DATA ANALYTICS FOR FINANCE: FINAL PROJECT

The Piotroski F-Score was developed by Joseph Piotroski, a Stanford accounting professor. The score is a discrete result between 0 and 9, and it evaluates the following elements (with sub-conditions):

Profitability:

1. ROA (+1 point if positive)

2. Operating Cash Flow (+1 if positive)

3. Δ ROA (+1 if positive)

4. CFO / Total Assets > ROA (+1 if satisfied)

Leverage, liquidity, and source of funds:

1. Δ Leverage (+1 if negative)

2. Δ Current Ratio (+1 if positive)

3. Δ Shares (+1 if negative)

Operating Efficiency:

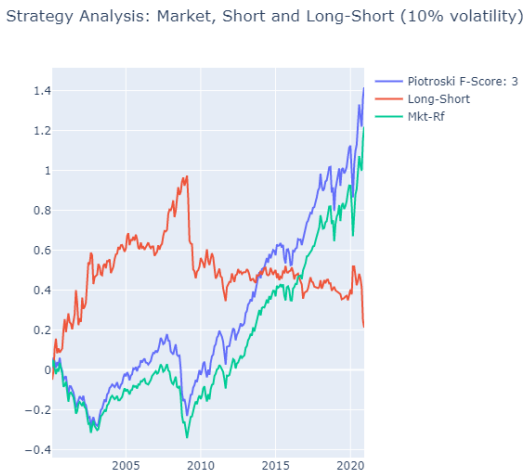
1. Δ Gross Margin (+1 if positive)

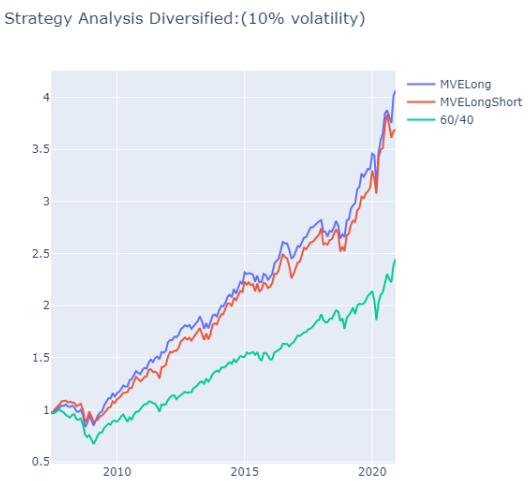
2. Δ Asset Turnover (+1 if positive)

A score of 0-2 is considered weak, a score of 3-7 is average, and a score of 8-9 is considered strong. The ratio takes three relevant factors to economic prosperity (profitability, liquidity, and efficiency) and allows us to see which companies perform well based on those characteristics. Profitability allows shareholders to be more rewarded and for the business to grow further. Leverage is an important factor, as when it is too high it can lead to shareholder value destruction. Liquidity allows for more flexible and dynamic reactions to the competitive environment. And efficiency is a testimonial of competitive advantage and positioning. Based on this, it is at least a predictor of an improving balance sheet, although it is not one of a cheap business.

In his paper "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers", Joseph Piotroski used a sample of high book-to-market stocks. These stocks could be either cheap, financially distressed or both. Therefore, working a formula based on the 3 previously shown factors gives a downside protection from bankruptcy or operational failures. In his analysis -despite potential data snooping bias-, high F-Score stocks scored an average return higher by 9.7% on average (over a 20-year period, with 12-month buy-and-hold).

One of the limitations of this strategy, is that it focuses on accounting captions and ratios. These are heavily affected by accounting standards (e.g. U.S.GAAP vs IFRS), faithful representation, relevance, and other quality characteristics of financial statements. Of course, these risks can be negated by appropriate diversification. Another criticism is that some industries will be more prone to have higher scores than others. For instance, a mature industry will have a harder time improving its margins than a DTC industry or software industry. Therefore, one needs to careful between economic significance and accounting or statistical significance.





When looking at the results of our study it does seem that the Piotroski-F Score strategy has a competitively sound aspect. There have been three crashes in our sample (2000, 2007–2008, and 2020) and this makes our sample hard to get statistically relevant info from. We propose:

- Another study for another sample.

