Optimal Portfolio Allocation

Osman Khan, Tudor Muntianu, Joe Gyorda, Sri Yenamandra, Oliver Levy

Math 70 Spring 2020 - Professor Demidenko

Contents

Context - Why Diversify?

- Imagine buying shares from a **single company**, X, with $500, and let's say in one year, the stock has an equal chance of being worth $1000 or nothing.
    - Its variance, Var(X), would be 250 $dollars^2$, and its expected payout is $500.
- If instead we could buy **two stocks**, X and Y, each with $250, and each of which has an equal chance to be worth $500 or nothing in one year.
    - The variance of the portfolio after one year would be Var(X)+Var(Y)+2Cov(X,Y), which would equal at least zero, and at most 250 $dollars^2$, while the expected payout is $500.
    - If the two stocks were anything less than perfectly correlated, the **portfolio's variance** would be **less than** the original portfolio's while having the same expected payout.
- The basis of optimal portfolio allocation is this premise: there exists a line, given different companies' returns and correlations, that defines the **minimal risk** a portfolio can take for any given return.

Background: Optimal Portfolio Allocation

- Given a set of financial assets (e.g. stocks) and a fixed investment amount (e.g. $1000), **in what proportion should we allocate our wealth towards buying each stock/asset in order to maximize our risk-adjusted returns?**[1]
- Is buying every asset in **equal proportion,** and thereby spreading risk, the best approach?
- If not, how can we determine the **optimal proportions (weights)** with which we should buy each asset?
- We will examine real stock data to answer these questions.
  - We neglect transaction costs of making investments.
  - We use daily returns of stocks in our computations $(P_{Day\ i} - P_{Day\ i-1}) / P_{Day\ i-1}$, where P is the closing price of a given stock.

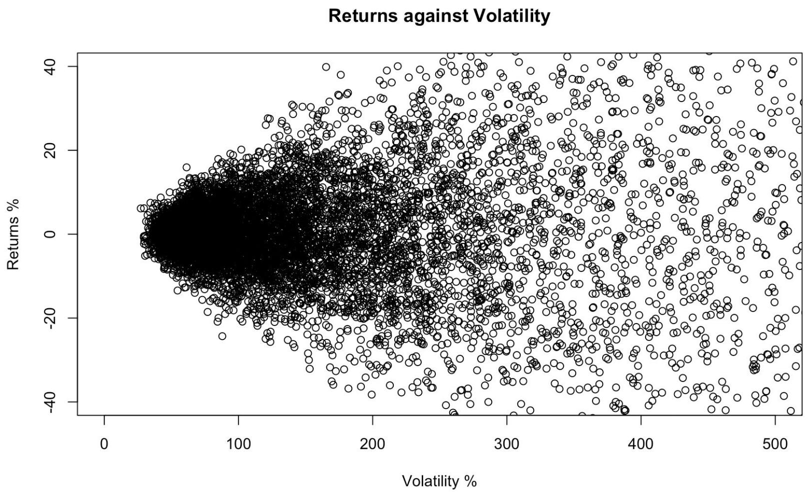
[1]Demidenko, E. (2020). Optimal Portfolio Allocation. In *Advanced statistics with applications in R* (pp. 230-236). Hoboken, NJ: Wiley.

Portfolio Allocation Models

1. Markowitz Model
2. Interval Portfolio Model
3. Hierarchical Clustering Model

Markowitz Model

- In 1952, Harry Markowitz developed a method to find the ideal minimum-volatility portfolio for any given return.
- The tradeoff between risk and return can be expressed as a curve called the “optimal frontier” where for any given level of risk, returns are maximized, and vice versa, with a specific weighting of assets.[1]
- The Markowitz Model assumes that investors are only allowed to go “long” or purchase assets.
  - In our version of the Markowitz model, we assumed that shares could also be shorted (negative returns are allowed).



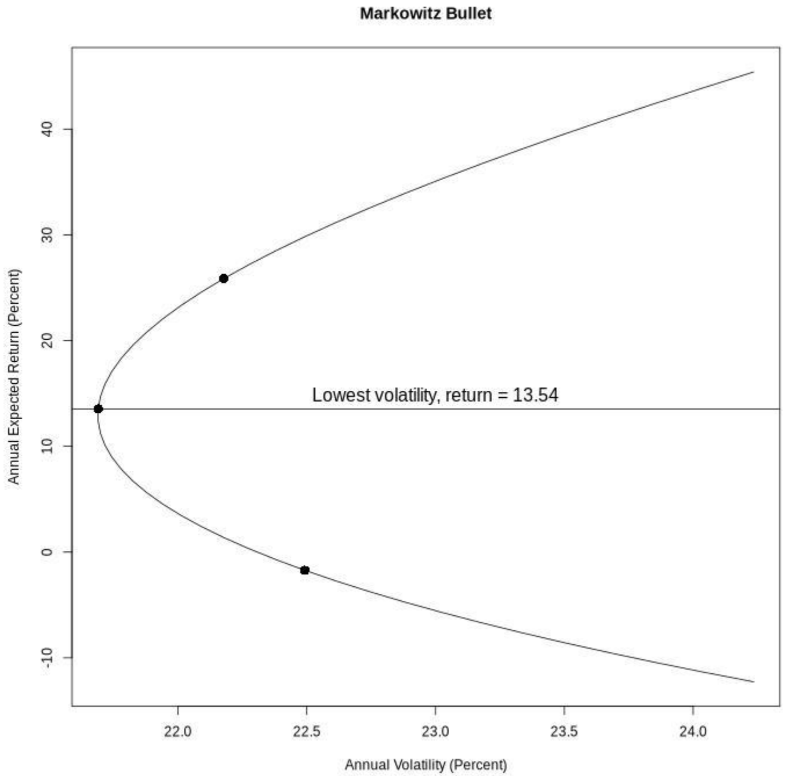
An example of a “Markowitz Bullet”
This data was taken from 10,000 simulations of a 30-stock portfolio over six months

