Market Efficiency Anomalies: A Study of January Effect in Chinese Stock Market

Contents

1 Introduction

Fama introduced **Efficient Market Theory** (EMH), which suggests that all available information on a stock market are fully reflected by stock price [1]. It is impossible for an investor to perform excess performance. A paramount assumption involved in this theory is that investors in the market are rational enough to respond to all market information quickly. This theory is important because it relates to how people react in the market. However, there are some anomalies which against this theory in subsequent investigations. Calendar effect is one of the most famous anomalies. It mainly includes seasonal effect, month effect, week effect and holiday effect. In this report, we mainly focus on the **January effect**.

The January effect is a popular market anomaly in both academic and professional literature. It means that the average yield of the stock market in January is higher than that of other months and is statistically significant. The January effect contradicts the efficient market hypothesis because excess returns in January are not attributed to any new and relevant information. Moreover, the theoretical basis of the efficient market hypothesis is derived from the randomness of stock price changes. But the January effect, as a kind of market anomalies, shows the regular pattern of stock price, which is the average return rate of January is higher than that of other months. This phenomenon can not be explained by the efficient market hypothesis.There are two main explanations about the January effect. One of them is tax-loss selling hypothesis. The hypothesis suggests selling securities that have already suffered significant losses before the end of the year, and then buying an equivalent. This brings losses to investors and thus reduces taxable income. The other explanation is window-dressing hypothesis. It implies that fund managers usually whitewash windows at the end of the season or at the end of the year to submit beautiful performance sheets to fund investors.

However, there are some debates about these two hypothesis.

In order to analyze whether the January effect exists, we choose two representative market portfolio indices in Chinese stock market called Shanghai Composite Index and Shenzhen Component Index as our object of study. By F-test, we can compare the expected returns in January and other months. The result is momentous for investors to make their investment decision.

2 Literature Review

Since January effect was introduced in the last century, scholars had set off a research upsurge on the January effect, which had been confirmed in the stock markets of many countries and regions. Therefore, in order to examine whether January effect is still prevalent and whether it exists in Chinese stock market, we researched the reports of last five years regarding January effect.

In the study of Sarangi, Kar & Mohanty[2], to examine whether January Effect is available, the author used observations of 15 years, from 1998 to 2013, of the two major indices reported by National Stock Exchange (NSE),and test the January effect by using Friedman's sum rank test, Mann-Whitney U-test and dummy variable regression analysis, which are tests for seasonality. This article proposes three important reasons to explain why the return on income in January was higher than in December. The first reason is, in order to avoid taxation, investors sell stocks in December. The second reason is that portfolio managers sell poorly performing stocks in order to improve their portfolios. The third reason is that portfolio managers target more small company stocks by selling certain stocks, and they hope to obtain profits that exceed their value.

In Kemal and Sinem's paper [1], January effect has been examined in 23 Borsa Istanbul sector and sub-sector indices. The method they use is OLS method to test the existence of the anomaly. Finally, they find that there is an evidence for January effect in leasing and sports factoring indices.

From Ullah and Ullan [3], this article explores whether the January effect exists, using regression models to test seasonality and measure the incremental effect on a monthly basis. The data set used by this study is the daily closing values of KSE-100 index from 1st January 2004 to 31st December 2014. The normality test and stationarity test are very enlightening to our data processing and testing. Intercept term analysis was carried out by GARCH, EGARCH and TGARCH results. Its greatest feature is that it passes many different tests.

Avdalovic and Milenkovic's [4] paper was to test historical data on evolution of stock exchange indexes and evaluated the possibility of certain calendar months profit beyond average during the year, also acknowledged as the January effect. However, in this study it had been justified that the traditional January effect exists only on the stock market in Macedonia.

The report written by Plastun *et al.* [5] comprehensively analyzed the movement of monthly anomalies by using several statistical techniques (average analysis, Student's t-test, ANOVA, the Mann–Whitney test). The results showed that the January effect was most pronounced in the US. However, the December effect and the Mark Twain effect were never prevalent in the US.

