



MA574 Portfolio Valuation and Risk Management

Fall 2017

Professor Marcel Blais

Final Course Report

Khasan Dymov

Alexander Shoop

Contents

1 Abstract	3
2 Introduction	3
2.1 Background	3
2.2 Methodology	4
3 Portfolio Management	9
3.1 General Process	9
3.2 Portfolio Analysis	10
3.2.1 General Performance Overview	10
3.2.2 Returns	14
3.2.3 Sharpe Ratio	15
3.2.4 Treynor Ratio	15
3.2.5 Information Ratio - S&P500	16
3.2.6 Sortino Ratio	16
3.2.7 Maximum Drawdown	17
3.2.8 Portfolio Alpha & Beta	18
3.2.9 Comparison to Industry Benchmarks	19
3.2.10 Margin Call Analysis	19
3.2.11 Leverage Analysis	19
4 Capital Asset Pricing Model	22
4.1 Traditional Approach	22
4.2 Least Trimmed Squares	24
4.3 Constant β Shrinkage Estimation	25
4.4 Exponentially Weighted Moving Average (EWMA)	25
4.5 Comparison and Analysis	26
5 Factor Modeling	27
5.1 CAPM	29
5.2 French & Fama	30
5.3 CAPM Extension - F_1	31
5.4 CAPM Extension - F_2	32
5.5 CAPM Extension - F_1 & F_2	33
5.6 French & Fama Extension - F_1	34
5.7 French & Fama Extension - F_2	35
5.8 French & Fama Extension - F_1 & F_2	36
5.9 Recap	37
5.10 Portfolio Optimization	37
5.10.1 Backtesting Results	38
5.11 Quantile-Based Portfolio Back-testing	41

5.11.1 Backtesting Results	43
6 Conclusion	45
Appendix A Portfolio Rebalancer.py	46
Appendix B FactorModels_wFamaFrench.py	50
Appendix C FactorModels_F1F2.py	54
Appendix D EF_Portfolio_multiWeek_plotting.py	62
Appendix E CAPM_altBeta.py	66
Appendix F Full Trade History	71

1 Abstract

The purpose of this project is to implement the mathematics and techniques learned in the WPI course MA 574: *Portfolio Valuation and Risk Management* to a virtual portfolio of assets. In particular, the topics focused on in this report are Markowitz Portfolio Optimization (also known as Markowitz Portfolio Theory, or MPT), the Capital Asset Pricing Model (CAPM), and Factor Modeling.

2 Introduction

2.1 Background

The inspiration for this project came from Harry Markowitz's Modern Portfolio Theory (MPT) covered in our class textbook *Statistics & Finance - An Introduction* by David Ruppert. This theory, which has served as a foundation for economists and investment advisors for over six decades, was published in the *Journal of Finance* in 1955 [1]. Interestingly enough, when Harry Markowitz presented what would become MPT while defending his dissertation at the University of Chicago School of Economics, economist Milton Friedman argued that portfolio theory was not economics, and thus, Markowitz should not be awarded his Ph.D [1]. Fortunately for Markowitz, the committee approved his work and he was awarded his Ph.D. In 1990, nearly 40 years after publication, Markowitz received the Nobel Prize in Economics for his work, indicating its incredible value in modern economics.

The central idea of Markowitz's Modern Portfolio Theory is actually relatively simple: for an investor constructing a portfolio, the ultimate goal (if we assume investors are rational...) is to maximize expected returns and minimize risk. Mathematically, this translates to maximizing a portfolio's mean return and minimizing the portfolio's return variance. The mathematics in MPT boil down to a non-linear optimization problem, where an investor seeks to find the optimal weights for n risky assets in their portfolio that satisfy their requirements for risk and return.

Our goal was to apply Markowitz's MPT to our own portfolio and analyze the results. From there, we would critique our portfolio performance through various portfolio metrics.

But we also took it further. William Sharpe, a student of Harry Markowitz, built upon Markowitz's MPT and developed his own model, known as the Capital Asset Pricing Model, or CAPM for short [2]. This model "describes the relationship between systematic risk and expected return for assets, particularly stocks" [2]. We will apply this model to our portfolio and data and analyze the results. Additionally, we will perform various factor modeling on our portfolio, to compare to William Sharpe's Capital Asset Pricing Model. In particular, we will focus on the French & Fama Three Factor Model, a model that "expands on the capital asset pricing model (CAPM) by adding size and value factors to the market risk

factor in CAPM" [3]. We will even make adjustments to the Fama and French Model with the addition of several new factors.

2.2 Methodology

As discussed in our prior report, Markowitz Portfolio Optimization - *Phase 1*, we selected 20 stocks based mostly on personal interest and familiarity. Below are the stocks we selected, with their respective tickers:

1. Amazon Inc - AMZN
2. AutoZone Inc - AZO
3. The Boeing Company - BA
4. Cigna Corporation - CI
5. Costco Wholesale Corp. - COST
6. Corbus Pharmaceuticals - CRBP
7. Tableau Software Inc - DATA
8. Ford Motor Company - F
9. Alphabet Inc - GOOG
10. Hubspot Inc - HUBS
11. JPMorgan Chase & Co - JPM
12. Manulife Financial Corp - MFC
13. 3M Company - MMM
14. Vail Resorts Inc - MTN
15. Netflix Inc - NFLX
16. NIKE Inc - NKE
17. Starbucks Corp - SBUX
18. iShares 20+ Year Treasury Bond ETF - TLT
19. Vanguard High Dividend Yield Fund - VYM
20. Waste Management Inc - WM

However, for *Phase 2* of this project, with implementation in Interactive Brokers, we also needed an asset to serve as our "risk-free" asset. As discussed in MA 574, a "risk-free" asset is an asset with some return and zero risk or volatility. This is the purely mathematical definition - in reality, there is no such thing, as some level of risk exists in *every* investment. However, United States government bonds (Treasuries) are typically regarded as the closest proxy to a risk-free asset. As such, we explored different options for U.S. government bonds.

The image below, sourced from TreasuryDirect.gov, shows a snapshot of U.S. Notes of varying maturity and yield that were available at the time we were shopping. For the purposes of our portfolio, and given the equities in our portfolio, we felt that a five year note would be appropriate. As such, we selected the following bond (indicated in red in the image):

1. CUSIP: 9128283C2
2. Yield: 2.000%
3. Maturity: 10/31/2022

Note: The following table displays data for the 20 most recently auctioned securities that have not yet matured, for each security type. If you would like to see data for additional securities, please use our [Auction Query](#).

Bills	CMBs	Notes	Bonds	TIPS	FRNs	
Security Term	CUSIP	Reopening	Issue Date	Maturity Date	High Yield	Interest Rate
7-Year	9128283D0	No	10/31/2017	10/31/2024	2.280%	2.250%
5-Year	9128283C2	No	10/31/2017	10/31/2022	2.058%	2.000%
2-Year	912828F62	Yes	10/31/2017	10/31/2019	1.596%	1.500%
3-Year	912828Z22	No	10/16/2017	10/15/2020	1.657%	1.625%
10-Year	912828R0	Yes	10/16/2017	08/15/2027	2.346%	2.250%
7-Year	912828Y5	No	10/02/2017	09/30/2024	2.130%	2.125%
5-Year	912828W9	No	10/02/2017	09/30/2022	1.911%	1.875%
2-Year	912828X7	No	10/02/2017	09/30/2019	1.462%	1.375%
10-Year	912828R0	Yes	09/15/2017	08/15/2027	2.180%	2.250%
3-Year	912828V1	No	09/15/2017	09/15/2020	1.433%	1.375%
7-Year	912828U3	No	08/31/2017	08/31/2024	1.941%	1.875%
2-Year	912828T6	No	08/31/2017	08/31/2019	1.345%	1.250%
5-Year	912828S8	No	08/31/2017	08/31/2022	1.742%	1.625%
10-Year	912828R0	No	08/15/2017	08/15/2027	2.250%	2.250%
3-Year	912828Q2	No	08/15/2017	08/15/2020	1.520%	1.500%
7-Year	912828N9	No	07/31/2017	07/31/2024	2.126%	2.125%
5-Year	912828P4	No	07/31/2017	07/31/2022	1.884%	1.875%
2-Year	912828K5	No	07/31/2017	07/31/2019	1.395%	1.375%
10-Year	912828X88	Yes	07/17/2017	05/15/2027	2.325%	2.375%
3-Year	912828J8	No	07/17/2017	07/15/2020	1.573%	1.500%

Figure 1: US Treasury Note Selection, sourced from Treasurydirect.gov [4]

Once we selected all of our assets, we constructed a spreadsheet (in Google Sheets) to keep track of our portfolio. This spreadsheet, *MarketPortfolio.xlsx* can be found in the *.zip*-file associated with this project report. In this spreadsheet, we tracked our check-in times, portfolio value (net liquidation value), excess liquidity, equity value, bond value, equity weight, bond weight, prices, sizes of positions, total position value, and individual equity weights. We also kept track of the Markowitz Optimization outputs computed by our Python script *Portfolio Rebalancer.py* (this script can be found in Appendix A), most notably, the expected return and risk of the tangency portfolio, as well as the individual asset weights. Additionally, we tracked the specific adjustments necessary to each asset position to satisfy the tangency portfolio weights.

Our script also computed and plotted the Markowitz Efficient Frontier, tangency portfolio, and Capital Market Line. We saved the plots for the initial portfolio setup, as well as the plots for every weekly rebalance. Below are the plots for the initial portfolio setup (computed based on a 2-year lookback):

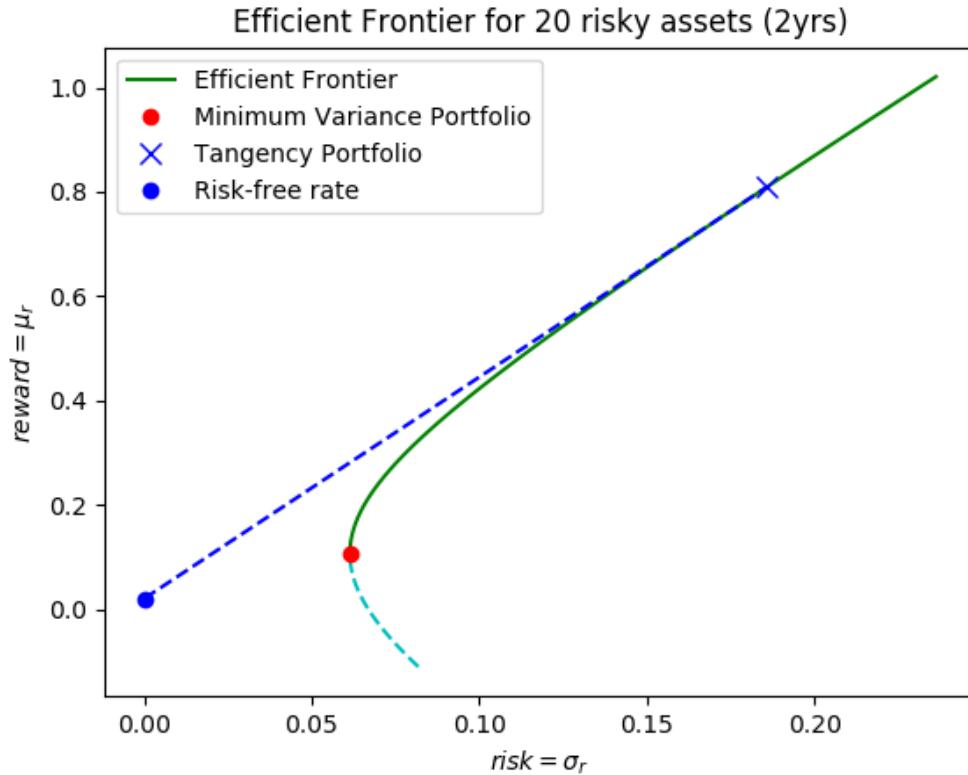


Figure 2

The tangency portfolio, indicated by the "X," had the following computed annual risk and return:

$$\sigma_r = 0.1857112219, \mu_r = 0.8089575445$$

This level of annual return is *extremely* high for an equity portfolio, but the corresponding level of volatility, approximately 18.6%, seems appropriate. Our bond, indicated by the blue dot, sits at the point (0,0.02), based on the risk-free assumption and the yield of 2.00%. Our initial portfolio then is equally split between the bond and the equity portfolio, with $\approx \$500,000$ invested in each. This then moves our overall portfolio down the Capital Market Line (indicated by the blue dashes) exactly equidistant between the bond and the tangency portfolio. Our overall portfolio then, has initial parameters:

$$\sigma_p = 0.09285561093, \mu_p = 0.4044787723$$

We also looked at the weights for the initial tangency portfolio - partially because we needed to know them to form our positions in Interactive Brokers, and partially because we were surprised at the values for risk and reward for the tangency portfolio. The bar chart below shows the initial tangency portfolio weights:

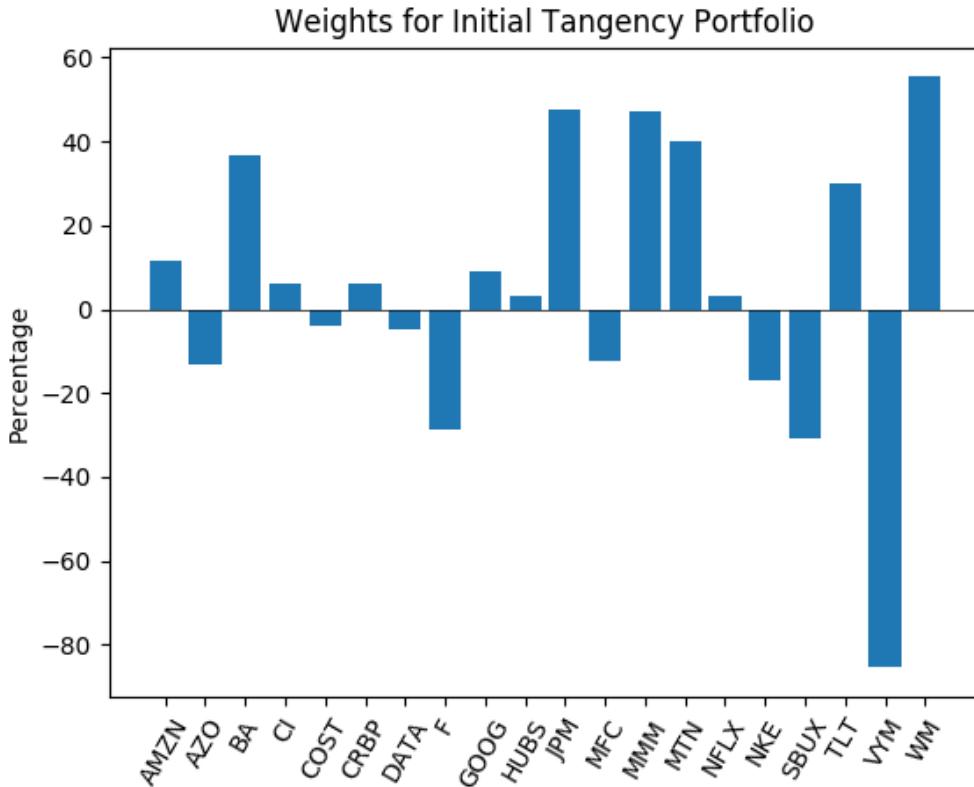


Figure 3

The results are rather interesting. First of all, the extremely high expected return now

makes more sense. Given that the Markowitz optimization is unconstrained, the routine decided to build a fair amount of leverage into the portfolio - as can be seen by some of the very large positions (the sum of the weights is in fact 1, so the results are valid). The most curious position is the extremely large short position in VYM, the Vanguard High Dividend Yield fund, at -85.44%. This fund has very high correlation with the S&P500, so it is interesting that the Markowitz routine chose to essentially short the market in order to place large bets on other assets.

Out of curiosity, we compared these asset weights to the weights for the minimum variance portfolio for these 20 assets, based on the same 2-year lookback. Here we see that the majority of the positions have positive weight, and most are fairly low - with the exception of the positions in TLT and VYM. This comparison does shed some light on the situation - it appears that the Markowitz routine believes a large long position in VYM is useful for reducing portfolio volatility. Thus, forming a large short position in VYM in order to place bets on other assets is the safest approach as the volatility itself is not very volatile. It is interesting however that this was not also the case for TLT - perhaps there is more going on that is immediately evident.

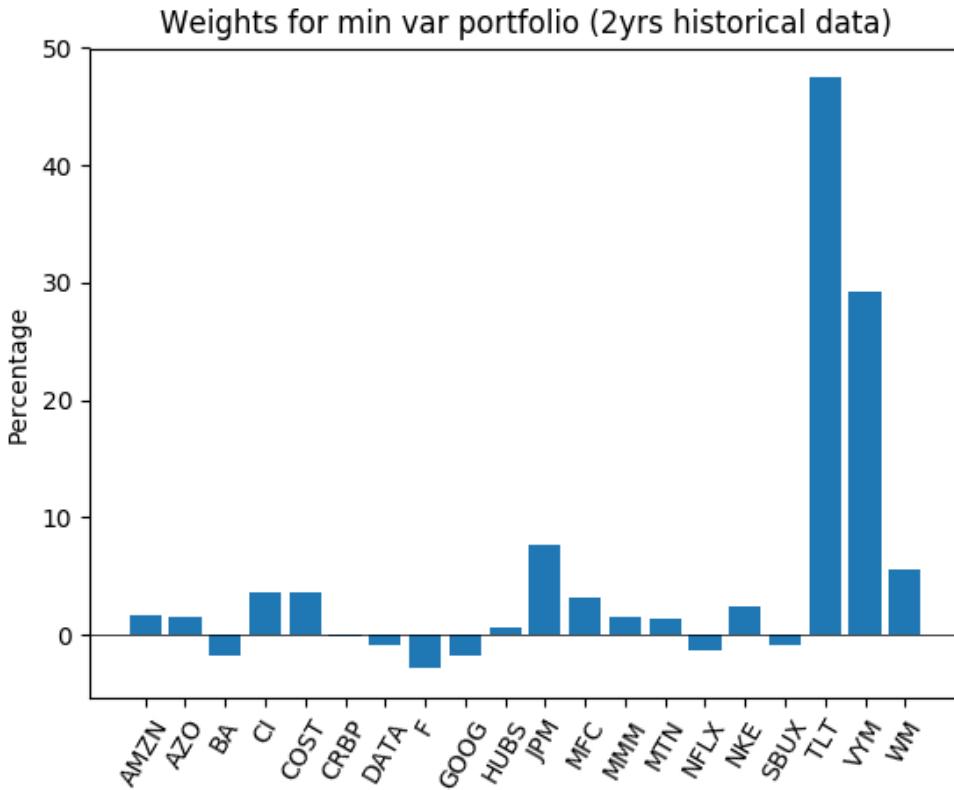


Figure 4

After the initial portfolio setup, rebalancing was done on a weekly basis and the portfolio was tracked in our MarketPortfolio spreadsheet.

