

# Research on Stock Return Rate

-- Focus on Relationship among Stock Return Rate, Market Return Rate, Risk-Free Rate and Other Stock Return Rate in Chinese Stock Market

Han Yu

Economic and Financial Research Institute in Nanjing Audit University, Nanjing, Jiangsu Province, 211800, China

**Abstract:** There is a certain relationship among stock return rate, market return rate and risk-free interest rate, which is worth discussing, and it is helpful for us to analyze stocks and evaluate their prices. I have found that the market return rate and risk-free rate have correlation through multiple regression, and other stock's return rate can affect the target stock to some extent. The stock return rate is positively related to the market interest rate and inversely related to the risk-free interest rate.

**Keywords:** Regression Test, Model building, Multiple Regression, CAPM.

## 1. Introduction

### 1.1. Research Question

Relationship among  $R_f$ ,  $R_m$ ,  $R_{it}$ ,  $R_{jt}$   
 I think there is a certain relationship among market interest rate, risk-free interest rate and stock return rate, and there may be other models to reflect this relationship, so I conducted multiple regression to explore the relationship among these data. I often find that different stocks often rise and fall together. The rate of return of a single stock is not only related to the market rate of return, but also related to the rate of return of other stocks. Therefore, I want to explore the relationship among the return rate of a stock, the market return rate, the risk-free interest rate and other stock that may affect the target stock.

## 2. Model Building

After considering the relationship between stock yield, market yield and risk-free interest rate separately, and whether there are other interference factors.

Let's consider the following general formulas that I build, which may reflect the relationship among these variables:

$$R_{it} = \alpha + \beta_1 R_m + \varepsilon$$

$$R_{it} = \alpha + \beta_2 R_f + \varepsilon$$

$$R_{it} = \alpha + \beta_1 R_m + \beta_2 R_f + \beta_3 R_f^2 + \beta_4 R_{jt} + \varepsilon$$

$$R_{it} - R_f = +\beta_1 (R_m - R_f) + \beta_2 R_{jt} + \varepsilon$$

$R_{it}$  is the return rate of the stock that I pick,  $R_{jt}$  is the stock that affects the picked stock.  $R_f$  is the risk-free rate that is often measured by the three-month time deposit rate.  $R_m$  is the market return.  $\alpha$  is the intercept term.  $\varepsilon$  is the error.

I select the Stock 600009 (Shanghai Airport) and Stock 600897(Xiamen Airport) to run the model.

The data comes from the official website data of Shanghai Stock Exchange and the website of Bank of China.

## 3. Data Characteristics

I chose Shanghai Airport (600009) as the target stock which to be researched. I collected the monthly return of this stock and market returns from January 2000 to December 2019. I also found the personal 3-month time deposit interest rate in the past 20 years as the value of risk-free interest rate on the official website of Bank of China. I take the difference between market rate of return and risk-free interest rate as a kind of variable and the difference between Stock 600009 monthly rate of return and risk-free interest rate as another variable, which will be tested in the following model that I build.

There is a table that describe the variables:

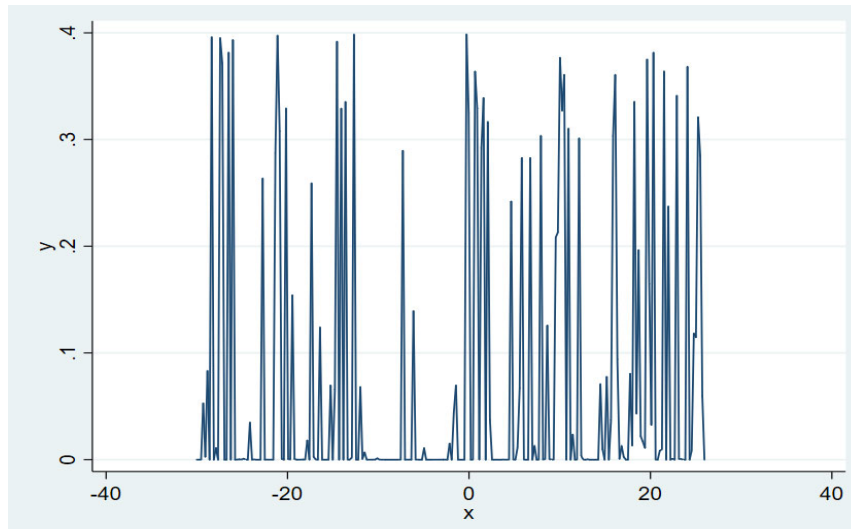
Table 1. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Excess return of the Stock 600009	240	-0.398	9.067	-41.67	41.43
Excess return of Shanghai Stock Market	240	-1.351	7.634	-27.51	25.65
Rt (600009)	240	1.575	8.937	-38.79	43.41
Rt (600897)	240	1.478	9.583	-40.98	39.96
Rm	240	0.622	7.547	-24.63	27.45
Rf	240	1.973	.614	1.1	3.33