In reading the article Shen *et al.* [6] we can see that, this article suggests that One reason for the January effect is that investors are only willing to sell at a premium, so the transaction price is unreasonably high.

3 Methodology

3.1 Data Description

Collecting historical data about daily rate of returns from 2010.1 to 2019.12 based on the Shanghai Composite Index (SSEC) and Shenzhen Component Index (SZI) from Wind Database, later the average returns of each month are calculated so as to make a linear regression.

The reasons why we select this time period are:

1) Since the implementation of the Limit System in Dec.1996, the maximum range of daily increase or decrease has been limited to 10% for every stocks, which leads to the relatively stable stage of China's stock market from then on.
2) We could apply the existed theorems on the latest data from 2010 to 2019 to test if they are still in compliance.

The sample data consists the daily closed price of the SSEC and SZI, notated as P_t (we deal with these two indices separately). Through this sequence of P_t, the value of the securities portfolio at time t, we define the daily rate of return R_t of

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{1}$$

where P_{t-1} stands for the close price last day.

3.2 Model of Daily Rate of Return

There are researches[7] suggest that the stock price follows a random walk, where

$$R_t = u + \epsilon_t \tag{2}$$

To better serve for study on January effect, we adjust this model into an OLS type:

$$R_t = \beta_1 + \sum_{i=2}^{11} \beta_i m_{2t} + \epsilon_t \tag{3}$$

Here m_{it} is a dummy variable. For instance, if $i = 3$, *i.e.* this rate of return takes place in March, then $m_{3t} = 1$, otherwise $= 0$. In addition, β_1 stands for the average monthly rate of return for January, and $\beta_i (i = 2, 3, ..., 12)$ each accounts for the difference between daily rate of return of that month and January's.

Hence, this OLS model is applied to examine if the average daily rate of return in these 12 months are significantly equal. The null hypothesis could be as the following

$$H_0 : \beta_2 = \beta_3 = ... = \beta_{12}$$

If not all the parameters $\beta_i (i = 2, 3, ..., 12)$ are equal to 0 at the same time, then it suggests the existence of so-called **January Effect**.

At present, most previous studies chose the OLS models, but considering that the sequence of return rates could be time series, which might be non-stationary or have heteroscedasticity problems that violates the assumption of OLS, this article would pre-process with the OLS model, and then choose an appropriate empirical model on the basis of the relevant tests.

3.3 Relevant Tests

3.3.1 Stationary Test (ADF)

When we collect our data, we should pay attention to the stationary of the data, because if the data is not stationary, we may find unit roots. Any regression relationship between independent variables and dependent variables with unit roots is deceptive, because any error in the residual sequence does not decline with increasing sample size. It means that the deviation in the model is permanent. Thus, we use R to do an ADF test on our data. Our null and alternative hypotheses are as follows:
H_0 : The yield data of SSEC has unit root, and the time series data is non-stationary series. H_1 : The yield data of SSEC does not have unit root, and the time series data is stationary series. The results are listed below

```
	Augmented Dickey-Fuller Test

data:  上证指数_ts
Dickey-Fuller = -12.966, Lag order = 13, p-value = 0.01
alternative hypothesis: stationary


	Augmented Dickey-Fuller Test

data:  副本副本深证成指_ts
Dickey-Fuller = -12.859, Lag order = 13, p-value = 0.01
alternative hypothesis: stationary
```

Figure 1: ADF test

Based on the result above, we could comment that, the Augmented Dickey-fuller t-statistic for SSEC is equal to -12.996, which is much greater than the critical value -2.873339. And the p-value here is 0.01, also smaller than 0.05. We reject the null hypothesis. Hence we have a stationary time series data. And we have a similar consequence for Shenzhen security index, with ADF test p-value = 0.01. Thus, the time series for SZI is also stationary.

3.3.2 ARCH Effect Test

A significant assumption of the classic linear regression model (OLS) is: the random errors in the overall regression function meet the condition of same variance, that is, they all have the same variance.