The portfolio was closed on December 4th, 2017. Once this was completed, all of the analysis, CAPM modeling, and Factor modeling began.

3 Portfolio Management

3.1 General Process

After our initial portfolio setup, we developed a routine for managing the portfolio with regard to the project guidelines. After every Friday market close, and before market open on Monday, we recorded all the necessary data and made any necessary adjustments. First, all positions were recorded in our *MarketPortfolio.xlsx* file - we tracked the closing price, the size, the net value of the position, and the resulting asset weights in the portfolio. We also recorded the portfolio net liquidation value as well as the excess liquidity, a habit we picked up from the liquidity and margin project from earlier in the course. Once this was complete, we ran our Markowitz optimization script (with a rolled-forward-by-1-week adjusted 2-year window), which can be found in Appendix A: Portfolio Rebalancer.py. This gave us the new weights for the tangency portfolio, as well as the expected return and volatility of the tangency portfolio. From this output, we were able to calculate the necessary adjustments to each position. Lastly, we recorded the adjustment to each position in the sheet, and placed the orders in Interactive Brokers by market-open on Monday.

We found that rebalancing the portfolio every week became very tedious, as we had to take the new weights computed for the tangency portfolio and manually calculate the adjustment to the position in each of our securities (except for the bond position). This was very laborious and had lots of room for human error, so we wrote a small function that computed the adjustments for us (note: the function PortfolioRebalancer is just one part of the script file PortfolioRebalancer.py):

```
1 def PortfolioRebalancer(currwts, newwts, currcashval, lastcloseprices, tickers
2   ↪ ):
3     #function takes as input current stock weights in portfolio,
4     #new calculated tangency portfolio weights, the current cash value
5     #of the portfolio, latest closing prices for all the stocks, and tickers
6
7     newcashvalue = np.array(np.divide(np.multiply(newwts, currcashval),
8       ↪ currwts))
9     cashvalchange = np.array(np.subtract(newcashvalue, currcashval))
10    sharechange = np.array(np.divide(cashvalchange, lastcloseprices))
11
12    for x in range(0,20):
13      print("Adjust position in " + tickers[x] + " by " + str(sharechange[x
14        ↪ ]) + " shares.")
```

Figure 2 below is a small sample of the output of this function.

```
Adjust position in AMZN by [ 6.19736942] shares.  
Adjust position in AZO by [ 3.88447357] shares.  
Adjust position in BA by [ 32.37483384] shares.  
Adjust position in CI by [ 42.51440625] shares.
```

Figure 5

And so our weekly rebalancing process was significantly streamlined - we simply had to update our Excel sheet, where we were keeping track of our portfolio adjustments, after market close on Friday, adjust the date range in our Python script, and then based on that output, place the orders in Interactive Brokers. As a matter of fact, it is even possible to automate this last step of order placement through the IB Excel API; however, we decided to not implement this due to time constraints (ironic...). Lastly, it should be noted that it is not possible in Interactive Brokers paper trading accounts to trade fractional shares - as such, we adjusted our positions by truncating the values produced by our rebalancing calculations (e.g. A computed adjustment of 6.19736 shares resulted in a transaction of 6 shares).

Our full trade history, sourced from Interactive Brokers' built-in reporting capabilities[5], can be found in Appendix D.

3.2 Portfolio Analysis

3.2.1 General Performance Overview

The screenshot below, sourced from Interactive Brokers, shows our final portfolio value, as well as the realized profit and loss for each asset, after we completely closed our portfolio. The final portfolio value of \$988,083 indicates a loss from \$1,000,000 of approximately 1.19%. The individual asset returns seemed to be a mix between profits and losses. Our largest loss came from VYM, which as noted before, had a very large short position in the beginning. This fact remained constant throughout the lifetime of the portfolio - the Markowitz optimization routine always wanted to massively short VYM. It is also worth noting that our bond position lost \$2,281.

MA574 Final Report

The screenshot shows the Interactive Brokers SIMULATED TRADING interface with the account number DU794875. The main window displays various financial metrics under different sections:

- Balances:** Shows Net Liquidation Value (988,083 USD), Equity with Loan Value (988,083 USD), Cash (989,790 USD), and MTD Interest (-1,707 USD) across US Securities and US Commodities.
- Margin Requirements:** Shows Current Initial Margin (0 USD) and Current Maintenance Margin (0 USD).
- Available for Trading:** Shows Current Available Funds (988,083 USD), Current Excess Liquidity (988,083 USD), SMA (2,032,554 USD), and Buying Power (3,952,332 USD).
- Market Value - Real FX Balance:** A table showing currency-wise breakdown of cash, MTD interest, stock, options, futures, FOPs, unrealized P&L, and realized P&L.
- FX Portfolio - Virtual FX Position:** A table showing positions for various stocks like WM, VYM, US-T GOVT Note, TLT, SBUX, NKE, NFLX, MTN, MMM, MFC, JPM, HUBS, GOOG, F, DATA, CRBP, COST, CI, BA, AZO, and AMZN.
- Portfolio:** A detailed table of positions with columns for Fnnd Instrmnt, Position, Currency, Market Value, Market Price, Average Price, Unrealized P&L, Realized P&L, and Liquidate Last.

At the bottom left, it says "Last updated at 12:52".

Figure 6: Final Portfolio Statistics & Realized P&L from Interactive Brokers

The plot below shows all of the efficient frontiers, our realized portfolios (mix of stocks and bond), and the capital market lines we computed throughout the five weeks of portfolio management. Opacity indicates the time - less opaque plot elements are further in the past. Based on this, we see that our realized portfolio (the red dot), consistently trended left and downwards from the capital market line each week. In fact, the overall portfolio expected return for our last week of management is nearly 50% down from the initial expected return. We believe this is due to the fact that as the stock portion of the portfolio lost more money than the bond, the resulting weight of the bond value in the portfolio dragged the return down the CML towards the point (0,0.02). It is worth mentioning that, as expected, none

of the red dots sit directly on their respective CMLs. This is because as prices fluctuated during the trading week, the individual asset weights in the portfolio changed, turning the portfolio "inefficient" based upon the definition in the Markowitz model.

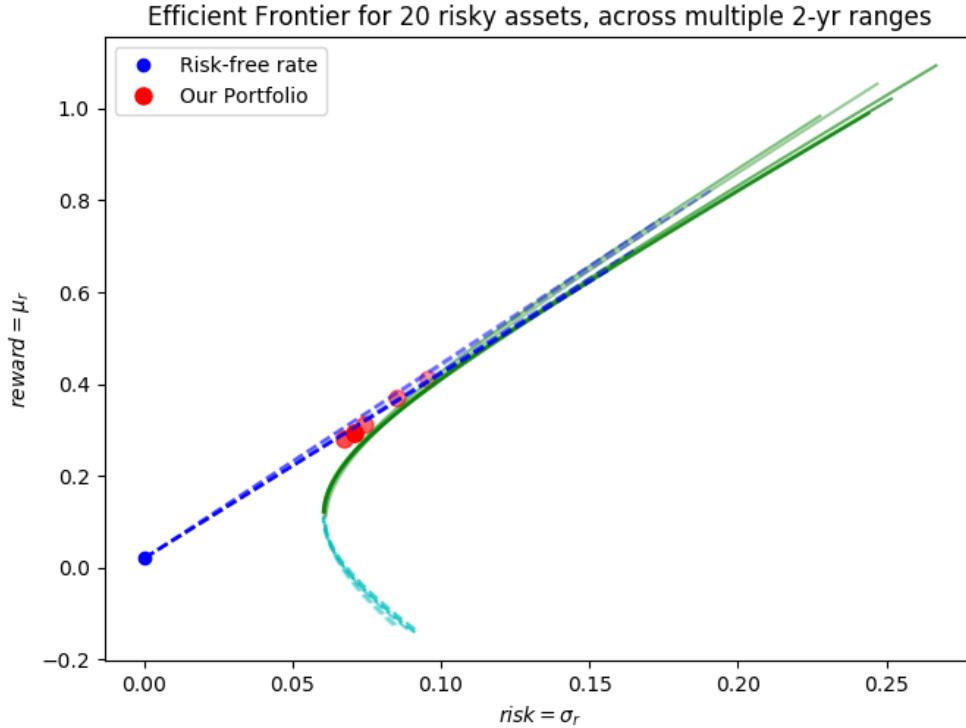


Figure 7

The two plots below indicate the weekly portfolio value and weekly portfolio return.

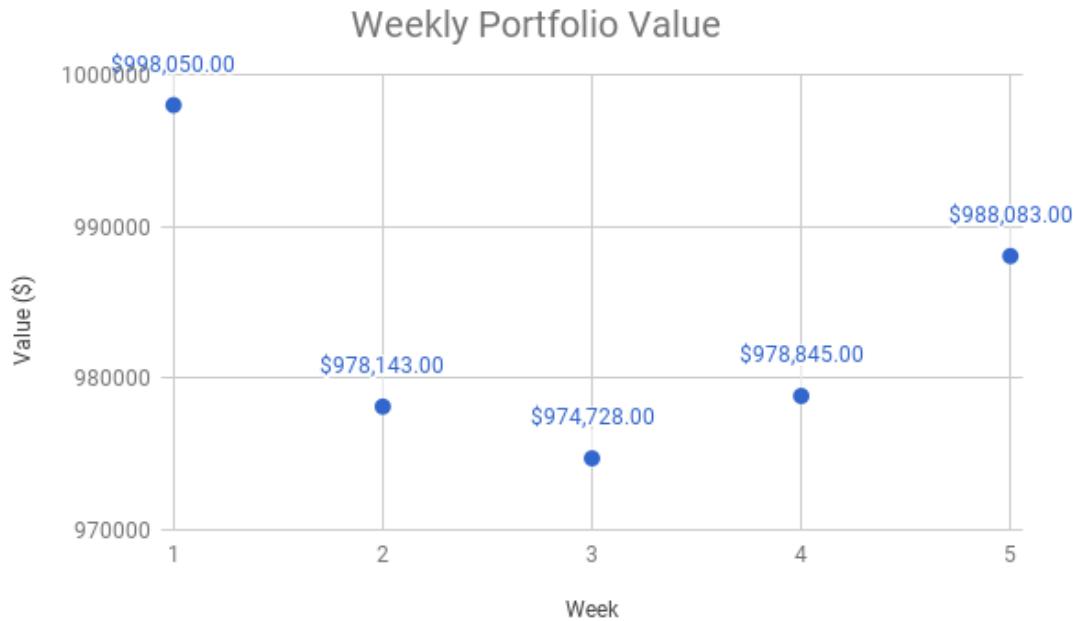


Figure 8

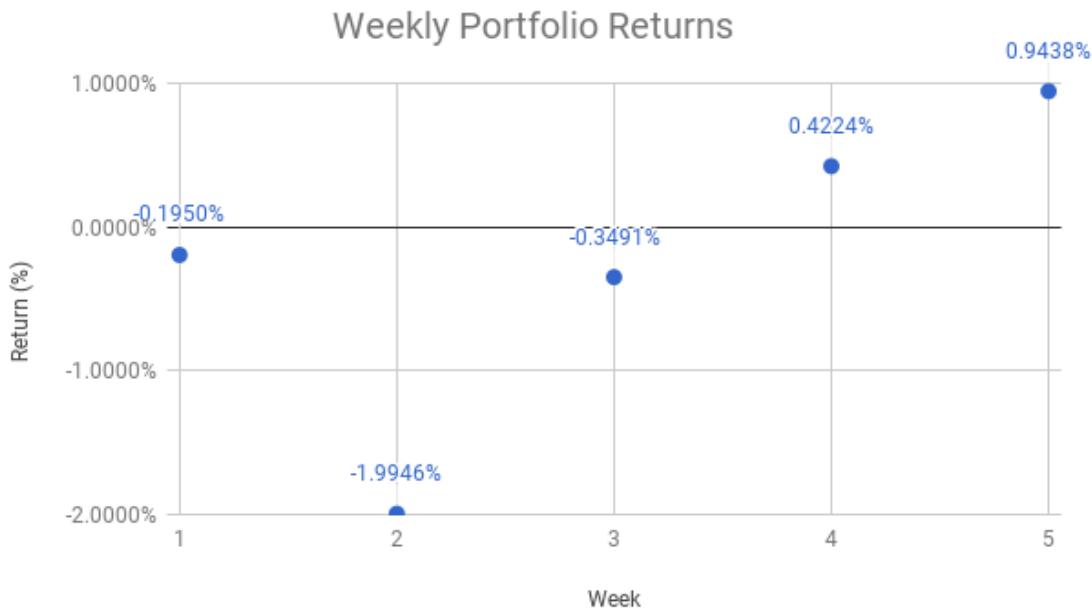


Figure 9

We see from these two figures that our portfolio lost money during the first three weeks, but began to climb substantially after week 3.

3.2.2 Returns

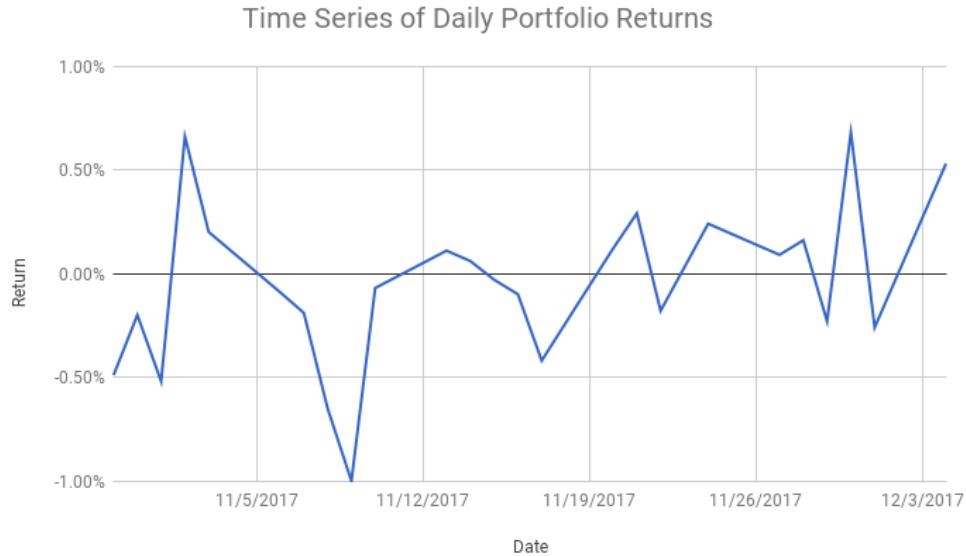


Figure 10

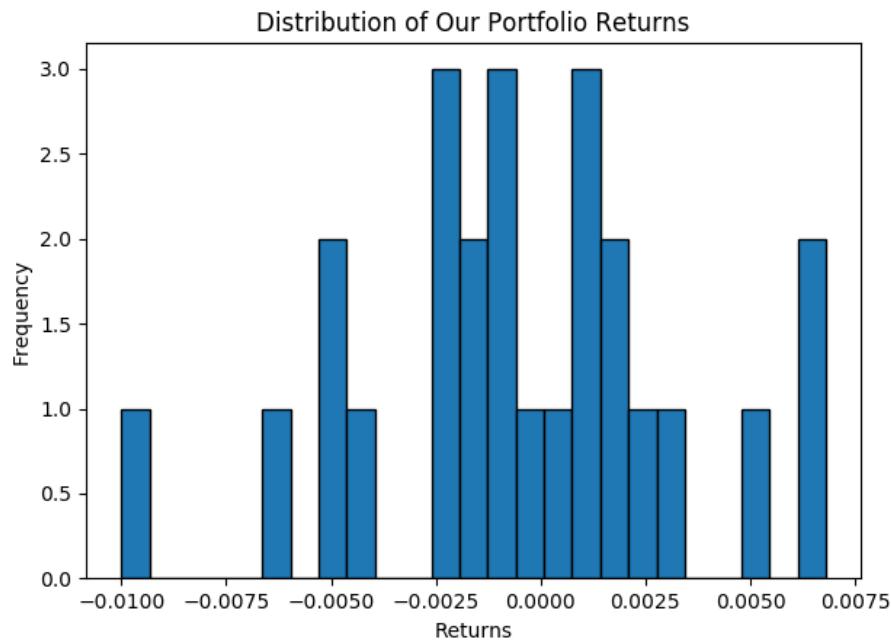


Figure 11: Daily Raw Returns (not percent returns)

The following table provides several summary statistics about our portfolio:

Portfolio Returns Summary	
Days	25
Mean Daily Return	-0.05%
Daily Return SD	0.39%
Annualized Return	-12.11%
Annualized SD	6.25%

Figure 12

Based on our 25 days of data, the annualized return is -12.11% and the annualized volatility is 6.25%. Although this result is disappointing, 25 trading days in general are not sufficient enough of a good sample for making any sort of long-term projections or statements about a portfolio.