Data source: official website of Bank of China and Shanghai Stock Exchange

The difference between market rate of return and risk-free interest rate is the excess return of Shanghai Stock Market and the difference between stock monthly rate of return and risk-free interest rate is the excess return of Stock 600009. There are the data summary. From the Table 1, for example, the mean of Stock 600009 is 1.575, and the mean of market return

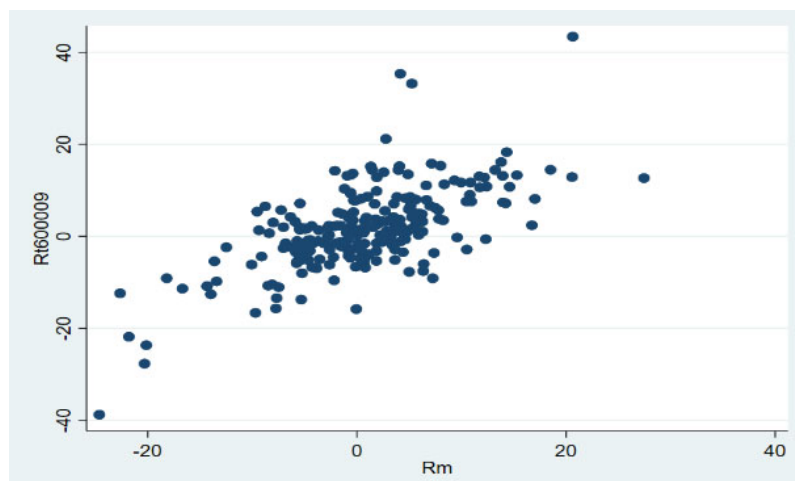
is 0.622. The maximum and minimum values of Stock 600009 are 43.41 and -38.79, respectively, and for market return, the maximum and minimum values are 27.45 and -24.63. The standard deviation of the two variables are 8.937 and 7.547.



**Figure 1.** Distribution Histogram of Stock (600009) Return Rate

Making the distribution histogram of Stock (600009) return rate in Figure 1, I found that these data do not meet the standard distribution well, and the data dispersion is large. It

shows that the return rate of stock may be random, which may embody the economic theory of random walk [1].



**Figure 2.** Scatter Chart of Single Stock (600009) Return Rate

After making scatter charts of single Stock (600009) return rate and market return rate, according to the distribution of these scatter charts, I can roughly see that there is a positive correlation between single stock return rate and market return rate. In order to further analyze their correlation, I make regression analysis on it.

#### 4. Models Regression

First, Test the formula:

$$R_{it} = \alpha + \beta_1 R_m + \varepsilon$$

Where  $R_{it}$  is the return rate of the Stock 600009.

**Table 2.** Linear Regression of Rt600009 and Rm

Rt600009	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Rm	0.771	0.058	13.19	0	0.656	0.886	***
Constant	1.095	0.442	2.48	0.014	0.225	1.966	**
Mean dependent var		1.575		SD dependent var		8.956	
R-squared		0.422		Number of obs		240.000	
F-test		173.914		Prob > F		0.000	
Akaike crit. (AIC)		1604.731		Bayesian crit. (BIC)		1611.693	