[1]Demidenko, E. (2010, November 22). Portfolio risk management via interval probability. Retrieved June 02, 2020, from https://content.iospress.com/articles/model-assisted-statistics-and-applications/mas00168

Statistical parameters in the context of stocks

- The mean vector represents the average returns of the stocks in the data set.
- Variance represents the volatility of stocks: the relationship is that variance = volatility^2
- The vector of stock returns can be constructed as a vector of random variables X_1, X_2, … such that each X_i is a normally-distributed random variable with the mean as the mean return of the stock and standard deviation as the volatility of the stock
- We want to define a weighting of stocks such that for any return, we have the minimum volatility portfolio
- Given a vector of weights that sum to 1, *w*, how do we calculate the portfolio's return and volatility?
  - $w^T\mu$ = mean returns; we can set a desired return r such that $w^T\mu = r$
  - $w^T\Omega w$ = variance = volatility^2

Markowitz Model Output



(Left) The "Markowitz Bullet" optimal frontier for our stock data (see data collection). The x-axis of the graph is the annual volatility of a portfolio, and the y-axis of the graph is the annual expected return of the portfolio. The point on the graph with the lowest volatility, denoted with a black dot and intersecting line, denotes the return (13.54%) that is obtainable with the lowest associated risk. This is the optimal portfolio as determined by the Markowitz model.

Portfolio Allocation Models

1. Markowitz Model
2. Interval Portfolio Model
3. Hierarchical Clustering Model

Interval Model

- The interval model is similar to the Markowitz Model, except instead of aiming to find the optimal frontier, we aim to maximize the probability of achieving a range of returns.
- We specify a lower bound, r, and an upper bound, R, of returns and calculate, using the Markowitz model, the relative probability of either occurring. Then, we can use optimization techniques to find the weighting required to achieve said returns.[1]

[1]Demidenko, E. (2010, November 22). Portfolio risk management via interval probability. Retrieved June 02, 2020, from https://content.iospress.com/articles/model-assisted-statistics-and-applications/mas00168

Interval Model Continued

The return of optimum-weighted portfolio will return $\mathbf{w}^T\mathbf{X}$, where $\mathbf{X}$ is the vector of normally distributed random variables representing returns, and $\mathbf{w}$ represents the vector of weights. To get the probability that the returns fall in the interval, we can find $P\ (r < \mathbf{w}^T\mathbf{X} < R)$. Given that we assume the return random variables are normally distributed, we know that the return of the portfolio must also be normally distributed.

Therefore, by standardizing the return, we get the following:

$$P\ (r < \mathbf{w}^T\mathbf{X} < R) = P(R < \mathbf{w}^T\mathbf{X}) - P(r < \mathbf{w}^T\mathbf{X}) = \Phi((R - \mathbf{w}^T\boldsymbol{\mu})/sd) - \Phi((r - \mathbf{w}^T\boldsymbol{\mu})/sd)$$

Where sd is the standard deviation of $\mathbf{w}^T\mathbf{X}$, which is equal to sqrt($\mathbf{w}^T\boldsymbol{\Omega}\,\mathbf{w}$)

Portfolio Allocation Models

1. Markowitz Model
2. Interval Portfolio Model
3. Hierarchical Clustering Model

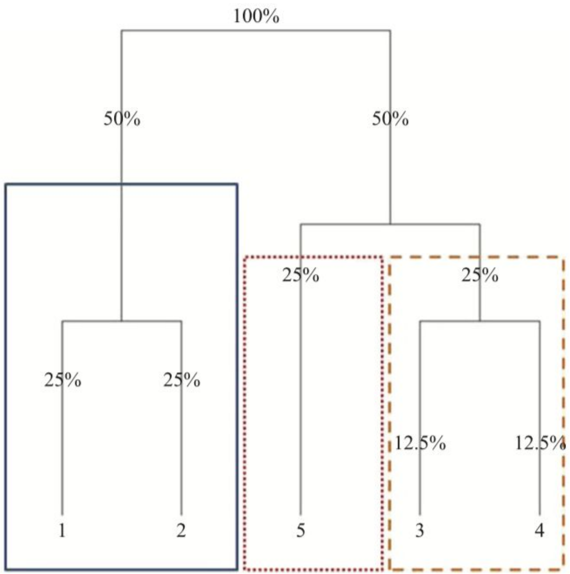
Hierarchical Clustering Model

Intuition:

- Some stocks may be close substitutes of one another, i.e. they may have similar trends in returns over time.
- We expect companies in the same industry to have correlated performance, and we want to diversify not just across companies, but across industries.[2]

Implementation:

- Ultrametric distance:[3] $\sqrt{2(1-\rho)}$
- Average linkage for hierarchical clustering
- Output is a vector of weights for our model
- This model aims to broaden exposure to all industries by allocating weights equally within clusters.

