Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/55/20230967

An Empirical Test Based on the Validity of the Capital Asset Pricing Model of American Firms

Binghuan Li[1,a,*]

[1]*School of Finance, Zhongnan University of Economics and Law, Wuhan, 430073, China*
a. libinghuan@stu.zuel.edu.cn
**corresponding author*

Abstract: Nowadays, quantitative finance has become a more and more mainstream research direction, and the origin model of quantitative finance is the capital asset pricing model (CAPM)created by William Sharp and others. This paper mainly studies the feasibility of CAPM model for the American market, in order to prove the universality and accuracy of CAPM model, and analyzes the relevant errors. The research object of this paper is 59 randomly selected listed companies in 11 industries in S&P500 index, and the regression is carried out by the ordinary least square method, and the goodness of fit is obtained, so as to prove its universality. The data comes from Yahoo Finance. Through data analysis, this paper believes that CAPM model is universal, but it has low adaptability for some special industries, so it is necessary to introduce more variable factors or carry out non-linear regression to improve the accuracy of its prediction.

Keywords: Capital Asset Pricing Model (CAPM), Ordinary Least Squares (OLS), linear regression

1. Introduction

CAPM model has attracted wide attention since its launch. Some of the empirical analysis results are considered effective, while others are considered invalid, among which some are considered effective. Black et al. selected the stocks listed on the New York Stock Exchange from 1931 to 1956 as the research object for empirical research, and the results showed that the expected return rate of high-risk return stocks was not high [1]. Low-risk stocks have higher returns; Banz proposed the small-company effect and found that companies with smaller market capitalization had higher returns on their portfolios [2]. Fama et al. conducted a cross-sectional inspection analysis of CAPM model and concluded that CAPM model may not be valid. Roll offers a Roll critique, pointing out that a fully efficient portfolio is essentially impossible to achieve in a real-world capitalist market [3].

This paper studies the effectiveness of Capital Asset Pricing Model (CAPM) in different industries and companies in the United States. Fifty-nine companies from 11 industries in S&P500 were randomly selected for research, including information technology, health care, consumer discretionary, finance, communication services, industrial, consumer staples, energy, utilities, real estate, and materials. Data from January 1, 2018 to August 1, 2023 were selected, with monthly steps. The regression function and Multiple R-squared are obtained by using the ordinary least square method (OLS) for linear regression of the data. The effectiveness of CAPM model in the American stock market is discussed, and the accuracy of predicting future returns is also discussed. The

DOI: 10.54254/2754-1169/55/20230967

significance of this paper is to verify the suitability of the CAPM model and find the defects of the CAPM model in the research process. In the subsequent research, the model can be further improved, such as increasing variable factors and relaxing hypothesis conditions, so as to make it applicable to more situations.

2. Theoretical Model

CAPM model is a forecasting model based on the expected return equilibrium of risk assets. Its theoretical origin is the mean-variance theory proposed by Harry Markowitz in 1952. Then, in 1964, Sharp et al. proposed the CAPM model on this basis, which provided the asset pricing model in theory for the first time, and presented the theoretical relationship between expected return and expected risk in a linear manner. The mathematical expression of the CAPM model is as follows:

$$E(Ri) = Rf + \beta i(E(Rm) - Rf) \quad (1)$$

Where, E (Ri) is the expected rate of return of a single asset or portfolio, Rf is the risk-free rate of return, E (Rm) is the expected rate of return of a market portfolio, and βi is the systematic risk coefficient of an asset, which describes the sensitivity of the rate of return of an asset to market changes and can be used to represent the systemic risk or non-diversifiable risk of a single asset or portfolio. The β coefficient reflects the sensitivity of an asset to changes in market price and classifies different assets by the difference of β values. The classification method is as follows: when $\beta<1$, the volatility of the asset is less than the volatility of the market price, and investment in the asset can not get a good return but can help investors reduce the probability of loss; When $\beta>1$, the price volatility of the asset is greater than that of the market, then the asset will face a greater loss than the market. When $\beta=1$, it indicates that the rate of return obtained by the purchase of the asset is the same as the market rate of return, that is, the same loss as the market.

3. Methodology

3.1. Data selection and Preprocessing

3.1.1. Sample Selection

Table 1: Number of codes and enterprises by industry.