However, the residual variance obtained by actual data often changes with time, which means the model does not meet the assumption of homoscedasticity. Therefore, when applying the OLS for regression, it will lead to errors in the results if we ignore the possibility of heteroscedasticity.

ARCH Langerand multiplier test (ARCH-LM test) is used to detect whether the data obtained by constructing the OLS model has heteroscedasticity issues. Use auxiliary detection to perform regression analysis on null hypothesis. In addition, the test results follow the χ_p^2 distribution.

H_0 : The time series has no ARCH effect at the lag order p (in this paper, we set $p = 1$). H_1 : The time series has ARCH effect at the lag order p.

$$u_t^2 = \beta_0 + \sum_{s=1}^{p} \beta_s u_{t-1}^2 + \epsilon_t \tag{4}$$

where u_t is the residual term, which is a regression of constant and lag-square residuals up to order p.

Two statistics in the regression:

1. The $N * R^2$ statistic is the Engle$'$s LM test statistic, which is the R^2 obtained by multiplying the number of observations N by the regression test;

2. The F statistic is an omitted variable test value for the joint significance of all lagged p-order residual squares.

After testing the known data and the OLS regression model, the data are:

df	Chi-square	p-value	heteroscedasticity or not
1	103.45	$< 2.2e-16$	yes
2	215.24	$< 2.2e-16$	yes
3	273.83	$< 2.2e-16$	yes
4	303.61	$< 2.2e-16$	yes
5	317.47	$< 2.2e-16$	yes

Table 1: SSEC ARCH-LM test

df	Chi-square	p-value	heteroscedasticity or not
1	73.095	$< 2.2e-16$	yes
2	187.73	$< 2.2e-16$	yes
3	234.43	$< 2.2e-16$	yes
4	262.98	$< 2.2e-16$	yes
5	278.59	$< 2.2e-16$	yes

Table 2: SZI ARCH-LM test

The p-value $< 2.2e-16$, both indicate that there is autoregressive condition in the residual after using OLS for regression. For heteroscedasticity, the OLS model does not fit the data well.

3.3.3 Likelihood Ratio Test

To examine the data, we use the **Likelihood Ratio Test**, which reflects both sensitivity and specificity of the model defined above. The idea of the likelihood ratio test is: if the parameter constraint is effective, then adding such a constraint would not cause a substantial decrease in the maximum value of its **likelihood function**.

That is to say, the essence of the likelihood ratio test is to compare the maximum likelihood function under constrained conditions with the maximum likelihood function under unconstrained conditions.

Our purpose of using the likelihood ratio test is to select a more suitable model,

$$H_0 : R_t \sim 1 \quad \text{(reduced model)}$$

$$\begin{aligned} H_1 : R_t \sim & Month_2 + Month_3 + Month_4 + Month_5 + Month_6 \\ & + Month_7 + Month_8 + Month_9 + Month_{10} + Month_{11} + Month_{12} \end{aligned}$$
(full model)

The difference between the two models is that the null hypothesis believes that there is no difference between the return rate respectively from February to December and January, that is, there is no January effect. The alternative hypothesis believes that the return rate for at least one month' return rate is the significantly different from January.

4 Empirical Results and Data Inference

4.1 Statistical Description

Within the time period 2010-2019, the rate of returns samples of two major indices: Shanghai Composite Index (SSEC), and Shenzhen Component Index (SZI), is shown in the **Table 3**. It is reported that the **average daily rate of return** of SSEC and SZI within sample period were 0.008% and 0.007% respectively, and statistically significant. SZI had larger standard deviation (1.615%) than SSEC (1.339%), suggesting a higher volatility. Additionally, the maximal and minimal values come from the time period of June to September for both two indices. This clustering tendency might indicate their relevance to some extent.

