sequences. By comparing the real data and the randomly generated data, this process can help us to find the features of gap and gap-filling in real data.

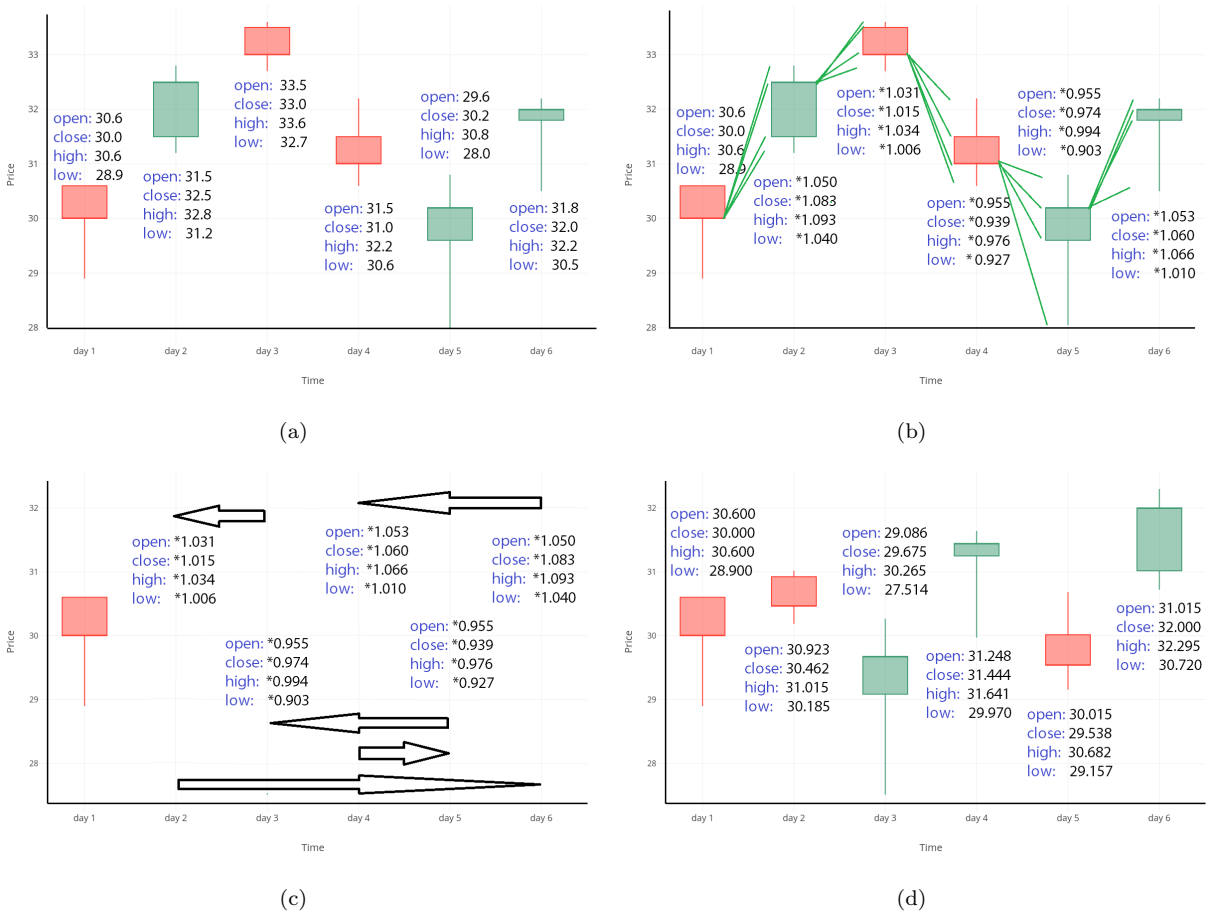

Figure 2. Randomly switching processing: (a) the original data; (b) calculate the daily growth to close price on previous trade day; (c) randomly switch daily growth; and, (d) recalculate the price data from earliest close price.