Industry Code	Industry Involved	Number of Enterprises
A	Information Technology	5
B	Health Care	5
C	Consumer Discretionary	5
D	Finance	6
E	Communication Services	5
F	Industrial	8
G	Consumer Staples	6
H	Energy	5
I	Utilities	5
J	Real Estate	5
K	Materials	4

The data in Table 1 are all from Yahoo Finance website, and the monthly closing price data of 59 listed companies in 11 industries of S&P 500 index is selected as the research object. Due to the widespread global impact of the epidemic, the data selection range is from January 1, 2018 to August 1, 2023, with a total of 4071 observations. Sample stocks were selected for A information technology, B health care, C consumer discretionary, D finance, E communication services, F industrial, G Consumer staples, H energy, I utilities, J real estate, and H materials. The company classification and company code are shown in Table 2.

Table 2: Information about the selected business.

Industry Involved	Stock Code				
A	AAPL	MA	MSFT	NVDA	V
B	ABBV	JNJ	LLY	PFE	UNH
C	AMZN	MCD	NKE	SBUX	TSLA
D	BAC	BRka	GS	JPM	MS
	WFC				
E	ATVI	DIS	GOOG	META	NFLX
F	BA	CAT	DE	HON	
	LMT	RTX	UNP	UPS	
G	COST	KO	PEP	PG	PM
	WMT				
H	COP	CVX	EOG	PXD	XOM
I	D	DUK	NEE	SO	SRE
J	AMT	CCI	EQIX	PLD	PSA
K	APD	CTVA	LIN	SHW	

3.1.2. Risk-Free Asset Selection

A risk-free asset is a bond that has no default risk, no inflation risk, no liquidity risk, no interest rate risk, and no arbitrary other risk. U.S. Treasury bills are often used as a proxy for the risk-free yield, Rf. In this article, the US 10-year bond is chosen as the risk-free bond, because the US dollar is the anchor of the global currency, and the US 10-year Treasury bond yield is the global risk-free interest rate indicator. By comparing this indicator with the earnings yield of the stock market (the inverse of the P/E ratio), we can calculate whether we should invest in the stock market or the bond market at present. According to the data and calculation, the average monthly yield of risk-free bonds is 1.76697%, and the average annual yield is 23.4304%.

3.1.3. Total Market Return Selection

In the US market, choose the total market return, choose the S&P 500 index as the total market return, because this index contains more companies and involves more industries, so it can better reflect the changes in the overall market. According to the data and calculation, the average monthly yield of risk-free bonds is 0.8937%, and the average annual yield is 11.2673%.

3.2. Test Method and Empirical Results

3.2.1. Test Method

In terms of data processing, the data will be processed using the ordinary least square method, the basic principle of which is to minimize the sum of distances from the fitted line to the actual point, even if the residual sum of squares is minimized. According to the CAPM model expression, investor

DOI: 10.54254/2754-1169/55/20230967

returns come from two parts. One part comes from unexpected returns, which is called excess returns. The other part comes from the expected return, that is, the market risk compensation return. For a given portfolio, if the expected return of the portfolio and the expected return of the market portfolio are known, the CAPM model can be tested by analyzing whether there is a linear relationship between the expected return and the β coefficient. Let the excess return of stock i be Zi, and the excess return of market portfolio be Zm, whose mathematical expressions are as follows:

$$\begin{cases} Zi = \alpha i + \beta i \times Zm + \varepsilon i \\ Cov[Zm, \varepsilon i] = 0 \end{cases} \tag{2}$$

In the expression, Zi is the explained variable, Zm is the explaining variable, β is the parameter to be estimated, αi is the intercept term, and random disturbance term.

According to Sharp's paper, the beta value can be expressed as[4]:

$$\beta = \frac{Cov(Rm,Ri)}{Var(Rm)} \tag{3}$$

In this paper, the difference between the return rate of the selected enterprises and the risk-free interest rate (US 10-year Treasury bond) is selected as the explained variable, and the difference between the market portfolio return rate (S&P 500 index) and the risk-free interest rate is selected as the explaining variable.

$$Rit - Rft = \alpha it + \beta i(Rmt - Rft) + \varepsilon i \tag{4}$$

Where Rit stands for the yield of i enterprise at time t, Rmt stands for the yield of S&P 500 Index at time t, and Rft stands for the yield of US 10-year Treasury bond at time t.

3.2.2. Empirical Result

The following data is processed, and the data of 59 companies from January 1, 2018 to August 1, 2023 are selected for fitting by ordinary least square method, and Table 3 is obtained according to the order of goodness of fit from the highest to lowest.

Table 3: Goodness of fit for 59 firms.

