



National Taiwan university

Institute of finance

Econometrics

Student number : M11318012

Multi-factor model in Taiwan's semiconductor industry

Student : 葉正恩 Ian

Professor : Dear Mr. Xue 薛博今教授

Date:113/12/22

Content

Chapter1-Introduction.....	3
1.1 Research background and purpose	3
1.2 Research objective	4
1.3 Research Framework and Process	4
Chapter2-literature discussion.....	5
2.1 Fama French's five factors model (2015)	5
Chapter3- Research Methodology	6
3.1 Research variable and sample	6
3.2 Theoretical Model	8
Chapter4- Research Methodology	8
4.1 Selection of variable and factors	8
4.2 Calculation of factors and Y	8
4.3 Result analysis	11
4.3.1 Final model	13
4.3.2 Explanation of Model Results	14
4.3.3 Checking Assumption	17
4.3.4 The regression of 30% of the sample.....	20
Chapter5- Research Conclusions and Recommendations	21
5.1 Conclusion:	21
5.2 Reference:	22

Chapter1-Introduction

1.1 Research background and purpose

The Fama-French Five-Factor Model represents a significant breakthrough in asset pricing theory. Proposed by Eugene Fama and Kenneth French in 2015, the model builds upon the original Three-Factor Model by adding two additional factors: profitability (RMW: Robust Minus Weak) and investment (CMA: Conservative Minus Aggressive). These new factors reflect key financial characteristics of firms, such as profitability and investment patterns, providing a more comprehensive explanation of market anomalies. Fama and French's research demonstrated that the Five-Factor Model outperforms the Three-Factor Model in capturing systematic influences on portfolio returns, particularly when measuring the performance of small-cap and high-value firms.

Insights:

- **Comprehensive Return Drivers:** This model enhances the explanation of stock returns by adding profitability (RMWt) and investment (CMAt) factors to the original three factors (market risk, size, and value).
- **Profitability Impact:** Firms with higher profitability (RMWt) are rewarded with higher returns, supporting the view that robust profit generation adds to a stock's appeal.
- **Investment Behavior:** Conservative firms (CMAt) tend to outperform, suggesting that aggressive investment strategies may lead to inefficiencies or higher risks.
- **Risk Exposure Analysis:** The coefficients provide insights into a stock's exposure to different types of risks, helping in portfolio construction and risk management.
- **Alpha and Model Fit:** The model accounts for more variability in stock returns compared to the three-factor model, but any significant alpha (ait) might point to inefficiencies, skill, or omitted factors.

In this study, we build on the Fama-French Five-Factor Model to explore the asset pricing characteristics of Taiwan's semiconductor industry. As a critical player in the

global semiconductor supply chain, Taiwan's semiconductor sector is characterized by its technological intensity and high capital investment. These features may challenge the applicability of traditional value-based investment factors. Therefore, beyond the fundamental structure of the Five-Factor Model, this study incorporates industry-specific factors, such as leverage and momentum, to provide a more holistic view of the long-term determinants of investment returns in this industry. Furthermore, we compare the performance of these factors in the Taiwanese and U.S. markets to offer practical insights for investors based on historical data.

1.2 Research objective

The objectives of this study are as follows: :

1. Based on the Fama French five-factor model, applied to the Taiwan market and adding more factors, and compare the Taiwan semiconductor industry with the US market
2. To provide insights for those planning to invest in Taiwan semiconductor industry based on the long-term historical database

1.3 Research Framework and Process

This study is divided into five chapters, with the arrangement and content summarized as follows:

Chapter 1: Introduction

Describes the research motivation, objectives, framework, and process.

Chapter 2: Literature Review

Explores domestic scholars' studies on financial crisis prediction models and the definitions of financial crises provided by relevant institutions.

Chapter 3: Research Methodology

Explains the theoretical foundation of the logistic regression model, the research subjects, data sources, research design, and research process, including variable selection, data cleaning, and screening criteria.

Chapter 4: Empirical Analysis

Interprets the logistic regression model's prediction results on corporate default probabilities, identifies significant variables and prediction accuracy, and analyzes the effectiveness of the model's validation results.

Chapter 5: Research Conclusions

Summarizes the empirical findings and research contributions, and provides concluding remarks.