\*\*\* p<.01, \*\* p<.05, \* p<.1

Data source: official website of Bank of China and Shanghai Stock Exchange

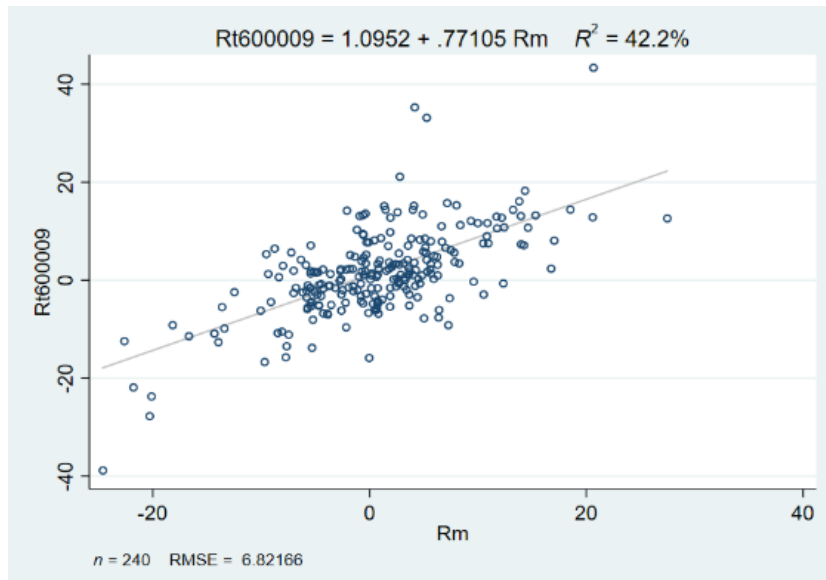


Figure 3. Linear Regression Diagram 1

There is a positive correlation between stock interest rate and market interest rate, and the regression coefficient is close to 1. From the perspective of t-value and p-value, the importance of regression coefficient is very high and cannot be ignored.

I make the null hypothesis that the value of coefficient is 1 and the alternative hypothesis that the value of coefficient is not 1. By using T test, the difference between the estimated value of 0.771 and 1 is divided by its standard deviation 0.058.

$$t_1 = \frac{\widehat{\beta}_1 - 1}{se(\widehat{\beta}_1)} = \frac{0.771 - 1}{0.058} = 3.945$$

$$t_{(0.95,240)}=1.96$$

The result of t-value is 3.945. Under 95% confidence level, when the value of variable reaches 240, the value of t is 1.96. Since 3.945 is greater than 1.96, indicating that it is in the rejection domain, we should reject the null hypothesis.

I make another null hypothesis that the value of coefficient is 0.5 and the alternative hypothesis that the value of coefficient not 0.5. By using T test, the difference between the

estimated value of 0.771 and 0.5 is divided by its standard deviation 0.058.

$$t_1 = \frac{\widehat{\beta}_1 - 0.5}{se(\widehat{\beta}_1)} = \frac{0.771 - 0.5}{0.058} = 4.672$$

$$t_{(0.95,240)}=1.96$$

The result of t-value is 4.672. Under 95% confidence level, when the value of variable reaches 240, the value of t is 1.96. Since 4.672 is greater than 1.96, indicating that it is in the rejection domain, we should reject the null hypothesis.

Compared with the previous hypothesis test, the value of regression coefficient is closer to 1 than 0.5, and the synchronization between market return rate and stock return rate is very high.

Then, I test the formula:

$$R_{it} = \alpha + \beta_2 R_f + \varepsilon$$

Table 3. Linear Regression of Rt600009 and Rf

Rt600009	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Rf	-2.157	0.935	-2.31	0.022	-3.998	-0.316	**
Constant	5.83	1.93	3.02	0.003	2.027	9.633	***
Mean dependent var		1.575		SD dependent var		8.956	
R-squared		0.022		Number of obs		240.000	
F-test		5.328		Prob > F		0.022	
Akaike crit. (AIC)		1731.069		Bayesian crit. (BIC)		1738.030	