Month	Mean		sd		t-value		Min		Max	
from	SH	SZ	SH	SZ	SH	SZ	SH	SZ	SH	SZ
Jan	-0.101	-0.165	1.687	1.929	-0.677	-1.012	-7.705	-8.227	4.745	4.946
Feb	0.145	0.236	1.385	1.726	1.304	1.682	-6.407	-7.340	5.601	5.588
Mar	0.068	0.080	1.199	1.494	0.639	0.738	-4.396	-5.289	4.263	4.767
Apr	0.023	-0.023	1.107	1.325	0.143	-0.219	-4.792	-6.220	3.043	4.090
May	-0.065	-0.044	1.382	1.675	-1.002	-0.502	-6.505	-7.556	3.480	4.034
Jun	-0.195	-0.193	1.495	1.814	-0.694	-1.429	-7.397	-8.244	5.532	5.688
Jul	0.003	0.016	1.500	1.764	-0.053	0.153	-8.483	-7.586	5.764	5.239
Aug	-0.063	-0.089	1.506	1.657	-0.307	-0.886	-8.491	-7.829	5.340	5.321
Sep	0.011	0.006	1.157	1.521	-0.086	0.090	-3.520	-6.551	4.895	6.454
Oct	0.155	0.166	1.247	1.582	1.389	1.136	-5.223	-6.074	4.094	4.888
Nov	0.013	-0.017	1.167	1.462	-0.062	-0.204	-5.481	-7.003	4.310	5.287
Dec	0.101	0.109	1.232	1.427	0.254	0.771	-5.430	-4.512	4.325	4.840
Total	0.008	0.007	1.339	1.615	0.071	0.027	-8.491	-8.244	5.764	6.454

Table 3: Average Monthly Rate of Return of SSEC & SZI, 2010-2019

4.2 Relevant Tests

The analysis process through R and result for SSEC data is shown below:

```
Call:
lm(formula = Return ~ ., data = data_SH)

Residuals:
     Min       1Q   Median       3Q      Max
-1.29577 -0.17655 -0.00289  0.18080  0.79928

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.10104    0.10102  -1.000   0.3195
Month_2      0.24626    0.14287   1.724   0.0876 .
Month_3      0.16867    0.14287   1.181   0.2403
Month_4      0.12404    0.14287   0.868   0.3872
Month_5      0.03623    0.14287   0.254   0.8003
Month_6     -0.09360    0.14287  -0.655   0.5137
Month_7      0.10417    0.14287   0.729   0.4675
Month_8      0.03833    0.14287   0.268   0.7890
Month_9      0.11157    0.14287   0.781   0.4366
Month_10     0.25633    0.14287   1.794   0.0756 .
Month_11     0.11423    0.14287   0.800   0.4257
Month_12     0.20216    0.14287   1.415   0.1599
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3195 on 108 degrees of freedom
```

```
Multiple R-squared:  0.09543,Adjusted R-squared:  0.003302
F-statistic: 1.036 on 11 and 108 DF,  p-value: 0.4205
```

Referring to the p-value for F-statistic, it is larger than 0.05, so that fails to reject the null hypothesis $H_0 : \beta_2 = \beta_3 = ... = \beta_{12}$. Equivalently, it is likely to happen that all β_is equal to 0 at the same time, which leads to the invalid January Effect phenomenon on the basis of the sample data.

```
> lrtest(model)
Likelihood ratio test

Model 1:
    Return ~ Month_2+ Month_3+ Month_4+ Month_5+ Month_6
 + Month_7+ Month_8+ Month_9+ Month_10+ Month_11+ Month_12
Model 2: Return ~ 1
  #Df  LogLik  Df  Chisq Pr(>Chisq)
1  13 -27.018
2   2 -33.036 -11 12.036     0.3609
```

According to the above-mentioned logarithm likelihood ratio on the Shenzhen index, it can be seen that the likelihood ratio and the p-value is so large that it fails to reject the null hypothesis. The larger the likelihood function, the more likely an unknown situation will occur, and the more reasonable the corresponding result will be. At this time, the null hypothesis H_0 should not be rejected.

