An Investment Strategy in a Portfolio of Stocks Using Regression Analysis

12/8/2022

Linear Econometrics for Finance: *Final Project*

Group 6: *Peng Liu, Hao Yuan, Chengyi Xu, Chenyu Zhao, Yuchen Wang*

1. Executive Summary:

The purpose of this report is to establish a stock portfolio through the CAPM model combined with a risk-free interest rate and market interest rate, analyze the overall income of the portfolio, and compare the overall performance of the portfolio before and after COVID-19. We also adopted the Fame-French model which tries to explain the excess returns observed in some portfolios over and above those predicted by the Capital Asset Pricing Model (CAPM). The assumption of our overall portfolio is that the main factors affecting investment decisions are expected return and risk, and investors follow the Dominance rule, that is, under the same risk level, they choose securities with higher returns. With the same yield level, select the securities with lower risk.

The securities we choose to form our portfolio are AAPL, BA, AMZN, GOOG, MSFT, and S&P500. According to our research, the portfolio we decided on has performed statistically significantly well compared to the market return in the aspects of various comparison metrics (Sharpe Ratio, Jensen's Alpha, etc), before and during the era of COVID-19.

2. Security Selection and Data Source:

There are a few reasons why we invest in stocks included in the S&P 500 and the index S&P 500 itself:

a. Diversification: The S&P 500 includes 500 of the largest publicly traded companies in the United States, representing a wide range of industries. By investing in the S&P 500, investors can diversify their portfolios and reduce their overall risk.

b. Potential for growth: Many of the companies included in the S&P 500 are well-established and have a track record of steady growth. By investing in these companies, investors can potentially earn a return on their investment through capital appreciation.

Data Source:

The data source to support this report is Yahoo Finance and the Fama-French database.

a. Yahoo Finance is a financial news and data website that provides information on a wide range of financial topics, including stocks, mutual funds, currencies, commodities, and more. Yahoo Finance allows users to track the performance of individual stocks and other securities, as well as to access financial news, analysis, and other resources.
b. The Fama-French database is a database of historical stock market data that was developed by Eugene Fama and Kenneth French, two economists who have made significant contributions to the field of finance. The database includes data on stock returns, risk measures, and other financial variables for many publicly traded companies. It is used by researchers and analysts to study financial markets and to test investment strategies.

For Pre-Covid Era. we adopted the data with the time period from 2015/01/01 to 2020/12/31, for the Post-Covid Era, we adopted the time period from 2021/01/01 to 2021/12/31 from Yahoo Finance.
For Fama-French data, we adopted the data with the time period from 2015/01/01 to 2021/12/31.

3. Pre-Covid Individual Stock CAPM analysis:

The coefficients of a CAPM regression on individual stocks represent the relationship between the returns on the stock and the returns on the overall market. In a CAPM regression, the dependent variable is the returns on the individual stock and the independent variables are the returns on the market and the risk-free rate of return.

The beta coefficient, which represents the risk associated with the stock, is one of the most important coefficients in a CAPM regression. It measures the stock's sensitivity to movements in the market.

Other coefficients in a CAPM regression include the intercept term, which represents the expected return on the stock when the market return is zero, and the coefficient for the market return, which represents the expected return on the stock for every 1% increase in the market return.

According to our regression result in Exhibit 1. We can see that all of our securities except for the S&P500 index which is the market return itself, are having a beta higher than 1, which suggests that the stock is more volatile than the market.

On the other hand, the securities all have an intercept (Alpha) value greater than 1, indicating that the minimum return an investor can expect to earn on the stock is positive. Moreover, all securities are having a P-value smaller than 0.05, which indicates that they are statistically significant at the individual level. Therefore, they could be valid and investor-favored securities to form a portfolio, and it won't be any improvement to ignore anyone from these securities.

One thing worth noting is that only securities of AMZN and S&P passed the F-test, which may indicate that the observed differences or relationships could have occurred by chance and are not statistically significant.

4. Post-Covid Individual Stock CAPM analysis:

According to the regression result in Exhibit 2. We can see some of the stocks (AAPL, BA, GOOG) are having a beta lower than 1, indicating that the stock is less volatile than the market. BA is underperforming since it has a negative intercept value and a P-Value slightly larger than 5%, and MSFT performed a little bit less than other securities except for BA since it has an intercept value lower than 1, while all other stocks and securities are performing better than the Pre-Covid era.