\*\*\* p<.01, \*\* p<.05, \* p<.1

Data source: official website of Bank of China and Shanghai Stock Exchange

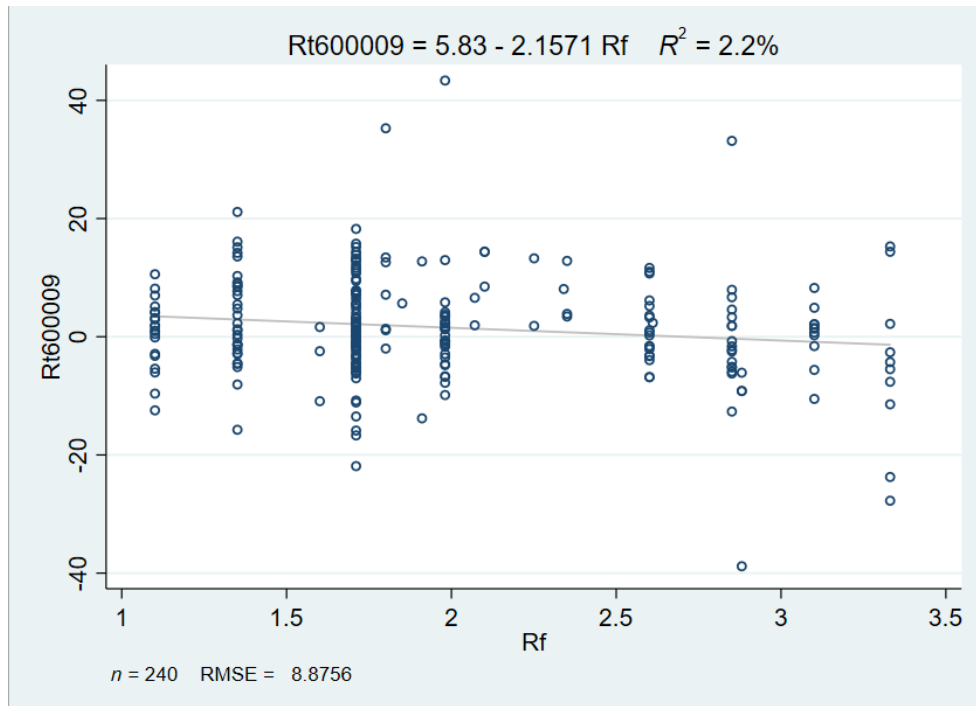


Figure 4. Linear Regression Diagram 2

From Figure 4, there is a negative correlation between stock return and risk-free interest rate. The regression coefficient is about -2.

I make another null hypothesis that the value of coefficient is -1 and the alternative hypothesis that the value of coefficient is not -1. By using T test, the difference between the estimated value of -2.15 and -1 is divided by its standard deviation 0.935.

$$t_2 = \frac{\widehat{\beta}_2 - 0.5}{se(\widehat{\beta}_2)} = \frac{-2.15 - (-1)}{0.935} = -1.236$$

$$t_{(0.95,240)}=1.96$$

The value of regression coefficient of Rf is -2.15 and its standard deviation of 0.935. The t-value calculated by hypothesis test is -1.236, which is more than the t-value of -1.96 under the confidence interval of 95%, indicating that it is outside the rejection domain. The null hypothesis should be accepted, that is, the regression coefficient is -1, showing that the value of regression coefficient of Rf is closed to -1.

From the point of view of p-value, p-value is low, close to

0.022, less than 0.05, indicating that the importance level of regression coefficient is high. However, if there are only variables:  $R_t$  and  $R_f$ , the fitting degree of their regression is low. Thus, this univariate regression cannot reflect their relationship well, indicating that there may be other variables that have a higher degree of influence.

Further discussion:

Multiple regression:

$$R_{it} = \alpha + \beta_1 R_m + \beta_2 R_f + \beta_3 R_f^2 + \beta_4 R_{jt} + \varepsilon$$

In the new model, try to add the square term of risk-free interest rate to further observe the correlation. At the same time, the return rate of other stock such as Stock (600897) is introduced as the influencing factor, which is  $R_{jt}$ .

From the previous regression results, the importance of market return rate is high, and it is positively correlated with the return rate of stock. However, there is a negative correlation between risk-free interest rate and stock return rate, but its importance level is low. So, In the new model, I will further test these variables.

After multiple regression:

Table 4. Multiple Regression 1

Rt600009	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Rm	0.774	0.06	12.83	0	0.655	0.893	***
Rf	-1.197	0.719	-1.67	0.378	-2.613	0.219	*
Rf2	-0.186	1.143	-0.56	0.694	-0.023	0.038	*
Rt600897	0.589	0.045	2.09	0.029	-0.099	0.078	*
Constant	3.515	1.49	2.36	0.019	0.58	6.451	**
Mean dependent var		1.584		SD dependent var		8.973	
R-squared		0.435		Number of obs		239.000	
F-test		60.423		Prob > F		0.000	
Akaike crit. (AIC)		1597.455		Bayesian crit. (BIC)		1611.360	

\*\*\* p<.01, \*\* p<.05, \* p<.1

Data source: official website of Bank of China and Shanghai Stock Exchange

Stock returns are positively correlated with market returns, but negatively correlated with risk-free interest rates. Judging from the value of  $t$  and  $p$ , the regression coefficient of market return has the highest importance level.

On the other hand, the returns of other stock also have influence on the returns of the stock that studied, from the view of its value of  $t$  and value of  $p$ . The importance level of regression coefficient of Stock 600897 is not low, so it cannot be ignored.