3.2.3 Sharpe Ratio

The Sharpe Ratio is one of the most famous measures used in portfolio management. A consequence of William Sharpe's work upon Modern Portfolio Theory, the Sharpe Ratio is "the average return earned in excess of the risk-free rate per unit of volatility or total risk" [6]. The formula for calculating the Sharpe Ratio, as shown in MA 574: Portfolio Valuation & Risk Management, is:

$$RS = \frac{\mu_p - \mu_f}{\sigma_p} \quad (1)$$

From Section 3.2.2, our parameters for μ_p and σ_p are -12.11% and 6.25% respectively. The value for μ_f is the yield on our bond, 2%. With these values, our computed Sharpe Ratio then, is

$$RS = -2.256915737 \quad (2)$$

3.2.4 Treynor Ratio

$$RT = \frac{\mu_p - \mu_f}{\beta_p} , \quad \beta_p = \frac{\sigma_{pM}}{\sigma_M^2} \quad (3)$$

As for the Treynor Ratio, our parameters for μ_p and μ_f are -12.11% and 2%, respectively. For our portfolio β_p , we utilized our CAPM script and computed a value of 1.050789, indicating that our portfolio is slightly more volatile than the market. With these values, our computed Treynor Ratio is:

$$RT = -0.1342375092 \quad (4)$$

3.2.5 Information Ratio - S&P500

The Information Ratio seeks to make a comparison between our portfolio and some other benchmark. For our purposes here, we utilized the S&P500 as our benchmark. As defined in MA 574, the Information Ratio RI can be calculated as such:

$$RI = \frac{\mu_p - \mu_B}{\sigma_{R_p - R_B}} , \quad B = \text{Benchmark} \quad (5)$$

As before, our value of μ_p is -12.11%, the annualized return of our portfolio based on daily returns from October 30th, 2017 to December 4th, 2017. In order to remain consistent in the use of this formula, we utilized the daily return data for the S&P500 index from this date range - data obtained from Yahoo Finance [7]. From this data, we calculated the annualized return on the S&P500, and the annualized standard deviation of the difference of returns between our portfolio and the S&P500; $\mu_B = 0.2527948234$ and $\sigma_{R_p - R_B} = 0.06583540257$ respectively. With these values, we calculated our Information Ratio to be:

$$RI = -5.678557537 \quad (6)$$

Based on the heuristic rule provided in lecture that an $RI > 1$ is an indicator of success, our calculated value shows that we are quite far from success. However, our result may simply be a consequence of the amount of data used.

3.2.6 Sortino Ratio

$$\text{Sortino Ratio} = \frac{\mu_p - r_o}{\sigma_{r_o^-}} \quad (7)$$

where r_o is a target or minimum accepted return, and $\sigma_{r_o^-}$ is the sample downside semistandard deviation. For our minimum accepted return, we used 2%, the yield on our U.S. Note. As before, μ_p is -12.11%. In order to compute the downside semistandard deviation, we computed the daily return equivalent to our minimum required annual return, and then compared the resulting value of 0.000079213646 (0.00792113646%) to all of our daily portfolio returns. The daily returns above this value were removed, and the daily standard deviation of the remaining values was computed, and then annualized. The result was $\sigma_{r_o^-} = 4.297855969\%$, which gave a Sortino Ratio of:

$$\text{Sortino Ratio} = -3.281992208 \quad (8)$$

As before, as this ratio is negative, it indicates that our portfolio has performed poorly. The data and calculations for the Sharpe Ratio, Treynor Ratio, Information Ratio and Sortino Ratio can be found in the Excel file *PortfolioReturns.xlsx*, located in the .zip-file associated with this project report.

3.2.7 Maximum Drawdown

Definition 3.1. *Maximum Drawdown:* The maximum drawdown is the max equity retraction over a holding period; usually quoted as a percentage.

Definition 3.2. The maximum drawdown over $[0, T]$ is

$$MDD(T) = \max_{0 \leq t \leq T} [M(t) - V(t)] \quad (9)$$

where $M(t)$ indicates the "high watermark" at time t , and $V(t)$ indicates the portfolio value at time t .

The figure below, provided during a MA 574 lecture by Professor Marcel Blais, illustrates this concept very well:



Figure 13: Maximum Drawdown Concept. *Source: Professor Marcel Blais*

We will evaluate our maximum drawdown on a weekly basis. The following plot (from before) shows our weekly portfolio values:

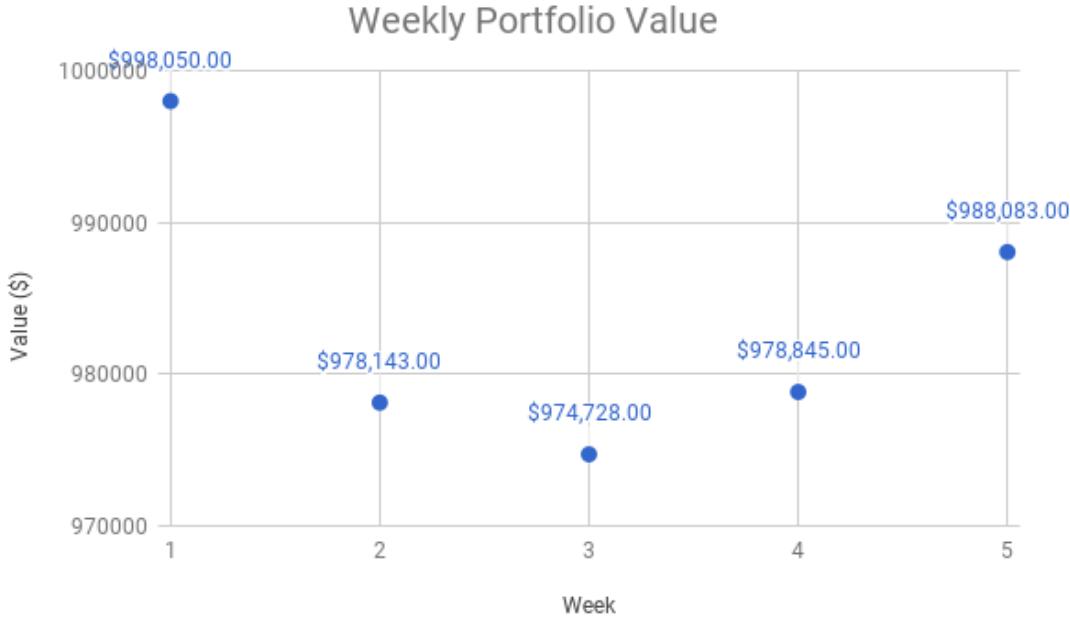


Figure 14: Time series of our weekly portfolio value

Based on this graphic, we see that our lowest portfolio value occurred in Week 3 - \$974,728.00. Our "high watermark" is the highest value our portfolio achieved, which is rather unfortunately our initial portfolio value of \$1,000,000. As such, our maximum drawdown is

$$MDD(T) = \$1,000,000 - \$974,728.00 = \$25,272.00 \quad (10)$$

which is equivalent to a 2.53% drop in portfolio value.

3.2.8 Portfolio Alpha & Beta

From the traditional CAPM model that we constructed (discussed in later sections), we computed the following portfolio alpha and beta:

Table 1: Portfolio Alpha and Beta

Alpha	Beta
-0.012	1.050789

The result for alpha is consistent with our overall level of return, and the value for beta is logical.

3.2.9 Comparison to Industry Benchmarks

Based on the performance of our portfolio over the 25 trading days, it is not particularly valuable to compare it to any standard industry benchmarks as the result will always be the same - our loss of -12.11% is indicative of poor performance. However, given a much longer time span, it may be interesting to compare the performance of the portfolio to various benchmarks such as the S&P500, the Vanguard Total Stock Market Index (VTSMX), and the Russell 2000. Benchmarking is a widely debated topic, and the methods and approaches used can often vary investor-to-investor.

3.2.10 Margin Call Analysis

Throughout the lifetime of our portfolio, we did not receive any margin calls. Our net liquidation value remained high enough at all times to satisfy the margin requirements within Interactive Brokers. As such, no adjustments needed to be made.

3.2.11 Leverage Analysis

Beyond computing ratios, we analyzed the leverage of our portfolio throughout its lifetime. Leverage can be a useful and informative metric about a portfolio's risk and margin requirements. It can also be utilized for "knowing the amount of debt held by a company in evaluating whether it can pay its debts off" [8]. For the purposes of this project, we considered the leverage ratio L to be:

$$L = \frac{\text{Portfolio Market Value}}{\text{Net Liquidation Value}} \quad (11)$$

We computed the leverage ratio of our portfolio on a daily basis throughout its lifespan, utilizing the values provided through the Interactive Brokers portfolio management features. The plot below shows the daily leverage ratio of our portfolio:

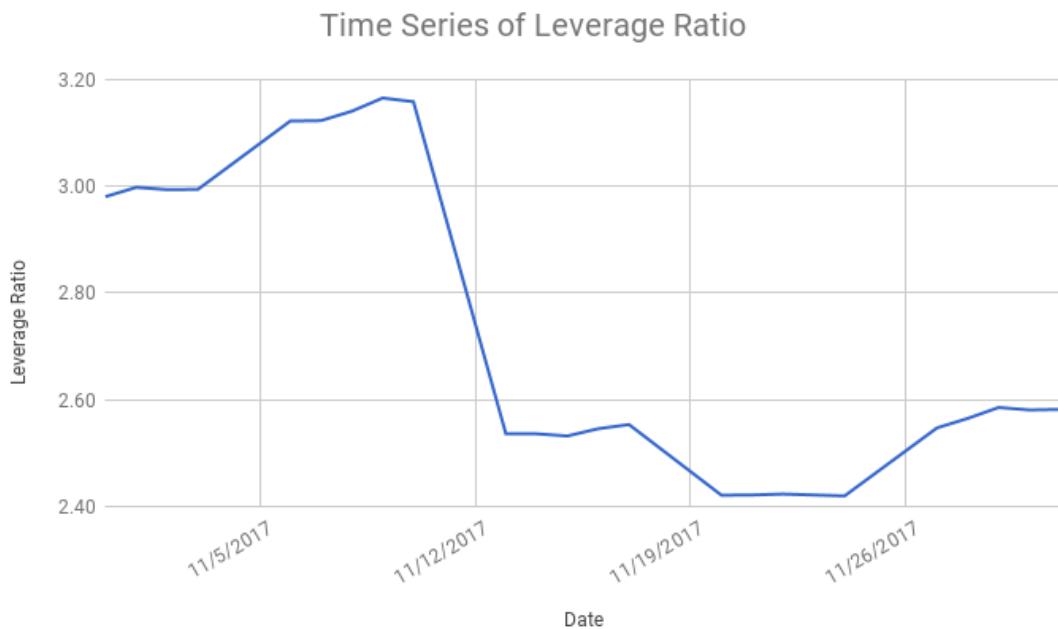


Figure 15: Time series of daily portfolio leverage ratio

This plot indicates that our portfolio is fairly highly leveraged, with a high of approximately 3.2x and a low of approximately 2.4x. If we compare this plot to our initial tangency portfolio weights...

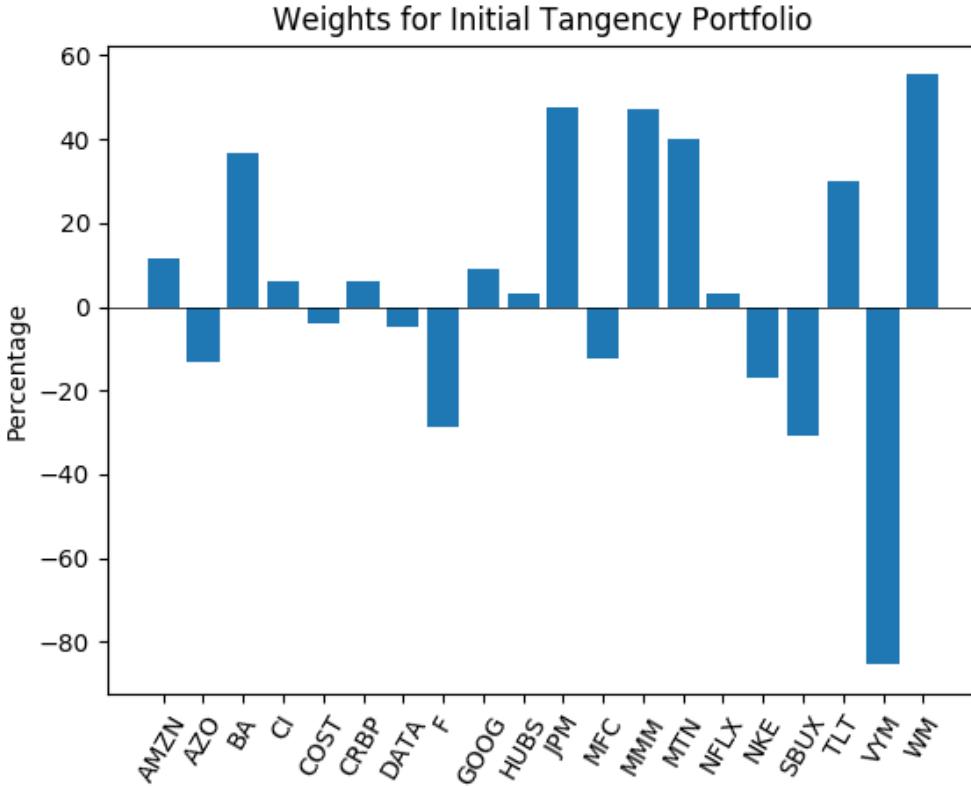


Figure 16

...we see that these results are logical. Although the weights in our assets did change from week to week, there was still nonetheless a high amount of leverage in our portfolio due to some very large long and short positions every week. Such a high leverage ratio makes our portfolio more risky and raises our margin requirements significantly. Although we never received any margin calls from Interactive Brokers (likely due to our large bond position serving as a hedge to our equity positions), such a high leverage ratio should be a cause for concern.

If we were to do this project again, we would like to form our portfolios using constrained Markowitz optimizations that significantly limit the size of short positions. This would help bring the "Portfolio Market Value" at any point much closer to the portfolio Net Liquidation Value, bringing the leverage ratio much closer to 1. This in turn would reduce our margin requirements and may also help reduce our losses as a result of small but adverse movements in any of our assets. From Section 3.2.1, we see that our position in VYM, which was consistently a *very* large short position, lost us a disproportionately large amount of money relative to all of our other positions. Reducing the portfolio leverage is one possible step to mitigate this issue in the future.

4 Capital Asset Pricing Model

The capital asset pricing model (aka. CAPM) is one of the most well-known theoretical models in the financial analysis world that is best used for "passive" investing known as indexing[9]. We aimed to calculate alternative β coefficients for the CAPM. For this project, we used the same 20 risky assets as in our Markowitz Portfolio Optimization project. We calculated the β coefficients from the traditional approach using historical returns as well as the alternative methods, such as least trimmed squares (LTS), constant β shrinkage estimation, and exponentially weighted moving average (EWMA). We decided to use a 5-year date-range (between 1/1/2012 and 1/1/2017) for the purposes of presenting these alternative β coefficients.

4.1 Traditional Approach

The traditional CAPM approach involves using the historical returns for each security and the risk-free asset (aka. S&P500 index for our case). We used the following simple asset returns formula to determine our asset returns:

$$R_{j,t} = \frac{P_{j,t}}{P_{j,t-1}} - 1 \quad (12)$$

Where $P_{j,t}$ is the monthly adjusted close price for the jth security. For our results and analysis, we primarily used the "excess" returns, which is the difference between the security/risk-free asset return and the risk-free rate μ_f ($\mu_f = 0.02$, from the previous section 2.2).

$$R_{j,t}^* = R_{j,t} - \mu_f, \quad R_{M,t}^* = R_{M,t} - \mu_f \quad (13)$$

After gathering our respective excess returns, we utilized the built-in Python ordinary least squares (OLS) estimator `statsmodels.api.OLS()` for the following CAPM regression model:

$$R_{j,t}^* = \beta_j(R_{M,t}^*) + \epsilon_{j,t} \quad (14)$$

The table above contains each of our assets and their respective CAPM (traditional method) β coefficients. Furthermore, our portfolio β was determined to be 1.050789.

At a glance, by definition from our MA 574 class, any β that is > 1 is aggressive. Thus, one can say for eg. that the securities JPM, HUBS, and CRBP are highly aggressive to the market. Therefore, since our portfolio $\beta = 1.050789$, our portfolio overall is indeed aggressive; this was anticipated especially considering our relatively volatile assets such as AMZN, HUBS, and NFLX.

It is interesting to note and point out the "negative" β coefficients in our results above, namely for NKE and TLT. By definition, when a stock has a negative β , this means it is inverse to the market. We looked at the segmented time analysis for NKE; as it turns out,

Table 2: Ordered Stock Betas

Stock	Beta
NKE	-0.56412
TLT	-0.456008
MTN	0.444818
CI	0.508103
AZO	0.60286
WM	0.70178
SBUX	0.774694
DATA	0.778643
VYM	0.90504
COST	0.913119
GOOG	0.979013
F	1.039445
MMM	1.04565
BA	1.08145
NFLX	1.268151
AMZN	1.399145
MFC	1.468536
JPM	1.505132
HUBS	2.162088
CRBP	2.442068

we can observe that at around July 2015, NKE returns not only jumped but some of the the peaks coincide with the troughs for the S&P 500 index returns, ie. opposite to the market.



