The empirical test of Asset Pricing Models in US Machinery market: Comparison of CAPM, FF3 and FF5 Models

Yeqiu Wang*

School of Natural of Science, Department of Mathematics, University of Manchester, Manchester, United Kingdom

*Corresponding author: yeqiu.wang@student.manchester.ac.uk

Abstract. As a sample, this study looks at data relating to the machinery sector of the US stock market from July 1926 to July 2022. This study's overarching objective is to investigate how three different financial models—the CAPM, the Fama-French 3 Factor Model, and the Fama-French 5 Factor Model, which have been implemented in the machinery sector over the course of time. This will allow the researchers to determine whether or not these models are applicable to the US stock market. In general, the three-factor model has better explanatory power for excess returns than the five-factor model, even with the addition of the investment and profitability factors for a single sector in the United States, the machinery sector and it explains almost all of the excess returns. This maybe because the three-factor model takes into account fewer variables than the multifactor model. Furthermore, the new profitability and investment factors do not have a single-factor effect that is as strong as that of the size and valuation factors.

Keywords: CAPM; Fama-French Three Factor Model; Fama-French Five Factor Model; US Stock Market; Machinery Industry.

1. Introduction

Currently in real life, the topic of how to reasonably use asset pricing models and portfolios is always hot for financial practitioners such as quantitative analysts, fund managers and others in the field of financial engineering in the capital management sector. For example, the application of the Fama-French five-factor model to the Chinese A-share market by Li et al. (2017) found that the five-factor model outperformed the CAPM, three-factor, and carhart four-factor models [1]. They found that the multifactor model performed better on the Chinese market than others. This was discovered when the models were compared to one another. This assisted in the disclosure of additional rules that govern the pricing of assets on the Chinese capital market and provided a road map for locating a factor model that is better suited to the Chinese market. In addition to that, this provided fresh insight into the criteria that govern the valuation of various assets. In addition to this, an empirical test of the three different pricing models will be carried out in order to determine which of the three is best suited for the machinery market in the United States.

The main purpose of the research is to study the impact of the candidate factors, such as MKT, SMB, HML, RMW, on the portfolio returns and claim which factors and financial models seem to be more relevant and applicable to U.S. machinery industry. Following that, different regressions using statistical methods based on the three different models were run to evaluate the efficiency and final results. Asset pricing models aims to determine economic and statistical factors for different assets, the well-known includes CAPM, Fama-French 3-factor and 5-factor models. In 1964, Sharpe proposed the Capital Asset Pricing Model [2], which states the expected return on any stock. Although the CAPM has been widely used, past research has shown that it is not a comprehensive model of stock returns. Equities with low market capitalization tend to have higher returns than the CAPM predicts, and value stocks likewise tend to have better returns than the CAPM predicts. This is why Fama and French (1992) added to CAPM, which primarily takes into account the effect that value and small-cap companies beat markets on a regular basis [3]. The approach incorporates an adjustment for outperforming tendency, which has been shown to yield more accurate assessments of managerial competence. 2016 saw Fama and French add two new elements to their three-factor model to better capture asset returns: RMW and CMA [4]. The study's methodology and the empirical findings based on these models will be illustrated below.

2. Data and method

2.1 Models

2.1.1 CAPM

The CAPM has been developed in order to quantify this systematic risk. Given the risk and cost of capital involved in the asset in question, it is often used in the financial industry to value risky securities and to calculate expected returns.

$$E[R_i] = R_f + \beta_i(E[R_m] - R_f) \quad (1)$$

2.1.2 Fama-French Model

(a) Three Factor Model

This model is an asset pricing model built on the CAPM that includes size risk and value risk variables in addition to market risk components. It was developed in 1992. This approach accounts for the fact that value and small-cap equities are frequently among the market's best performers.

$$R_{it} - R_{ft} = \alpha_{it} + \beta_1\left(R_{Mt} - R_{ft}\right) + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_{it} \quad (2)$$