The possible economic explanation I think is that because

Stock 600009 and Stock 600897 are both airport stocks, and they will be affected by the same market conditions, so they have a positive relationship. If stocks with conflicting types are selected, the possible relationship coefficient between them can be negative.

I also find that the value of regression coefficient of risk-free interest rate and the square of risk-free interest rate are both not very important. Does that mean that the impact of risk-free interest rate on stock return rate can be ignored?

Then I do the F-test:

**Table 5.** Residual Analysis under Unconstrained and Constrained Circumstances

Source	SS	df	MS
Model	8093.112	1	8093.11158
Residual	11275.34	238	46.535028
Total	19168.45	239	80.2027123
Source	SS	df	MS
Model	8344.987	3	2781.662
Residual	10818.62	235	46.03669
Total	19163.61	238	80.51937

Data source: official website of Bank of China and Shanghai Stock Exchange

When I take out the two variables  $R_f$  and  $R_f^2$ , this is constrained situation. When the two variables are still in the model, this is the unconstrained situation.

We assume that the coefficient of risk-free interest rate and the return rate of other stocks in the above formula are zero, and the sum of error squares obtained by unconstrained model and constrained model is obtained respectively.

$$F = \frac{(SSE_R - SSE_U)/J}{SSE_U/(N - K)} = \frac{(11275 - 10818)/2}{10818/(240 - 3)} = 5.07$$

$$F_{(2,237)}=3.02$$

Since  $5.07 > 3.02$ , We reject the original hypothesis. We can draw the conclusion that the risk-free interest rate does have an impact on the stock yield, and this impact cannot be ignored.

I want to further test whether there is multicollinearity between variables, and I test variance expansion factor:

**Table 6.** VIF Test

Variables	VIF	1/VIF
$R_f^2$	53.930	0.019
$R_f$	53.440	0.019
$R_m$	1.080	0.928
$R_{t600897}$	3.796	0.263
Mean VIF	36.150	

From Table 6, I find that the variance expansion factor is very large, which shows that there is multicollinearity. There is serious multicollinearity between  $R_f$  and  $R_f^2$ .

More accurate analysis

Some analysis based on CAPM:

$$\text{CAPM: } E(r_i) = \beta_i [E(r_m) - r_f] + r_f$$

CAPM (Capital Asset Pricing Model), developed based on Markowitz's modern portfolio selection theory, assumes that each investor uses a series of portfolios to obtain the finite boundary including all risky assets based on the maximization principle of Markowitz. It obtains the asset allocation line through the tangent of risk-free return, and then develops into the securities market line, which is used to price a single asset or portfolio. However, Gao Yan (2018) said "This model is not fully applicable in the Chinese market" [2]. Drew, Naughton and Veeraraghavan tested the multifactor model on Shanghai Stock Exchange market factor alone is not sufficient to describe the cross-section of average stock returns in China [3]. Based on CAPM, I have further processed the data in order to analyze the more exact

correlation among the variables.

Where  $E(r_i)$  is the expected return of Stock 600009,  $r_f$  is the risk-free rate,  $E(r_m)$  is the market return and  $\beta$  measures the stock's sensitivity to market risk.

According to the previous analysis, the correlation of market returns is high, but the correlation of risk-free interest rates is not high, while the returns of other stocks have an impact on the stocks studied. Therefore, in order to study the coefficient better, I simply deform and change the formula as:

$$R_{it} - R_f = \alpha + \beta_1 (R_m - R_f) + \beta_2 R_{jt} + \varepsilon$$

$R_{it}$  is the return on Stock 600009.  $R_{jt}$  is the return rate of other stocks, in this case, is Stock 600897.  $R_f$  is the risk-free rate that is often measured by the three-month time deposit rate.  $R_m$  is the market return.  $\alpha$  is the intercept term.  $\varepsilon_i$  is the error.

$R_{it} - R_f$  is the excess return of Stock 600009,  $R_{mt} - R_f$  is the excess return of Shanghai Stock Market.