Figure 17: NKE & S&P500 Returns of Last 5 Years. *Source: Yahoo Finance [10]*

4.2 Least Trimmed Squares

For the least trimmed squares (LTS) approach, this involved "trimming" the values that are smaller the α -quantile and the values that are greater than the $(1 - \alpha)$ – quantile. Essentially, the asset returns that are seen before the α -quantile and those after the $(1 - \alpha)$ -quantile are set to the value at the α -quantile or $(1 - \alpha)$ -quantile. This can also be called "feeding in the residuals." The list of estimated β coefficients can be found in the comparison table in section 4.5, but we chose VYM as our representative asset to demonstrate the original-to-trimmed returns distribution. We chose α to be 0.05. The original returns distribution as well as trimmed returns distribution for VYM is as follows:

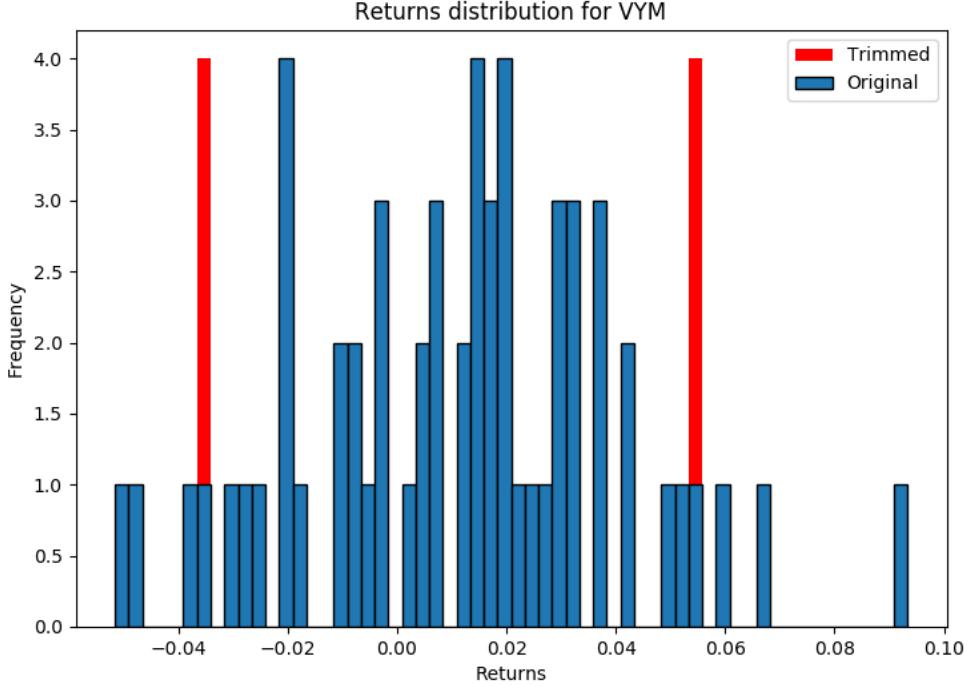


Figure 18: Returns Distribution (original & trimmed) of VYM

4.3 Constant β Shrinkage Estimation

This shrinkage approach is useful because our estimated $\hat{\beta}$ coefficients could be far from the true β 's. This method involves using the previously estimated $\hat{\beta}_{i,OLS}$ coefficients from the traditional CAPM method, calculating the average $\bar{\beta}$, and then using a shrinkage parameter. We chose $\alpha_i = \frac{2}{3}$ because it is common in practice. We used the James-Stein estimator[11] formula, detailed below. Our results can be found in section 4.5:

$$\hat{\beta}_{i,JS} = \bar{\beta} + \alpha_i(\hat{\beta}_{i,OLS} - \bar{\beta}) \quad (15)$$

4.4 Exponentially Weighted Moving Average (EWMA)

This EWMA method is an extremely useful technique because it puts more emphasis (ie, weight) to more recent data points versus older data. One reason why older data may not be as reliable is due to mergers and acquisitions; there should be more leverage on newer data since it is more relevant to current time. EWMA method uses a smoothing parameter, namely $\lambda \in (0, 1)$. For demonstration purposes, we chose $\lambda = 0.94$ (because it is one of the standard values for monthly data) and $\lambda = 0.8$ (to have a value less than our first λ). The larger the λ , the more weight is put to the past. In fact, if $\lambda = 1$, then we just have the CAPM traditional model. We were able to calculate the new β coefficients via:

$$\hat{\beta}_{i,t} = \frac{\sigma_{i,M,t}}{\sigma_{M,t}^2} \quad (16)$$

Where the numerator is the i th asset covariance estimate and the denominator is the market variance estimate. Their iterative formulas are:

$$\sigma_{i,M,t} = (1 - \lambda)R_{i,t-1}^* R_{M,t-1}^* + \lambda\sigma_{i,M,t-1}, \quad \sigma_{M,t}^2 = (1 - \lambda)R_{M,t}^2 + \lambda\sigma_{M,t-1}^2 \quad (17)$$

4.5 Comparison and Analysis

Finally, after accomplishing all the above alternative methods along with the traditional CAPM approach for calculating the various β coefficients, we have our final β comparison table. We performed a conditional color format, where large β 's are colored green and small β 's are colored red:

Stocks	Beta CAPM	Beta LTS	Shrinkage Beta	EWMA (Lambda = 0.94)	EWMA (Lambda = 0.8)
AMZN	1.399145	1.194054	1.249423	1.303481	0.780364
AZO	0.60286	0.569738	0.718567	0.645602	0.625685
BA	1.08145	0.926644	1.037627	1.005153	0.77169
CI	0.508103	0.51499	0.655395	0.856552	0.865074
COST	0.913119	0.839198	0.925406	1.103	1.151821
CRBP	2.442068	2.498855	1.944706	2.409043	2.647842
DATA	0.778643	0.81902	0.835755	-0.137976	-1.677926
F	1.039445	0.946539	1.009623	0.573273	0.232377
GOOG	0.979013	0.824539	0.969335	1.192994	1.233285
HUBS	2.162088	1.888545	1.758052	2.230947	2.211438
JPM	1.505132	1.215495	1.320081	2.107107	2.809843
MFC	1.468536	1.163366	1.295684	1.840141	2.283879
MMM	1.04565	0.951167	1.01376	1.104079	0.841187
MTN	0.444818	0.401589	0.613205	0.643012	0.182447
NFLX	1.268151	0.92305	1.162094	2.175395	1.479845
NKE	-0.56412	0.435616	-0.05942	0.186329	0.109066
SBUX	0.774694	0.706595	0.833122	1.00854	1.042099
TLT	-0.456008	-0.353081	0.012655	-0.641323	-0.974508
VYM	0.90504	0.795745	0.92002	0.891551	0.793794
WM	0.70178	0.684433	0.784514	0.728559	0.483231

Figure 19: All alternative β coefficients.

We can see, at least across the traditional, LTS, and shrinkage methods that the estimated β coefficients are relatively the same. We also see that CRBP is the most aggressive (with a $\beta > 2.0$) and that NKE and TLT are the most inverse to the market.

More importantly though, we see the differing β coefficients for the exponentially weighted

moving average method, especially for the DATA asset. Why might this be? The answer becomes clear when one looks at the time-series price figure for DATA:



Figure 20: 5-year time-series of adjusted close price for DATA [12].

The EWMA approach very much took into consideration the July 2015 - February 2016 extreme drop that happened with Tableau Software (DATA). Some reasons for this drop were because of the over-crowded data analytics competition, namely Microsoft, and its "own sales execution"[13]. Thankfully, EWMA (especially with $\lambda = 0.80$, thus more consideration into the past) took this into account and thus assigned a negative β coefficient.

If we were hired as financial analysts for a firm, we would choose the exponentially weighted moving average approach thanks to its usefulness when considering more relevant new data.

5 Factor Modeling

For the factor model portion of our project, we still used the same exact 20 risky assets from our Markowitz Portfolio Project and we used the S&P500 index as the market portfolio. We implemented the traditional CAPM method, the French-and-Fama model, and two

additional factors.

Our first factor (F_1) was the Momentum Factor for Global ex USA portfolio from Ken French's website. This factor is the overall stock return portfolio of major country stock returns MINUS USA's stocks. In more detail, "the global portfolios use global size breaks, but [we] use the momentum breakpoints for the four regions to allocate the stocks of these regions." [14] We wanted to choose this factor because it appeared to give the overall global stock return trend, but without USA's global portfolio in consideration. This way we could see if our assets followed the non-USA global portfolio trend.

Our second factor (F_2) was the RSTAMOM index retrieved from Bloomberg Terminal which measures the adjusted total monthly % change in "retail and food services sales in USA" [15]. This factor was specific to USA, and we so believed that the retail & food sales would have some significance to our 20 risky assets which involve consumer sales.

For this section of our project, we used a 10-month holding period, where any date past 1/1/2017 was considered "new." In other words, our historical analysis was between 3/1/2016 and 1/1/2017.

We first implemented the traditional CAPM historical returns approach to determine our β coefficients, expected asset returns, variance of asset returns, and covariance matrix between all 20 of our assets. We performed these steps for each of our factor models, including the extensions with our chosen F_1 and F_2 factors. The results for each model follow. For our visuals, we utilized a red-green color shading scheme - green indicates positive values, and the more positive the value, the more green it becomes. Red indicates negative values, and the more negative the value, the more red it becomes.

5.1 CAPM

stocks	Beta	Expected Asset Returns	Asset Returns Variance
AMZN	0.3882996761	0.05862408491	0.04653771858
AZO	1.368458362	0.1561202576	0.02605462627
BA	-0.3334384091	-0.01316704653	0.02570371607
CI	2.069167984	0.2258196924	0.04783724458
COST	1.726060103	0.1916908256	0.03148173099
CRBP	1.000302784	0.1194999018	1.019769336
DATA	2.381244777	0.2568619036	0.1012313381
F	1.100985359	0.1295147759	0.01197206546
GOOG	1.226849848	0.1420344894	0.03890423958
HUBS	3.930907599	0.4110065298	0.1794704381
JPM	1.529645798	0.1721535371	0.04238492542
MFC	1.688330885	0.1879379084	0.06504423972
MMM	1.275724638	0.1468960542	0.0139290697
MTN	-0.2050561288	-0.0003968888414	0.01862031354
NFLX	-2.75434536	-0.253974138	0.1826653003
NKE	0.8264901644	0.1022107981	0.02235244801
SBUX	1.36008683	0.1552875432	0.02751813557
TLT	-0.1938543273	0.0007173519343	0.02090183911
VYM	0.8717926501	0.1067170266	0.003691830407
WM	0.2226501202	0.04214695936	0.01544017372

Figure 21

5.9 Recap

To compare the different factor models, we utilized the matrix norm approach provided in the project guidelines. This was implemented using the `numpy.linalg.norm` function from the `numpy` library in Python, which uses the Frobenius norm for unordered norms [16].

	CAPM + F1	CAPM + F2	CAPM + F1 + F2	FF + F1	FF + F2	FF + F1 + F2
CAPM + F1	-	0.01612839349	0.01580362711	0.04727571084	0.04526575934	0.04844141244
CAPM + F2	0.01612839349	-	0.008747349727	0.04640838613	0.04287255482	0.04610362816
CAPM + F1 + F2	0.01580362711	0.008747349727	-	0.04529557746	0.04159542112	0.04491693634
FF + F1	0.04727571084	0.04640838613	0.04529557746	-	0.01190094569	0.00975880674
FF + F2	0.04526575934	0.04287255482	0.04159542112	0.01190094569	-	0.01024218435
FF + F1 + F2	0.04844141244	0.04610362816	0.04491693634	0.00975880674	0.01024218435	-

Figure 37

5.10 Portfolio Optimization

Based on the results we have seen, the French-Fama models seemed to be unreasonable and there may be some error in our code implementation. Therefore, we chose the factor models CAPM & F1 and CAPM & F2 for demonstration purposes in order to continue our work in portfolio optimization.

For the purpose of backtesting, we utilized the portfolio backtesting capabilities on the website PortfolioVisualizer.com. This tool allows you to "construct one or more portfolios based on the selected mutual funds, ETFs and stocks to analyze and backtest portfolio returns, risk characteristics, standard deviation, annual returns and rolling returns. The results include a visualization of the portfolio growth chart and rolling returns, CAGR, standard deviation, Sharpe ratio, Sortino ratio, annual returns and inflation adjusted returns" [17].

Taking our chosen models of "CAPM & F1" and "CAPM & F2" (and the associated expected asset returns as well as covariance returns) and also using our previous work in our Markowitz Efficient Frontier project, we formed optimal portfolios and observed their respective weights for a target portfolio overall return of 10%; we decided on a target return of 10% because this is the average annual return for the S&P 500 index [18].

Our code that calculates the optimal weights takes in a vector of expected asset returns, covariance matrix of asset returns, and a target return value. Due to the nature of our "CAPM & F1" and "CAPM & F2" models, we got a weird assortment of expected asset returns; the results can be found below in Table 3. We believe it is because of our chosen factors and especially also because of how the covariance matrices of asset returns gets calculated.

Table 3: Factor Based Portfolio Weights

Stock	CAPM+F1 Weights	CAPM+F2 Weights
AMZN	-1.28360	-0.03588
AZO	-2.56679	-0.69166
BA	-2.56025	0.36103
CI	-0.76878	0.33432
COST	0.41914	-0.15293
CRBP	-1.69795	-0.03873
DATA	2.10227	-0.91175
F	-0.19794	-0.39646
GOOG	0.42695	1.20205
HUBS	-0.21813	0.05860
JPM	0.07241	-0.08883
MFC	-1.20424	0.49460
MMM	1.95534	-0.10846
MTN	0.51758	-0.65460
NFLX	2.01097	-0.22027
NKE	-1.75463	0.92387
SBUX	3.70834	0.15667
TLT	-2.66854	1.38948
VYM	1.82389	-0.42545
WM	2.88396	-0.19560

As we can see, the CAPM + F1 model has a very large amount of leverage in it, while the CAPM + F2 still has some large, but much more reasonable weights. We believe that the excessive leveraging may be a result of the covariance matrix estimates. We took these weights and plugged them into Portfolio Visualizer to complete the backtesting.

5.10.1 Backtesting Results

The following table provides a brief summary of the performance of our two portfolios, as well as several metrics used previously in this report:

Table 4: CAPM+F1 & CAPM+F2 Portfolio Performance Overview

Portfolio	Initial Balance	Final Balance	Stdev	Max.DD	Sharpe Ratio	Sortino Ratio
CAPM+F1	\$500,000	\$1,658,145	96.22%	-36.64%	1.96	5.6
CAPM+F2	\$500,000	\$416,926	40.22%	-34.59%	-0.37	-0.51

As we can see, there is a large difference in the performance of the two portfolios. The extremely high leveraged CAPM+F1 model had an extremely high return, while the CAPM+F2 model ended up losing money. This is reflected in the Sharpe and Sortino Ratios

for the two portfolios. The following plots show the growth of the two portfolios over the backtesting period, as well as the overall annualized returns:

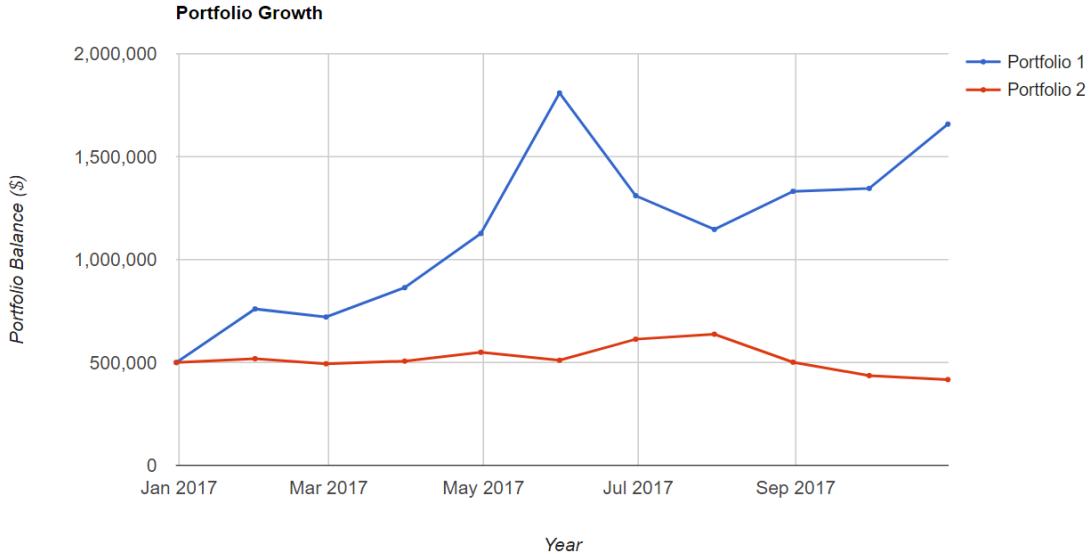


Figure 38: Monthly Portfolio Value

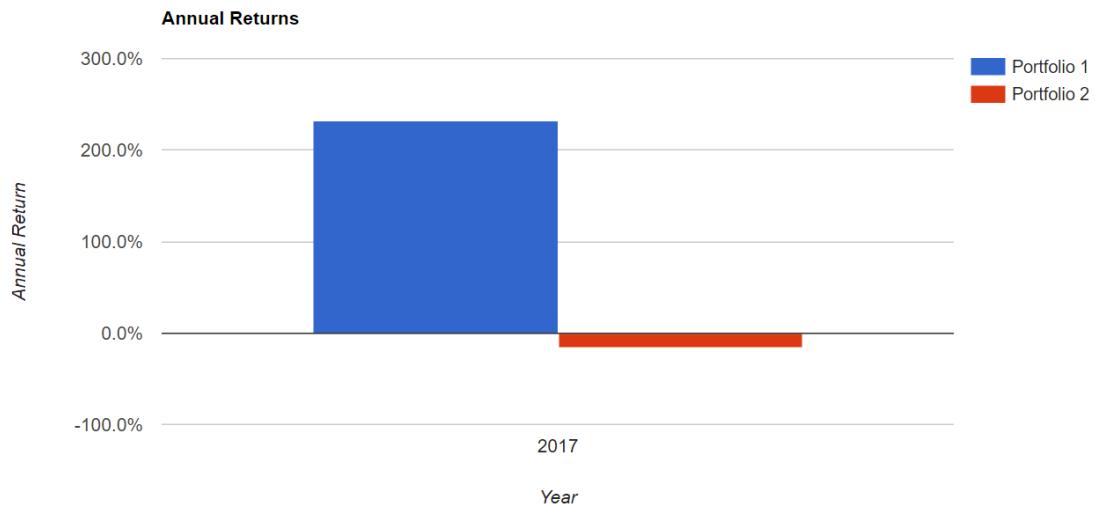


Figure 39: Annualized Portfolio Returns

The full results of our portfolio backtesting can be accessed here: <https://goo.gl/p2vHi9>, as well as in the Excel file *PV_Backtesting.csv* included in the *.zip*-file associated with this report. It is very interesting and we recommend taking a look.

To compare our performance and models, we looked at the actual annualized expected returns and covariance returns for our 20 risky assets, for a holding period between Jan

	Actual Cov	CAPM + F1	CAPM + F2
Actual Cov	-	0.4376514213	0.4233969401
CAPM + F1	0.4376514213	-	0.2054495368
CAPM + F2	0.4233969401	0.2054495368	-

Figure 41

Finally, after analyzing all of the above tables and comparisons, one thing we can immediately deduce is that our "CAPM & F1" chosen factor model's covariance of asset returns is more aligned with the actual 10-month holding-period covariance of returns.

Without a doubt, the worst performing portfolio was the "CAPM & F2" factor model while the best performing was the "CAPM & F1" model. We believe this is the case due to our chosen factors of "global ex US momentum factor" and "adjusted retail US sales." It is conceivable that because of the volatility for certain assets in our portfolio (namely, NFLX, HUBS, and CRBP) and because of the global international presence of some of our assets (namely, JPM, MFC, and AMZN) that our factor models gave the results that they did.

