Capital Asset Pricing Model (CAPM) and its Realization in Stata

Zhaokun Zhang*

March 30, 2019

Abstract

It's well known to us all that CAPM model is relatively important in the field of finance. It's the centerpiece of financial economics. In this article I will give a brief introduction to CAPM model containing its definition, assumptions and formulas. Then I will use data from stock exchanges to simulate and realize it with the help of Stata, which may help you better understand this model.

1 CAPM Model

1.1 Definition

In finance, the capital asset pricing model (CAPM) is a model used to determine a theoretically appropriate required rate of return of an asset, to make decisions about adding assets to a well-diversified portfolio[1]. It mainly describes the relationship between systematic risk (also known as non-diversifiable risk) and expected return for some assets, particularly for stocks.

In brief, CAPM is the centerpiece of financial economics, which relates asset risk to expected return. It helps to identify those mis-priced securities, as well as helps firms and individuals to make investment decisions.

1.2 Assumption

Since CAPM is a theoretical model, it must be satisfied with some assumptions when applying. It can be divided into three parts:

1. There exist enough investors in the market, that is, individual wealth is much less than the collective wealth. Then, each investors are price-takers and the prices are unaffected by trade;
2. All relevant assets are publicly traded, and untradables like human capital don't matter. People borrow or lend any sum at risk-free rate and don't need to pay taxes and transaction costs;
3. As for investors themselves, they are Mean-Variance optimizers and use Markowitz portfolio selection. Besides, they also have "homogeneous expectations".

1.3 Formula

The CAPM is a model for pricing an individual security or portfolio. The ideal formula is as follows:

$$E(R_i) = r_f + \beta_i \left[E(R_M) - r_f\right]$$

*2016 WISE IUEC, Student ID: 15220162202515

[1] Wikipedia,https://en.wikipedia.org/wiki/Capital_asset_pricing_model#Overview

where $E(R_i)$ is the expected return of capital asset i, r_f is the risk-free rate of interest in the market, $E(R_M)$ is the expected return of market, and β_i is the sensitivity of the expected excess asset returns to the expected excess market returns.

However, in the market there exists some stock performing better than the prediction based on CAPM model, and there are some stocks performing worse than prediction. Thus, we define $\alpha_i$as the excess return for each stock compared to the predicted return. Now the actual formulas is :

$$E(R_i) = \alpha_i + r_f + \beta_i \left[E(R_M) - r_f \right]$$

When $\alpha_i > 0$, the realized return is higher than expected return based on CAPM, meaning now the price is too low, and you should buy it; when $\alpha_i < 0$, the realized return is lower than expected return, meaning now the price is too high, and you should sell it in the market; and when $\alpha_i = 0$, the realized return is equal to the expected return, showing the price is fair.

2 Empirical Analysis

2.1 Data and Variables

Now in order to have a further and better understanding of CAPM model, I want to use some data in stock exchanges to simulate it. Firstly, as CSI300 index contains about 300 different stocks, we can regard it as the market index. Then I choose about 10 stocks randomly from index, whose codes are 603259, 002925, 601066, 600588, 601888, 002032, 600436, 002007, 002153, 603288. Aftering selecting stocks, I search for the settlement price and yield rate. The data chosen is recorded in days from 2018/06/20 to 2019/03/08. By using Stata, we can summarize some characteristics of these 10 stocks and CSI300 index:

Variable	Obs	Mean	Std. Dev.	Min	Max
S603259	175	.0002171	.0329994	-.0951	.1
S002925	175	.0011223	.0354487	-.1	.1001
S601066	175	.0108623	.0559869	-.1003	.4391
S600588	175	.0036017	.0343257	-.1001	.1
S601888	175	-.0000811	.0302209	-.1	.0999
S002032	175	.0011697	.0255023	-.0761	.1
S600436	175	.0000246	.0265125	-.0964	.0769
S002007	175	.0014566	.0307878	-.1001	.1
S002153	175	.0021869	.0316162	-.1001	.0891
S603288	175	.0002091	.0250935	-.1	.1
CSI300	175	.0001754	.0153643	-.048	.0595

Figure 1: Summary statistics of variables

Apart from the expected return, what we also need is the risk-free rate of interest. We've already known that the 1-year treasury bill can be used to represent the risk-free rate, r_f. So the interest rate per year is

2.447% with annual compounding[2]. But the expected return is in days, then we should use the equation to calculate the risk-free rate per day:

$$1 + EAR = [(1 + r_f(T)]^{\frac{1}{T}}$$

that is,

$$1 + 2.447\% = \left[(1 + r_f(\frac{1}{365})\right]^{365}$$

$$r_f(\frac{1}{365}) = 0.006624\%$$

Then the risk-free rate per day is 0.006624%.