Similarly, one thing worth noting is that only securities of AMZN and S&P passed the F-test, which may indicate that the observed differences or relationships could have occurred by chance and are not statistically significant.

5. Portfolio Weight Selection

To construct a portfolio, it is necessary to assign a weight for each security in the portfolio since one's investment quota is limited. In this research, we also consider the short strategy for investors which indicates investors can have negative weight on some certain securities and weight larger than 1 on some certain securities.

We assign random weights between (-2, and 2) for each security in the portfolio since normally people are able and willing to short and long 2 times higher or lower than their limit. We created 3000 portfolios with the randomly assigned weight in the range (-3 and 3) which adds up to 1. Then we ranked the portfolio in the criteria of expected return and obtain the weights for our further research.
Our decided weights for the portfolio are AAPL: -1.120543, BA: 0.437896, AMZN:1.814199, GOOG: -0.491866, MSFT: 1.913571, S&P:-1.553257 according to Exhibit 3.

6. Portfolio CAPM Analysis

After obtaining the weights of each security in the portfolio, we are able to construct the portfolio returns with regards to each weight. Then constructing a CAPM regression model with the portfolio returns in Exhibit 4. In this model, the Beta is 2.0723, indicating quite high volatility compared to the market return. The alpha is 5.9501, which is also considered high, indicating an excess return on this portfolio given the volatility. Moreover, the coefficients are all significant ($P<0.05$).

7. Portfolio Fama French Analysis

With the Fama French data, we also constructed a Fama French regression model with the factor of Mkt-Rf, SMB, HML, and MOM.
Market Risk Premium (Mkt-Rf): This is the excess return of the market portfolio over the risk-free rate, similar to the beta in the CAPM. A coefficient of 2.239 indicates high volatility, and 1 more unit increase of the Mkt-Rf rate representing a 2.239 units increase on the required return, other factors controlled.
Size (SMB): This factor captures the excess return of small-cap stocks over large-cap stocks. A coefficient of -1.355 indicates a portfolio favoring large-cap companies.
Value (HML): This factor captures the excess return of value stocks (stocks that are cheap based on measures such as price-to-book ratio) over growth stocks (stocks that are expected to grow at a faster rate than the market), a coefficient of -2.069 indicating a portfolio favoring growth stocks.
Momentum (MOM): This factor captures the excess return of stocks that have had strong past performance (momentum) over those with weak past performance. A coefficient of -0.065 slightly indicates that stocks with weak past performance have outperformed those with strong past performance.

8. OLS assumptions

The two models are based on OLS regression, which highlights the need of making sure the assumptions of OLS are met.

a. Linearity: The relationship between the predictor variables and the response variable should be linear.
b. Independence of errors: The errors (residuals) should be independent of each other, which means that the value of one error should not affect the value of another error.
c. Constant variance / Homoscedasticity: The variance of the errors should be constant across different fitted values.
d. Normality of errors: The errors should be normally distributed.

In order to check the assumptions, this report includes the residuals plots, the normality of variance plots, the normality of error plots, partial regression plots, the Breusch-Pagan heteroscedasticity test, the Rainbow linearity test, Jungbox Autocorrelation test (Exhibit 5). All assumptions are satisfied to support our previous analysis.

9. Ratio Comparison (CAPM / Fama French)

We obtained 3 metrics to compare the two models: Sharpe Ratio, Jensen's Alpha, and Treynor ratio. The portfolio has a Shape ratio of 0.3835, Jensen's Alpha of 5.76 - 5.86, and Treynor ratios of 3.198 - 3.455. Indicating outstanding performance in the aspects of risk-adjusted performance. While the Fama-French model has lower ratios of Jensen's Alpha and Treynor ratio due to a smaller beta1 (Mkt-rf). Exhibit 6 includes the ratios of the different methods of calculating the ratios for comparison.

10. Portfolio Performance Comparison (Pre / Post Covid)

Exhibit 7 showed greater volatility of our portfolio, but it is compensated by a higher possibility of return.