5.11 Quantile-Based Portfolio Back-testing

For this portion of the project, we first ranked our 20 risky assets in order from best to worst for each of our previously chosen factor models. How did we rank our assets? We ordered in decreasing order the "expected asset returns" and then looked at the increasing order of "variance asset returns." This means we consider both the reward and risk aspects of our 20 assets. We ended up with the following ranking:

Table 6: Asset Ranking (best-to-worst)

CAPM+F1	CAPM+F2
HUBS	WM
DATA	BA
CI	TLT
MFC	AZO
JPM	MMM
COST	COST
SBUX	AMZN
AZO	VYM
MMM	NKE
GOOG	SBUX
F	MTN
CRBP	CI
NKE	F
VYM	MFC
AMZN	NFLX
WM	JPM
MTN	DATA
BA	GOOG
TLT	CRBP
NFLX	HUBS

From this table then, we deduce the assets used in the quintile-based backtesting, by taking the five best and five worst assets from each model. The following table shows the assets used and their respective positions in our backtested models:

Table 7: Top and Bottom Quintile Assets for Each Model

	CAPM+F1	CAPM+F2
Position	Asset	Asset
Long	HUBS	WM
Long	DATA	BA
Long	CI	TLT
Long	MFC	AZO
Long	JPM	MMM
Short	WM	JPM
Short	MTN	DATA
Short	BA	GOOG
Short	TLT	CRBP
Short	NFLX	HUBS

Given that each of these portfolios has 10 assets and the project guidelines call for equal weights in each of the long and short positions, we will backtest each of these portfolios with 25% in each long asset and 5% in each short asset. This will allow us to invest 100% of our portfolio, and are not unreasonable figures for equity weights. Once more, we will utilize PortfolioVisualizer for backtesting. The backtesting date range is also the same as before: January 1, 2017 to October 31, 2017.

5.11.1 Backtesting Results

Table 8: CAPM+F1 & CAPM+F2 Quintile Backtesting Summary

Portfolio	Initial Bal.	Final Bal.	Stdev	Return	Max.DD	Sharpe Ratio	Sortino Ratio
CAPM+F1	\$500,000	\$754,189	12.70%	50.84%	-2.86%	3.59	10.91
CAPM+F2	\$500,000	\$622,149	10.74%	24.43%	-1.82%	2.23	10.04

This table has several interesting results. First of all, both portfolios performed very well, with the CAPM+F1 portfolio returning 50.84% annually, and the CAPM+F2 portfolio returning 24.43% annually. Both of these are admirable results. The CAPM+F1 portfolio is more volatile than the CAPM+F2 portfolio, as indicated by the higher standard deviation. Furthermore, the CAPM+F1 portfolio retracted significantly more than the CAPM+F2 portfolio, with a maximum drawdown of -2.86% versus -1.82%. Both the Sharpe and Sortino Ratios indicate that these were successful portfolios, but the Sortino Ratio tells us a bit more. Despite the fact that the return on the CAPM+F1 portfolio was double that of the CAPM+F2 portfolio, the Sortino Ratios are very similar. This indicates that the CAPM+F2 portfolio achieves its returns with more "efficiency" than the CAPM+F1 portfolio, as it has much less downside volatility. So for a more risk-averse approach, the CAPM+F2 portfolio is significantly better than the CAPM+F1 portfolio.

The plots below show the growth throughout the backtesting period of our two quintile CAPM+F1 portfolio and the CAPM+F2 portfolio, respectively.

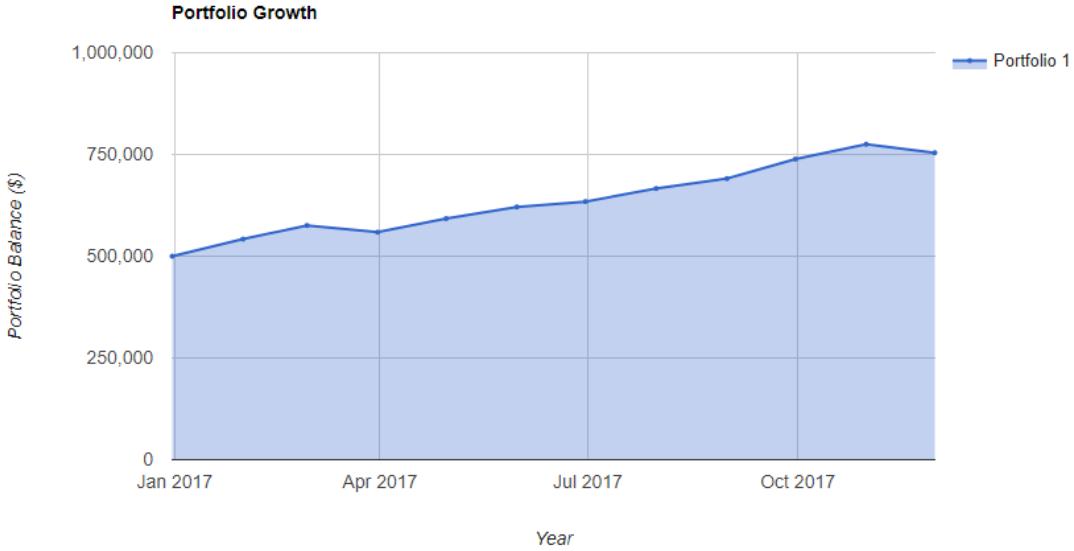


Figure 42: CAPM+F1 Quintile Portfolio Growth



Figure 43: CAPM+F2 Quintile Portfolio Growth

The full results of our CAPM+F1 quintile backtesting can be found here: <https://goo.gl/PeUA2F> and in the Excel file *PV_Backtesting2.csv*. The full results of our CAPM+F2 quintile backtesting can be found here: <https://goo.gl/pvy2UH> and in the Excel file *PV_Backtesting3.csv*.

Lastly, if we compare these two models to the Markowitz optimized portfolios from the previous section, we see that both of our quintile models underperform the CAPM+F1 model,

but outperform the CAPM+F2 model. However, that said, the CAPM+F1 results from the prior section are rather extreme, both in terms of return and volatility. This is primarily due to the highly leveraged weighting scheme computed by the Markowitz optimization. As such, we would be very apprehensive about utilizing this model in the future as it has an extreme amount of downside potential. The CAPM+F2 model from before did lose money, so we are not particularly thrilled with it, considering the state of the stock market during the backtesting period. As such, if we had to choose a model, we would be fairly happy with either of the two quintile models.

If we were seeking to maximize returns, we could use the CAPM+F1 model, and if we wanted to be a bit more conservative, we could use the CAPM+F2 model. Overall, both of these models seemed more reasonable than the models in the prior section. We ultimately believe this to be due to the fact that we are using an unconstrained Markowitz optimization, and would be curious to see how the results change if we were to implement a constrained Markowitz routine to the model construction.

6 Conclusion

Overall, this set of projects was extremely educational and we are happy to have had the opportunity to get hands-on experience in this area. There are certainly areas we could improve in our implementation and analysis- however, we do believe that the fundamentals of our work are solid and substantial.

All related project materials, including code, plots, and data, can be found in the .zip-file associated with this project. For any questions, please contact the authors of this paper at akshoop@wpi.edu (Alexander Shoop) and kmdymov@wpi.edu (Khasan Dymov).

Appendix A Portfolio Rebalancer.py

```

1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """
4
5  @author: Khasan Dymov and Alex Shoop
6  """
7
8  import numpy as np
9  import matplotlib.pyplot as plt
10 from pandas_datareader import data
11 from datetime import datetime
12 #####
13 ##### #setting output np options
14 np.set_printoptions(threshold = np.inf)
15
16 #Import specified date range of prices data,
17 #and calculated returns (log returns)
18
19 #Specified tickers
20 tickers = [ 'AMZN' , 'AZO' , 'BA' , 'CI' , 'COST' , 'CRBP' , 'DATA' , 'F' , 'GOOG' , 'HUBS' , \
21             'JPM' , 'MFC' , 'MMM' , 'MIN' , 'NFLX' , 'NKE' , 'SBUX' , 'TLT' , 'VYM' , 'WM' ]
22
23
24 #Specified date range
25 startDate = datetime(2015,11,24)
26 endDate = datetime(2017,11,24)
27
28 #Use pandas_reader.data.DataReader to load desired data
29 securityRawData = data.DataReader(tickers, 'yahoo', startDate, endDate)
30
31 #Getting adjusted close prices
32 close = securityRawData[ 'Adj Close' ]
33 #fixing date ordering
34 close = close.iloc [:-1]
35 #getting log returns
36 dailyLogReturns = np.log(close).diff()
37 #removing the 1st row, of empty values for first date
38 dailyLogReturns = dailyLogReturns.drop(startDate)
39 #Pulling just the last day closing prices for rebalancing function
40 finalprices = np.matrix([close.iloc [:-1].iloc [0]]).T
41
42 #Gather expected returns, and covariance matrix
43 mu2yrsdaily = dailyLogReturns.mean()
44 var2yrsdaily = dailyLogReturns.var()
45 sd2yrsdaily = np.sqrt(var2yrsdaily)
46 cov2yrsdaily = dailyLogReturns.cov()
47
48 #Annualize parameter estimates

```

```

50
51 #2 year lookback:
52 muannual = np.sqrt(((1+mu2yrsdaily)**505))-1
53 varannual = np.sqrt(252)*var2yrsdaily
54 sdannual = 252*sd2yrsdaily
55 covannual = 252*cov2yrsdaily
56
57 # OPTIMAL PORTFOLIO FUNCTION
58 def optimalPortfolio(mu, omega, muP):
59     # We follow by example
60     # the process as explained in Statistics & Finance by David Ruppert
61     # Necessary coefficients for risk-efficient portfolio
62     n1 = np.matrix(np.ones_like(mu)).T
63     n1t = n1.T
64     invomega = np.linalg.inv(omega)
65     muR = np.matrix(mu).T
66     mut = muR.T
67
68     A = np.matrix.item(n1t*invomega*muR)
69     B = np.matrix.item(mut*invomega*muR)
70     C = np.matrix.item(n1t*invomega*n1)
71     D = B*C-A**2
72
73
74     # Fixed vectors for efficient weights calculation
75     g = (B*invomega*n1 - A*invomega*muR)/D
76     h = (C*invomega*muR - A*invomega*n1)/D
77
78     # The final formula:
79     wts = g + muP*h
80
81     # Our efficient weights
82     return wts
83
84 #Step 1: Find the minimum variance portfolio for our problem
85
86 n1 = np.matrix(np.ones_like(muannual)).T
87 n1t = n1.T
88 invomega = np.linalg.inv(covannual)
89 muR = np.matrix(muannual).T
90 mut = muR.T
91
92 A = np.matrix.item(n1t*invomega*muR)
93 B = np.matrix.item(mut*invomega*muR)
94 C = np.matrix.item(n1t*invomega*n1)
95 D = B*C-A**2
96
97 #Need to use specific g,h for minimum variance portfolio
98
99 gmin = (B*invomega*n1 - A*invomega*muR)/D
100 hmin = (C*invomega*muR - A*invomega*n1)/D

```

```

101 g1 = gmin.T*np.matrix(covannual)*gmin
102 h1 = hmin.T*np.matrix(covannual)*hmin
103 g1h1 = gmin.T*np.matrix(covannual)*hmin
104
105 mumin = np.matrix.item(- g1h1/h1)
106 sigmamin = np.matrix.item(np.sqrt(g1 - g1h1**2/h1))
107
108 #Weights associated with this minimum variance portfolio
109 wtsmin = np.array(gmin + mumin*hmin)
110 barX = np.arange(np.size(wtsmin))
111
112 #Step 2: Calculate the efficient frontier
113
114 muP = np.linspace(min(muannual), max(muannual), 100)
115 sigmaP = np.zeros(0)
116
117 for target in muP:
118     wtsP = optimalPortfolio(muannual, covannual, target)
119     sigmaP = np.append(sigmaP, np.sqrt(wtsP.T*np.matrix(covannual)*wtsP))
120
121 effindices = np.where(muP>mumin)
122 ineffindices = np.where(muP<mumin)
123
124 #Step 2.5: Calculate the tangency portfolio
125 muF = 0.02
126 omegabar = invomega*(muR - muF*n1)
127 omegaT = omegabar/(n1t*omegabar)
128 sigmaT = np.matrix.item(np.sqrt(omegaT.T*np.matrix(covannual)*omegaT))
129 muT = np.matrix.item(muT*omegaT)
130
131 #Step 3: Plot minimum variance portfolio and upper section of efficient
132 #→ frontier
133 plt.plot(sigmaP[effindices], muP[effindices], 'g', label = "Efficient Frontier")
134 plt.plot(sigmaP[ineffindices], muP[ineffindices], 'c--')
135 plt.plot(sigmamin, mumin, 'ro', label = "Minimum Variance Portfolio")
136 plt.plot(sigmaT, muT, 'bx', markersize = 8, label = "Tangency Portfolio")
137 plt.plot([0, sigmaT], [muF, muT], 'b--')
138 plt.plot(0, muF, 'bo', label = "Risk-free rate")
139 plt.title("Efficient Frontier for 20 risky assets (2yrs)")
140 plt.xlabel(r"$risk = \sigma_r$")
141 plt.ylabel(r"$reward = \mu_r$")
142 plt.legend()
143
144
145 #Step 4: Bar chart of weights
146
147 #plt.bar(barX, (wtsmin)*100)
148 #plt.axhline(y = 0, linewidth = 0.5, color = 'black')
149 #plt.xticks(barX, stocks, rotation=60)

```

```

150 #plt.ylabel("Percentage")
151 #plt.title("Weights for min var portfolio (2yrs)")
152
153 #Step 5: Bar chart of weights in tangency portfolio
154 wtstangent = optimalPortfolio(muannual, covannual, muT)
155
156 #plt.bar(barX, (wtstangent)*100)
157 #plt.axhline(y = 0, linewidth = 0.5, color = 'black')
158 #plt.xticks(barX, stocks, rotation=60)
159 #plt.ylabel("Percentage")
160 #plt.title("Weights for Initial Tangency Portfolio")
161
162 print("Expected return = " + str(muT))
163 print("Portfolio std. dev = " + str(sigmaT))
164 print("The weights of the assets in the tangency portfolio are " + str(
    ↪ wtstangent))
165
166
167 #Rebalancing function
168
169 def PortfolioRebalancer(currwts, newwts, currcashval, lastcloseprices, tickers
    ↪ ):
    #function takes as input current stock weights in portfolio,
    #new calculated tangency portfolio weights, the current cash value
    #of the portfolio, latest closing prices for all the stocks, and tickers
170
171     newcashvalue = np.array(np.divide(np.multiply(newwts, currcashval),
    ↪ currwts))
172     cashvalchange = np.array(np.subtract(newcashvalue, currcashval))
173     sharechange = np.array(np.divide(cashvalchange, lastcloseprices))
174
175
176     for x in range(0,20):
177         print("Adjust position in " + tickers[x] + " by " + str(sharechange[x
    ↪ ]) + " shares.")
178
179 #Current weights and cash value pasted from Excel file (may automate in future
    ↪ )
180 #MUST CONVERT TO MATRIX, THEN TRANSPOSE!
181 currwts = np.matrix([0.1209828446,
182 -0.1262697252,
183 0.2820707935,
184 0.109136186,
185 -0.004279442142,
186 0.05482424766,
187 -0.0438397217,
188 -0.2276068822,
189 -0.02123034674,
190 0.01040946642,
191 0.3374382686,
192 -0.01833587159,
193
194
195

```

```

196 0.3876116784 ,
197 0.3434809172 ,
198 -0.009318535294 ,
199 -0.03576921161 ,
200 -0.2293183649 ,
201 0.3503882697 ,
202 -0.8579834182 ,
203 0.5776088474]) .T
204
205 currwts = np .matrix ([53370 ,
206 -55702.23824 ,
207 124431.8423 ,
208 48144.00144 ,
209 -1887.819945 ,
210 24185 ,
211 -19339.32 ,
212 -100405.8 ,
213 -9365.489865 ,
214 4592 ,
215 148856.48 ,
216 -8088.63 ,
217 170989.8237 ,
218 151522.1154 ,
219 -4110.75 ,
220 -15779.12 ,
221 -101160.7982 ,
222 154569.2037 ,
223 -378488.1663 ,
224 254804.5905]) .T
225
226
227 PortfolioRebalancer (currwts , wtstangent , currwts , finalprices , tickers )

```

Appendix B FactorModels_wFamaFrench.py

```

1
2 #!/usr/bin/env python3
3 # -*- coding: utf-8 -*-
4 """
5 @author: Khasan Dymov and Alex Shoop
6 """
7
8 import os
9 import datetime
10 import numpy as np
11 import pandas as pd
12 import matplotlib.pyplot as plt
13 import pandas_datareader as pdr
14 import statsmodels.api as sm

```

```

15 import fix_yahoo_finance as yf #yahoo finance data fix
16 ##### yf.pdr_override() # yahoo finance data fix
17
18 #Specified tickers
19 tickers = [ 'AMZN' , 'AZO' , 'BA' , 'CI' , 'COST' , 'CRBP' , 'DATA' , 'F' , 'GOOG' , 'HUBS' , \
20             'JPM' , 'MFC' , 'MMM' , 'MIN' , 'NFLX' , 'NKE' , 'SBUX' , 'TLT' , 'VYM' , 'WM' ]
21
22 #Market portfolio is S&P 500
23 sp = " ^GSPC "
24
25
26 #Date range
27 startDate = datetime.datetime(2010,1,1)
28 endDate = datetime.datetime(2017,1,1)
29
30 #secDF = pd.DataFrame()
31 #i = 0
32 #for tick in tickers:
33 #    secDF[tick] = pdr.get_data_yahoo(tick, startDate, endDate) [ 'Adj Close ' ]
34 #    print(i)
35 #    i += 1
36
37 #Fetch Yahoo Finance data
38 # NB: If the data reader below breaks, just run again.
39 #     Yahoo data reader can be wacky sometimes.
40 #Using only adj close data
41 print("Attempting to import monthly YF data.....")
42 secData = pdr.get_data_yahoo(tickers, startDate, endDate, interval = 'm') [ 'Adj \
43     Close ' ]
44 SP500Data = pdr.get_data_yahoo(sp, startDate, endDate, interval = 'm') [ 'Adj \
45     Close ' ]
46 print(".\n")
47 print(".\n")
48 print(".\n")
49 print("Successfully imported all YF data.")
50
51 #Simple monthly returns, for securities and market portfolio
52 returnsSecuDF = pd.DataFrame(index = secData.index)
53 returnsSecuDF = secData/secData.shift(1) - 1
54 returnsMarketDF = pd.DataFrame(index = SP500Data.index)
55 returnsMarketDF = SP500Data/SP500Data.shift(1) - 1
56
57 #Expected returns of jth security, E[R_j]
58 expReturnsSecu = returnsSecuDF.mean()
59 #Expected returns of market portfolio, E[R_M]
60 expReturnsMarket = returnsMarketDF.mean()
61 #Annualize the parameters
62 muSecu = expReturnsSecu*12
63 muMarket = expReturnsMarket*12
64
65 #risk-free rate (same as in Markowitz Proj)