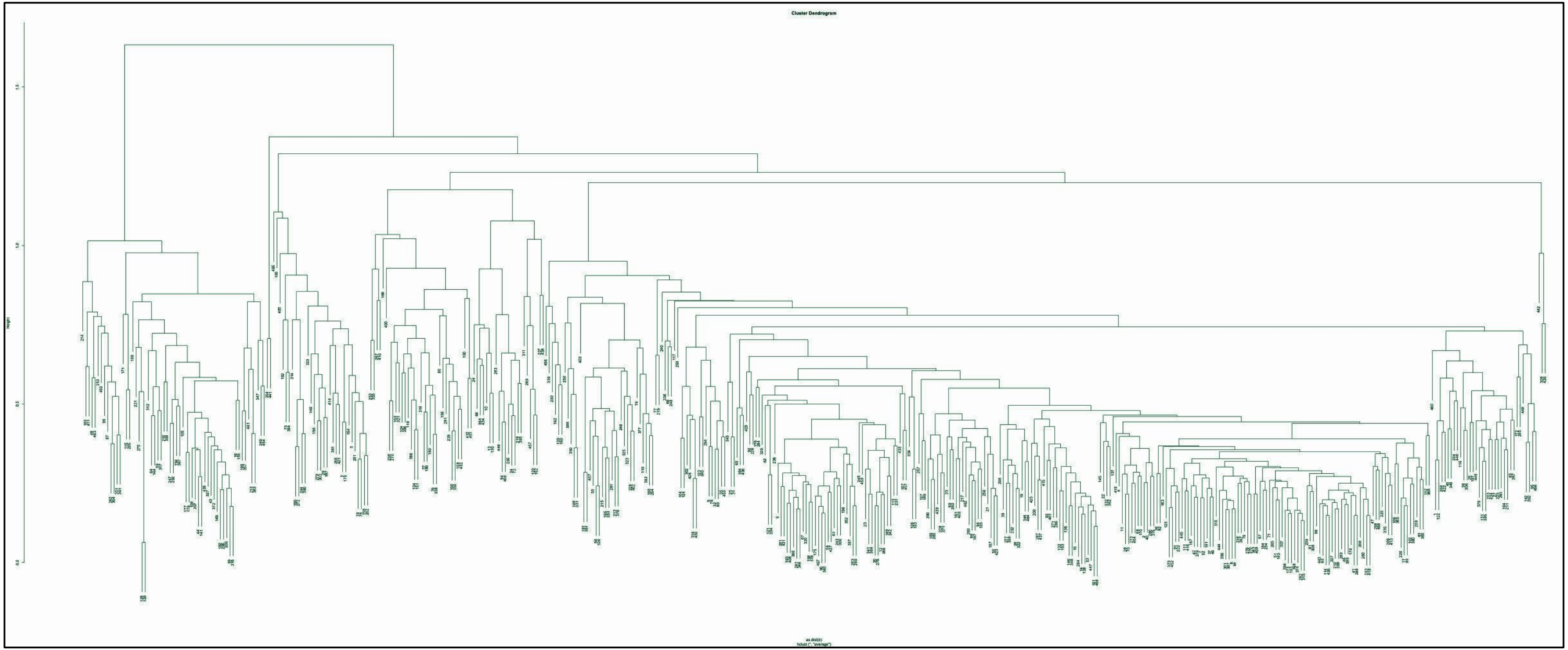


Example of equal-weights allocation within clusters[2]

[2]Raffinot, T. (2017). Hierarchical Clustering-Based Asset Allocation. The Journal of Portfolio Management, 44 (2), 89-99. doi:10.3905/jpm.2018.44.2.089.

[3]Mantegna, R. N., Bonanno, G., & Lillo, F. (n.d.). Hierarchical structure of correlations in a set of stock prices. Retrieved May 29, 2020, from http://ocs.unipa.it/pdf/santafe.pdf

Hierarchical Clustering Output



(Above) The dendrogram produced through hierarchical clustering of our data. Each branch represents one stock; we use the branching here to calculate stock weights as described on the previous slide.

Data collection

- We retrieved five years of daily S&P 500 stock prices from 2013 to 2018.
  - Taken from Yahoo! Finance
- We then dropped all NAs, leaving us with **468 stocks**.
- Then, we calculated daily returns for each stock.

	Name	2013-02-08	2013-02-11	2013-02-12	2013-02-13	2013-02-14	2013-02-15	2013-02-19	2013-02-20	2013-02-21	2013-02-22
1	AAL	15.0700	14.8900	14.4500	14.3000	14.9400	13.9300	14.3300	14.1700	13.6200	13.5700
2	AAPL	67.7142	68.0714	68.5014	66.7442	66.3599	66.9785	65.8714	65.3842	63.7142	64.1785
3	AAP	78.3400	78.6500	78.3900	78.9000	78.6600	78.8300	79.1200	80.4200	79.2600	79.2000
4	ABBV	36.3700	36.1300	35.7700	35.4200	35.0500	36.9300	37.5400	38.1600	38.4900	38.7900
5	ABC	46.5200	46.8500	46.7000	46.7400	46.6700	46.7700	46.6700	47.2200	46.4800	46.4800
6	ABT	34.3900	34.4200	34.2700	34.2900	34.2800	34.8400	35.1800	34.7000	34.4700	34.1800
7	ACN	73.0100	73.0900	72.8900	73.3200	73.2100	73.0800	74.2800	75.4000	74.6900	74.1500
8	ADBE	38.3100	38.9900	38.5500	38.9000	38.7000	38.5100	38.5800	39.0400	38.8200	38.5400
9	ADI	44.7200	45.9900	46.1500	46.2900	46.0200	46.3900	46.4000	46.9000	45.7000	45.1905
10	ADM	30.3100	30.2600	30.2200	30.9100	31.0300	31.7200	32.5800	33.0900	32.3200	32.4000
11	ADP	60.7100	60.7900	60.2400	60.5600	60.2500	61.2600	61.2200	61.4300	61.1600	60.7100
12	ADSK	38.7900	39.0700	38.9100	38.9100	38.5400	38.6500	39.1000	39.1700	38.5800	38.1000
13	ADS	153.8800	154.0600	153.4900	152.3800	152.2700	152.7500	155.6400	154.9400	152.2700	152.2700
14	AEE	32.7900	32.6500	32.8300	33.2000	33.2000	33.1800	33.1200	33.2300	33.3800	33.2900
15	AEP	44.5500	44.5700	44.7300	44.8700	44.7600	44.8400	45.2000	45.6300	45.4300	45.5600
16	AES	11.1200	11.0600	11.2300	11.3100	11.3100	11.2200	11.1700	11.3600	11.3500	11.1200
17	AET	49.8300	50.5200	50.9600	50.2600	49.3300	49.5700	47.5700	48.7500	48.2600	47.8700
18	AFL	50.3800	50.3600	50.1100	49.4300	48.4300	48.9400	49.2600	50.0300	49.8300	49.9300
19	AGN	87.2600	87.1300	86.7300	86.0400	86.2300	86.2800	87.4000	84.5200	82.7000	83.7200
20	AIG	38.7600	38.8900	39.5000	38.9300	38.6400	39.3100	38.3600	38.6000	37.3600	39.2900