11. Conclusion

The portfolio we decided on has performed statistically significantly outstanding compared to the market return in the aspects of various comparison metrics (Sharpe Ratio, Jensen's Alpha, etc), before and during the era of COVID-19.

Exhibits:

Exhibit 1: Pre-Covid Individual CAPM results:

Stock	Intercept (Alpha)	Beta	P-Value (Beta)	F-test
AAPL	0.9344	1.3033	0.0000	0.1768
BA	0.6636	1.3198	0.0000	0.1761
AMZN	2.0517	1.5354	0.0000	0.0177
GOOG	0.9679	1.0137	0.0000	0.9351
MSFT	1.8696	1.1132	0.0000	0.4755
S&P	0.0000	1.0000	0.0000	0.0000

Exhibit 2: Post-Covid Individual CAPM results:

Stock	Intercept (Alpha)	Beta	P-Value (Beta)	F-test
AAPL	4.1058	0.9923	0.0000	0.9530
BA	-1.3718	0.7206	0.057	0.4197
AMZN	2.3517	1.2232	0.0000	0.1082
GOOG	1.0411	0.8820	0.0000	0.1902
MSFT	0.9816	1.0267	0.0000	0.7777
S&P	0.0000	1.0000	0.0000	0.0183

Exhibit 3: Portfolio Weights on Each security (Ranked by Expected Returns)

Returns	Volatility	AAPL weight	BA weight	AMZN weight	GOOG weight	MSFT weight	S&P weight
4.471090	2.963237	1.128386	-1.000564	1.631787	-0.946793	1.673242	-1.486058
4.441013	3.679425	1.886860	-1.573090	1.824442	0.019487	0.740035	-1.897733
4.415020	2.926533	-0.585150	-0.755135	1.680142	0.759467	1.591007	-1.690331
4.233448	3.682932	-0.724377	-1.530423	1.844859	1.911526	1.080872	-1.582457
4.001257	2.798526	-0.589406	-0.893570	1.275253	1.064601	1.649846	-1.506723
...	...	...	...	...	...	...	...
-1.449852	2.095892	-1.027735	0.950802	-1.660919	1.524639	-0.143328	1.356541
-1.455006	2.322513	-0.189110	1.552003	-1.044935	-0.271244	-0.754678	1.707965
-1.681400	3.204142	1.479719	1.930374	-1.196340	-0.513126	-1.773056	1.072429
-2.043581	2.550931	0.589834	1.469570	-1.605463	0.783273	-1.507866	1.270653
-2.149761	3.657058	1.960102	1.763350	-1.760352	-1.041238	-1.413305	1.491442