```

```

64 muF = 0.02
65
66 #set up the final output dataframe
67 finalAssetsDF = pd.DataFrame(columns = secData.columns, \
68                               index = [ 'Beta' , 'Expected Asset Returns' , 'Asset \
69 → Returns Variance' ])
70 #Beta calculation
71 for stock in finalAssetsDF:
72     tempDF = pd.concat ([ returnsSecuDF [ stock ] , returnsMarketDF ] , axis = 1)
73     tempDF = tempDF.dropna ()      # dropping any NaN row entries
74
75     # Returns for stock security , and market portfolio
76     R_sec = tempDF . iloc [ : , 0]
77     R_market = tempDF . iloc [ : , 1]
78
79     # Excess returns , for regression model usage
80     R_excessSec = R_sec - muF # excess returns for jth security
81     R_excessMarket = R_market - muF # excess returns for market portfolio
82
83     #Beta estimate , via covariance formula between R_j and R_M
84     covariR = np.cov (R_sec , R_market)
85     beta_formula = covariR [ 0 , 1 ] / covariR [ 1 , 1 ]
86     #alpha = expReturnsSecu . AMZN - beta_formula * (expReturnsMarket)
87     #print ("Beta estimate via covariance formula is: ")
88     #print (beta_formula)
89
90     #CAPM beta estimate , via OLS regression model
91     # The formula we are estimating is :
92     #    R*_j = beta (R*M) + eps
93
94     # Setup for OLS fit
95     Y = R_excessSec
96     X = R_excessMarket . rename ( 'ExcessMarketReturn' )
97     X = sm.add_constant (X)
98     modelFit = sm.OLS (Y , X) . fit ()
99     #print ("OLS results are: ")
100    #print (modelFit . summary ())
101
102    #Output results
103    betaCAPM = modelFit . params [ 'ExcessMarketReturn' ]
104    expAssetReturn = muF + betaCAPM * (muMarket - muF)
105    resid = modelFit . resid
106    varAssetReturn = betaCAPM ** 2 * returnsMarketDF . var () + resid . var ()
107
108    #print statement
109    print ("Beta from covariance formula is: ")
110    print (beta_formula)
111    print ("Beta from OLS regression model fit is: ")
112    print (betaCAPM)
113
114    finalAssetsDF [ stock ] = [ modelFit . params [ 'ExcessMarketReturn' ] , \

```

```

114                                         expAssetReturn , \
115                                         varAssetReturn ]
116
117 #Print results
118 finalAssetsDF
119
120
121 # below is for Fama-French
122 #os . chdir (os . path . dirname (os . path . dirname (__file__)))
123 #os . chdir ('Data')
124 file = 'F-F_SML_HML.csv'
125 dataCSV = pd.read_csv( file )
126 FFdata = dataCSV[ [ 'SMB' , 'HML' ] ]
127 FFdata.index = secData.index      # setting correct indices
128 FFdata = FFdata[1:]      # removing 1st row for correct dimensions
129
130 #set up the final output dataframe
131 finalAssetsDFTWO = pd.DataFrame( columns = secData.columns , \
132                                     index = [ 'Beta Mkt' , 'Beta SMB' , \
133                                     'Beta HML' , 'Expected Asset Returns' , \
134                                     'Asset Returns Variance' ] )
135
136 #Beta calculation
137 for stock in finalAssetsDFTWO:
138     tempDF = pd.concat ([ returnsSecuDF [ stock ] , returnsMarketDF , FFdata ] , axis =
139     ↪ 1)
140     tempDF = tempDF.dropna()      # dropping any NaN row entries
141
142     # Returns for stock security , and market portfolio , and Fama French
143     ↪ factors
144     R_sec = tempDF . iloc [ : , 0 ]
145     R_market = tempDF . iloc [ : , 1 ]
146     FFnewdata = pd.concat ([ tempDF . iloc [ : , 2 ] , tempDF . iloc [ : , 3 ] ] , axis = 1 )
147
148     # Excess returns , for regression model usage
149     R_excessSec = R_sec - muF # excess returns for jth security
150     R_excessMarket = R_market - muF # excess returns for market portfolio
151
152     #Beta estimate , via OLS regression model
153     # The formula we are estimating is :
154     #    R*_j = beta1(R*_M) + beta2(SMB) + beta3(HML) + eps
155
156     # Setup for OLS fit
157     Y = R_excessSec
158     X = pd.concat ([ R_excessMarket , FFnewdata ] , axis = 1 )
159     X = X.rename (columns = { 'Adj Close' : 'ExcessMarketReturn' })
160     factorsCOV = X . cov ()
161     X = sm.add_constant (X)
162     modelFit = sm.OLS (Y , X) . fit ()
163     #print ("OLS results are: ")
164     #print (modelFit . summary ())

```

```

163
164 #Output results
165 betas = modelFit.params[ 'ExcessMarketReturn' ], \
166     modelFit.params[ 'SMB' ], \
167     modelFit.params[ 'HML' ]
168 expAssetReturn = muF + betas[0]*(muMarket - muF) + \
169             betas[1]*FFdata[ 'SMB' ].mean() + \
170             betas[2]*FFdata[ 'HML' ].mean()
171 resids = modelFit.resid
172 betasVec = np.array([ betas ])
173 # diff formula to capm version
174 varAssetReturn = np.dot(np.dot(betasVec, factorsCOV), betasVec.T) + resids.
175 ↪ var()
176
177 finalAssetsDFTWO[ stock ] = [ betas[0] , betas[1] , betas[2] , expAssetReturn ,
178 ↪ varAssetReturn.item() ]

```

Appendix C FactorModels_F1F2.py

```

1#!/usr/bin/env python3
2# -*- coding: utf-8 -*-
3"""
4@author: Khasan Dymov and Alex Shoop
5"""
6
7
8import os
9import datetime
10import numpy as np
11import pandas as pd
12import matplotlib.pyplot as plt
13import pandas_datareader as pdr
14import statsmodels.api as sm
15import fix_yahoo_finance as yf #yahoo finance data fix
16#####
17yf.pdr_override() # yahoo finance data fix
18
19#Specified tickers
20tickers = [ "AMZN" , 'AZO' , 'BA' , 'CI' , 'COST' , 'CRBP' , 'DATA' , 'F' , 'GOOG' , 'HUBS' , \
21            'JPM' , 'MFC' , 'MMM' , 'MIN' , 'NFLX' , 'NKE' , 'SBUX' , 'TLT' , 'VYM' , 'WM' ]
22
23#Market portfolio is S&P 500
24sp = "^GSPC"
25
26#Date range
27startDate = datetime.datetime(2016,3,1)
28endDate = datetime.datetime(2017,1,1)
29
30#Fetch Yahoo Finance data

```

```

31 # NB: If the data reader below breaks, just run again.
32 #     Yahoo data reader can be wacky sometimes.
33 #Using only adj close data
34 print("Attempting to import monthly YF data....")
35 secData = pdr.get_data_yahoo(tickers, startDate, endDate, interval = 'm')[ 'Adj
    ↪ Close' ]
36 secData = secData[::-1]
37 SP500Data = pdr.get_data_yahoo(sp, startDate, endDate, interval = 'm')[ 'Adj
    ↪ Close' ]
38 print(".\n")
39 print(".\n")
40 print(".\n")
41 print("Successfully imported all YF data.")

42
43 #Importing F1 (Global ex US Momentum factor) and F2 (Bloomberg RSTAMOM index)
44 file1 = 'Global_ex_US_MOM_Factor_10mo.csv'
45 file2 = 'bloomberg_RSTAMOM_10mo.csv'
46 dataCSV = pd.read_csv(file1)
47 F1data = dataCSV[['MOM']]
48 F1data.index = secData.index      # setting correct indices
49 dataCSV = pd.read_csv(file2)
50 F2data = dataCSV[['RSTAMOM']]
51 F2data.index = secData.index      # setting correct indices

52
53
54 #Simple monthly returns, for securities and market portfolio
55 returnsSecuDF = pd.DataFrame(index = secData.index)
56 returnsSecuDF = secData/secData.shift(1) - 1
57 returnsMarketDF = pd.DataFrame(index = SP500Data.index)
58 returnsMarketDF = SP500Data/SP500Data.shift(1) - 1

59
60 #Expected returns of jth security, E[R_j]
61 expReturnsSecu = returnsSecuDF.mean()
62 #Expected returns of market portfolio, E[R_M]
63 expReturnsMarket = returnsMarketDF.mean()
64 #Annualize the parameters
65 #muSecu = expReturnsSecu*12
66 #muMarket = expReturnsMarket*12m
67 muMarket = (1+expReturnsMarket)**12 - 1

68
69 #risk-free rate (same as in Markowitz Proj)
70 muF = 0.02

71
72 #set up the final output dataframe
73 finalAssetsF1 = pd.DataFrame(columns = secData.columns, \
    index = [ 'Beta CAPM', 'Beta F1', 'Expected Asset
    ↪ Returns', 'Asset Returns Variance'])
74 finalAssetsF2 = pd.DataFrame(columns = secData.columns, \
    index = [ 'Beta CAPM', 'Beta F2', 'Expected Asset
    ↪ Returns', 'Asset Returns Variance'])
75 finalAssetsF1F2 = pd.DataFrame(columns = secData.columns, \

```

```

78         index = [ 'Beta CAPM' , 'Beta F1' , 'Beta F2' , '
79             ↪ Expected Asset Returns' , 'Asset Returns Variance' ])
80
81 # dropping NaN rows
82 tempDF = pd.concat([returnsSecuDF , returnsMarketDF , F1data , F2data] , axis = 1)
83 tempDF = tempDF.dropna()
84 # Returns for ALL securities , and market portfolio
85 # and F1 F2 factors
86 R_sec = tempDF.iloc [:, :20]
87 R_market = tempDF.iloc [:, 20]
88 newF1 = tempDF.iloc [:, 21]
89 newF2 = tempDF.iloc [:, 22]
90 newF1F2 = pd.concat([tempDF.iloc [:, 21] , tempDF.iloc [:, 22]] , axis = 1)
91 # Excess returns , for regression model usage
92 R_excessSec = R_sec - muF # excess returns for jth security
93 R_excessMarket = R_market - muF # excess returns for market portfolio
94
95 # X variable setup , for OLS regression purposes
96 # Also , getting CovF matrix for each factor model
97 XF1 = pd.concat([R_excessMarket , newF1] , axis=1)
98 XF1 = XF1.rename(columns = { 'Adj Close' : 'ExcessMarketReturn' })
99 factorsF1COV = XF1.cov()
100 XF1 = sm.add_constant(XF1)
101
102 XF2 = pd.concat([R_excessMarket , newF2] , axis=1)
103 XF2 = XF2.rename(columns = { 'Adj Close' : 'ExcessMarketReturn' })
104 factorsF2COV = XF2.cov()
105 XF2 = sm.add_constant(XF2)
106
107 XF1F2 = pd.concat([R_excessMarket , newF1F2] , axis=1)
108 XF1F2 = XF1F2.rename(columns = { 'Adj Close' : 'ExcessMarketReturn' })
109 factorsF1F2COV = XF1F2.cov()
110 XF1F2 = sm.add_constant(XF1F2)
111
112 # big beta matrix , for covariance calculation purposes
113 bigBetaF1 = np.zeros(0)
114 bigBetaF2 = np.zeros(0)
115 bigBetaF1F2 = np.zeros(0)
116
117 #Beta calculation FOR CAPM
118 for stock in finalAssetsF1:
119     #CAPM beta estimate , via OLS regression model
120     # The formula we are estimating is :
121     #      R*_j = beta (R*_M) + eps
122
123     # Setup for OLS fit
124     Y = R_excessSec[stock]
125     modelFitF1 = sm.OLS(Y, XF1).fit()
126     modelFitF2 = sm.OLS(Y, XF2).fit()
127     modelFitF1F2 = sm.OLS(Y, XF1F2).fit()

```

```

128 #Output results
129 betasF1CAPM = modelFitF1.params[ 'ExcessMarketReturn' ], modelFitF1.params[ ' 
130   ↪ MOM' ]
131 if stock == 'AMZN':
132     bigBetaF1 = np.append( bigBetaF1 , betasF1CAPM)
133 else:
134     bigBetaF1 = np.vstack([ bigBetaF1 , betasF1CAPM])
135
136 # remember, need to annualize expected returns and variance returns
137 #expAssetF1Return = muF + betasF1CAPM[0]*( muMarket - muF) + betasF1CAPM
138   ↪ [1]*((1+F1data.mean())**12 - 1)
139 expAssetF1Return = (1 + muF + betasF1CAPM[0]*( expReturnsMarket - muF) +
140   ↪ betasF1CAPM[1]*(F1data.mean().item()))**12 - 1
141 residsF1 = modelFitF1.resid
142 betasF1Vec = np.array([betasF1CAPM])
143 varAssetF1Return = (np.dot(np.dot(betasF1Vec,factorsF1COV), betasF1Vec.T)
144   ↪ + residsF1.var())*12
145 finalAssetsF1[stock] = [betasF1CAPM[0] , betasF1CAPM[1] , expAssetF1Return.
146   ↪ item() , varAssetF1Return.item()]
147
148 betasF2CAPM = modelFitF2.params[ 'ExcessMarketReturn' ], modelFitF2.params[ ' 
149   ↪ RSTAMOM' ]
150 if stock == 'AMZN':
151     bigBetaF2 = np.append( bigBetaF2 , betasF2CAPM)
152 else:
153     bigBetaF2 = np.vstack([ bigBetaF2 , betasF2CAPM])
154
155 #expAssetF2Return = muF + betasF2CAPM[0]*( muMarket - muF) + betasF2CAPM
156   ↪ [1]*((1+F2data.mean())**12 - 1)
157 expAssetF2Return = (1 + muF + betasF2CAPM[0]*( expReturnsMarket - muF) +
158   ↪ betasF2CAPM[1]*(F2data.mean().item()))**12 - 1
159 residsF2 = modelFitF2.resid
160 betasF2Vec = np.array([betasF2CAPM])
161 varAssetF2Return = (np.dot(np.dot(betasF2Vec,factorsF2COV), betasF2Vec.T)
162   ↪ + residsF2.var())*12
163 finalAssetsF2[stock] = [betasF2CAPM[0] , betasF2CAPM[1] , expAssetF2Return.
164   ↪ item() , varAssetF2Return.item()]
165
166 betasF1F2CAPM = modelFitF1F2.params[ 'ExcessMarketReturn' ], modelFitF1F2.
167   ↪ params[ 'MOM' ],modelFitF1F2.params[ 'RSTAMOM' ]
168 if stock == 'AMZN':
169     bigBetaF1F2 = np.append( bigBetaF1F2 , betasF1F2CAPM)
170 else:
171     bigBetaF1F2 = np.vstack([ bigBetaF1F2 , betasF1F2CAPM])
172
173 #expAssetF1F2Return = muF + betasF1F2CAPM[0]*( muMarket - muF) +
174   ↪ betasF1F2CAPM[1]*((1+F1data.mean().item())**12 - 1) + betasF1F2CAPM
175   ↪ [2]*((1+F2data.mean().item())**12 - 1)
176 expAssetF1F2Return = (1 + muF + betasF1F2CAPM[0]*( expReturnsMarket - muF)
177   ↪ + betasF1F2CAPM[1]*(F1data.mean().item()) + betasF1F2CAPM[2]*( F2data.
178   ↪ mean().item()) )**12 - 1

```

```

164 residsF1F2 = modelFitF1F2.resid
165 betasF1F2Vec = np.array([betasF1F2CAPM])
166 varAssetF1F2Return = (np.dot(np.dot(betasF1F2Vec, factorsF1F2COV),
167 ↪ betasF1F2Vec.T) + residsF1F2.var())*12
168 finalAssetsF1F2[stock] = [betasF1F2CAPM[0], betasF1F2CAPM[1],
169 ↪ betasF1F2CAPM[2], expAssetF1F2Return, varAssetF1F2Return.item()]
170
171 # calculating covariance between asset returns
172 covAssetReturnsCAPM_F1 = np.empty((20, 20))
173 covAssetReturnsCAPM_F1[:] = np.nan
174 covAssetReturnsCAPM_F2 = np.empty((20, 20))
175 covAssetReturnsCAPM_F2[:] = np.nan
176 covAssetReturnsCAPM_F1F2 = np.empty((20, 20))
177 covAssetReturnsCAPM_F1F2[:] = np.nan
178 for i in range(20):
179     for j in range(20):
180         covAssetReturnsCAPM_F1[i][j] = np.dot(np.dot(bigBetaF1[i],
181 ↪ factorsF1COV), bigBetaF1[j])
182         covAssetReturnsCAPM_F2[i][j] = np.dot(np.dot(bigBetaF2[i],
183 ↪ factorsF2COV), bigBetaF2[j])
184         covAssetReturnsCAPM_F1F2[i][j] = np.dot(np.dot(bigBetaF1F2[i],
185 ↪ factorsF1F2COV), bigBetaF1F2[j])
186 covPD_CAPM_F1 = pd.DataFrame(covAssetReturnsCAPM_F1, index = secData.columns,
187 ↪ columns=secData.columns)
188 covPD_CAPM_F2 = pd.DataFrame(covAssetReturnsCAPM_F2, index = secData.columns,
189 ↪ columns=secData.columns)
190 covPD_CAPM_F1F2 = pd.DataFrame(covAssetReturnsCAPM_F1F2, index = secData.
191 ↪ columns, columns=secData.columns)
192
193 #####
194
195 # below is for Fama-French
196 file = 'F-F_SML_HML_10mo.csv'
197 dataCSV = pd.read_csv(file)
198 FFdata = dataCSV[['SMB', 'HML']]
199 FFdata.index = secData.index      # setting correct indices
200 FFdata = FFdata[1:]      # removing 1st row for correct dimensions
201
202 #set up the final output dataframe
203 finalAssetsFF_F1 = pd.DataFrame(columns = secData.columns, index = [
204 'Beta Mkt',
205 ↪ 'Beta SMB', 'Beta HML', 'Beta F1', 'Expected Asset Returns', 'Asset Returns',
206 ↪ 'Variance'])
207
208 finalAssetsFF_F2 = pd.DataFrame(columns = secData.columns, \
209 index = [
210 'Beta Mkt', 'Beta SMB', \
211 'Beta HML', 'Beta F2', 'Expected Asset',
212 ↪ 'Returns', \
213 'Asset Returns Variance'])