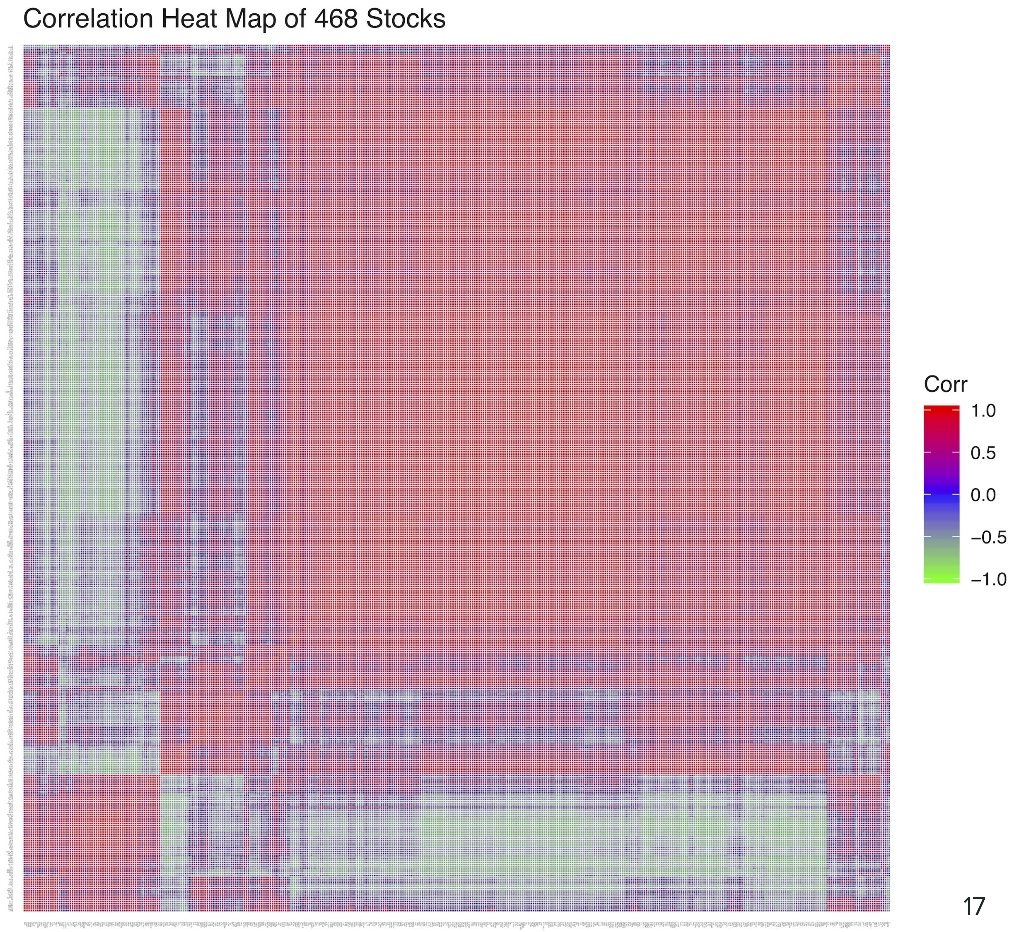
(Above) First 20 rows and 11 columns of our stock dataframe in R - actual dimensions are 468 x 1260

Correlation Heat Maps

- We created correlation heat maps of the stock data to determine if there were any interesting relationships between certain stocks.
- We used the ggplot2 package to help create the heat maps
- We apply hierarchical clustering to group the stocks on the heat map to get a better visualization of correlation within and across clusters.
- We also create a partial correlation heat map to remove the influence of other stocks when computing correlations.