Chapter2-literature discussion

2.1 Fama French's five factors model (2015)

The equation of the Fama five factors model:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_iSMB_t + h_iHML_t + r_iRMW_t + c_iCMA_t + e_{it}.$$

Variables and Their Meanings:

1. $R_{it} - R_{Ft}$: The excess return of stock i over the risk-free rate ($R_{it} - R_{Ft}$).
 - o This measures the return of the stock relative to a risk-free benchmark, isolating the impact of risk factors.
2. $R_{Mt} - R_{Ft}$: The market risk premium, which is the return of the market portfolio minus the risk-free rate.
 - o It reflects the compensation investors demand for taking on market risk.
3. SMB_t : The size premium (Small Minus Big).
 - o Captures the return difference between small-cap stocks and large-cap stocks, highlighting the additional return investors expect from small firms due to their higher risk.
4. HML_t : The value premium (High Minus Low).

- o Captures the return difference between high book-to-market (value) stocks and low book-to-market (growth) stocks, indicating the tendency of value stocks to outperform growth stocks.
- 5. RMWt: The profitability premium (Robust Minus Weak).
 - o Represents the return difference between stocks of companies with high profitability (robust) and those with low profitability (weak), reflecting that more profitable companies tend to deliver higher returns.
- 6. CMAt: The investment premium (Conservative Minus Aggressive).
 - o Captures the return difference between firms that invest conservatively (low asset growth) and those that invest aggressively (high asset growth). Conservative firms tend to outperform aggressive ones over time.
- 7. Coefficients: The sensitivities of the stock's return to each factor:
- 8. ait: The alpha or the portion of the stock's return unexplained by the model.
 - o A positive ait indicates potential outperformance or unique factors not captured by the five factors.
- 9. Eit: The error term, representing idiosyncratic risks or noise.

Chapter3- Research Methodology

3.1 Research variable and sample

3.1.1 Research Samples

The research subjects and sampling period are as follows:

1. **Data Source:** Taiwan Economic Journal (TEJ) database:半導體 N
2. **Research Industries:** This study follows the industry classification of the Taiwan Stock Exchange, focusing on listed companies in the semiconductor.
3. **Number of Sample Companies:**

Date	Number of companies
Apr-2005	39
...	...

Sep-2024	83
----------	----

4. **Sampling Period:** The study identifies companies with default records during the period from April 1, 2005, to September 30, 2024.
5. **Sample size:** 219
6. **Dependent Variable (Y):** portfolio return

3.1.2 Research data and factors setup

1. Retrieve stock price information for all companies during the study period.
2. Obtain financial reports for all companies during the study period.
3. Convert quarterly financial report data into monthly data.

該月結束日	名稱_right	年/月	季別	合併(Y/N)/單半	月份	幣別	稅前息前(A)	稅後	稅後息前	綜合
2024/09/30	麗正	2024/09	3 Y	Q	9	NTD	1.18	0.53	0.93	0.82
2024/08/30	麗正	2024/09	3 Y	Q	9	NTD	1.18	0.53	0.93	0.82
2024/07/31	麗正	2024/09	3 Y	Q	9	NTD	1.18	0.53	0.93	0.82
2024/06/28	麗正	2024/06	2 Y	Q	6	NTD	2.15	1.33	1.71	1.26
2024/05/31	麗正	2024/06	2 Y	Q	6	NTD	2.15	1.33	1.71	1.26
2024/04/30	麗正	2024/06	2 Y	Q	6	NTD	2.15	1.33	1.71	1.26

4. Construction of the factors.
5. Standardize the data.
6. Perform univariate analysis on each independent variable with the dependent variable.
7. Build a correlation matrix to analyze the relationships between independent variables. Remove one variable from pairs with a correlation coefficient greater than 0.7. The variable with the smaller p-value in the univariate analysis is retained. Variables such as 'PEG', 'RMW', 'ROA', and 'LIQ' are removed.
8. Change 'RD', 'WML', 'INV_TURN', 'QUICK', 'COV' as dummy variables, because they are not very significant in the first try of regression model
9. Conduct stepwise regression on the selected independent variables to identify those that have a significant impact on the multiple regression model.

3.2 Theoretical Model

3.2.1 Research Model

This study employs a multiple regression model to see the sensitivity of each factor with the portfolio return.