Where:

$$R_{it} = sum\ of\ a\ stock's\ or\ portfolio's\ gains\ as\ of\ a\ specific\ date\ and\ time\ t$$
$$R_{ft} = risk - free\ rate\ of\ return\ at\ time\ t$$
$$R_{Mt} = cumulative\ portfolio\ returns\ from\ the\ market\ as\ of\ time\ t$$
$$R_{it} - R_{ft} = expected\ excess\ return$$
$$R_{Mt} - R_{ft} = positive\ portfolio\ returns\ above\ market\ average\ (index)$$
$$SMB_t = size\ premium(small\ minus\ big)$$
$$HML_t = value\ premium(high\ minus\ low)$$
$$\beta_{123} = linear\ correlation$$

(b) Five Factor Model

The model was proposed in 2014 with the foundation of three-factor model, as reported by (Fama and French, 2015). It expands the idea of dividend discounting by asserting that a company's current value is based on future payments. Adding two new variables—investment and profitability—Fama and French refined the discounted dividend model's capacity to reflect the link between risk and return [5].

$$R = \alpha + \beta_m MKT + \beta_s SMB + \beta_h HML + \beta_r RMW + \beta_c CMA \quad (3)$$

RMW is returns of companies with high and low operating profitability.

CMA is returns for companies with low versus high investment policies.

2.2 Construction of factors

As summarised by Ke Zhang in 2018, the concept underlying the model's factor return construction is to classify stocks along one or more factor dimensions and calculate the returns of various stock portfolios based on those classifications, which are then calculated by some arithmetic rule [6]. Considering the importance of the size factor, a fixed size factor can be chosen for classification.

Each stock is ranked on a scale from one to three based on the size of its corresponding factor at the close of every month. Then mean rate of return of the stocks within each tier is calculated for the following month. The dissimilarity in typical yields between different layers is used to measure the return effect of the factor. This process of factor construction allows for uniformity in the scale of each factor and comparability of returns between factors

This paper uses the 2 × 3 approach used in Fama - French (2015a) to construct factors along two factor sub-dimensions, dividing stocks into 2 × 3 = 6 portfolios. One of the factors is fixed to size and is divided by median into two tiers, and the other is BP, OP or Inv, divided into 3 tiers according to the 30% and 70% quartiles.

(1) First, based on their median market capitalization, small-cap (S) stocks and large-cap (B) equities are separated.

(2) Next, we divide the data set into three groups based on how high (H), medium (N), and low (L) the sample's book-to-market ratio was, using the quartiles at 30% and 70%.

(3) The group can be broken down into 6 subsets (SH, SN, SL, BH, BN, and BL) by crossing the two indications.

(4) Similarly, operating margin and investment style replace book-to-market ratio, while robust (R), concentrated (N), and more concentrated (C) replace book-to-market ratio (L).

(5) Instead of the book-to-market ratio, we can use operating profitability and investment style, and classify each according to conservative (C), moderate (N), or aggressive (A) behaviour (A). There are twelve distinct portfolios that can be created from the full portfolio.

(6) To do this, you'll need to calculate the market value weighted average return for each of the aforementioned portfolios, and then utilise the variance in those average returns to create four factors labelled SMBt, HMLt, RMWt, and CMAt. The calculation method for factors is displayed in Table 1.

Table 1. The Calculation Method for Factors

Factor Name	Calculation Method
SMB_t	$SMB_{BM} = \frac{SH + SN + SL}{3} - \frac{BH + BN + BL}{3}$ $SMB_{OP} = \frac{SR + SN + SW}{3} - \frac{BR + BN + BW}{3}$ $SMB_{Inv} = \frac{SC + SN + SA}{3} - \frac{BC + BN + BA}{3}$ $SMB_t = \frac{SMB_{BM} + SMB_{OP} + SMB_{Inv}}{3}$
HML_t	$HML_t = \frac{BH + SH}{2} - \frac{BL + SL}{2}$
RMW_t	$RMW_t = \frac{BR + SR}{2} - \frac{BW + SW}{2}$
CMA_t	$CMA_t = \frac{BC + SC}{2} - \frac{BA + SA}{2}$

2.3 Data

2.3.1 Sample selection: USA Machinery Industry

In 1911, the United States supported the development of the earliest combination machine tools for machining automobile components. In 1953, common components for combination machine tools were standardised in consultation with American machine tool manufacturers. As machine tools and tool materials advance, the United States of America's manufacturing process has vastly improved.