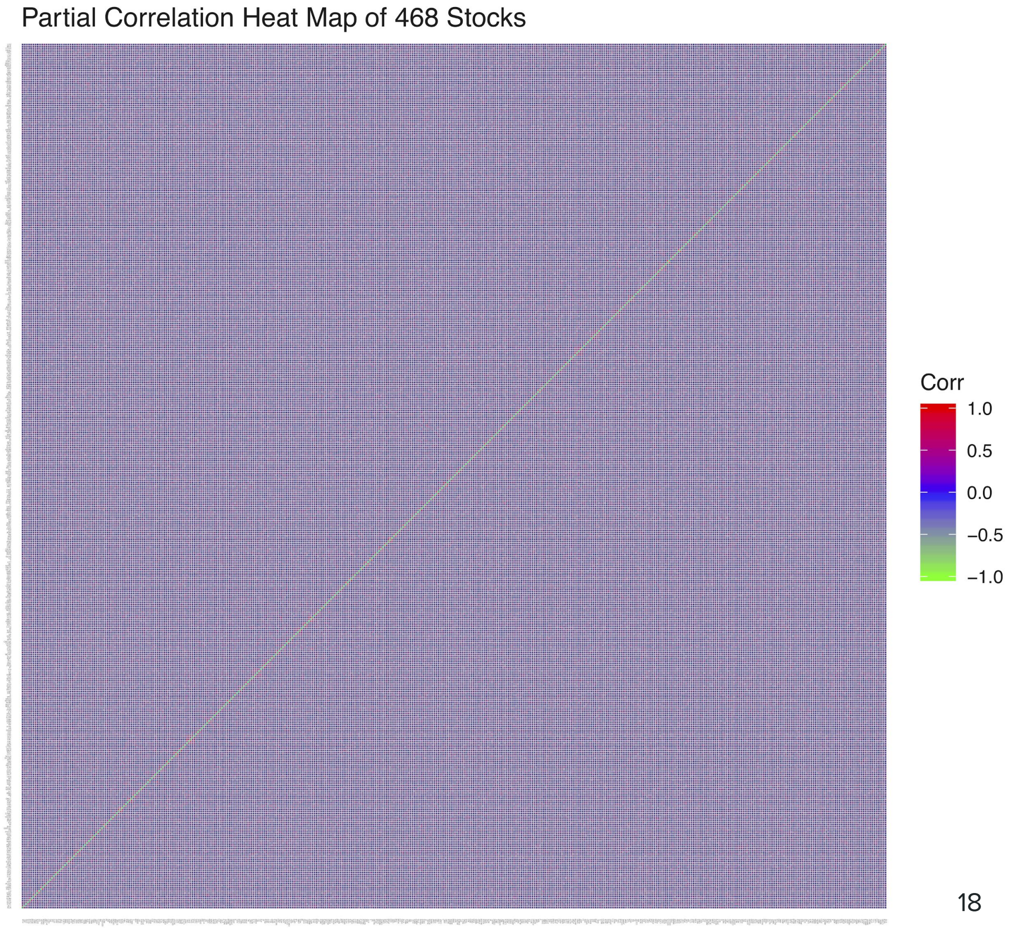
Correlation Heat Map

- Most stocks exhibit a strong positive correlation (red), while many others exhibit a strong negative correlation (green).
- Few stocks exhibit little to no correlation (blue).



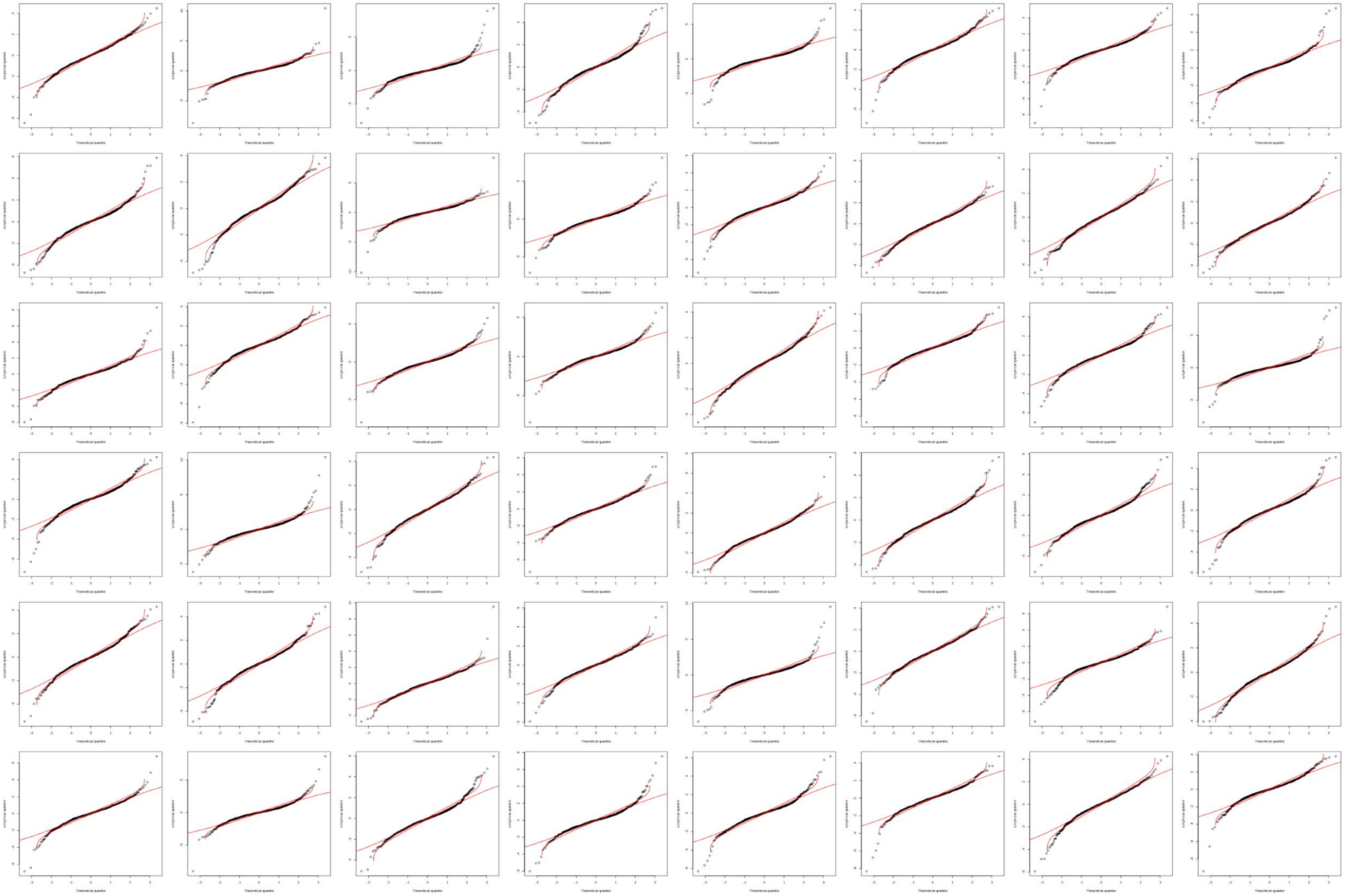
Partial Correlation Heat Map

- After removing the influence of other stocks, nearly all stocks exhibit little to no correlation (blue).
- Note that this is only representative of individual stocks, not industries.