As shown in Figure 2, the process is as follows

- **Get real data and make statistics**

  Figure 2(a) is the candlestick plot of original empirical data. There are two gaps, one is up gap which occurred on day 2, and another one is down gap on day 4. The up gap is filled on day 4, but the down gap is not filled until day 6. We define $X_{t,open}$, $X_{t,close}$, $X_{t,high}$ and $X_{t,low}$ to indicate the opening, closing price, highest and lowest price of a stock on t trading

day respectively. For $t \in [1, 2, 3, ..., T]$ and T represents length of the data.

- **Calculate ratios to last close**

  Calculating the ratios of opening price, closing price, and the highest and lowest price per day to the previous day's closing price in turn.

$$\begin{aligned} r_{t+1,open} &= \frac{X_{t+1,open}}{X_{t,close}}, \quad r_{t+1,close} = \frac{X_{t+1,close}}{X_{t,close}}, \\ r_{t+1,high} &= \frac{X_{t+1,high}}{X_{t,close}}, \quad r_{t+1,low} = \frac{X_{t+1,low}}{X_{t,close}}. \end{aligned} \tag{1}$$

  For the data of the first day, it does not require calculation since there are no earlier data. In this way, we can get the relative changes of the price indices of every day. As shown in figure 2(b), this relative change is given on a daily mark as vector $\boldsymbol{r}_{t+1} = (r_{t+1,open}, r_{t+1,close}, r_{t+1,high}, r_{t+1,low})$.

- **Randomly shuffle ratio vectors**

  Keep the data of the first day, and randomly rearrange the other daily's ratio vectors. i.e. randomly shuffle the sequence $(\boldsymbol{r_2, r_3, r_4, ..., r_t, ...r_T})$.

  From figure 2(c), current $\boldsymbol{r_2^*, r_3^*, r_4^*, r_5^*, r_6^*}$ are moved from old $\boldsymbol{r_3, r_4, r_6, r_4, r_2}$ respectively. For example, $r^*_{2,open} = r_{3,open}$, $r^*_{2,close} = r_{3,close}$, $r^*_{2,high} = r_{3,high}$, and $r^*_{2,low} = r_{3,low}$ and so on.

- **Consequently recalculate prices**

  The daily prices are recalculated from the first day's original data and the rearranged ratio data. This process should be carried out in order, the second day, the third day, and so on.

$$\begin{aligned} X^*_{t+1,open} &= r^*_{t+1,open} X^*_{t,close}, \quad X^*_{t+1,close} = r^*_{t+1,close} X^*_{t,close}, \\ X^*_{t+1,high} &= r^*_{t+1,high} X^*_{t,close}, \quad X^*_{t+1,low} = r^*_{t+1,low} X^*_{t,close} \end{aligned} \tag{2}$$

  There are no gaps in the new candlestick plot 2(d). Notice that it is also possible to generate gaps by randomly adjusting the original data that do not have a gap.

If there is an internal mechanism of gap generation and re-filling in the time series, the stock prices before and after will have a strong correlation, such as between the lowest price in the next few days and the highest price in the previous day. If we switch the price sequences randomly, the process will eliminate the internal structure and see different statistics results.

4. Results of Empirical Statistics on Original Data

4.1. *The statistics of original data*

Table 2 exhibits some meaningful statistical norms, such that (N_{100}) the number of gaps in the average 100 trade date per stock; the average and standard deviation of (TD_{fill}) which presents the number of the trade days required to fill the gap; the average and standard deviation of (ND_{fill}) which presents the number of the natural days to fill the gap; and $N_{unfilled}$ the average number of gaps that occurred during this period but were not closed until the end of the period. Besides, S_{gap} defined as the gap size or gap width to distinguish how big one is the gap. For an up gap occurred at trade date t, the size is $S_{t,gap} = \frac{X_{t,low} - X_{t-1,high}}{X_{t-1,close}}$ and for a down $S_{t,gap} = \frac{X_{t-1,low} - X_{t,high}}{X_{t-1,close}}$ respectively. There are some clear evidence that we could find from table 2.