**Table 7.** Multiple Regression 2

ER of the stock 600009	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ER of the Shanghai Stock Market	1.122	0.125	13.50	0	0.668	0.896	***
Rt600897	0.467	0.674	1.05	0.078	-0.098	0.092	*
Constant	0.659	0.448	1.47	0.143	-0.225	1.542	
Mean dependent var		-0.398			SD dependent var	9.067	
R-squared		0.687			Number of obs	240.000	
Adjusted R Square		0.685					
F-test		182.232			Prob > F	0.000	
Akaike crit. (AIC)		1605.861			Bayesian crit. (BIC)	1612.822	

\*\*\* p<.01, \*\* p<.05, \* p<.1

Data source: official website of Bank of China and Shanghai Stock Exchange.

**Table 8.** Residual Statistics

	df	SS	MS	F	Significance F
Regression analysis	1	8520.186	8520.186	182.2321118	3.23396E-31
Residual	238	11127.59	46.75458		
Total	239	19647.78			

Data source: official website of Bank of China and Shanghai Stock Exchange

From Table 7&8, according to the regression analysis, from the tables above, the regression coefficient is 1.122. The standard deviation of coefficient of regression is only 0.125, which is relatively small.

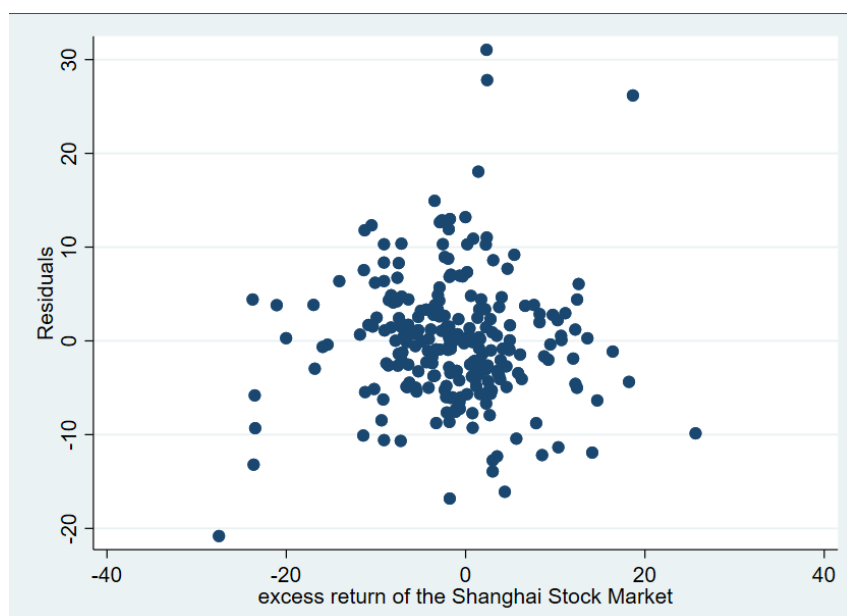
I carried out relevant hypothesis testing to explore the importance level of regression coefficient. The null hypothesis is that the value of regression coefficient is 0, and the alternative hypothesis is that the value of regression coefficient is not 0. After carrying out relevant hypothesis testing, I focused on the test results of t-value and p-value.

The value of t reaches 13.5, which is far greater than the value of 1.96 at 95% confidence level, indicating that the value of t is within rejection domain. Therefore, the null hypothesis that the value of regression coefficient is 0 should be rejected. In addition, from the point of view of p-value, at

95% confidence level, p-value is very small, which is less than 0.05, indicating that regression coefficient is of high importance.

From the length of the confidence interval, 0.668 to 0.896 is a relatively small range, and the data concentration is relatively high at 95% confidence level. However, from the residual diagram, the residual value is 46.754, and from the value of R-square, the R-square is 0.687, the modified R-square is 0.685, which is higher than the models that tested before, which shows that the fitting degree of this model is higher.

Next, I analyze the residual error of the Rm in order to see if there is heteroscedasticity in the model:

**Figure 5.** Residual Plot

From Figure 5, the greater the market excess return rate, the variance of the disturbance term has no obvious change,

which indicates that there is no heteroscedasticity.

I use Breusch-Pagan-Godfrey test to further analyze

heteroscedasticity:

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: excessR<sub>m</sub>

chi2(1) = 0.52

Prob > chi2 = 0.4705

The equation has been estimated:

$$\begin{aligned} ExcessR_{t600009} &= 0.659 + 1.122 ExcessR_m + 0.467 Rt_{600897} \\ &\quad (0.488) \quad (0.125) \quad (0.674) \\ R^2 &= 0.687; \quad n = 240 \end{aligned}$$

The value of p is greater than 0.05, indicating that there is no heteroscedasticity, which shows that there is no obvious correlation between the variance of the random error term and the observed value of the explanatory variable.