Table 2. Machinery Industry and its sub-sectors

Industry	Sub-Sectors
Machinery	Combustion devices and turbines
	Tools and machinery for the garden and farm
	Tools, machines, and transport mechanisms for building, mining, and moving materials
	Vehicles and tools for building construction
	Mechanical & Electronic Equipment for the Mining Industry Other Than Oil Field
	Mechanical and electrical apparatus for the petroleum and natural gas industries
	Lifts and revolving stairways
	Equipment for moving materials from one place to another
	The use of mechanical lifting devices like cranes, hoists, and monorails
	Machinery
	The Tools and Machines Needed to Work Metal
	Machines designed specifically for use in industrial settings
	Machines and tools used in manufacturing in general
	machinery for the service and refrigeration industries
	Automated merchandising systems
	Industrial washing and drying apparatuses
	Temperature control, warm air heating, and cooling devices
	Pumping and measuring
	Appliances for the service sector
	Various types of commercial and industrial machinery

Referring to the data and information shown in the official website of the United States government: Bureau of Economic Analysis [7], In 2011, the domestic and international sales of the US machinery industry were $413.7 billion. Dollar bills on conveyor belts The USA is the third-largest supplier of machinery in the world and the biggest consumer for it. The domestic US market is made up of US manufacturers to the tune of 58.5%.

For US manufacturers to maintain their dominance in a very competitive international market, they must remain committed to technical innovation. It is crucial to integrate information technology with machinery in order to increase production, efficiency, and sustainability. Many US institutions are also conducting cutting-edge research in related fields of engineering and science, which helps the US equipment industry remain competitive.

Many other manufacturing and service businesses rely on the extremely complex technology that is produced by the machinery and plant manufacturing sector. End users can increase the productivity of their equipment by using process control and other automation technologies. Sales of various kinds of machinery frequently come with a range of high-value services, such as expert engineering, building, and logistics.

In August 2013, the manufacturing of machinery and equipment directly employed more than 1.3 million Americans. These positions are virtually entirely in high-skill, high-paying.

fields and professions. Numerous additional manufacturing and service sectors employ hundreds of thousands of Americans, including the machinery industry.

A vast industry producing machinery and equipment exists in the United States. The vast majority of American firms with a global presence are small and medium-sized enterprises, although there are also numerous huge publicly traded companies and recognisable American brands. The manufacturing of machinery and equipment is widespread in the United States but centred in the industrial Midwest. California and Texas are also key areas for the machinery industry.

In the Machinery portfolio selected for this paper, the following sub-sectors are included in the Machinery sector, shown in Table 2 Machinery Industry and its sub-sectors.

2.3.2 Descriptive statistics

To begin, get the information from Kenneth R. French's data library [8]. The response variable is the excess return of the machine industry portfolio ('Mach_RF'). Difference between market yield and risk-free rate('Mkt_RF'), sum of nine large stock portfolios' returns minus nine small stock portfolios' returns ('SMB'), the disparity between the mean returns of a pair of value and a pair of growth portfolios('HML'),The difference between the average returns of two portfolios with high operating profitability and two portfolios with low operating profitability('RMW'), Additionally to the average returns of the twelve portfolios ('avg') are all among the interested regressors. The summary of variables statistics is calculated in Table 3.