(i) On average, there will be a gap of about 3.5 to 6.8 in 100 trading days. Either up gaps or down gaps in the US stock exchange markets have a larger amount than Chinese markets do. The less occurrence of the gaps in Chinese markets perhaps due to the "T+1" trading rules and 10% restrictions imposed by the Chinese markets. Besides, the number of unfilled gaps is also significantly larger in America than in China. The proportion of down gaps in the US market is greater than up gaps and almost twice that of up gaps.

(ii) The average covering time is about 27.6-50.5 trading days, but both China and the United States have a large variance. In detail, the gap-fill time of down gaps is much longer than that of up gaps in China except SZSE-2, while it is almost the same in the U.S. This may indicates that the development of the Chinese stock market is relatively weak, and it requires more time to recover once it is reduced by the unexpected environment impacts.

(iii) The average gap size is between 0.9% and 1.6%. Whether there are up gaps or down gaps, the NYSE has the smallest gap size, and the NASDAQ has the largest gap size. The order is the same, NASDAQ $\succ$ SZSE-2 $\succ$ SHSE $\succ$ SZSE-1 $\succ$ NYSE

Table 2. Gap Statistics of Actual Data From Some Stock Exchanges

Gap type	**Index**	**SHSE**	**SZSE-1**	**SZSE-2**	**NYSE**	**NASDAQ**
up gap	N_{100}	1.5579	1.6637	1.8118	3.6070	2.4547
	TD_{fill}	30.171 (101.634)	27.626 (89.562)	36.780 (124.639)	33.417 (124.921)	33.452 (119.584)
	ND_{fill}	47.443 (158.724)	44.657 (143.611)	61.423 (210.110)	48.498 (181.419)	48.533 (173.702)
	S_{gap}	0.01464 (0.02501)	0.01446 (0.04997)	0.01577 (0.01889)	0.01021 (0.39952)	0.01638 (0.09832)
	$N_{unfilled}$	1.3721	1.7355	0.5681	6.6599	4.2136
down gap	N_{100}	1.9666	2.0534	2.1424	3.2272	2.1723
	TD_{fill}	50.531 (161.448)	47.893 (150.297)	27.767 (90.522)	30.140 (108.777)	33.237 (120.140)
	ND_{fill}	80.175 (254.046)	76.911 (238.986)	44.973 (143.906)	43.789 (158.284)	48.262 (174.473)
	S_{gap}	0.01137 (0.01391)	0.01127 (0.01492)	0.01271 (0.01506)	0.00943 (0.01860)	0.01630 (0.04216)
	$N_{unfilled}$	1.7211	1.6032	3.4128	1.5172	1.4469
total	N_{100}	3.5245	3.7171	3.9542	6.8342	4.6270
	TD_{fill}	41.531 (138.606)	38.822 (127.162)	31.896 (107.598)	31.869 (117.585)	33.351 (119.845)
	ND_{fill}	65.707 (217.741)	62.474 (202.580)	52.511 (177.521)	46.274 (170.901)	48.405 (174.064)
	S_{gap}	0.01282 (0.01968)	0.01271 (0.03537)	0.01403 (0.01689)	0.00985 (0.29481)	0.01634 (0.07790)
	$N_{unfilled}$	3.0932	3.3387	3.9809	8.1771	5.6605

4.2. *Statistical on shuffling data*

This paper uses randomly shuffling technology to reveal how far in difference of the actual data to random series. We believe that random shuffling will destroy the original pattern, and the differences can reflect the influence of the internal pattern on the gap-filling phenomenon. The results of statistics on shuffling data as shown in table 3. Compare table 3 to table 2, It is can be found as follows.