Chapter4- Research Methodology

4.1 Selection of variable and factors

Factor	Variable
MKT_RF (market)	Market returns minus risk free rate
SMB (Size)	Market value of firms
HML (Value)	B/M ratio
RMW (profitability)	ROE
LEV (leverage)	D/E ratio
WML (momentum)	Average return of pass 12months
ROA	ROA
GM	Gross margin
COV	Coverage ratio
INV_TURN	Inventory Turnover
PEG	Price/Earnings to Growth
RD	Research & Development
TURN	Turnover Ratio
LIQ	Liquidity Ratio
QUICK	Quick ratio
Y (Portfolio return-Rf)	

4.2 Calculation of factors and Y

Y= Portfolio return-Rf

Definition:

The dependent variable (Y) represents the monthly stock returns of companies in the Taiwanese semiconductor industry.

1. Method1

(consideration into size effect, based on the fama's calculation)

Example: HML

Step 1: Remove missing data

Step 2: Group by Size (Small vs. Big)

Step 3:

Categorize companies into three BM ratio groups for each time point:

- **Low BM:** Bottom 30%.
- **Middle BM:** Middle 40%.
- **High BM:** Top 30%

Step 4: Combine size and BM ratio

Step 5: Calculate value-weighted returns

Step6: Calculate the Value Factor (HML):

$$R_{\text{High}} = \frac{R_{\text{Small_High}} + R_{\text{Big_High}}}{2}$$

$$R_{\text{Low}} = \frac{R_{\text{Small_Low}} + R_{\text{Big_Low}}}{2}$$

$$\text{HML} = R_{\text{High}} - R_{\text{Low}}$$

Applied factors: HML, RMW

2. Method2

(Without consideration of size effect, divide by 30%, 40%, 30%)

Step 1: Remove missing data

Step 2: Categorize Gross margin (GM)

Low: Bottom 30% (lowest gross margin).

Middle: Middle 40% (moderate gross margin).

High: Top 30% (highest gross margin).

Step3: Calculate Value-Weighted Returns for the Low Gross Margin Group

$$R_{\text{High}} = \frac{\sum (\text{Return} \times \text{Market Cap})}{\sum \text{Market Cap}}$$

$$R_{\text{Low}} = \frac{\sum (\text{Return} \times \text{Market Cap})}{\sum \text{Market Cap}}$$

Step4: Compute the Gross Margin Factor (GM)

$$\text{GM} = R_{\text{High}} - R_{\text{Low}}$$

Applied factors: SMB, QUICK, LIQ, COV, GM, RMW, PEG, INV_TURN, SMB, ROA, RD, TURN, WML, HML

3. Method3

(Without consideration of size effect, divided by 5 groups)

Example: LEV

Step 1: Remove missing data

- Exclude entries with missing debt ratios.

Step 2: Categorize by leverage (Debt ratio)

- Divide companies into 5 groups using quintiles: "Low," "L2," "M," "H2," and "High."
- Use only the **highest 20% (High)** and **lowest 20% (Low)** groups for the calculation.

Step 3: Calculate value-weighted returns

- **High Leverage Group**
- **Low Leverage Group**

Step 4: Calculate LEV

Applied factor: LEV

4.3 Result analysis

Table1. single factor analysis

Variable	P-Value
MKT_RF	1.01E-57
QUICK	7.92E-11
LEV	6.69E-10
LIQ	1.17E-08
COV	7.87E-06
GM	2.27E-05
RMW	0.000953
PEG	0.002435
INV_TURN	0.018341
SMB	0.013504
ROA	0.02167
RD	0.030695
TURN	0.042819
WML	0.261677
HML	0.419913