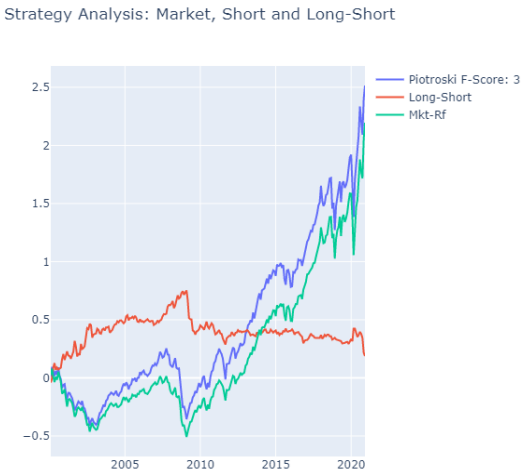
- Combining our signal with other signals that cover us from accounting and misinformation risk such as the Beneish M-Score (Valaskova and Fedorko 2021).

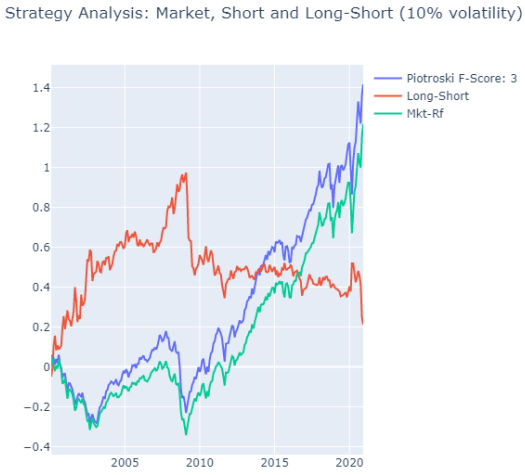
- Studying the effect of leverage and transactions costs further to see the feasibility of this strategy.

Strategy analysis

For this analysis the data will be first cleaned, and the missing signals will be eliminated. After that, the investment universe will be mapped. Filter conditions will be set for our study to ensure that there are no logical flaws in our data. Our sample ranges from the id stock data's lowest date to the dataset's last month. The monthly risk-free rates, as well as the excess returns, will be calculated and set away for the remainder of the study. The *assignSignalSort* method also helped us split our strategy data into 3 terciles based on the signal.

Afterwards, portfolios will be assigned based on a value-weighted approach. The reasons why a value-weighted approach is chosen are liquidity (access to stocks with better data) and lower transaction costs. The top tercile will be the basis for our long portfolio, and a combination of the lowest tercile and the top tercile will be the basis for our long-short portfolio. A value-weighted market portfolio will also be used to have an idea about the general market.





The graphs give us the following conclusions:

- The Long Strategy has the best returns (superior to the market). The Long-Short the worst returns.
- There is a clear schism between the first half of our data (which has been affected by 2 crises) and the second half (good market conditions and covid-19 crisis).
- When standardizing the volatility across portfolios the conclusions remain similar.

Finally, we can generate the following tables to get a quick summary of our backtests:

Piotroski F-Score Long	1	2	Full
Annualized Excess Return	0.84%	**13.59%**	7.19%
Annualized Sharpe Ratio	0.05	**0.94**	0.47
Alpha CAPM	**0.15%**	-0.08%	0.06%
Alpha FF3	**0.19%**	-0.16%	0.07%
T-stat Alpha CAPM	1.60	-1.22	1.10
T-stat Alpha FF3	**2.02**	**-2.58**	1.28
CAPM IR	**0.50**	-0.39	0.24
FF3 IR	**0.65**	-0.87	0.28
Piotroski F-Score Long Short			
Annualized Excess Return	**3.95%**	-1.67%	1.15%
Annualized Sharpe Ratio	**0.41**	-0.27	0.14
Alpha CAPM	**0.30%**	-0.04%	0.23%
Alpha FF3	**0.33%**	-0.21%	0.25%
T-stat Alpha CAPM	1.52	-0.25	1.76
T-stat Alpha FF3	1.71	-1.37	**1.98**
CAPM IR	**0.47**	-0.08	0.39

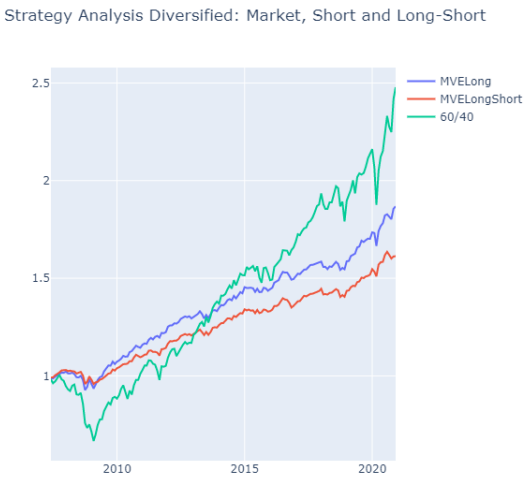
FF3 IR	0.55	-0.46	**0.98**
Market Portfolio			
Annualized Excess Return	-1.06%	**14.78%**	6.83%
Annualized Sharpe Ratio	-0.06	**1.02**	0.43