(i) Almost in all markets shuffling data have more gaps than empirical data. Especially, the down gap in each market has some considerable increment. The increment degree is larger in the US stock market than that in China. It indicates that the transaction behavior in the actual market would avoid the appearance of gaps, especially in the US.

(ii) Time to fill an up gap change significantly. Almost all the US markets take more time to fill the gaps, which indicates that there is a certain phenomenon of early gap-filling. However, the situation in China is quite complicated. The average gap-filling time of down gaps in SHSE and SZSE-1 has decreased by about 40% and the gap-filling time of up gaps has increased by around 25%. This may indicate that the up gaps are tend to be refilled while

the refilling of the down gaps is impeded. Overall, the difference on gap-filling time has been narrowed down.

(iii) Apart from NYSE, the size of the gap has been reduced to varying degrees. The overall size of the gap is slightly reduced and there is also a consistent development trend of the size of the gaps.

Table 3. Gap Statistics on Data From Randomly Shuffling Process

Gap type	Index	SHSE	SZSE-1	SZSE-2	NYSE	NASDAQ
up gap	N_{100}	1.6138	1.6660	1.7014	4.4009	3.2939
	TD_{fill}	38.693* (139.629)	35.967* (129.187)	29.788 (96.372)	34.442* (130.213)	34.026* (125.903)
	ND_{fill}	61.118* (221.774)	57.823* (207.958)	48.962 (160.567)	49.998* (189.075)	49.391* (182.776)
	S_{gap}	0.01212* (0.02400)	0.01216* (0.04986)	0.01243* (0.01573)	0.01049* (0.36743)	0.01593* (0.08766)
	$N_{unfilled}$	1.3171	1.6062	1.1744	5.0613	3.6147
down gap	N_{100}	2.3995	2.3731	2.6045	3.7991	2.8963
	TD_{fill}	39.447* (137.351)	39.023* (132.376)	30.937 (98.115)	38.052* (132.823)	35.545* (129.311)
	ND_{fill}	62.368* (217.707)	62.983* (214.584)	50.664 (161.983)	55.295* (193.047)	51.588* (187.646)
	S_{gap}	0.00905* (0.01150)	0.00918* (0.01252)	0.01052* (0.01263)	0.009813* (0.01835)	0.01538 (0.03994)
	$N_{unfilled}$	1.7934	1.6466	1.6485	2.6528	2.2409
total	N_{100}	4.0133	4.0391	4.3060	8.2001	6.1902
	TD_{fill}	39.144* (138.271)	37.763* (131.078)	30.484 (97.431)	36.114* (131.441)	34.737* (127.511)
	ND_{fill}	61.865* (219.351)	60.855* (211.891)	49.992 (161.425)	52.452* (190.943)	50.419* (185.074)
	S_{gap}	0.01028* (0.01770)	0.01042* (0.03537)	0.01128* (0.01397)	0.01018* (0.2708)	0.01567* (0.06971)
	$N_{unfilled}$	3.1105	3.2527	2.8229	7.7142	5.8556

[1] * means we can reject the null hypothesis that the random result is the same with the original one at 5% significance level using Kruskal-Wallis rank sum test.

[2] The red letters mean bigger and blue letters mean smaller than the original results.

5. Result of Statistics on No-trend Data

5.1. *Detrending processing*

As we all know, the stock markets of the United States and China have performed differently in the ten years we studied; the US stock market was on an upward trend while the Chinese stock market was on a relative downward trend. It is well known that the overall trend of the stock market does have influence on the time of gap filling. For example, an upward gap is difficult to be filled when prices are rising, because rising prices are difficult to replace the previous low level. In addition, the market trend is often affected by many external factors, such as economic recession, interest rate, policy changes, leading to the collective response of the capital market. Will this overall trend on earth affect the whole thing of gap-filling? In order to find it out, we use the detrend technology to get rid of the trend of the whole market from the price changes of all individual stocks. Based on the respective market data, we use the following steps to eliminate the trend of each stock.