Then I do the hypothesis of the regression coefficient of excessR<sub>m</sub>.

$$t_1 = \frac{\widehat{\beta}_1 - 1}{se(\widehat{\beta}_1)} = \frac{1.122 - 1}{0.125} = 1.76$$

$t_{(0.95,240)}=1.96$

The value of regression coefficient of excessR<sub>m</sub> is 1.122 and its standard deviation of 0.125. The t-value calculated by hypothesis test is 1.76, which is less than the t-value of 1.96 under the confidence interval of 95%, indicating that it is outside the rejection domain. The null hypothesis should be accepted, that is, the regression coefficient is 1.

Compared with the hypothesis test results of previous models, the correlation coefficient between stock excess return and market excess return is closer to 1, with higher positive correlation and better fitting results.

## 5. Conclusion

According to the conclusion of the above analysis of different models, in the stock market, the return rate of stock is positively correlated with the market return rate, and negatively correlated with the risk-free interest rate. But the fitting degree of this model alone is not high. After the variables are properly processed, the regression coefficient obtained by using stock excess return rate and market excess return rate as variables has better fitting effect.

On the other hand, I have found that the returns of other stocks will have a certain impact on the returns of the stocks analyzed, especially the stocks of the same type, which may have positive correlation characteristics because their reactions to the market are in the same direction. When designing the model, it is best to add the return rate of these other stocks.

The risk-free interest rate alone has a low impact on stock returns, but it cannot be ignored. China's stock and securities trading market is still immature, and there will be other factors affecting stock returns. After adding other stock return variables, some stocks with strong correlation are related and will influence each other. When we buy a target stock, we need to consider these factors such as the influence of other stocks on it.

Since Chinese market is still immature, CAPM may not be fully applicable in Chinese market. Compared with the capital markets of Western developed countries, the mechanism of China's stock market is not mature enough, and there are some

phenomena such as imperfect information management. At the same time, information transparency is not high, and there is information asymmetry among participants in the market, which affects the operational efficiency of the market. Moreover, system risk is not the only factor that determines returns, and stock returns are also closely related to the size of the company's share capital, future development decisions, mergers and acquisitions and other factors. These factors will also reduce the explanatory power of CAPM. Therefore, we need to have a deeper understanding of the Chinese market and consider more influencing factors in order to make the analysis more comprehensive.

## References

- [1] Fama, Eugene. F. (1970) "Efficient capital markets: a review of theory and empirical work" *Journal of Finance*, vol.25, pp.383-417.
- [2] Gao Yan, Yang Xin (2018) "A Study on the Relationship Between CAPM and China Stock Market".
- [3] Drew, M. E., Naughton, T. and Veeraraghavan, M. (2003) "Firm size, book-to-market equity and security returns: evidence from the shanghai stock exchange" *Australian Journal of management*, vol. 28:2, September.

Stata Code

```
asdoc summarize Rt600009 Rf Rm excessRt excessRm
```

```
scatter Rt600009 Rm
```

```
twoway scatter excessRt excessRm
```

```
twoway function y=normalden(Rt600009), range(-30 40)
```

```
asdoc reg excessRt excessRm
```

```
aaplot excessRt excessRm
```

```
aaplot Rt600009 Rf
```

```
scalar t1=(_b[excessRm]-0/-se[excessRm])
```

```
scalar p1=2*ttail(72,abs(t1)) scalar list t1 p1
```

```
regress Rt600009 Rm Rf Rm-Rf
```

```
asdoc reg Rt600009 Rm Rf Rt600897
```

```
testparm Rm Rf
```

```
(1) Rm = 0
```

```
(2) Rf = 0
```

```
F (2, 237) = 89.02
```

```
Prob > F = 0.0000
```

```
. test (_b[Rm]=1) (_b[Rf]=0)
```

```
(1) Rm = 1
```

```
(2) Rf = 0
```

```
F (2, 237) = 9.13
```

```
Prob > F = 0.0002
```

```
scalar sser=e(rss) scalar df_rest=e(df_r)
```

```
scalar J=df_rest-df_unrest
```

```
asdoc estat vif
```

```
swilk Rt600009 Rm Rf
```

```
asdoc estat hettest,rhs iid
```

(File Myfile.doc already exists, option append was assumed)

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance  
Variables: excessRm

chi2(1) = 0.52  
Prob > chi2 = 0.4705