Exhibit 4: Portfolio CAPM regression model

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                ex_ind1   R-squared:                       0.159
Model:                            OLS   Adj. R-squared:                  0.145
Method:                 Least Squares   F-statistic:                     11.17
Date:                Mon, 19 Dec 2022   Prob (F-statistic):            0.00145
Time:                        01:44:07   Log-Likelihood:                -259.30
No. Observations:                  61   AIC:                             522.6
Df Residuals:                      59   BIC:                             526.8
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      5.9501      2.244      2.652      0.010       1.460      10.440
mktrf          2.0723      0.620      3.342      0.001       0.831       3.313
==============================================================================
Omnibus:                        5.589   Durbin-Watson:                   2.301
Prob(Omnibus):                  0.061   Jarque-Bera (JB):                4.677
Skew:                           0.638   Prob(JB):                       0.0965
Kurtosis:                       3.459   Cond. No.                         3.68
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
<F test: F=2.9898036664285024, p=0.08901997086535143, df_denom=59, df_num=1>
```

Exhibit 5

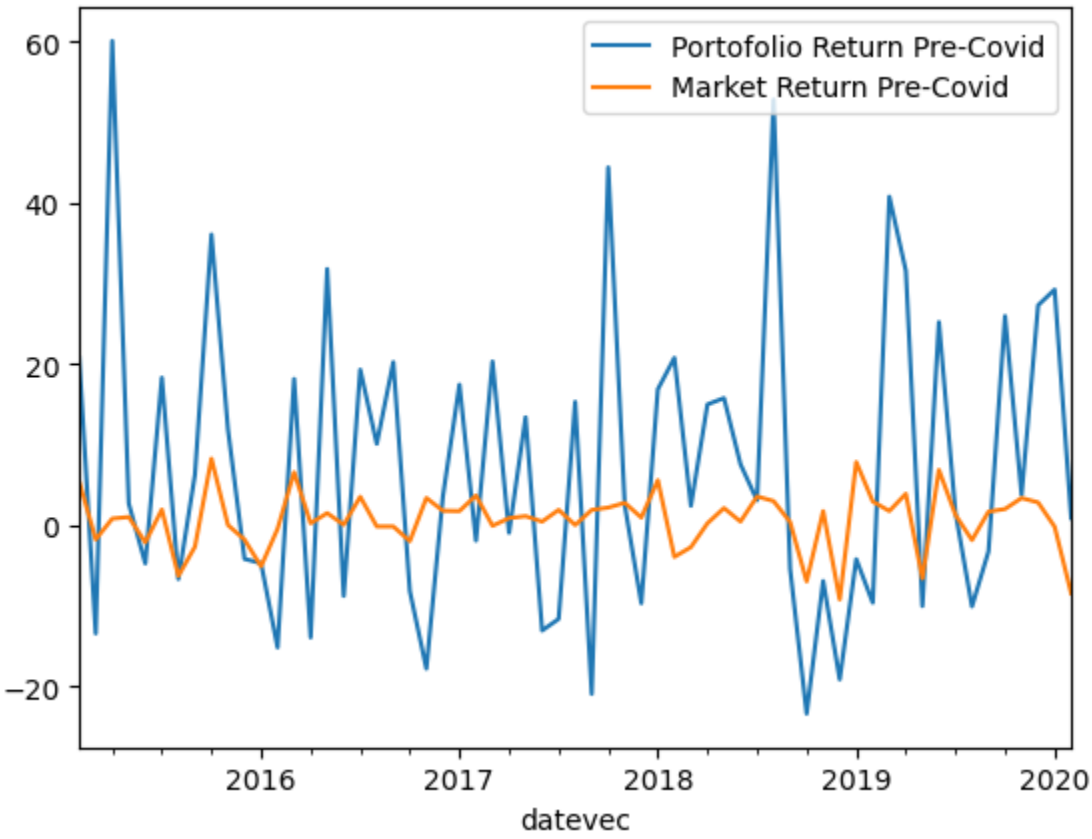
Check Attached file

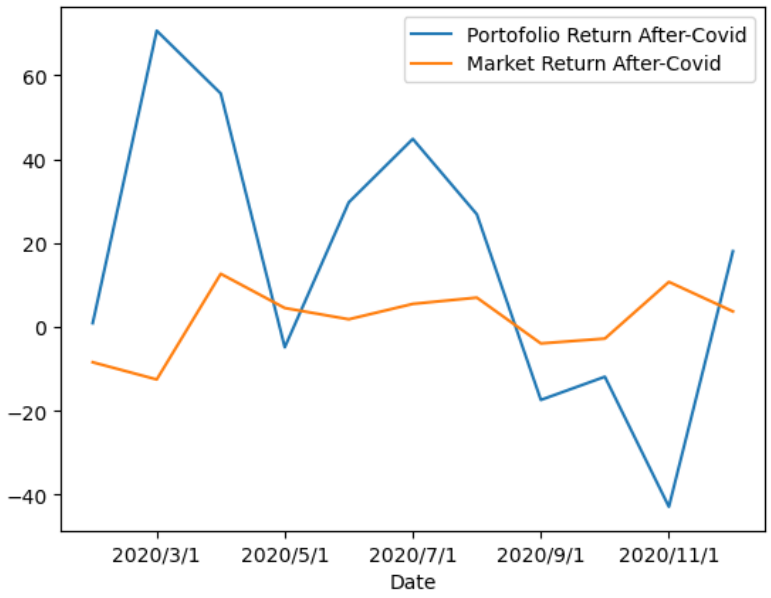
Exhibit 6.

	SR	JA	TI
CAPM_portfolio	0.383505	5.865028	3.4549

	SR	JA	TI
FF_portfolio	0.383505	5.760994	3.197892

Exhibit 7:




Portofolio Return After-Covid
Market Return After-Covid
60
40
20
0
−20
−40
2020/3/1
2020/5/1
2020/7/1
2020/9/1
2020/11/1
Date