Table 3. Summary of variables statistics

	n	*mean*	*sd*	*median*	*min*	*max*	*skew*	*kurtosis*
Mkt_RF	709	0.563	4.477	0.92	-23.24	16.10	-0.505	1.805
SMB	709	0.225	3.026	0.10	-15.35	18.34	0.342	3.083
HML	709	0.298	2.968	0.23	-13.97	12.75	0.106	2.331
RMW	709	0.280	2.209	0.24	-18.73	13.09	-0.294	11.572
CMA	709	0.283	2.041	0.09	-6.89	9.05	0.304	1.389
RF	709	0.363	0.268	0.38	0.00	1.35	0.655	0.606
Mach_RF	709	0.660	6.280	1.07	-31.91	23.02	-0.400	2.116

For the key variable Mach_RF, its mean is 0.660 with a standard deviation of 6.280. The minimum value of it is -31.91 while the maximum of it is 23.02. Compared to the risk-free rate, the Mach_RF fluctuates violently since it has a far larger standard deviation. The skewness of Mach_RF is negative, which means the distribution of Mach_RF skews to the left. This index indicates that in the sample period the excess return of machine industry portfolio has more negative values than positive values. The kurtosis of Mach_RF is larger than 1, which means the distribution of portfolio excess return is peaked. Both the skewness and the kurtosis suggest that the distribution of the excess return of machine industry portfolio is tended to be nonnormal.

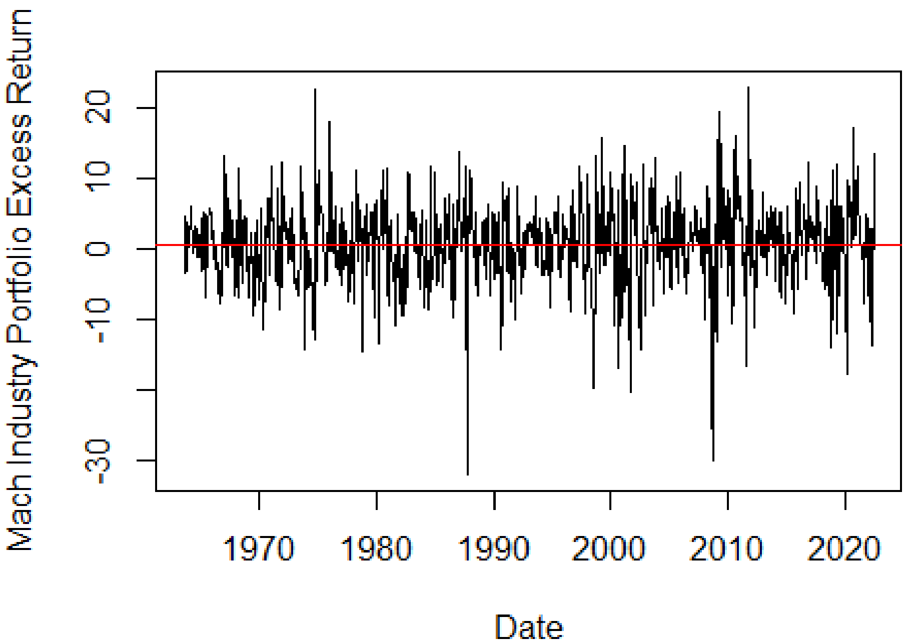

Figure 1. Time series plot of Mach_RF

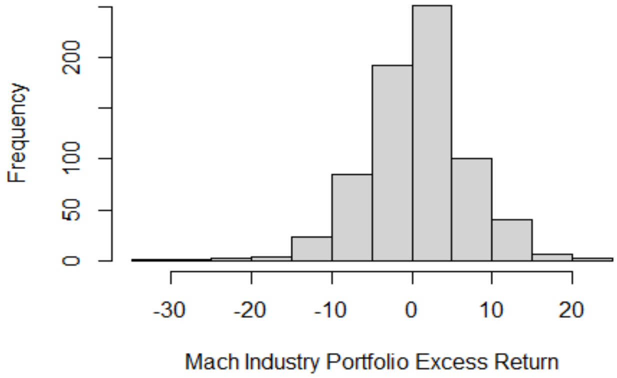

Figure 2. Histogram of Mach_RF

In the time series plot of Mach_RF (Figure 1), which can be seen that the excess return fluctuates violently around the red mean line. Some extreme negative values appeared around October 1987 and October 2008. This phenomenon is reasonable because two famous economic crises, the Black Monday [9] and the Subprime mortgage crisis [10] happened respectively. The histogram of Mach_RF also verifies that the distribution of Mach_RF is negatively skewed, which mean bad times are more than good times in stock market.