Table2. Correlation matrix

	WML	RMW	HML	SMB	LEV	ROA	GM	COV	INV_TURN	PEG	RD	TURN	LIQ	QUICK	MKT_RF
WML	1.000	0.302	-0.292	-0.084	-0.079	0.309	0.203	0.241	0.069	-0.290	0.032	0.107	-0.057	0.028	-0.179
RMW	0.302	1.000	-0.457	-0.266	-0.292	0.871	0.694	0.693	0.315	-0.843	0.072	-0.218	0.047	0.189	-0.286
HML	-0.292	-0.457	1.000	0.088	0.045	-0.376	-0.320	-0.290	0.032	0.356	-0.209	0.009	-0.117	-0.151	0.029
SMB	-0.084	-0.266	0.088	1.000	-0.320	-0.341	-0.139	-0.024	-0.113	0.330	-0.129	0.615	0.460	0.334	0.183
LEV	-0.079	-0.292	0.045	-0.320	1.000	-0.293	-0.523	-0.612	-0.414	0.350	0.153	-0.187	-0.551	-0.686	0.130
ROA	0.309	0.871	-0.376	-0.341	-0.293	1.000	0.760	0.738	0.361	-0.913	0.019	-0.236	0.012	0.162	-0.224
GM	0.203	0.694	-0.320	-0.139	-0.523	0.760	1.000	0.868	0.445	-0.756	0.080	-0.109	0.219	0.356	-0.216
COV	0.241	0.693	-0.290	-0.024	-0.612	0.738	0.868	1.000	0.476	-0.717	0.031	-0.009	0.254	0.389	-0.213
INV_TURN	0.069	0.315	0.032	-0.113	-0.414	0.361	0.445	0.476	1.000	-0.330	0.223	-0.239	0.133	0.323	-0.066
PEG	-0.290	-0.843	0.356	0.330	0.350	-0.913	-0.756	-0.717	-0.330	1.000	0.071	0.254	-0.095	-0.234	0.240
RD	0.032	0.072	-0.209	-0.129	0.153	0.019	0.080	0.031	0.223	0.071	1.000	0.019	-0.202	-0.087	0.034
TURN	0.107	-0.218	0.009	0.615	-0.187	-0.236	-0.109	-0.009	-0.239	0.254	0.019	1.000	0.069	-0.003	0.268
LIQ	-0.057	0.047	-0.117	0.460	-0.551	0.012	0.219	0.254	0.133	-0.095	-0.202	0.069	1.000	0.844	-0.071
QUICK	0.028	0.189	-0.151	0.334	-0.686	0.162	0.356	0.389	0.323	-0.234	-0.087	-0.003	0.844	1.000	-0.164
MKT_RF	-0.179	-0.286	0.029	0.183	0.130	-0.224	-0.216	-0.213	-0.066	0.240	0.034	0.268	-0.071	-0.164	1.000

4.3.1 Final model

(1) Description: Model incorporates six quantitative variables and one qualitative variable (1 for having positive momentum).

(2) Result of multiple regression :

Dep. Variable:	Portfolio_return_rf	No. Observations:	219
Model:	OLS	Df Residuals:	211
Method:	Least Squares	Df Model:	7
Date:	12/22/2024	R-squared:	0.862
Time:	02:46:50	Adj. R-squared:	0.858
Log-likelihood	-93.529	AIC:	203.1
Covariance Type:	nonrobust	BIC:	230.2

	coef	std err	t	P > z	[0.025	0.975]
Const	-0.05	0.036	-1.380	0.169	-1.783	0.128
HML	-0.0702	0.028	-2.541	0.012	0.562	1.719
SMB	-0.3604	0.035	-10.222	0.000	-2.063	-0.192
LEV	0.1721	0.035	-2.227	0.000	-2.062	-0.131
GM	-0.0721	0.035	-2.086	0.038	-1.682	-0.027
TURN	0.1543	0.034	4.582	0.000	-1.624	0.449
MKT_RF	0.8274	0.028	30.022	0.000	0.773	0.882
WML_dummy	0.1033	0.053	1.943	0.053	-0.001	0.208

(3) Model Equation :

Portfolio_return_Rf = -0.05 -0.0702 HML (value factor) - 0.3604SMB (size factor) + 0.1721 LEV (leverage factor) - 0.0721 GM (gross margin factor) - 0.1543TURN (turnover ratio) + 0.8274 Mkt_Rf + 0.1033 WML_dummy (momentum factor)

(4) Insight of the variable

Factors	Variable inside	Coef.	Economic meaning
HML	B/M ratio	-0.0702	V
SMB	Value size	-0.3604	V
LEV	Debt ratio	0.1721	V
GM	Gross profit/Revenue	-0.0721	V
TURN	COGS/Av. Inv	-0.1543	V
Mkt_Rf	Market premium	0.8274	V
WML_dummy	Last 12 months Average return	0.1033	V

4.3.2 Explanation of Model Results

(1) HML

Description: This factor measures the return difference between high book-to-market (value) stocks and low book-to-market (growth) stocks.