```

```

204 finalAssetsFF_F1F2 = pd.DataFrame(columns = secData.columns, \
205                                     index = [ 'Beta Mkt' , 'Beta SMB' , \
206                                               'Beta HML' , 'Beta F1' , 'Beta F2' , 'Expected \
207                                               ↪ Asset Returns' , \
208                                               ↪ Asset Returns Variance '])
209 # dropping NaN rows
210 tempDF = pd.concat ([ returnsSecuDF , returnsMarketDF , FFdata , F1data , F2data ] , \
211                      ↪ axis = 1)
212 tempDF = tempDF.dropna()
213 # Returns for ALL securities, and market portfolio, and Fama–French factors
214 # and F1 F2 factors
215 R_sec = tempDF . iloc [:, :20]
216 R_market = tempDF . iloc [:, 20]
217 FFnewdata = pd.concat ([ tempDF . iloc [:, 21] , tempDF . iloc [:, 22]] , axis = 1)
218 newF1 = tempDF . iloc [:, 23]
219 newF2 = tempDF . iloc [:, 24]
220 newF1F2 = pd.concat ([ tempDF . iloc [:, 23] , tempDF . iloc [:, 24]] , axis = 1)
221 # Excess returns, for regression model usage
222 R_excessSec = R_sec - muF # excess returns for jth security
223 R_excessMarket = R_market - muF # excess returns for market portfolio
224
225 # X variable setup, for OLS regression purposes
226 # Also, getting CovF matrix for each factor model
227 XF1 = pd.concat ([ R_excessMarket , FFnewdata , newF1] , axis = 1)
228 XF1 = XF1.rename (columns = { 'Adj Close' : 'ExcessMarketReturn' })
229 factorsFF_F1COV = XF1 . cov()
230 XF1 = sm.add_constant (XF1)
231
232 XF2 = pd.concat ([ R_excessMarket , FFnewdata , newF2] , axis = 1)
233 XF2 = XF2.rename (columns = { 'Adj Close' : 'ExcessMarketReturn' })
234 factorsFF_F2COV = XF2 . cov()
235 XF2 = sm.add_constant (XF2)
236
237 XF1F2 = pd.concat ([ R_excessMarket , FFnewdata , newF1 , newF2] , axis = 1)
238 XF1F2 = XF1F2.rename (columns = { 'Adj Close' : 'ExcessMarketReturn' })
239 factorsFF_F1F2COV = XF1F2 . cov()
240 XF1F2 = sm.add_constant (XF1F2)
241
242 # big beta matrix, for covariance calculation purposes
243 bigBetaFF_F1 = np.zeros (0)
244 bigBetaFF_F2 = np.zeros (0)
245 bigBetaFF_F1F2 = np.zeros (0)
246
247 #Beta calculation FOR FRENCH-FAMA
248 for stock in finalAssetsFF_F1:
249     #Beta estimate, via OLS regression model
250     # The formula we are estimating is:
251     #   R*_j = beta1(R*_M) + beta2(SMB) + beta3(HML) + eps
252
253     # Setup for OLS fit
254     Y = R_excessSec [stock]
```

```

253 modelFitF1 = sm.OLS(Y, XF1).fit()
254 modelFitF2 = sm.OLS(Y, XF2).fit()
255 modelFitF1F2 = sm.OLS(Y, XF1F2).fit()
256
257 #Output results
258 betasFF_F1 = modelFitF1.params[ 'ExcessMarketReturn' ], modelFitF1.params[ ,
259   ↪ 'SMB' ], \
260   modelFitF1.params[ 'HML' ], modelFitF1.params[ 'MOM' ]
261 if stock == 'AMZN':
262     bigBetaFF_F1 = np.append( bigBetaFF_F1, betasFF_F1 )
263 else:
264     bigBetaFF_F1 = np.vstack( [ bigBetaFF_F1, betasFF_F1 ] )
265
266 #expAssetReturnFF_F1 = muF + betasFF_F1[0]*( muMarket - muF ) + \
267 #           betasFF_F1[1]*((1+FFdata[ 'SMB' ].mean())**12 - 1) + betasFF_F1
268 ↪ [2]*((1+FFdata[ 'HML' ].mean())**12 - 1) + \
269 #           betasFF_F1[3]*((1+F1data.mean().item())**12 - 1)
270 expAssetReturnFF_F1 = (1 + muF + betasFF_F1[0]*( expReturnsMarket - muF ) +
271 ↪ \
272     betasFF_F1[1]*( FFdata[ 'SMB' ].mean() ) + betasFF_F1[2]*( FFdata[ ,
273   ↪ 'HML' ].mean() ) + \
274     betasFF_F1[3]*( F1data.mean().item() ) )**12 - 1
275
276 residsF1 = modelFitF1.resid
277 betasF1Vec = np.array([ betasFF_F1 ])
278 varAssetReturnFF_F1 = (np.dot(np.dot(betasF1Vec,factorsFF_F1COV),
279   ↪ betasF1Vec.T) + residsF1.var())*12
280 finalAssetsFF_F1[stock] = [ betasFF_F1[0], betasFF_F1[1], betasFF_F1[2],
281   ↪ betasFF_F1[3], expAssetReturnFF_F1, varAssetReturnFF_F1.item() ]
282
283
284 betasFF_F2 = modelFitF2.params[ 'ExcessMarketReturn' ], modelFitF2.params[ ,
285   ↪ 'SMB' ], \
286   modelFitF2.params[ 'HML' ], modelFitF2.params[ 'RSTAMOM' ]
287 if stock == 'AMZN':
288     bigBetaFF_F2 = np.append( bigBetaFF_F2, betasFF_F2 )
289 else:
290     bigBetaFF_F2 = np.vstack( [ bigBetaFF_F2, betasFF_F2 ] )
291
292 #expAssetReturnFF_F2 = muF + betasFF_F2[0]*( muMarket - muF ) + \
293 #           betasFF_F2[1]*((1+FFdata[ 'SMB' ].mean())**12 - 1) + betasFF_F2
294 ↪ [2]*((1+FFdata[ 'HML' ].mean())**12 - 1) + \
295 #           betasFF_F2[3]*((1+F2data.mean().item())**12 - 1)
296 expAssetReturnFF_F2 = (1 + muF + betasFF_F2[0]*( expReturnsMarket - muF ) +
297 ↪ \
298     betasFF_F2[1]*( FFdata[ 'SMB' ].mean() ) + betasFF_F2[2]*( FFdata[ ,
299   ↪ 'HML' ].mean() ) + \
300     betasFF_F2[3]*( F2data.mean().item() ) )**12 - 1
301 residsF2 = modelFitF2.resid
302 betasF2Vec = np.array([ betasFF_F2 ])
303 varAssetReturnFF_F2 = (np.dot(np.dot(betasF2Vec,factorsFF_F2COV),
304   ↪ betasF2Vec.T) + residsF2.var())*12

```

```

293     finalAssetsFF_F2[stock] = [betasFF_F2[0], betasFF_F2[1], betasFF_F2[2],  

294     ↪ betasFF_F2[3], expAssetReturnFF_F2, varAssetReturnFF_F2.item()]  

295  

296     betasFF_F1F2 = modelFitF1F2.params['ExcessMarketReturn'], modelFitF1F2.  

297     ↪ params['SMB'], \  

298         modelFitF1F2.params['HML'], modelFitF1F2.params['MOM'],  

299     ↪ modelFitF1F2.params['RSTAMOM']  

300     if stock == 'AMZN':  

301         bigBetaFF_F1F2 = np.append(bigBetaFF_F1F2, betasFF_F1F2)  

302     else:  

303         bigBetaFF_F1F2 = np.vstack([bigBetaFF_F1F2, betasFF_F1F2])  

304  

305     #expAssetReturnFF_F1F2 = muF + betasFF_F1F2[0]*(muMarket - muF) + \  

306     #           betasFF_F1F2[1]*((1+FFdata['SMB'].mean())**12 - 1) +  

307     ↪ betasFF_F1F2[2]*((1+FFdata['HML'].mean())**12 - 1) + \  

308     #           betasFF_F1F2[3]*((1+F1data.mean().item())**12 - 1) +  

309     ↪ betasFF_F1F2[4]*((1+F2data.mean().item())**12 - 1)  

310     expAssetReturnFF_F1F2 = (1 + muF + betasFF_F1F2[0]*(expReturnsMarket - muF  

311     ↪ ) + \  

312         betasFF_F1F2[1]*(FFdata['SMB'].mean()) + betasFF_F1F2[2]*(  

313     ↪ FFdata['HML'].mean()) + \  

314         betasFF_F1F2[3]*(F1data.mean().item()) + betasFF_F1F2[4]*(  

315     ↪ F2data.mean().item())**12 - 1  

316     residSF1F2 = modelFitF1F2.resid  

317     betasF1F2Vec = np.array([betasFF_F1F2])  

318     varAssetReturnFF_F1F2 = (np.dot(np.dot(betasF1F2Vec, factorsFF_F1F2COV),  

319     ↪ betasF1F2Vec.T) + residSF1F2.var())*12  

320     finalAssetsFF_F1F2[stock] = [betasFF_F1F2[0], betasFF_F1F2[1], betasFF_F1F2  

321     ↪ [2], betasFF_F1F2[3], betasFF_F1F2[4], expAssetReturnFF_F1F2,  

322     ↪ varAssetReturnFF_F1F2.item()]  

323  

324     # calculating covariance between asset returns  

325     covAssetReturnsFF_F1 = np.empty((20, 20))  

326     covAssetReturnsFF_F1[:] = np.nan  

327     covAssetReturnsFF_F2 = np.empty((20, 20))  

328     covAssetReturnsFF_F2[:] = np.nan  

329     covAssetReturnsFF_F1F2 = np.empty((20, 20))  

330     covAssetReturnsFF_F1F2[:] = np.nan  

331     for i in range(20):  

332         for j in range(20):  

333             covAssetReturnsFF_F1[i][j] = np.dot(np.dot(bigBetaFF_F1[i],  

334             ↪ factorsFF_F1COV), bigBetaFF_F1[j])  

335             covAssetReturnsFF_F2[i][j] = np.dot(np.dot(bigBetaFF_F2[i],  

336             ↪ factorsFF_F2COV), bigBetaFF_F2[j])  

337             covAssetReturnsFF_F1F2[i][j] = np.dot(np.dot(bigBetaFF_F1F2[i],  

338             ↪ factorsFF_F1F2COV), bigBetaFF_F1F2[j])  

339     covPD_FF_F1 = pd.DataFrame(covAssetReturnsFF_F1, index = secData.columns,  

340     ↪ columns=secData.columns)  

341     covPD_FF_F2 = pd.DataFrame(covAssetReturnsFF_F2, index = secData.columns,  

342     ↪ columns=secData.columns)

```

```

327 covPD_FF_F1F2 = pd.DataFrame(covAssetReturnsFF_F1F2, index = secData.columns,
328                                columns=secData.columns)
329
329 # norm length covariance comparisons, etc
330 np.linalg.norm(covPD_CAPM_F1 - covPD_CAPM_F2)

```

Appendix D EF_Portfolio_multiWeek_plotting.py

```

1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """
4  @author: Khasan Dymov and Alex Shoop
5  """
6
7
8  import numpy as np
9  import matplotlib.pyplot as plt
10 from pandas_datareader import data
11 import datetime
12 from time import sleep
13
14 #####
15 #setting output np options
16 np.set_printoptions(threshold = np.inf)
17
18 # OPTIMAL PORTFOLIO FUNCTION
19 def optimalPortfolio(mu, omega, muP):
20     # We follow by example
21     # the process as explained in Statistics & Finance by David Ruppert
22     # Necessary coefficients for risk-efficient portfolio
23     n1 = np.matrix(np.ones_like(mu)).T
24     n1t = n1.T
25     invomega = np.linalg.inv(omega)
26     muR = np.matrix(mu).T
27     mut = muR.T
28
29     A = np.matrix.item(n1t*invomega*muR)
30     B = np.matrix.item(mut*invomega*muR)
31     C = np.matrix.item(n1t*invomega*n1)
32     D = B*C-A**2
33
34
35     # Fixed vectors for efficient weights calculation
36     g = (B*invomega*n1 - A*invomega*muR)/D
37     h = (C*invomega*muR - A*invomega*n1)/D
38
39     # The final formula:
40     wts = g + muP*h
41

```

```

42 # Our efficient weights
43     return wts
44
45 #Import specified date range of prices data,
46 #and calculated returns (log returns)
47
48 #Initial beginning date range
49 startDate = datetime.datetime(2015,10,27)
50 endDate = datetime.datetime(2017,10,27)
51
52 #Specified tickers
53 tickers = [ 'AMZN' , 'AZO' , 'BA' , 'CI' , 'COST' , 'CRBP' , 'DATA' , 'F' , 'GOOG' , 'HUBS' , \
54             'JPM' , 'MFC' , 'MMM' , 'MIN' , 'NFLX' , 'NKE' , 'SBUX' , 'TLT' , 'VYM' , 'WM' ]
55
56 #Money weights (risk-free + securities) of our portfolio for plotting purposes
57 ourPortfolio = [500000/1000000, 487595/998050, 439878/978143, 437527/974728, \
58                 441136/978516]
59
60 # BIG WHILE LOOP FOR MULTIPLE WEEKS, IN ORDER TO DO BIG PLOT
61 count = 0
62 # change count to 6, for finale where historical data range should be:
63 # startdate = 2015, 12, 1
64 # enddate = 2017, 12, 1
65 while count < 5:
66     #Rolling forward the date range if we should be Oct 27 or later
67     if count != 0:
68         # add 1 more week
69         startDate += datetime.timedelta(7)
70         endDate += datetime.timedelta(7)
71
72     #Use pandas_reader.data.DataReader to load desired data
73
74     ## this could be more efficiently implemented
75
76     # if the data reader below breaks, just run again.
77     # Yahoo data reader can be wacky sometimes.
78
79     securityRawData = data.DataReader(tickers, 'yahoo', startDate, endDate)
80
81
82     #Getting adjusted close prices
83     close = securityRawData['Adj Close'].astype('float')
84     #fixing date ordering
85     close = close.iloc[:-1]
86     #getting log returns
87     dailyLogReturns = np.log(close).diff()
88     #removing the 1st row, of empty values for first date
89     dailyLogReturns = dailyLogReturns.drop(startDate)
90     #Pulling just the last day closing prices for rebalancing function
91     finalprices = np.matrix([close.iloc[:-1].iloc[0]]).T
92

```

```

93 #Gather expected returns, and covariance matrix
94 mu2yrsdaily = dailyLogReturns.mean()
95 var2yrsdaily = dailyLogReturns.var()
96 sd2yrsdaily = np.sqrt(var2yrsdaily)
97 cov2yrsdaily = dailyLogReturns.cov()
98
99 #Annualize parameter estimates
100
101 #2 year lookback:
102 muannual = np.sqrt(((1+mu2yrsdaily)**505))-1
103 varannual = np.sqrt(252)*var2yrsdaily
104 sdannual = 252*sd2yrsdaily
105 covannual = 252*cov2yrsdaily
106
107
108 #Step 1: Find the minimum variance portfolio for our problem
109
110 n1 = np.matrix(np.ones_like(muannual)).T
111 n1t = n1.T
112 invomega = np.linalg.inv(covannual)
113 muR = np.matrix(muannual).T
114 mut = muR.T
115
116 A = np.matrix.item(n1t*invomega*muR)
117 B = np.matrix.item(mut*invomega*muR)
118 C = np.matrix.item(n1t*invomega*n1)
119 D = B*C-A**2
120
121 #Need to use specific g,h for minimum variance portfolio
122
123 gmin = (B*invomega*n1 - A*invomega*muR)/D
124 hmin = (C*invomega*muR - A*invomega*n1)/D
125 g1 = gmin.T*np.matrix(covannual)*gmin
126 h1 = hmin.T*np.matrix(covannual)*hmin
127 g1h1 = gmin.T*np.matrix(covannual)*hmin
128
129 mumin = np.matrix.item(-g1h1/h1)
130 sigmamin = np.matrix.item(np.sqrt(g1 - g1h1**2/h1))
131
132 #Weights associated with this minimum variance portfolio
133 wtsmin = np.array(gmin + mumin*hmin)
134 barX = np.arange(np.size(wtsmin))
135
136 #Step 2: Calculate the efficient frontier
137
138 muP = np.linspace(min(muannual), max(muannual), 100)
139 sigmaP = np.zeros(0)
140
141 for target in muP:
142     wtsP = optimalPortfolio(muannual, covannual, target)
143     sigmaP = np.append(sigmaP, np.sqrt(wtsP.T*np.matrix(covannual)*wtsP)))