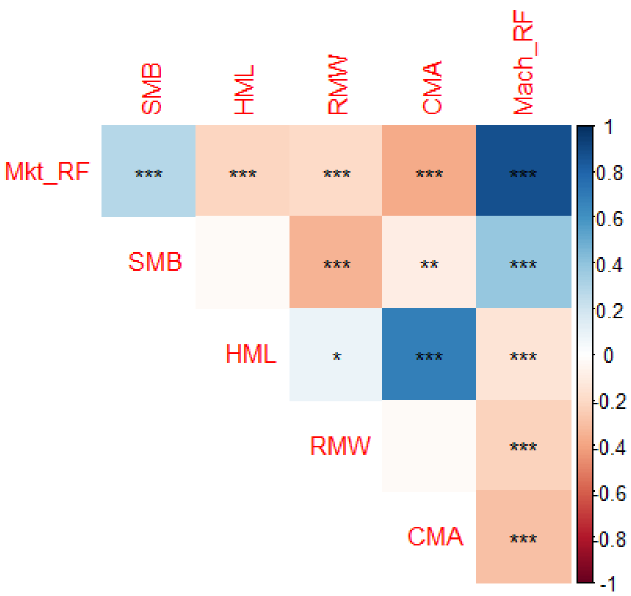

Figure 3. Heat map of variables

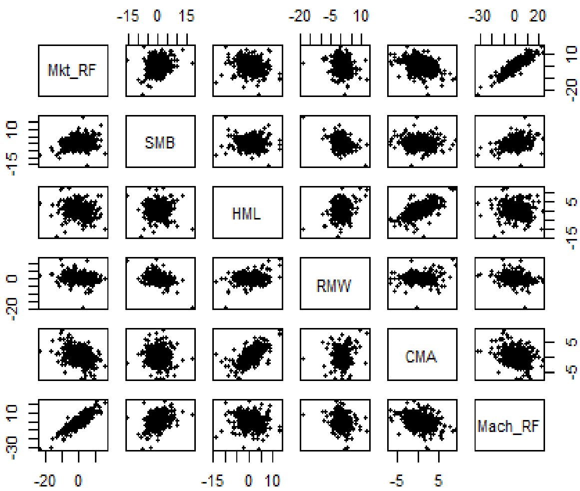

Figure 4. Pair-wise scatter plots

2.3.3 Correlation analysis

Before the regression analysis, it is efficient to conduct a primary correlation analysis to detect the relationships among these variables. The correlation result is depicted in the upper triangular heat map (Figure 3). The Mkt_RF, SMB and Mach_RF are positively correlated while the HML, RMW, CMA and Mach_RF have a negative correlation. On a one-percent significance level, all of the relationships hold. Such correlations are also shown in the format of pair-wise scatter plots (Figure 4).

3. Results

3.1 Regression results

In this step, it is straightforward to present the result through running three different models.

Table 4. Estimated results

	Dependent variable:		
	Mach_RF		
	(1)	(2)	(3)
Mkt_RF	1.225***	1.179***	1.182***
	(0.026)	(0.026)	(0.028)
SMB		0.310***	0.305***
		(0.038)	(0.040)
HML		0.089**	0.072
		(0.038)	(0.052)
RMW			-0.021
			(0.055)
CMA			0.042
			(0.079)
Constant	-0.029	-0.100	-0.102
	(0.116)	(0.112)	(0.115)
Observations	709	709	709
R^2	0.762	0.785	0.785
Adjusted R^2	0.762	0.784	0.783
Residual Std. Error	3.065	2.920	2.923
F Statistic	2,265.553***	856.981***	513.238***
Note:			*p<0.1; **p<0.05; ***p<0.01

Model (1) represents the CAPM, model (2) is the three-factor Fama-French model and model (3) is the five-factor Fama-French model. The overall effectiveness of the models should be evaluated first. The provided F statistics are statistically significant at the 1% level, indicating that all three models have overall statistical significance. Model (2) has the highest adjusted R-squared value 0.784, indicating that the Fama-French 3 factors account for 78.4% of the variance in the response variable. As this number is greater than the one obtained from model (1), it follows that Model (2) is superior towards the Model (1). Model (3) has a slightly lower adjusted R-squared of 0.783 compared to model (2). This suggests Model (3) contains some noise and that some of the factors may be redundant.