- Step 1: Take the first trading day of the research period as the base period, and calculate the adjustment coefficient according to the following formula;

- Step 2: For the five stock markets, the adjustment coefficients are calculated respectively using its market index. Specifically, we use SSE Component Index, Shenzhen Component, CHINEXT Price Index, NASDAQ Composite and NYSE Composite.
- Step 3: Divide the daily opening price, maximum price, minimum price and closing price of each stock by the adjustment coefficient of the corresponding date, and obtain the stock data after the trend adjustment.

After the trend adjustment, we eliminate the impact of the market trend on each stock. Now the gaps appearance and filling phenomenon are mainly caused by the stock price movement itself. We can reach a more general conclusion in this way.

5.2. *Statistical result on detrending data and shuffling data*

Table 4 shows the statistic result of no-trend data and the statistics on shuffling no-trend data is shown in table 5.

Table 4. Gap Statistics of No-trend Data From Some Stock Exchanges

Gap type	**Index**	**SHSE**	**SZSE-1**	**SZSE-2**	**NYSE**	**NASDAQ**
up gap	N_{100}	2.1725	3.0157	2.6894	4.2063	3.3948
	TD_{fill}	23.039 (103.802)	17.871 (75.305)	17.976 (75.612)	28.351 (118.555)	25.705 (107.116)
	ND_{fill}	36.464 (164.133)	28.729 (122.013)	29.685 (127.049)	41.165 (172.032)	37.312 (155.528)
	S_{gap}	0.01250 (0.02356)	0.01174 (0.02678)	0.01352 (0.01858)	0.00839 (0.03122)	0.01511 (0.08566)
	$N_{unfilled}$	1.1185	2.0154	0.8595	2.3461	1.6823
down gap	N_{100}	2.2705	3.2461	2.8684	4.0237	3.4649
	TD_{fill}	17.818 (73.020)	17.208 (72.198)	14.654 (61.776)	23.315 (88.606)	23.573 (94.618)
	ND_{fill}	28.417 (116.132)	28.034 (116.431)	23.903 (101.689)	33.718 (129.397)	34.215 (137.435)
	S_{gap}	0.00811 (0.01078)	0.00785 (0.01082)	0.00889 (0.01136)	0.00796 (0.01643)	0.01261 (0.03356)
	$N_{unfilled}$	1.3208	1.4798	1.3656	4.7450	4.8997
total	N_{100}	4.4430	6.2618	5.5578	8.2300	6.8597
	TD_{fill}	20.371 (89.443)	17.527 (73.711)	16.262 (68.839)	25.889 (105.374)	24.628 (101.002)
	ND_{fill}	32.352 (141.707)	28.369 (119.152)	26.701 (114.698)	37.536 (152.784)	35.748 (146.676)
	S_{gap}	0.01025 (0.01831)	0.00974 (0.02029)	0.01111 (0.01543)	0.00818 (0.02501)	0.01381 (0.06425)
	$N_{unfilled}$	2.4394	3.4952	2.2251	7.0911	6.5820

Comparing table 4 to table 2, we can find some changes before and after the detrending process.

(i) The number of both kinds of gaps increases in all 5 markets, while the changes of unfilled gaps are not significant. The increment of the US stock markets is somewhat larger than the Chinese markets, whereas the SZSE-1 experience the most sharp increment. It may indicate that the real market will avoid the appearance of gaps.

(ii) The time to fill a gap is significantly shorter after the detrending process and the difference between the up gaps and down gaps are narrow. After detrending, it takes even less time to close a down gap than it does to close a up gap in SHSE and SZSE-1, which means the overall trend in these two markets hinder the refilling process.

(iii) As to the gap size, the changes in both Chinese and US stock markets are not evident. The changes go in the same smaller direction and the order of the five markets remains the same.

Then we focus on the changes on the no-trend data before and after the random exchange. Comparing table 5 to table 4, it can be found as follows.