Industry involved	Stock Code	Multiple R-squared	Industry involved	Stock Code	Multiple R-squared
A	MSFT	0.9407	I	NEE	0.8547
D	BRKa	0.9364	D	GS	0.8523
G	COST	0.9217	I	D	0.8513
A	V	0.9154	I	DUK	0.8507
K	LIN	0.9138	D	MS	0.8490
F	HON	0.9117	J	PSA	0.8367
J	PLD	0.9085	F	UPS	0.8359
G	PEP	0.9043	I	SO	0.8359
K	SHW	0.9023	F	LMT	0.8236

DOI: 10.54254/2754-1169/55/20230967

Table 3: (continued).

F	UNP	0.9015	F	RTX	0.8140
C	MCD	0.8976	E	DIS	0.8119
E	GOOD	0.8952	G	PM	0.8091
J	EQIX	0.8937	B	PFE	0.8080
J	CCI	0.8887	F	DE	0.7900
A	MA	0.8882	F	CAT	0.7782
K	APD	0.8868	H	CVX	0.7746
G	KO	0.8840	D	WFC	0.7576
B	JNJ	0.8819	B	ABBV	0.7378
J	AMT	0.8780	B	LLY	0.7327
I	SRE	0.8777	E	NFLX	0.7291
D	BAC	0.8773	E	META	0.7128
G	WMT	0.8760	A	NADA	0.7127
D	JPM	0.8725	H	XOM	0.7091
C	NKE	0.8720	E	ATVI	0.7026
A	AAPL	0.8612	F	BA	0.6096
C	SBUX	0.8605	H	PXD	0.6018
G	PG	0.8580	H	COP	0.5513
B	UNH	0.8555	H	EOG	0.5139
C	AMZN	0.8554	C	TSLA	0.4035
K	CTVA	0.8549			

As can be seen from Table 3, the 43 randomly selected companies have different goodness of fit. Now, the companies with the highest goodness of fit in each industry are selected for analysis. The codes of these companies are MSFT, JNJ, MCD, BRKa, GOOD, HON, COST, CVX, SRE, PLD, LIN. A total of 11 companies. The regression results are shown in Table 4.

Table 4: Enterprise sample data regression results.

Stock Code	α	β	t statistic	P-value	Significance F
MAFT	0.0132	1.0652	2.7722	0.0072	3.1987E-42
JNJ	0.0051	0.9119	-0.8836	0.3801	2.4975E-32
MCD	-0.004	0.9496	-0.0677	0.9462	2.2398E-34
BRKa	0.0009	0.9302	0.2202	0.8264	3.2673E-41
GOOG	0.0078	1.0309	1.2488	0.2161	4.9065E-34
HON	0.0033	0.9550	-0.6282	0.5321	1.6794E-36
COST	0.0104	1.0845	1.8459	0.0694	3.2416E-38
CVX	-0.0023	0.8694	-0.2807	0.7798	4.9065E-23
SRE	-0.0036	0.9259	-0.5849	0.5606	7.9864E-32
PLD	0.0039	1.0851	0.6383	0.5255	5.4087E-36
LIN	0.0063	1.0018	1.1511	0.2539	7.4724E-37

DOI: 10.54254/2754-1169/55/20230967

Combined with the t statistic, p-value and Significance F, the equation set is proved to be effective as a whole, and the CAPM model is proved to be effective for these 11 enterprises. FIG. 1 to FIG. 11 show the scatter plot and regression line of the difference between the expected rate of return and the risk-free interest rate and the difference between the S&P 500 index and the risk-free interest rate of each enterprise, respectively.

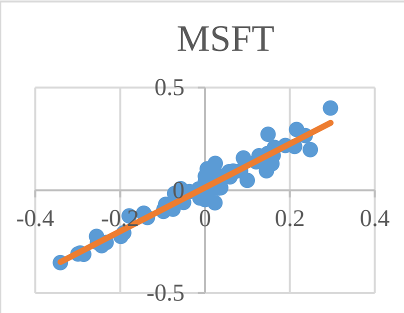

Figure 1: MSFT regression scatter plot.

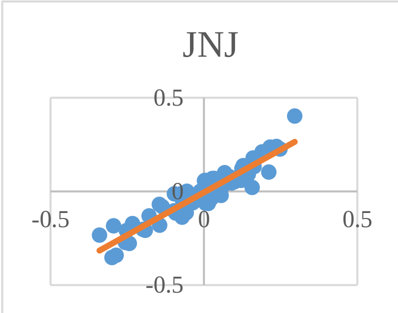

Figure 2: JNJ regression scatter plot.