2.2 Regression and Hypothetical test

With the risk-free rate and expected return, we can use CAPM model to calculate α_i and β_i. Before that we should get the risk premium for every stock and CSI300 index–$E(R_i) - r_f$, referred to RS in our data set. Then assume the formula is:

$$RS(R_i) = \alpha_i + \beta_i RS(R_M) + e$$

Then the results are as follows:

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10
VARIABLES	RCSI300	RCSI300	RCSI300	RCSI300	RCSI300	RCSI300	RCSI300	RCSI300	RCSI300	RCSI300
RS603259	0.296***									
	(0.0317)									
RS002925		0.216***								
		(0.0354)								
RS601066			0.118***							
			(0.0417)							
RS600588				0.243***						
				(0.0389)						
RS601888					0.298***					
					(0.0376)					
RS002032						0.347***				
						(0.0437)				
RS600436							0.389***			
							(0.0368)			
RS002007								0.234***		
								(0.0428)		
RS002153									0.261***	
									(0.0404)	
RS603288										0.403***
										(0.0447)
Constant	0.000065	-0.000118	-0.00116	-0.000748	0.000153	-0.000273	0.000125	-0.000216	-0.000444	0.000052
	(0.000898)	(0.00101)	(0.00102)	(0.000964)	(0.000944)	(0.000953)	(0.000864)	(0.00102)	(0.000970)	(0.000875)
Observations	175	175	175	175	175	175	175	175	175	175
R-squared	0.405	0.247	0.185	0.294	0.342	0.331	0.450	0.220	0.288	0.434
Robust standard errors in parentheses										
*** p<0.01, ** p<0.05, * p<0.1										

Figure 2: Regression results of 10 stocks

[2]China Bond,http://yield.chinabond.com.cn/cbweb-mn/yield_main?locale=zh_CN

The α and β for each stock is:

	RS603259	RS002925	RS601066	RS600588	RS601888	RS002032	RS600436	RS002007	RS002153	RS603288
beta	0.2962969	0.2155398	0.1179068	0.2425427	0.2975034	0.3465664	0.388845	0.2338423	0.2606348	0.4034934
alpha	0.000065	-0.000118	-0.001164	-0.000748	0.000153	-0.000273	0.0001254	-0.000216	-0.000444	0.0000515

Figure 3: The α and β for each stock

Draw the scatter plot between α and β by Stata:

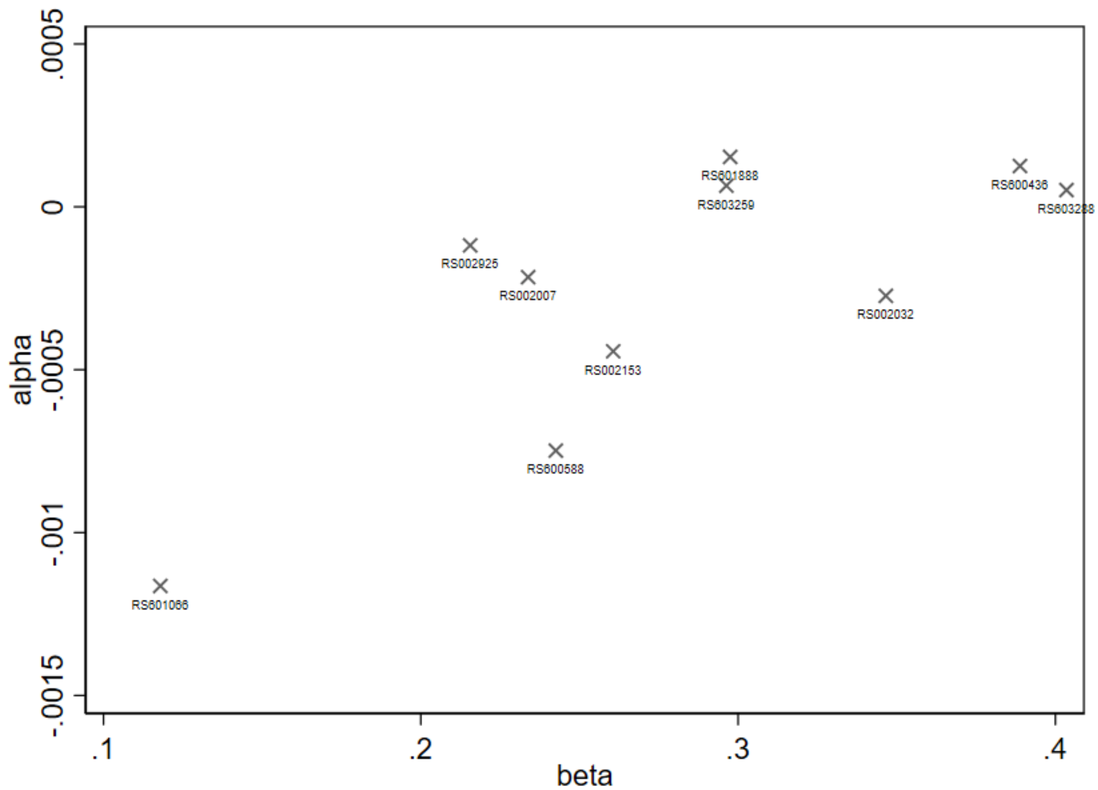

Figure 4: Scatter plot of α and β

Then I randomly choose one stock(S002153) to test the hypothesis. Firstly we want to know whether S002153 has an "alpha" over the market, that is,

$$\mathbb{H}_0 : \alpha = 0 \; VS \; \mathbb{H}_1 : \alpha \neq 0$$

Using the test of coefficients, we get:

```
. test _cons ==0

 ( 1)  _cons = 0

       F(  1,   173) =    0.21
            Prob > F =    0.6481
```

Figure 5: Test of α

Thus, the coefficient is NOT statistically significant, then we cannot reject the null hypothesis and draw a conclusion that the stock 002153 doesn't have an "alpha" over the market.
Furthermore, we also want to test beta:

$$\mathbb{H}_0 : \beta = 1 \ VS \ \mathbb{H}_1 : \beta \neq 1$$

where $\beta = 1$ means that S002153 is just the market and $\beta \neq 1$ means that S002153 softens or exaggerates market moves. The result is as follows:

```
. test RS002153 ==1

 ( 1)  RS002153 = 1

       F(  1,   173) =  335.21
            Prob > F =    0.0000
```

Figure 6: Test of β

The figure above shows that it's statistically significant and then we reject the null hypothesis. So the stock 002153 is impossible to be the same as market.

References

[1] Wikipedia, "Capital asset pricing model", https://en.wikipedia.org/wiki/Capital_asset_pricing_model

[2] Bodie, Z.; Kane, A.; Marcus, A.J.(2017). "Investments"(10th Edition). Global Education

[3] Jiaming Mao, "Regression", https://jiamingmao.github.io/data-analysis/assets/Lectures/Regression.pdf

Appendix

The Stata code is contained in the following do file:

```
// The data is the expected return of 10 stocks in CSI300 from 2018/06/20 to 2019/03/08
// And also include the expected return of CSI300, which referrs to market portfolio
use "C:\Users\Acezh\Desktop\CAPM.dta"
// Firstly summarize them
summarize S603259 S002925 S601066 S600588 S601888 S002032 S600436 S002007 S002153 S603288 CSI300, sep
// Now we know the risk-free rate is 0.006624% per day
gen r = 0.00006624
// Calculate the risk premium for 10 stocks and market
gen RS603259 = S603259-r
gen RS002925 = S002925-r
gen RS601066 = S601066-r
gen RS600588 = S600588-r
gen RS601888 = S601888-r
gen RS002032 = S002032-r
gen RS600436 = S600436-r
gen RS002007 = S002007-r
gen RS002153 = S002153-r
gen RS603288 = S603288-r
gen RCSI300  = CSI300 -r
// Through regression to calculate alpha and beta
reg RCSI300 RS601066 ,r
est store m1
reg RCSI300 RS002925,r
est store m2
reg RCSI300 RS601066 ,r
est store m3
reg RCSI300 RS600588 ,r
est store m4
reg RCSI300 RS601888 ,r
est store m5
reg RCSI300 RS002032 ,r
est store m6
reg RCSI300 RS600436 ,r
est store m7
reg RCSI300 RS002007 ,r
est store m8
reg RCSI300 RS002153 ,r
est store m9
reg RCSI300 RS603288 ,r
est store m10
// Then combine them together
outreg2 [m1 m2 m3 m4 m5 m6 m7 m8 m9 m10] using capm.xls,replace
// Draw a scatter plot to show the relationship between alpha and beta
twoway (scatter alpha beta, mlabel(var1)),scheme(s1mono)
// Test whether the coefficient of RS002153 is 1
// That is, test whether RS002153 is the market
reg RCSI300 RS002153 ,r
test RS002153 ==1
// Test whether RS002153 have an "alpha" over the market
test _cons ==0
```