Model Insight: The coefficient of -0.0702 indicates that growth stocks (low B/M) slightly outperform value stocks (high B/M) in this analysis.

Comparison to Fama-French: In the original Fama-French model, HML often shows a positive relationship, reflecting value stocks' tendency to outperform growth stocks. However, the sign and magnitude of this factor can vary by market and sector.

Possible Reasons:

- i The semiconductor industry is innovation-driven, where growth stocks often represent companies investing heavily in R&D and future technology, making them more attractive to investors.
- ii High B/M firms may be older, more mature companies with slower growth prospects, leading to lower returns in this sector.
- iii Investor sentiment in the Taiwanese semiconductor market may favor high-growth, low B/M companies, particularly during periods of technological advancements or market optimism.

(2) SMB

Description: This factor captures the return difference between small-cap and large-cap stocks.

Model Insight: The coefficient of -0.3604 suggests that large-cap stocks outperform small-cap stocks in the context of this analysis.

Comparison to Fama-French:

The Fama-French model often finds a positive SMB coefficient globally, indicating that small-cap stocks outperform. This deviation could reflect specific preferences or dynamics in the Taiwanese semiconductor industry.

- i Large-cap semiconductor companies in Taiwan (e.g., TSMC) often dominate the global market, benefiting from economies of scale, stronger supply chains, and better access to capital.
- ii Small-cap firms may lack the resources to compete in a highly capital-intensive industry, leading to lower investor confidence and returns.
- iii The market may perceive large-cap firms as safer investments during periods of economic uncertainty or industry volatility.

(3) LEV

Description: This factor measures the return difference between high-leverage (high debt-to-equity ratio) and low-leverage companies.

Model Insight: A positive coefficient of 0.1721 indicates that high-leverage firms tend to perform better than low-leverage firms.

Comparison to Fama-French: Leverage is not a part of the Fama-French five factors. Including this factor provides additional insights into the influence of capital structure on stock returns.

Possible Reasons:

- I. High leverage may indicate aggressive expansion strategies, which can lead to higher returns if the borrowed capital is used effectively.
- II. Semiconductor firms with high debt levels might be viewed as taking calculated risks to maintain or expand market share, which investors reward.

- III. Low-leverage companies might represent firms that are too conservative, missing out on growth opportunities in a competitive, rapidly evolving industry.

(4) GM

Description: This factor compares the performance of firms with high operating gross margins against those with low margins.

Model Insight: The coefficient of -0.0721 suggests that lower-margin companies slightly outperform higher-margin companies

Possible Reasons:

Companies with lower gross margins may be pricing their products competitively, driving higher revenue growth and market share, which appeals to investors. High-margin firms might face slower growth due to mature product lines or less competitive pricing strategies.

- In the semiconductor industry, lower-margin firms may represent contract manufacturers or suppliers operating at high volumes, translating into higher overall

(5) TURN

Description: This factor reflects the relationship between turnover efficiency and returns.

Model Insight: A positive coefficient of -0.1543 indicates that firms with lower inventory turnover rates tend to generate higher returns.

In general:

- higher turnover ratio indicates efficient inventory management, suggesting that the company quickly sells and replenishes inventory.
- A lower turnover ratio may signal inefficiencies, such as overstocking or slow-moving inventory.

For semiconductor companies:

- A high turnover ratio might indicate rapid product cycles and strong demand for components, but it can also reflect commoditized, low-margin products.

- A low turnover ratio could signal a focus on high-value, innovative products that take longer to sell but generate higher profitability.

I think this can explain how firms with lower inventory turnover rates tend to generate higher returns.

(6) Mkt_Rf

Description: This factor represents the market risk premium, calculated as the return difference between the market and the risk-free rate.

Model Insight: A strong positive coefficient of 0.8274 suggests that the overall market return significantly influences portfolio returns.

(7) WML_dummy

Description: This factor measures the return difference between high-momentum stocks and low-momentum stocks.

Model Insight: The positive coefficient of 0.1033 indicates that high-momentum stocks tend to outperform low-momentum stocks in this analysis.