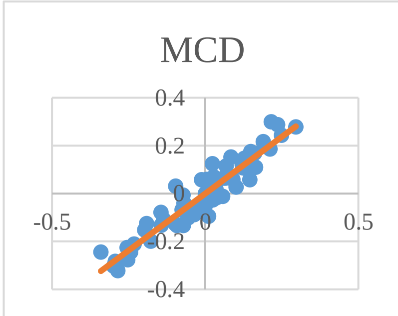

Figure 3: MCD regression scatter plot.

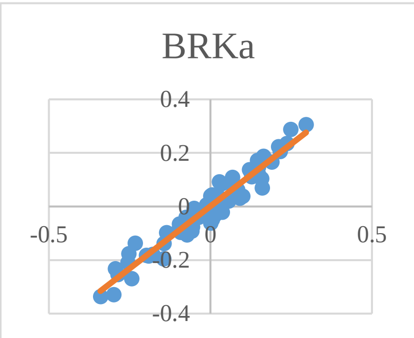

Figure 4: BRKa regression scatter plot.

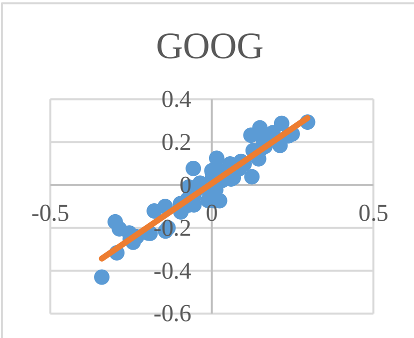

Figure 5: GOOG regression scatter plot.

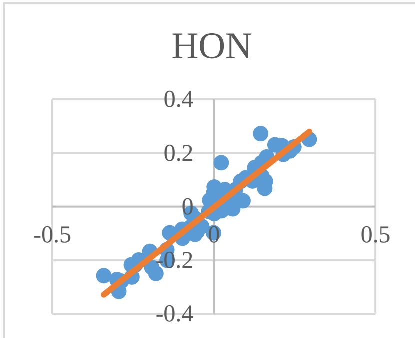

Figure 6: HON regression scatter plot.

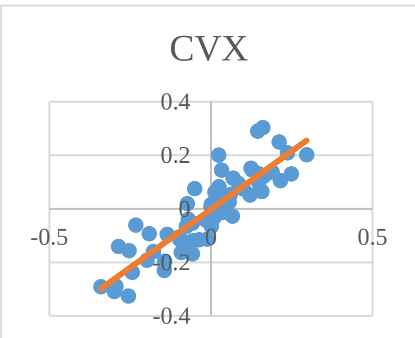

Figure 7: CVX regression scatter plot.

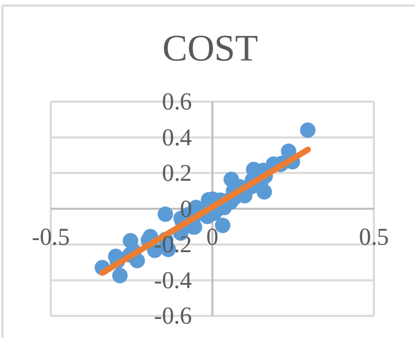

Figure 8: COST regression scatter plot.

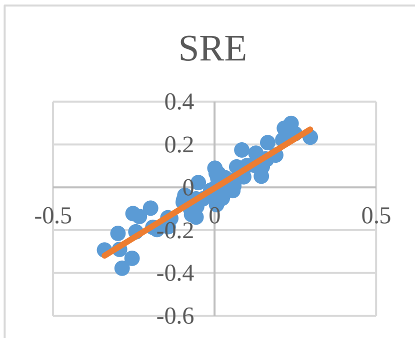

Figure 9: SRE regression scatter plot.

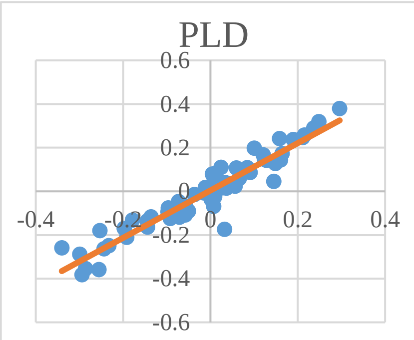

Figure 10: PLD regression scatter plot.