Table 5. Gap Statistics on Data From Randomly Shuffling Process

Gap type	Index	SHSE	SZSE-1	SZSE-2	NYSE	NASDAQ
up gap	N_{100}	2.0299	2.7262	2.5611	4.4780	3.9218
	TD_{fill}	24.220* (98.993)	21.745 (94.686)	16.908 (70.481)	26.276 (110.075)	27.231* (110.232)
	ND_{fill}	38.434* (157.376)	35.261* (155.349)	27.729 (117.035)	38.183 (159.997)	39.522* (159.973)
	S_{gap}	0.01165 (0.02312)	0.01124 (0.02748)	0.01191 (0.01630)	0.00902 (0.03119)	0.01580 (0.08114)
	$N_{unfilled}$	1.4400	1.9662	1.3070	2.6325	1.7134
down gap	N_{100}	2.6938	3.5377	3.3660	4.4492	4.0792
	TD_{fill}	18.310* (90.705)	16.145 (80.146)	12.951 (61.944)	27.634 (116.768)	24.896* (101.430)
	ND_{fill}	29.144* (144.452)	25.987* (129.430)	21.259 (104.044)	40.162 (169.686)	36.149* (147.212)
	S_{gap}	0.007445 (0.009444)	0.007812 (0.01015)	0.008441 (0.009921)	0.008459 (0.01674)	0.01298 (0.03299)
	$N_{unfilled}$	0.7449	0.9104	0.8336	3.4283	4.4272
total	N_{100}	4.7237	6.2639	5.9271	7.1105	5.6352
	TD_{fill}	20.850* (94.401)	18.582* (86.818)	14.661 (65.797)	26.953* (113.461)	26.041* (105.842)
	ND_{fill}	33.136* (150.212)	30.023* (141.370)	26.701* (114.698)	39.169* (164.899)	37.802* (153.608)
	S_{gap}	0.00927 (0.01697)	0.00932 (0.01985)	0.00996 (0.01321)	0.00874 (0.02503)	0.01433 (0.06107)
	$N_{unfilled}$	2.1849	2.8766	2.1426	6.0608	6.1406

[1] *means we can reject the null hypothesis that the random result is the same with the original one at 5% significance level using Kruskal-Wallis rank sum test.
[2] The red letters mean bigger and blue letters mean smaller than the original results.

(i) The changes of statistics before and after random exchange are reduced. The total number of the gaps did not change too much after the shuffling process, even there are less gaps in the US stock market.
(ii) Time to fill an up gap doesn't change significantly. SZSE-1 has the most time increment which is 22%. There is a significant time difference between the Chinese markets and the US markets in down gap filling process. All the Chinese markets require less time to fill the gaps in the randomly exchange models except SHSE that the gap filling time almost did not change, whereas all the US markets take more time to fill the down gap.
(iii) The changes of the gap sizes are statistically insignificant.

These may indicate that after the detrending process, the patterns in the original data series have been broken, so the statistic results are quite the same before and after random shuffling process, which means that the gap-filling phenomenon has something to do with the market.

6. Conclusions and Discussions

In this paper, we discussed the gap phenomena in stock price. Specifically, we focused on the statistic characteristics of gap and gap-filling time. Different markets were slightly different in some statistic characteristics, such as the the number of gaps, the gap-filling time, the size of the gap and so on, but generally they are similar. For example, the U.S. stock markets are likely to generate more gaps than China and the average time to fill different types of gaps is almost the same, while the Chinese stock markets need more time to fill a downward gap than an upward one. However the

whole gap-filling thing is quite the same in both China and the U.S. Beyond the statistic results, to reveal that the existed gap can be recovered in a specific pattern, we proposed a randomly shuffling process. After the Randomly Shuffling Process, we found that there were some significant differences in each characteristic, especially for the gap-filling time. In Chinese stock market, the differences between up gaps and down gaps have been narrowed down after the process, while the gap-filling time of both gap types has been increased in the US stock market. Overall, we can reject that most of the random results are the same with the original one at 5% significance level by using Kruskal-Wallis rank sum test. In short, we found similar gap-filling phenomenon in China and the US, which was that the stock market had some internal pattern in terms of gap-filling. We may recognize that this phenomenon is widespread in these countries and is more intense in China.