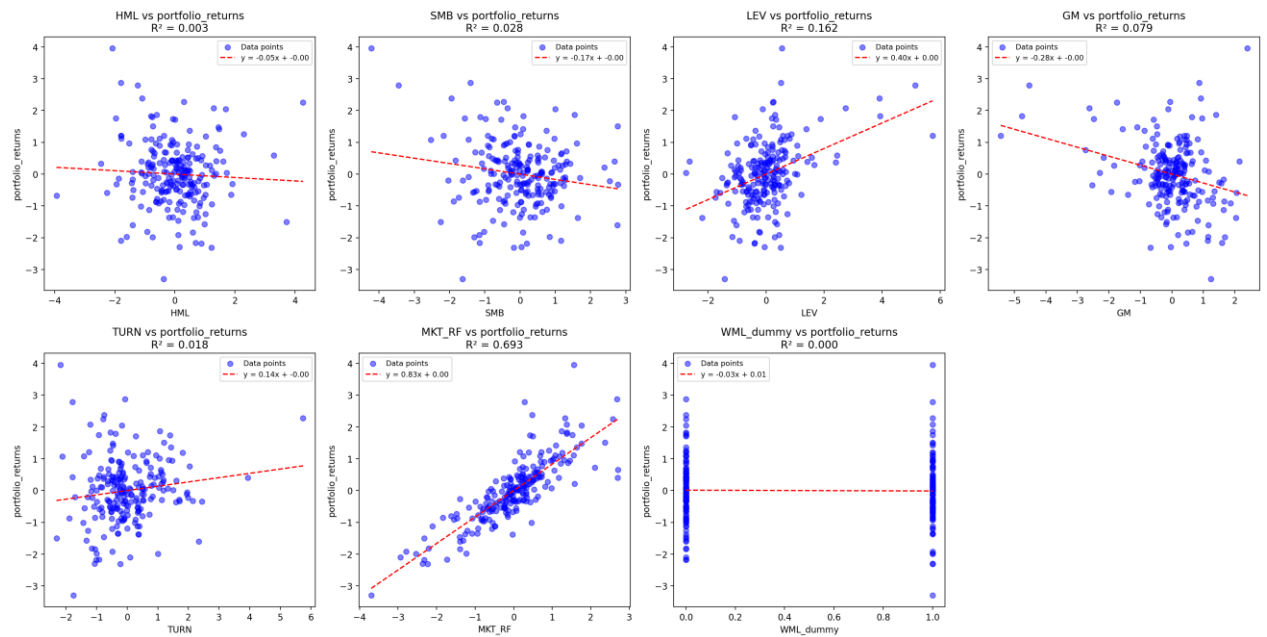
Possible Reasons:

- i In the semiconductor industry, companies with recent strong performance may signal innovation, successful product launches, or favorable market conditions, attracting more investor attention.
- ii Investors may have higher confidence in firms with positive recent trends, reinforcing the momentum effect.

4.3.3 Checking Assumption

(1) Checking Linearity

Scatter Plot of Independent Variables vs Dependent Variable with Regression Lines



(2) Checking multicollinearity

Variable	VIF
Const	2.015684
HML	1.169000
SMB	1.906922
LEV	1.827382
GM	1.831181
TURN	1.739764
MKT_RF	1.165082

All variables' VIF values are within acceptable range.

Conclusion:

(3) Checking autocorrelation

Durbin-Watson statistic: 1.9947 (close to 2)