Next, perform the LR test to select the best model. The likelihood-ratio test (LR test) compares the relative plausibility of two competing statistical models (nested and complicated models) to determine which one best fits the data. The two likelihoods should not differ too much if the constraint (the null hypothesis that the nested model is good enough) is satisfied by the data. The difference could be reflected in the p-value. For the comparison between model (1) and model (2), the p-value is close to zero, at the 1% significance level, the null hypothesis should be rejected. This means that the complex model is significantly more accurate. By same logic, the model (2) is better than model (3).

After checking the overall significance, it suggests it is better to see the factors' performance. Among the three models, the coefficients of Mkt_RF are statistically significant at 1% level. This means the excess return on the market is always important. For other factors, the SMB factor is statistically significant in model (2) and model (3). The HML factor is statistically significant in model (2) but not significant in model (3). Such a phenomenon suggests that the HML factor's effect might be cannibalized by RMW and CMA. The RMW and CMA factors are not significant in model (3), which means they might be redundant. The constant here is not significant, which means Fama-French 3 factors are good enough that little other information is ignored. To rank the factors importance, we conducted a forward stepwise regression. The term "stepwise regression" refers to a method of iteratively building a regression model, with each iteration narrowing down the pool of potential independent variables. The output above verify that the rank of the factors is Mkt_RF, SMB, HML, CMA and RMW.

For simplicity, this study will make the economic explaining of model (2) which is the winner of last section. The coefficient of Mkt_RF means keeping other conditions constant, if the Mkt_RF increase one unit, the excess return of portfolio six will increase 1.179 units. The coefficient of SMB means keeping other conditions constant, if the SMB increase one unit, the excess return of portfolio six will increase 0.310 units. The coefficient of HML means keeping other conditions constant, if the HML increase one unit, the excess return of portfolio six will increase 0.089 units.

3.2 Regression diagnosis

Here, the model's configurations will be validated by means of an automated test (Figure 5). To see if the residuals exhibit non-linear behaviour, plot them against the fitted values. Most of the time, linearity is met, as seen by the red line's nearly horizontal midsection and slightly bent tails. Whether or not the residuals follow a normal distribution can be seen in a Normal Q-Q plot. The normalcy assumption is supported by the observation that the majority of points lie on the 45 degree line. If the residuals are evenly distributed across the ranges of the predictors, this will be evident in the Scale-Location plot. Assuming equal variance, you can test it in this way (homoscedasticity). On the graph, there is a tiny curved line with rather randomly distributed points, so it is not terrible. If there are any influential cases, it is possible to find them using the Residuals vs. Leverage plot. A grey dashed line represents Cook's distance, therefore cases outside of it should be investigated. Cases have an outsized impact on the regression outcomes when they fall outside of the Cook's distance, as indicated by high Cook's distance scores. Having no influential point in this test is a huge stroke of luck since all of the examples fall well inside the Cook's distance lines.