Beyond that, we use detrending technology to remove the operating trend of the whole market from the price changes of all individual stocks. We apply the same analysis on the no-trend data and found the changes of statistics before and after random exchange are reduced, which means that the gap-filling phenomenon may possibly be related to the overall trend. Moreover, regardless of the trend, the number of gaps in the U.S. market and the number of gaps that have not been repaired are larger than those in China, reflecting the high activity of the U.S. stock market, which is also related to China's "T+1" trading rules and the daily 10% restrictions imposed by the Chinese markets. Both Chinese and American stocks have the desire to make up for the gap. However, some differences between real data and no-trend data suggest that the overall trend hinders the gaps' generation and slows down the gaps' refilling process to a certain extent.

Acknowledgement

We appreciate comments and helpful suggestions from Prof. Yougui Wang. This work was supported by Chinese National Natural Science Foundation (71701018) and (61673070), National Social Sciences Fund, China (14BSH024), and Beijing Normal University Cross-Discipline Project.

References

Bouchaud, J.P., Mézard, M., Potters, M. *et al.*, Statistical properties of stock order books: empirical results and models. *Quantitative finance*, 2002, **2**, 251–256.

Caporale, G.M. and Plastun, A., Price gaps: Another market anomaly?. *Investment Analysts Journal*, 2017, **46**, 279–293.

Cont, R., Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 2001, **1**, 223–236.

Cristelli, M., Universal relation between skewness and kurtosis in complex dynamics. In *Complexity in Financial Markets*, pp. 141–150, 2014, Springer.

Dahlquist, J.R. and Bauer, R.J., *Technical analysis of gaps: identifying profitable gaps for trading*, 2012, FT Press.

Filimonov, V. and Sornette, D., Spurious trend switching phenomena in financial markets. *The European Physical Journal B*, 2012, **85**, 155.

Hommes, C.H., Modeling the stylized facts in finance through simple nonlinear adaptive systems. *Proceedings of the National Academy of Sciences*, 2002, **99**, 7221–7228.

Hommes, C.H., Heterogeneous agent models in economics and finance. *Handbook of computational economics*, 2006, **2**, 1109–1186.

Lehoczky, J. and Schervish, M., Overview and History of Statistics for Equity Markets. *Annual Review of Statistics and Its Application*, 2018, **5**.

Lima, L., Modeling of the financial market using the two-dimensional anisotropic Ising model. *Physica A: Statistical Mechanics and its Applications*, 2017, **482**, 544–551.

Morris, G.L., *Candlestick Charting Explained: Timeless Techniques for Trading Stocks and Futures: Timeless Techniques for Trading stocks and Sutures*, 2006, McGraw Hill Professional.

Peacock, A.T., *The political economy of economic freedom*, 1997, Edward Elgar Publishing.

Schmitt, N. and Westerhoff, F., Heterogeneity, spontaneous coordination and extreme events within large-scale and small-scale agent-based financial market models. *Journal of Evolutionary Economics*, 2017, **27**, 1041–1070.

Sirignano, J. and Cont, R., Universal features of price formation in financial markets: perspectives from Deep Learning. *Arxiv*, 2018.

Wang, J., Yang, C.X., Zhou, P.L., Jin, Y.D., Zhou, T. and Wang, B.H., Evolutionary percolation model of stock market with variable agent number. *Physica A: Statistical Mechanics and its Applications*, 2005, **354**, 505–517.

Guo, Y., The guiding significance of the gap theory in portfolio investment technique analysis. *Sci/Tech infomation Development & Economy*, 2004, **14**, 80–82.