QQ-Plots

- To examine the distribution of the stocks in our data set, we use QQ-plots under the assumption that the returns of each stock follow a normal distribution
- We examine the daily returns within stocks.
- A random sample of 48 stocks is selected from our data set.
- Some of the stocks appear slightly skewed, but overall we can make the normal assumption for our models.



(Above) QQ Plots of Daily Returns Within 48 Stocks

Training Markowitz Bullet and Interval Portfolio Models

- Essentially, this required finding $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$, as they were the parameters that are used to construct the values in **w**.
- Vector $\boldsymbol{\mu}$ was estimated by averaging daily returns over six month training period
- Matrix $\boldsymbol{\Omega}$ was estimated by finding the covariance matrix of the returns over the same period
- We then used the $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ matrices as the mathematical basis of the models.

Testing Each Model

- Rolling test windows of **6 months**, **2 years**, and **3 years**.
    1. Train models on **rolling 6 month** period, test on **rolling 6 mo. windows** (slides 28/29).
    2. Train models on **rolling 2 year** period, test on **rolling 2 year windows** (slides 30/31).
    3. Train models on **rolling 1 year** period, test on **rolling 3 year windows** (slides 32/33).

    ➔ For each window, we test the models **once without rebalancing** and **once with rebalancing** at the halfway point.
- Train on **6 months**, test on following **3.5 years** - rebalance after various equally-spaced intervals (0-10 times) (slide 34).
- We use a starting investment of $100,000 in each case.

Testing Each Model Continued

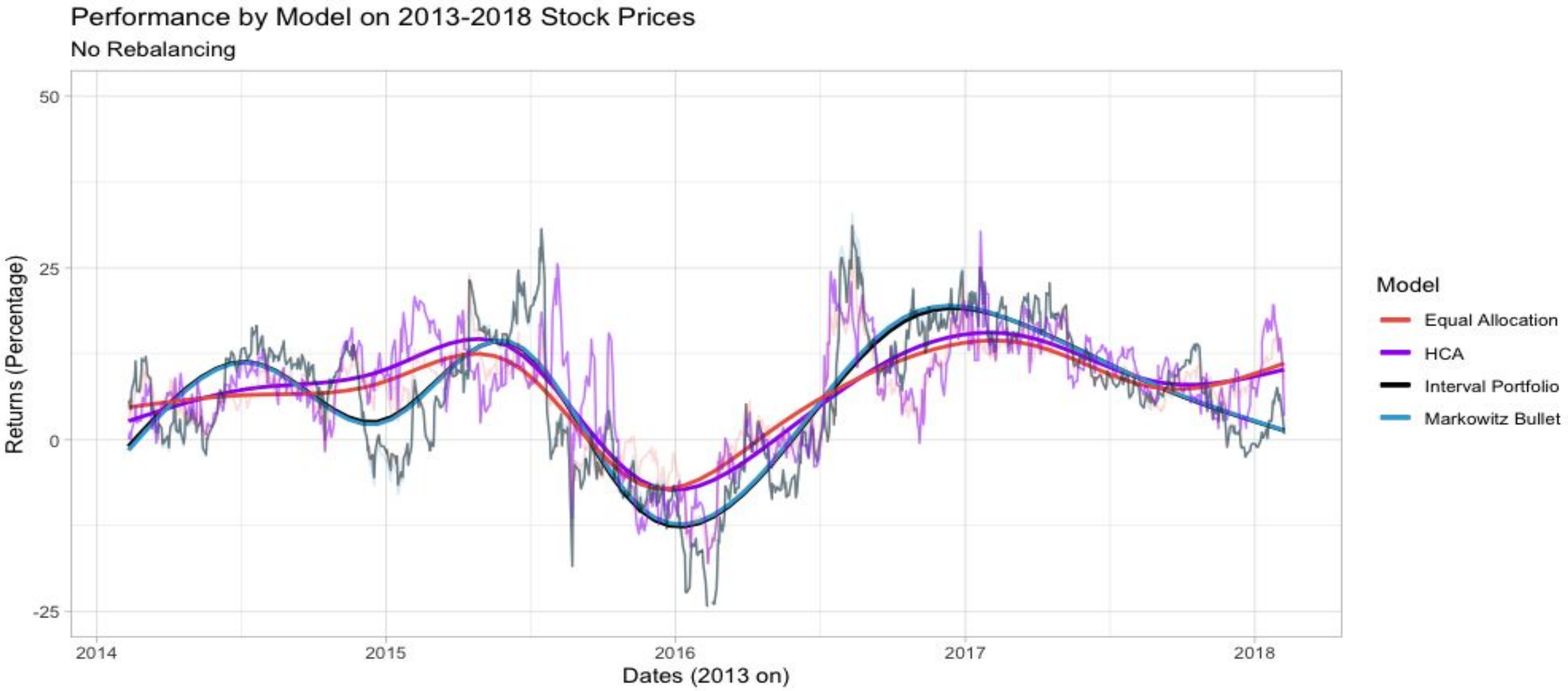
- To elaborate, the **“rolling” window** means simply that we train on a certain number of months from the start of the data, then test on the following months to see how the weights we learned from the first six perform monetarily at the end of the six months. Then we “roll” by moving our training and testing period forward by one day.
- Once we have rolled through the whole data set, we have several data points, or “returns” that are the sum of the returns from each stock from **the end of each testing period**.
- We do this for each model, and then analyze the returns over time.
- We select **30 stocks at random** from our data set to construct our portfolios.