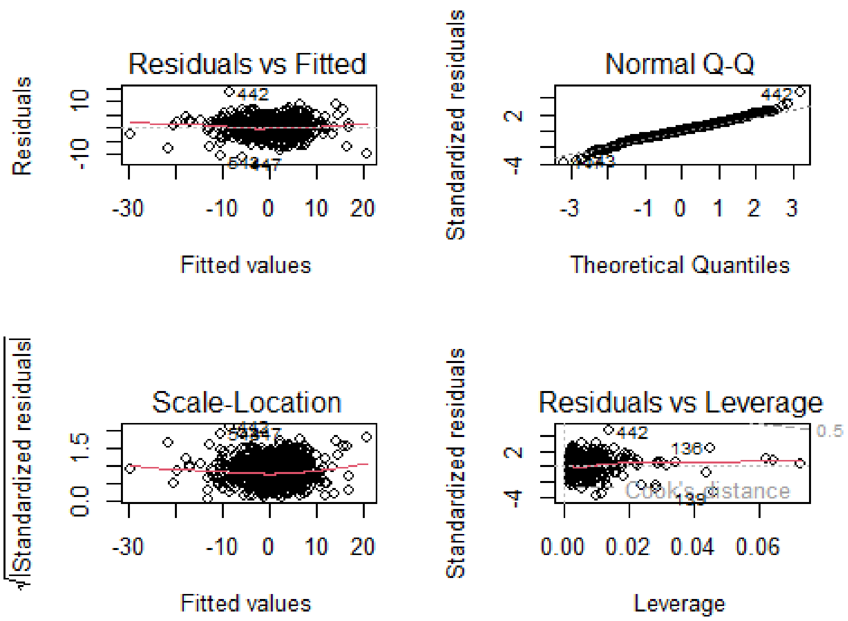

Figure 5. Regression diagnosis plots

4. Conclusion

This study, which selects monthly data from 1926.7 to 2022.7 for listed companies in the machinery sector of the US stock market and tests the applicability of the CAPM, three-factor, and five-factor models in the machinery sector of the US stock market according to the Fama and French (2015a; 2015b; 2016) methodology, provides analytical ideas to further reveal the pricing rules of the US capital market and find a more suitable factor model for the US market. This study demonstrates that first, all three models are applicable to the industry; second, the fitness of the three-factor model is better compared to that of the CAPM; and third, the five-factor model may contain factors that are redundant. Next, for the single-factor performance test , market excess returns are very important for the single-factor performance test. SMB performs well in both the three-factor and the five-factor models, whereas HML performs well in the three-factor model but not in the five-factor model. This may indicate that RMW and CMA have an effect on HML. In the five-factor model, RMW and CMA might be considered redundant based on the results of the tests. Because of this, the three factors, despite the fact that they lack some variables, are sufficient to apply to the data from the US machinery industry. After that, the importance of each of the factors is determined by placing them in the following order: Market Excess Returns, SMB, HML, CMA, and RMW.

References

[1] LI Zhibing, Yang Guangyi, FENG Yongchang, JING Liang. Fama-French Five Factor Model in China Stock Market. Journal of Financial Research, 2017, 444 (6): 191-206.

[2] Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. The journal of finance, 19 (3), 425-442.

[3] Fama, Eugene F. ; French, Kenneth R., Common risk factors in the returns on stocks and bonds [J]. Journal of Financial Economics,1993,33 (1):3-56.

[4] Fama, Eugene F. ; French, Kenneth R. , A five-factor asset pricing model [J]. Journal of Financial Economics, 2015,116 (1): 1-22.

[5] Zhang Shao-Lin. An empirical test of Fama-French five-factor model in A-share market and its extension[D]. Sichuan: Southwest University of Finance and Economics,2018.

[6] Ke Zhang, Fama-French five-factor model A-share empirical study, http://fdjpkc.fudan.edu.cn/_upload/article/files/77/f7/4af4e037434dbc7ebf6e1376f6c1/fb3e42b9-5f04-4041-a28c-380e785e5434.pdf [Access Date : 20 September 2022]

[7] An official website of the United States government: Bureau of Economic Analysis https://apps.bea.gov/industry/factsheet/factsheet.html#31G

[8] Data Base: Dartmouth, Data Library, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html [Access Date: 20 September 2022]

[9] Ke Limin, Jin Xuejun. The fourth in a series of securities revelations: "Black Monday"[J]. Zhejiang Finance, 1998 (5):43-44.

[10] Yang Jianping. The origin of the subprime mortgage crisis in the United States[J]. Productivity Research,2010 (2):33-36.