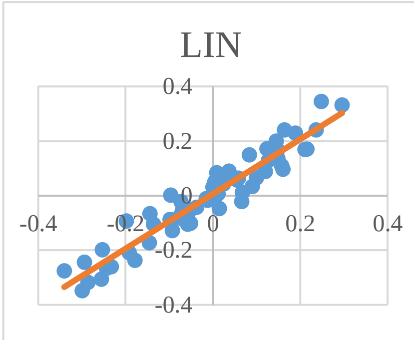

Figure 11: LIN regression scatter plot.

3.2.3. Data Feedback Analysis

According to the above regression results, it is not difficult to see that most of the selected enterprises can well verify the feasibility of CAPM model in real life, but there are still exceptions. For example, the regression coefficients of LLY, TSLA, BA, COP, EOG, and PXD are 0.7327, 0.4035, 0.6096, 0.5513, 0.5139, and 0.6018, respectively, indicating that if these enterprises use CAPM model to forecast, there is a large gap between the predicted value and the actual value. In view of the inapplicability of CAPM model in practice, the following points are put forward: Stocks in different industries may have different risk characteristics, including profit fluctuation, balance sheet structure, etc. Due to these industry-specific risk factors, a single market beta cannot fully explain the expected returns of stocks in different sectors[5]; Volatility in stock prices can be better explained by considering the impact of other economic factors, such as inflation expectations and interest rate changes, on expected stock returns[6];The effects of economic factors (market risk, size, value, investment, and profitability) on stock returns are proposed, and in subsequent research it is found that considering these factors can more accurately explain the expected return of stocks, especially across different industries[7].

In view of the low applicability of CAPM model to the US energy industry, the following possibility analysis is given:

Non-systemic risk: It is mentioned in the premise assumption of CAPM that only systemic risk is considered, but in the actual situation, non-systemic risk still exists objectively and has a greater impact on specific industries. The U.S. energy industry may be affected by a variety of factors, such as policy influences, geopolitical influences, changes in supply and demand, etc., which may lead to the existence of industry-specific risks unrelated to overall market risks.

Long-term factors: CAPM model is built on the basis of long-term portfolio stability, but for the US energy industry, there may be changes in supply and demand that have a huge impact on the actual price, resulting in large fluctuations in portfolio returns, which makes the CAPM model not applicable;

Limitations of assumptions: There are many assumptions in the establishment of CAPM model, such as investors are rational individuals who avoid risks and maximize utility, the market is frictionless, there are no transaction fees or taxes, investors all make decisions for the same period, investors have homogeneous expectations and beliefs, all investments are infinitely separable, and investors are price takers. These assumptions are intended to make investors' behavior follow mean-variance analysis. However, if the assumption is too ideal, there will be various unexpected situations outside the assumption in reality, and there may be some special circumstances and restrictions in the energy industry that are contrary to the assumptions of the CAPM model, so the applicability of the model will be affected.

Nevertheless, even if the CAPM model is not applicable in a particular industry, other risk assessment models or methods can be used to better analyze the risks and rewards of that industry. This may include the use of multifactor models (such as the Fama-French three-factor model), industry-specific models, or other quantitative and qualitative analytical methods to more accurately assess the risk and return relationship in the U.S. energy industry.

4. Conclusion

This paper mainly discusses whether CAPM model is applicable in the American market and analyzes the causes of error. It is concluded that CAPM model is applicable and can be predicted more accurately in most cases. There are some areas that can be improved in this paper: a. When the data is selected in month, the data accuracy is lacking. If the step size is changed to day, the data will be more and more accurate. b. The CAPM model itself has defects and is too idealized, which may be inconsistent or even contrary to the actual situation. Research on CAPM may focus on adding more impact factors to it to make it closer to the actual situation and make the prediction more accurate.

References

[1] *Black F, Jensen M C, Scholes M. The capital asset pricing model:some empirical tests[M]. In michael jensen, 1972.*
[2] *Banz R W. The relationship between return and market value of common stocks[J]. Journal of Financial Economics, 1981,9(1):3-18*
[3] *Fama E, James D M.Risk, return and e-quilibrium: empirical tests[J]. Journal of Finance, 1973,25(2):383-417*
[4] *Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance, 19(3), 425-442.*
[5] *Fama, E. F., & French, K. R. (1997). Industry costs of equity. Journal of Financial Economics, 43(2), 153-193.*
[6] *Chen, N., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. Journal of Business, 59(3), 383-403.*
[7] *Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1-22.*