```

```

144
145 effindices = np.where(muP>mumin)
146 ineffindices = np.where(muP<mumin)
147
148 #Step 2.5: Calculate the tangency portfolio
149 muF = 0.02
150 omegabar = invomega*(muR - muF*n1)
151 omegaT = omegabar/(n1t*omegabar)
152 sigmaT = np.matrix.item(np.sqrt(omegaT.T*np.matrix(covannual)*omegaT))
153 muT = np.matrix.item(muT*omegaT)
154
155 #Step 3: Plot minimum variance portfolio and upper section of efficient
156 # frontier
157 plt.plot(sigmaP[effindices], muP[effindices], 'g', alpha = 2/5 + (1/10)*
158 count)
159 plt.plot(sigmaP[ineffindices], muP[ineffindices], 'c—', alpha = 2/5 +
160 (1/10)*count)
161 #plt.plot(sigmaT, muT, 'rx', markersize = 8, alpha = 3/4 - (1/10)*count)
162 if count == 4:
163     plt.plot(sigmaT*ourPortfolio[-1], muT*ourPortfolio[-1], 'ro',
164 markersize = 8, label = "Our Portfolio")
165 plt.plot(sigmaT*ourPortfolio[count], muT*ourPortfolio[count], 'ro',
166 markersize = 8, alpha = 2/5 + (1/10)*count)
167 plt.plot([0, sigmaT], [muF, muT], 'b—', alpha = 2/5 + (1/10)*count)
168 # Just one plot function for the risk-free rate
169 if count == 0:
170     plt.plot(0, muF, 'bo', label = "Risk-free rate")
171 # increment loop
172 count += 1
173
174 plt.title("Efficient Frontier for 20 risky assets, across multiple 2-yr ranges
175 ")
176 plt.xlabel(r"$risk = \sigma_r$")
177 plt.ylabel(r"$reward = \mu_r$")
178 plt.legend()
179 plt.show()
180
181 #Step 4: Bar chart of weights
182
183 #plt.bar(barX, (wtsmin)*100)
184 #plt.axhline(y = 0, linewidth = 0.5, color = 'black')
185 #plt.xticks(barX, stocks, rotation=60)
186 #plt.ylabel("Percentage")
187 #plt.title("Weights for min var portfolio (2yrs)")
188
189 #Step 5: Bar chart of weights in tangency portfolio
190 wtstangent = optimalPortfolio(muannual, covannual, muT)
191
192 #plt.bar(barX, (wtstangent)*100)
193 #plt.axhline(y = 0, linewidth = 0.5, color = 'black')

```

```

189 #plt.xticks(barX, stocks, rotation=60)
190 #plt.ylabel("Percentage")
191 #plt.title("Weights for Initial Tangency Portfolio")
192
193 print("Expected return = " + str(muT))
194 print("Portfolio std. dev = " + str(sigmaT))
195 print("The weights of the assets in the tangency portfolio are " + str(
    ↪ wtstangent))

```

Appendix E CAPM_altBeta.py

```

1 #!/usr/bin/env python3
2 # -*- coding: utf-8 -*-
3 """
4 @author: Khasan Dymov and Alex Shoop
5 """
6
7
8 import os
9 import datetime
10 import numpy as np
11 import pandas as pd
12 import matplotlib.pyplot as plt
13 import pandas_datareader as pdr
14 import statsmodels.api as sm
15 import fix_yahoo_finance as yf #yahoo finance data fix
16 ######
17 yf.pdr_override() # yahoo finance data fix
18
19 #Specified tickers
20 tickers = ["AMZN", 'AZO', 'BA', 'CI', 'COST', 'CRBP', 'DATA', 'F', 'GOOG', 'HUBS', \
21             'JPM', 'MFC', 'MMM', 'MTN', 'NFLX', 'NKE', 'SBUX', 'TLT', 'VYM', 'WM']
22
23 #Market portfolio is S&P 500
24 sp = "^GSPC"
25
26 #Date range, 5 years
27 startDate = datetime.datetime(2012,1,1)
28 endDate = datetime.datetime(2017,1,1)
29
30 #Fetch Yahoo Finance data
31 # NB: If the data reader below breaks, just run again.
32 #     Yahoo data reader can be wacky sometimes.
33 #Using only adj close data
34 print("Attempting to import monthly YF data.....")
35 secData = pdr.get_data_yahoo(tickers, startDate, endDate, interval = 'm')[ 'Adj \
    ↪ Close' ]
36 #below is only necessary if date range is smaller, like 2 years.
37 #secData = secData[::-1]      # fixing ordering, so that it's oldest to newest

```

```

38 SP500Data = pdr.get_data_yahoo(sp, startDate, endDate, interval = 'm')[ 'Adj
  ↪ Close' ]
39 print(".\n")
40 print(".\n")
41 print(".\n")
42 print("Successfully imported all YF data.")
43
44
45 #Simple monthly returns, for securities and market portfolio
46 returnsSecuDF = pd.DataFrame(index = secData.index)
47 returnsSecuDF = secData/secData.shift(1) - 1
48 returnsMarketDF = pd.DataFrame(index = SP500Data.index)
49 returnsMarketDF = SP500Data/SP500Data.shift(1) - 1
50
51 #Expected returns of jth security, E[R_j]
52 expReturnsSecu = returnsSecuDF.mean()
53 #Expected returns of market portfolio, E[R_M]
54 expReturnsMarket = returnsMarketDF.mean()
55 #Annualize the parameters
56 muSecu = expReturnsSecu*12
57 muMarket = expReturnsMarket*12
58
59 #risk-free rate (same as in Markowitz Proj)
60 muF = 0.02
61
62 #set up the final output dataframe
63 finalAssets_a = pd.DataFrame(columns = secData.columns, \
64                               index = [ 'Beta CAPM' ])
65
66 #a, Historical returns data Beta calculation
67 for stock in finalAssets_a:
68     tempDF = pd.concat([returnsSecuDF[stock], returnsMarketDF], axis = 1)
69     tempDF = tempDF.dropna()    # dropping any NaN row entries, ie not all
  ↪ historical data available for this security
70
71 # Returns for stock security, and market portfolio
72 R_sec = tempDF.iloc[:, 0]
73 R_market = tempDF.iloc[:, 1]
74
75 # Excess returns, for regression model usage
76 R_excessSec = R_sec - muF # excess returns for jth security
77 R_excessMarket = R_market - muF # excess returns for market portfolio
78
79 #CAPM beta estimate, via OLS regression model
80 # The formula we are estimating is:
81 #   R*_j = beta (R*_M) + eps
82
83 # Setup for OLS fit
84 Y = R_excessSec
85 Xhist = R_excessMarket
86 Xhist = Xhist.rename('ExcessMarketReturn')

```

```

87 Xhist = sm.add_constant(Xhist)
88 modelFit_a = sm.OLS(Y, Xhist).fit()
89
90 #Output results
91 betas_a = modelFit_a.params['ExcessMarketReturn']
92 finalAssets_a[stock] = [betas_a]
93
94 # print results
95 orderedBetas_a = finalAssets_a.transpose().sort_values(by='Beta CAPM')
96
97 portfolioReturns = [-0.00490, -0.00200, -0.00520, 0.00660, 0.00200, -0.00090, \
98                 -0.00190, -0.00650, -0.01000, -0.00070, 0.00110, 0.00060, \
99                 -0.00030, -0.00100, -0.00420, 0.00120, 0.00290, -0.00180, \
100                0.00240, 0.00090, 0.00160, -0.00230, 0.00680, -0.00260, 0.00530]
101 SP500portfolioRelated = [-0.003192, 0.000944, 0.001592, 0.000190, 0.003097, \
102             ↪ 0.001271, \
103                 -0.000189, 0.001444, -0.003762, -0.000898, 0.000984, \
104                 -0.002310, -0.005526, 0.008196, -0.002626, 0.001276, \
105                 0.006541, -0.000750, 0.002056, -0.000384, 0.009848, \
106                 -0.000369, 0.008191, -0.002024, -0.001052]
107 plt.hist(portfolioReturns, 25, edgecolor = 'black')
108 plt.xlabel("Returns")
109 plt.ylabel("Frequency")
110 plt.title("Distribution of Our Portfolio Returns")
111
112 tempDF = pd.concat([pd.DataFrame(portfolioReturns), pd.DataFrame( \
113             ↪ SP500portfolioRelated)], axis = 1)
114 # Returns for our portfolio, and market portfolio
115 R_port = tempDF.iloc[:, 0].rename("Portfolio")
116 R_market = tempDF.iloc[:, 1].rename("Market")
117 # Excess returns, for regression model usage
118 newR_excessPort = R_port - muF
119 R_excessMarket = R_market - muF
120 modelFitPortfolio = sm.OLS(newR_excessPort, R_excessMarket).fit()
121 portfolioBETA = modelFitPortfolio.params['Market']
122 #####
123 # b, least trimmed squares
124
125 #set up the final output dataframe
126 finalAssets_b = pd.DataFrame(columns = secData.columns, \
127                               index = ['Beta LTS'])
128 #a, Historical returns data Beta calculation
129 for stock in finalAssets_b:
130     tempDF = pd.concat([returnsSecuDF[stock], returnsMarketDF], axis = 1)
131     tempDF = tempDF.dropna()    # dropping any NaN row entries, ie not all \
132     ↪ historical data available for this security
133
134     # Returns for stock security, and market portfolio
135     R_sec = tempDF.iloc[:, 0]

```

```

135 R_market = tempDF.iloc[:, 1]
136 q_alpha = R_sec.quantile(0.05)
137 q_alphaOpp = R_sec.quantile(1 - 0.05)
138 # setting the alpha-quantiles to the actual values. Thus, trimming it.
139 R_sec[R_sec < q_alpha] = q_alpha
140 R_sec[R_sec > q_alphaOpp] = q_alphaOpp
141
142 # Excess returns, for regression model usage
143 R_excessSec = R_sec - muF # excess returns for jth security
144 R_excessMarket = R_market - muF # excess returns for market portfolio
145
146 # Setup for OLS fit
147 Y = R_excessSec
148 Xtrim = R_excessMarket
149 Xtrim = Xtrim.rename('ExcessMarketReturn')
150 Xtrim = sm.add_constant(Xtrim)
151 modelFit_b = sm.OLS(Y, Xtrim).fit()
152
153 #Output results
154 betas_b = modelFit_b.params['ExcessMarketReturn']
155 finalAssets_b[stock] = [betas_b]
156
157 # choose VYM as representative asset
158 newVYM = returnsSecuDF['VYM']
159 newVYM = newVYM.dropna()
160 q_alpha = newVYM.quantile(0.05)
161 q_alphaOpp = newVYM.quantile(1 - 0.05)
162 # setting the alpha-quantiles to the actual values. Thus, trimming it.
163 newVYM[newVYM < q_alpha] = q_alpha
164 newVYM[newVYM > q_alphaOpp] = q_alphaOpp
165
166 # new regression
167 R_newMarket = returnsMarketDF.iloc[1:]
168 # Excess returns, for regression model usage
169 R_excessSec = newVYM - muF # excess returns
170 R_excessMarket = R_newMarket - muF # excess returns for market portfolio
171 # Setup for OLS fit
172 Y = R_excessSec
173 Xtrim = R_excessMarket
174 Xtrim = Xtrim.rename('ExcessMarketReturn')
175 Xtrim = sm.add_constant(Xtrim)
176 modelFit_b = sm.OLS(Y, Xtrim).fit()
177
178 betas_b_VYM = modelFit_b.params['ExcessMarketReturn']
179
180 # histogram plotting
181 vymReturns = returnsSecuDF['VYM'].dropna()
182 binBounds1 = np.linspace(min(newVYM), max(newVYM), 59)
183 binBounds2 = np.linspace(min(vymReturns), max(vymReturns), 59)
184

```

```

185 plt.hist(newVYM, bins = binBounds2, label = 'Trimmed', color = 'red') #
    ↪ trimmed distribution
186 plt.hist(vymReturns, bins = binBounds2, label = 'Original', edgecolor = 'black'
    ↪ ') # original distribution
187 plt.title("Returns distribution for VYM")
188 plt.xlabel("Returns")
189 plt.ylabel("Frequency")
190 plt.legend()
191 plt.show()
192 #####
193 # c, constant beta shrinkage, alpha_i = 2/3
194 betaBar = orderedBetas_a.mean().item()
195 alphaShrink = 2/3
196 # James–Stein estimator
197 betasJS = betaBar + alphaShrink*(orderedBetas_a - betaBar)
198 #####
199 # d, exponentially weighted moving average
200 lamb1 = 0.94      # the standard for 'monthly data', according to Professor
    ↪ Blais
201 lamb2 = 0.8
202 # need to make sure returns are ordered from newest to oldest
203 newToOldSecReturns = returnsSecuDF[::-1]
204 newMarketReturns = returnsMarketDF[::-1]
205 #####
206 # covariance estimate
207 def covExpWeighted(assetReturns, marketReturns, t, lam):
208     if t == 0:
209         return 0
210     else:
211         return (1 - lam)*assetReturns[t-1]*marketReturns[t-1] + lam*(
    ↪ covExpWeighted(assetReturns, marketReturns, t-1, lam))
212 # variance estimate
213 def varExpWeighted(marketReturns, t, lam):
214     if t == 0:
215         return 0
216     else:
217         return (1 - lam)*marketReturns[t]**2 + lam*varExpWeighted(
    ↪ marketReturns, t-1, lam)
218 #####
219 #set up the final output dataframe
220 finalAssets_d = pd.DataFrame(columns = newToOldSecReturns.columns, \
221                               index = [ 'EWMA Beta (Lambda=0.94)', 'EWMA Beta ('
    ↪ Lambda=0.8) '])
222 #####
223 # exponentially weighted avg beta calculation for each stock
224 for stock in finalAssets_d:
225     tempDF = pd.concat([newToOldSecReturns[stock], newMarketReturns], axis =
    ↪ 1)

```

```
228 tempDF = tempDF.dropna()      # dropping any NaN row entries, ie not all
→ historical data available for this security
229
230 # Returns for stock security, and market portfolio
231 R_sec = tempDF.iloc[:, 0]
232 R_market = tempDF.iloc[:, 1]
233
234 # function values
235 cov_lamb1 = covExpWeighted(R_sec, R_market, len(R_sec), lamb1)
236 cov_lamb2 = covExpWeighted(R_sec, R_market, len(R_sec), lamb2)
237 var_lamb1 = varExpWeighted(R_market, len(R_sec)-1, lamb1)
238 var_lamb2 = varExpWeighted(R_market, len(R_sec)-1, lamb2)
239
240 beta_lamb1 = cov_lamb1 / var_lamb1
241 beta_lamb2 = cov_lamb2 / var_lamb2
242
243 finalAssets_d [stock] = [beta_lamb1, beta_lamb2]
244
245 # print results
246 finalAssets_d
247 orderedBetas_d = finalAssets_d.transpose().sort_values(by='EWMA Beta (Lambda
→ =0.94)')
```

Appendix F Full Trade History

DEMO ACCOUNT - SIMULATED TRADING

VYM	2017-11-27, 09:30:26	-	-456	83.0000	83.0400	37,848.00	-3.21	-37,844.79	0.00	-18.24	O:P
VYM	2017-12-04, 09:30:10	-	5,015	85.681336	85.1900	-429,691.90	-25.08	414,942.20	-14,774.77	-2,464.05	C:P
Total VYM			0			-14,290.89	-80.69	0.00	-14,371.58	-831.80	
WM	2017-10-30, 09:30:07	-	3,340	82.5700599	81.8700	-275,784.00	-16.70	275,800.70	0.00	-2,338.20	O:P
WM	2017-11-06, 09:30:04	-	818	81.6893399	81.2400	-66,821.88	-4.09	66,825.97	0.00	-367.56	O:P
WM	2017-11-13, 09:30:08	-	-990	81.3876768	82.1400	80,573.80	-6.93	-81,676.95	-1,110.08	-744.80	C:P
WM	2017-11-20, 09:30:05	NYSE	2	80.9800	80.9200	-161.96	-1.00	162.96	0.00	-0.12	O
WM	2017-11-27, 09:30:22	BEX	-10	80.3900	80.8500	803.90	-1.02	-825.05	-22.17	-4.60	C
WM	2017-12-04, 09:30:17	-	-3,160	82.3493038	83.1500	260,223.80	-22.19	-260,287.63	-86.02	-2,530.20	C:P
Total WM			0			-1,166.34	-51.93	0.00	-1,218.27	-5,985.48	
Total						-11,233.10	-565.76	-11,279.18	-11,798.86	2,204.76	
Symbol	Date/Time	Exchange	Quantity	T. Price	C. Price	Proceeds	Comm/Fee	Basis	Realized P/L	MTM P/L	Code
Bonds											
USD											
T 2 10/31/22 0.0%	2017-10-30, 10:55:32	BARCBONDG	500,000	99.95703	100.015625	-499,785.15	-100.00	499,885.15	0.00	292.98	O
T 2 10/31/22 0.0%	2017-12-04, 08:00:22	BARCBONDG	-500,000	99.34766	99.367188	496,738.30	-100.75	-499,885.15	-3,247.60	-97.64	C
Total T 2 10/31/22			0			-3,046.85	-200.75	0.00	-3,247.60	195.34	
Total						-3,046.85	-200.75	0.00	-3,247.60	195.34	

Generated: 2017-12-10, 10:33:05 EST

References

- [1] Alan Levine. Harry markowitz, father of modern portfolio theory still diversified. <http://post.nyssa.org/nyssa-news/2011/12/harry-markowitz-father-of-modern-portfolio-theory-still-diversified.html>, Dec 2011.
- [2] Capital asset pricing model. <https://www.investopedia.com/terms/c/capm.asp>.
- [3] Fama and french three factor model. <https://www.investopedia.com/terms/f/famaandfrenchthreefactormodel.asp>.
- [4] Announcements, data & results. <https://www.treasurydirect.gov/instit/annceresult/annceresult.htm>.
- [5] Activity statements. https://gdcdyn.interactivebrokers.com/Universal/servlet/AccountAccess AuthenticateSSO?action=RM_VIEW_ACTIVITY&clt=1&mid=001.
- [6] Sharpe ratio. <https://www.investopedia.com/terms/s/sharperatio.asp>.
- [7] S&p 500 (). <https://finance.yahoo.com/quote/%5EGSPC?p=%GSPC>.
- [8] Leverage ratio. <https://www.investopedia.com/terms/l/leverageratio.asp>.
- [9] David Ruppert. *Statistics and finance: an introduction*. Springer, 2009.
- [10] Yahoo finance nke. <https://goo.gl/YBLJNb>.
- [11] W. James and Charles Stein. Estimation with quadratic loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 361–379. University of California Press, 1961.
- [12] Yahoo finance data. <https://goo.gl/yvjd4>.
- [13] Why tableau stock dropped. <http://fortune.com/2016/03/23/tableau-software-stock-drop/>.
- [14] Description of momentum factors for developed markets. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f Developed_mom.html.
- [15] Bloomberg terminal computer. <https://www.bloomberg.com/professional/solution/bloomberg-terminal/>.
- [16] numpy.linalg.norm. <https://docs.scipy.org/doc/numpy-1.13.0/reference/generated/numpy.linalg.norm.html>, journal=SciPy.

MA574 Final Report

- [17] Backtest portfolio asset allocation. <https://www.portfoliovisualizer.com/backtest-portfolio>, journal=.
- [18] Average annual return of s&p 500 index. <https://www.investopedia.com/ask/answers/042415/what-average-annual-return-sp-500.asp>, journal=.