Similar codes were applied for SZI data:

```
Call:
```

```
lm(formula = Return ~ ., data = data_SZ)

Residuals:
    Min      1Q  Median      3Q     Max
-1.2317 -0.2300 -0.0449  0.2249  1.0467

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.16511    0.11849  -1.393   0.1663
Month_2      0.40121    0.16756   2.394   0.0184 *
Month_3      0.24540    0.16756   1.465   0.1460
Month_4      0.14229    0.16756   0.849   0.3977
Month_5      0.12158    0.16756   0.726   0.4697
Month_6     -0.02839    0.16756  -0.169   0.8658
Month_7      0.18063    0.16756   1.078   0.2834
Month_8      0.07635    0.16756   0.456   0.6496
Month_9      0.17152    0.16756   1.024   0.3083
Month_10     0.33124    0.16756   1.977   0.0506 .
Month_11     0.14776    0.16756   0.882   0.3798
Month_12     0.27371    0.16756   1.633   0.1053
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3747 on 108 degrees of freedom
Multiple R-squared:  0.1043, Adjusted R-squared:  0.0131
F-statistic: 1.144 on 11 and 108 DF,  p-value: 0.3354

> lrtest(model)
Likelihood ratio test
```

```
Model 1:
    Return ~ Month_2+ Month_3+ Month_4+ Month_5+ Month_6
+ Month_7+ Month_8+ Month_9+ Month_10+ Month_11+ Month_12
Model 2: Return ~ 1
  #Df  LogLik  Df  Chisq Pr(>Chisq)
1  13 -46.151
2   2 -52.762 -11 13.221     0.2791
```

The above process tells that, all p-values of other months are greater than 0.05, which means the difference between the return of January and that of the rest month in these years are insignificant. For the F-statistic of SZI, it has the p-value of $0.3354 > 0.05$ where H_0 is failed to be rejected.

By using the likelihood ratio test, the p-value is given by,

$$Pr(T_n > \text{realized } Tn)$$

By the result of likelihood test, realized Tn is given by,

$$2\ln\frac{L_1}{L_2} = 13.221$$

L_1 and L_2 respectively refer to the log-likelihood function of full model and reduced model.

$$Pr(T_n > \text{realized } T_n) = Pr(> 13.222) = 0.3609$$

The p-value is too large, so we do not reject H_0.

Hence, we can conclude that there is no significant ‘January effect’ in China.

5 Conclusion

From **Section** 4, we could draw the conclusion that the January effect could not be observed from the 2010-2019 daily rate of return data for SSEC and SZI, or generally the Chinese stock market.

5.1 Possible Causes of The January Effect

In this section, we try to explain some of the causes of January Effect, and the reasons why it is not obvious in Chinese stock market based on our selected data.

1. According to Ligon[8], higher January returns are attributed to higher trading volumes and lower real interest rates. In practice, the abnormal consequence may be a result for the Lunar New Year Holiday, one of the most celebrated festivals in China, usually arrives in February, where the amount of consumption increases sharply and contributes to the rise of stock prices through market activities.

   For instance, the rising demand of liquid payments gives pressure to the stock market before Holiday, especially in January. However, the demand of liquidity drops after the holiday, such as in March, then the stock market would have a certain degree of withdrawing or catch up as a consequence.

2. China Securities Regulatory Commission requires listed companies to disclose their financial position and audit report for last year before April 30. The disclosure enables the investors to assess and trade on those stocks, which promote the appearance of the post-Spring Festival effect[9].

3. Besides, the movement of the Spring Festival holiday date also influence the degree of January effect in China. To be more specific, there are 7 among 10 Lunar New Year holiday take place in February instead of January. Gao[10] argued that this could possibly lead to contradiction about the January effect in domestic stock market.

The discussion above briefly provides some possible explanations and further exploration is needed on this topic.