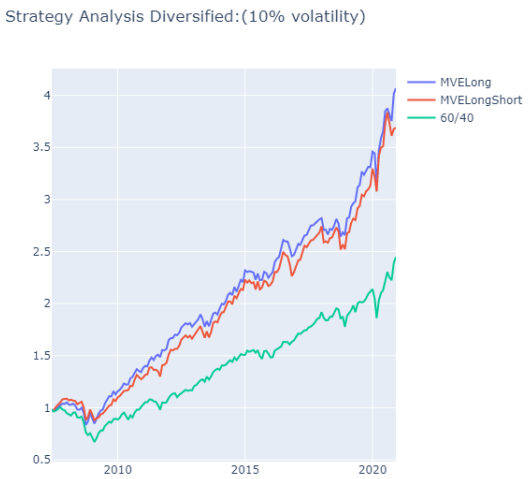
The conclusions and remarks that can be established after looking at the table are:

- There is a lack of uniform statistical significance (at a 95% level) for our portfolios overall (only 25 % of the alpha data is significant). Using another sample or a bigger one might allow us to obtain more statistically significant results. The FF3 model shows to be more explicative as it has three factors, and this is confirmed by results closer to statistical significance.
- The highest Sharpe Ratio over the full sample is the long portfolio one, meaning that the long strategy did beat the market. This is again confirmed by the positive alpha. However, this alpha is not significant.
- The long, short portfolio has the highest information ratio due to the low volatility thanks to the low denominator weight. Since stock betas converge to 1 (Brigham Young University 2022), idiosyncratic risk will be closer to 1.

Strategy as part of a diversified portfolio

The tangency portfolio is a portfolio that uses our investing universe to optimize the Sharpe Ratio (Ramirez-Hassan and Guerra-Urzola 2020). We'll have to identify this portfolio utilizing our long-only portfolio and other assets, as well as our long-short portfolio and other assets, for this section of the assignment. The VTI (Vanguard Total Stock Market) and the BND (Vanguard Total Bond Market Index Fund ETF) ETFs make up the other asset category. To establish our tangency portfolios, we will need to prepare a covariance matrix and a mean dataset with our asset cocktail. We will also compare this portfolio to a combination of the ETF's BND and VTI (40%, 60% respectively).





The MVELong weights are: "BND": 0.8365107775; "Long": -0.04267721623; "VTI": 0.2061664164 and the MVELongShort weights are: "BND": 0.6194846034; "Long Short": 0.2142294347; "VTI": 0.1662859619. One aspect of mean variance investing that might not be economically coherent is that it can push the investor or speculator to short an economically sound strategy, as it is the case here.

The graphs give us the following conclusions:

- The 60/40 portfolio has the highest returns since it has the highest market weight.
- When standardizing the volatility, we can observe that the MVELong and the MVELongShort dominate. However, the MVELong is better since shorting usually involves transaction costs.
- An out of sample test is necessary as the mean variance portfolios can be misleading. Using the weights set in a specific year over the rest of a sample might give more information.

Finally, we can generate the following tables to get a quick summary of our backtests:

	Annualized Excess Return	Annualized Sharpe Ratio
MVELong	4.70%	**1.09**
MVELongShort	3.59%	0.83
60/40	**7.22%**	0.71

The highest Sharpe ratio belongs to MVELong portfolio; however, this portfolio is an inverse of the economically sound strategy that we proposed as it shorts the top tercile. This portfolio is more of an 80/20 portfolio, meaning that higher weight in bonds although lowering returns allows for a more efficient (per volatility) portfolio.

References

Brigham Young University. 2022. *Bloomberg Guide: Beta.* January 14. https://guides.lib.byu.edu/c.php?g=216390&p=1428678.

Ramirez-Hassan, Andres, and Rosember Guerra-Urzola. 2020. "Optimal portfolio choice: a minimum expected loss approach." *Mathematics and Financial Economics* 97–120.

Valaskova, Katarina, and Richard Fedorko. 2021. "Beneish M-score: A measure of fraudulent transactions in global environment." *Globalization and its Socio-Economic Consequences.*