Assessing Each Model

- We compare each of the three previous models to the most simple weight allocation model: the **equal weights model.**
  - Each stock is allocated an equal weight of 1/N, where N = the number of stocks in our data set.
- We examine the **standard deviations** of the returns at the ends of each testing period for each model.
- We conduct a **difference-in-means paired t-test** to compare the returns at the ends of each testing period of each of our three models to the baseline equal weights model.

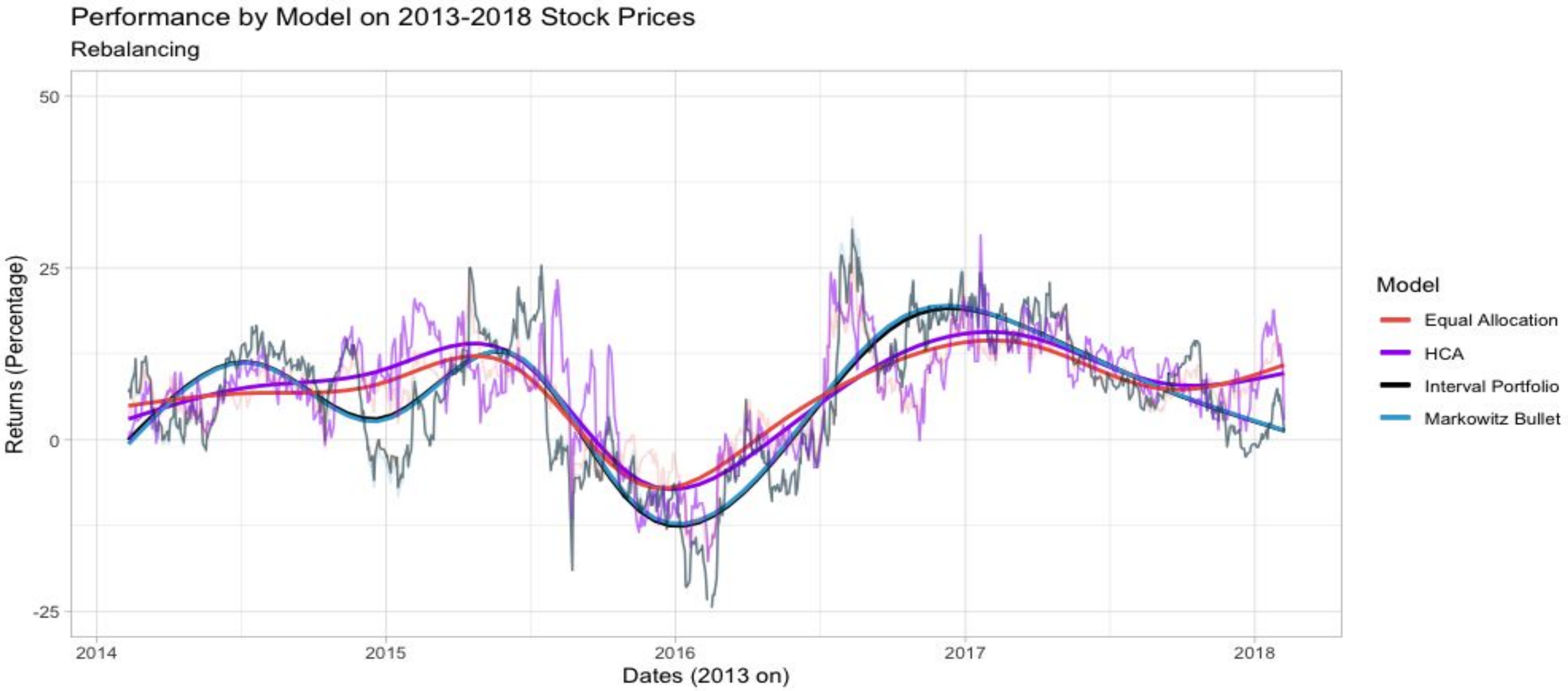
6-Month Rolling Window (No Rebalancing):