5.2 Further Exploration about This Topic

There is often a strong stationary tendency for the rate of return sequence in the stock market. It can hardly make efficient estimations for parameters without using time series model under this condition. An accessible model proposed by [11] is called the **GARCH** model, whose general expression is

$$\sigma_t^2 = \alpha_0 + \lambda_1 \sigma_{t-1}^2 + ... + \lambda_p \sigma_{t-p}^2 + \alpha_1 u_{t-1}^2 + ... + \alpha_q u_{t-q}^2 \tag{5}$$

where $\alpha_0 > 0$; $\alpha_i \geq 0,\ i = 1, 2, ..., q$; $\lambda_j \geq 0,\ j = 1, 2, ..., q$. Other articles use this method to build and test the outliers of monthly returns. Liu[12] found strong cluster of volatility in domestic stock market through the maximum likelihood estimation for the parameters of ARCH and GARCH in the model, suggesting the January effect is insignificant in China.

Besides, there are also reports applying the CAPM model to verify the January effect. Jiang[13] and Li[14] regarded that the January effect has no significant influence on the excess yield of the industries.

Other journals also demonstrate a certain powerful tests. For instance, Kar[2] applied non-parametric tests including Wilcoxion Mann-Whitney U-Test, and ANOVA for testing the seasonality of the data sequence, and eventually reached the conclusion that the seasonal effect did not exist in Indian stock market.

References

[1] Kemal EYUBOGLU and Sinem EYUBOGLU. Examining the january effect in borsa istanbul sector and sub-sector indices. *International Journal of Economic Perspectives*, 10(2):102 – 109, 2016.

[2] PRAGYAN PARIMITA SARANGI, N. C. KAR, and MUNMUN MOHANTY. Existence of january effect: An empirical investigation on indian stock market. *CLEAR International Journal of Research in Commerce & Management*, 6(6):1 – 9, 2015.

[3] Ullah. I and Ullah. S. Market efficiency anomalies: A study of january effect in karachi stock market. *Journal of Managerial Sciences.*, 10(1):31 – 44, 2016.

[4] Avdalovic S. and Milenkovic I. January effect on stock returns: Evidence from emerging balkan equity markets. *Economics institute, Belgrade, 2017.*, 45(4):7–21, 2017.

[5] Snežana Milošević-Avdalović and Ivan Milenković. January effect on stock returns: Evidence from emerging balkan equity markets. *Industrija*, 45(4):7 – 21, 2017.

[6] Chiou J. Shen K. Shen Y., Hung C. The january effect and prospect theory in taiwan. *Financial Analysts Journal*, 62(5):1113–1123, 2020.

[7] Zhang J. and Gu J. 中国股票市场春节效应的实证研究 / empirical study of spring festival effect in chinese stock market. 武汉理工大学学报（信息与管理工程版） / *JOURNAL OF WUHAN UNIVERSITY OF TECHNOLOGY(INFORMATION MANAGEMENT ENGINEERING)*, 27(6):141 – 144, 2005.

[8] James A. Ligon. A simultaneous test of competing theories regarding the january effect. *Journal of Financial Research*, (1):13, 1997.

[9] 孙丽敏 李红兵. 中国股票市场一月效应的实证研究 / empirical study on january effect of chinese stock market. 财贸研究 / *FINANCE AND TRADE RESEARCH*, 20(2):96 – 100, 2009.

[10] 高俊. 中国股票市场的春节效应. 2017.

[11] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307 – 327, 1986.

[12] 刘鹏. 我国股市”一月效应”和”年末效应”的实证分析 / the demonstration of january's and december's effects in china stock market. 广西财政高等专科学校学报 / *JOURNAL OF GUANGXI FINANCIAL COLLEGE*, 17(6):29 – 41, 2004.

[13] 蒋莹诗. CAPM模型对上海股票市场一月效应的验证. 时代经贸 / *Economic & Trade Update*, (5):80 – 81, 2014.

[14] 李呈娇. CAPM模型对中国上证a股的实证分析 / empirical analysis of shanghai stock exchange a shares by CAPM model. 武汉冶金管理干部学院学报 / *JOURNAL OF WUHAN METALLURGICAL MANAGER'S INSTITUTE*, 20(4):14 – 16, 2010.