Conclusion: No significant autocorrelation

(4) Checking Homoscedasticity

Breusch-Pagan Test Results

Lagrange multiplier statistic: 16.5328

p-value: 0.0207

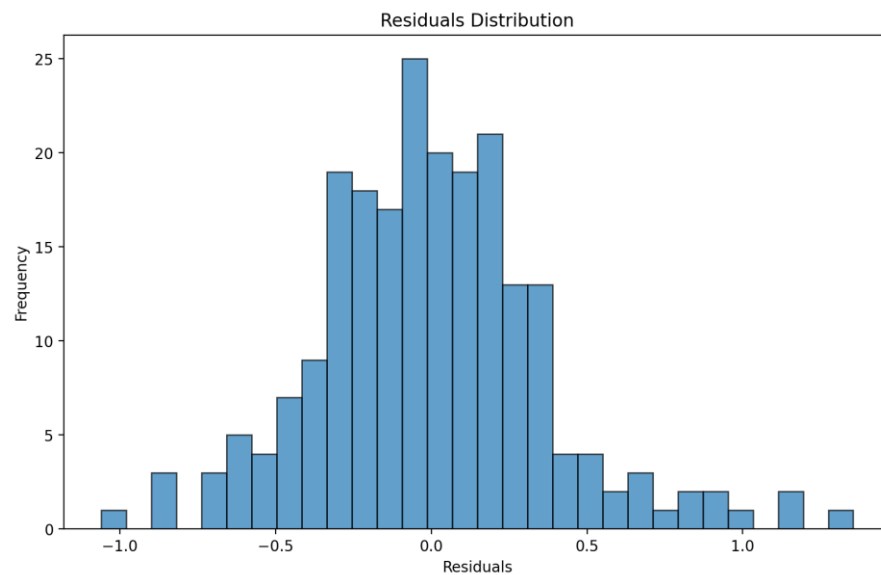
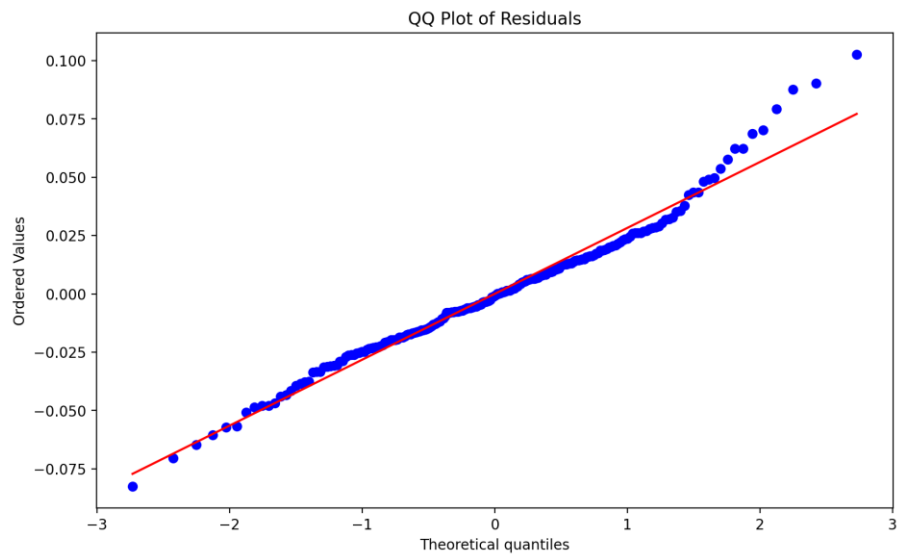
f-value: 2.4614

f p-value: 0.0190

Breusch-Pagan Test Interpretation:

Result: Heteroscedasticity is present (reject H0)

(5) Checking Normally distributed in residuals



Kolmogorov-Smirnov Test Interpretation:

Conclusion: Residuals are NOT normally distributed

4.3.4 The regression of 30% of the sample

4.3.4.1 The original model

```
=== ANOVA Table for Full Model ===
```

Sample Size: 219.0

ANOVA Summary:

	SS	df	MS	F	p-value
Source					
Regression	188.8755	7.0	26.9822	188.9908	0.0
Residual	30.1245	211.0	0.1428	NaN	NaN
Total	219.0000	218.0	NaN	NaN	NaN

Model Summary Statistics:
R-squared: 0.8624
Adjusted R-squared: 0.8579

4.3.4.2 Sample of model

```
=== ANOVA Table for 30% Sample Regression ===
```

ANOVA Summary:

	SS	df	MS	F	p-value
Source					
Regression	59.3811	7.0	8.4830	82.1213	0.0
Residual	8.1606	79.0	0.1033	NaN	NaN
Total	67.5416	86.0	NaN	NaN	NaN

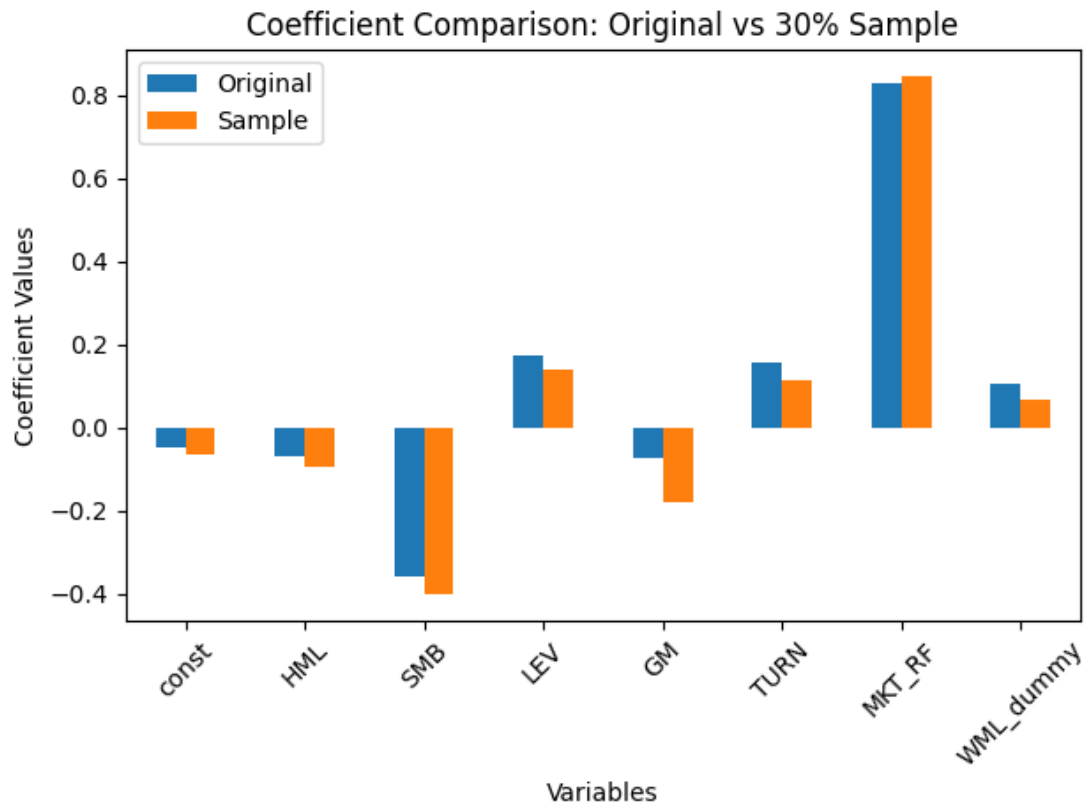
Additional Model Statistics:
R-squared: 0.8792
Adjusted R-squared: 0.8685

Mean Squared Error (MSE) for 30% Sample: 0.1030

Percentage change:

MSE change: $(0.1428 - 0.103) / 0.1428 = 27.87\%$

4.3.4.3 Coefficient Change:



Chapter5- Research Conclusions and Recommendations

5.1 Conclusion:

5.1.1 Problem remains:

- Doesn't meet the Assumption
- Not enough database, hard to get the long-term data like the US
- Hard to interpret for all factors

5.1.2 Comparison with the original five model

- RMW (not significant in Taiwan's semiconductor industry)
- CMA (based on the asset growth ratio) the missing value is remained to solve
- Market and Momentum are needed to consider

5.1.3 Recommendation for those who wants to invest in semiconductor industry

i Focus on Market Leaders

Insight: The model indicates that large-cap companies (negative SMB) in the Taiwanese semiconductor industry tend to outperform.

Recommendation:

Prioritize investments in established, large-cap firms with strong market dominance, as they are likely to offer more stable and consistent returns.

ii Consider Leverage as an Opportunity

Insight: High-leverage firms (positive LEV) perform better, potentially reflecting aggressive expansion strategies.

Recommendation:

Look for companies that effectively utilize leverage to fund growth and innovation, as these are more likely to generate higher returns.

iii Be Cautious of Value and Growth Dynamics

Insight: Growth stocks (negative HML) outperform value stocks in this sector.

Recommendation: Focus on companies with low book-to-market ratios (growth-oriented firms), which are more likely to thrive in the innovation-driven semiconductor industry.

iv Momentum Trends

Insight: High-momentum stocks (positive WML_dummy) outperform low-momentum ones.

5.2 Reference:

“Fama and French, “A Five-Factor Asset Pricing Model.”

(Fama & French, 2015)