Smoothing Function Applied For Legibility, Actual Returns in Background



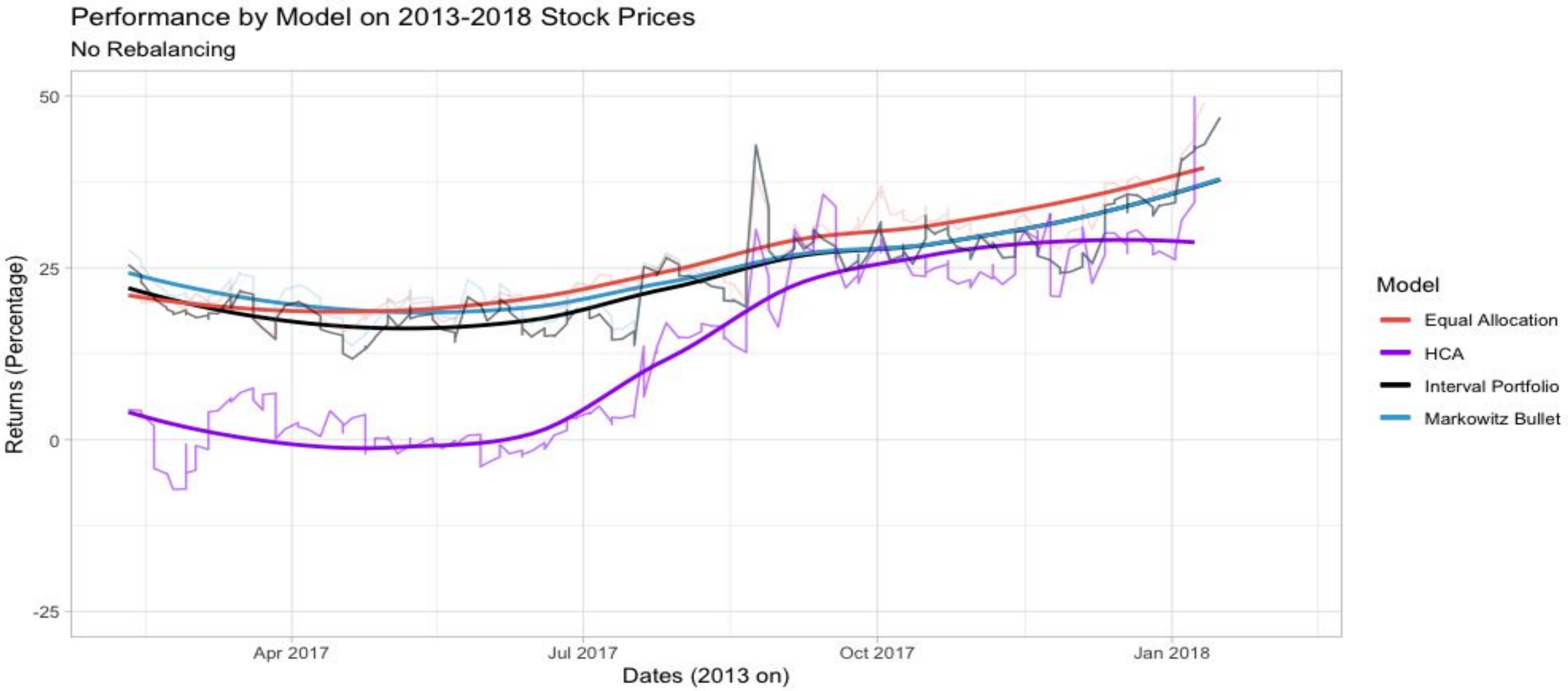
6-Month Rolling Window (Rebalance Halfway):

Smoothing Function Applied For Legibility, Actual Returns in Background



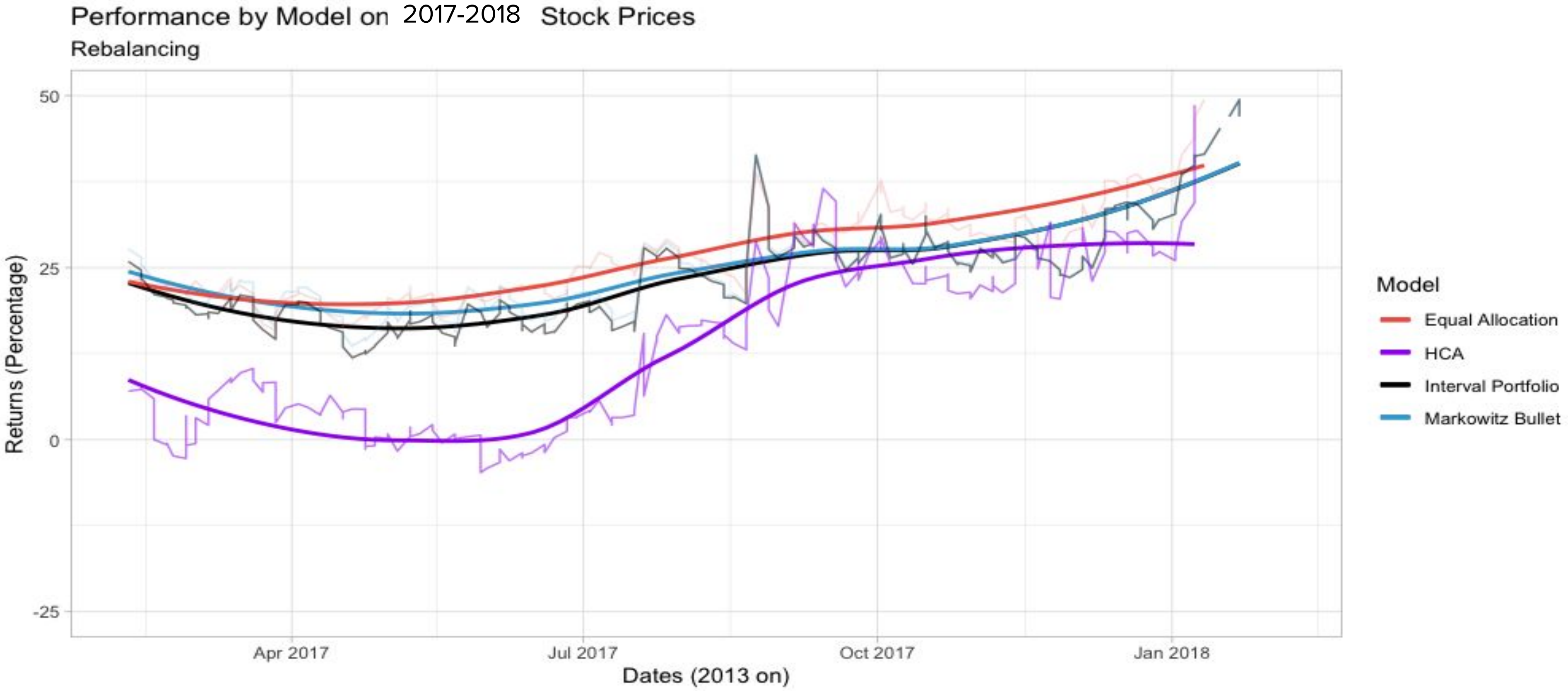
2-Year Rolling Window (No Rebalancing):

Smoothing Function Applied For Legibility, Actual Returns in Background



2-Year Rolling Window (Rebalance Halfway):

Smoothing Function Applied For Legibility, Actual Returns in Background



3-Year Rolling Window (No Rebalancing):

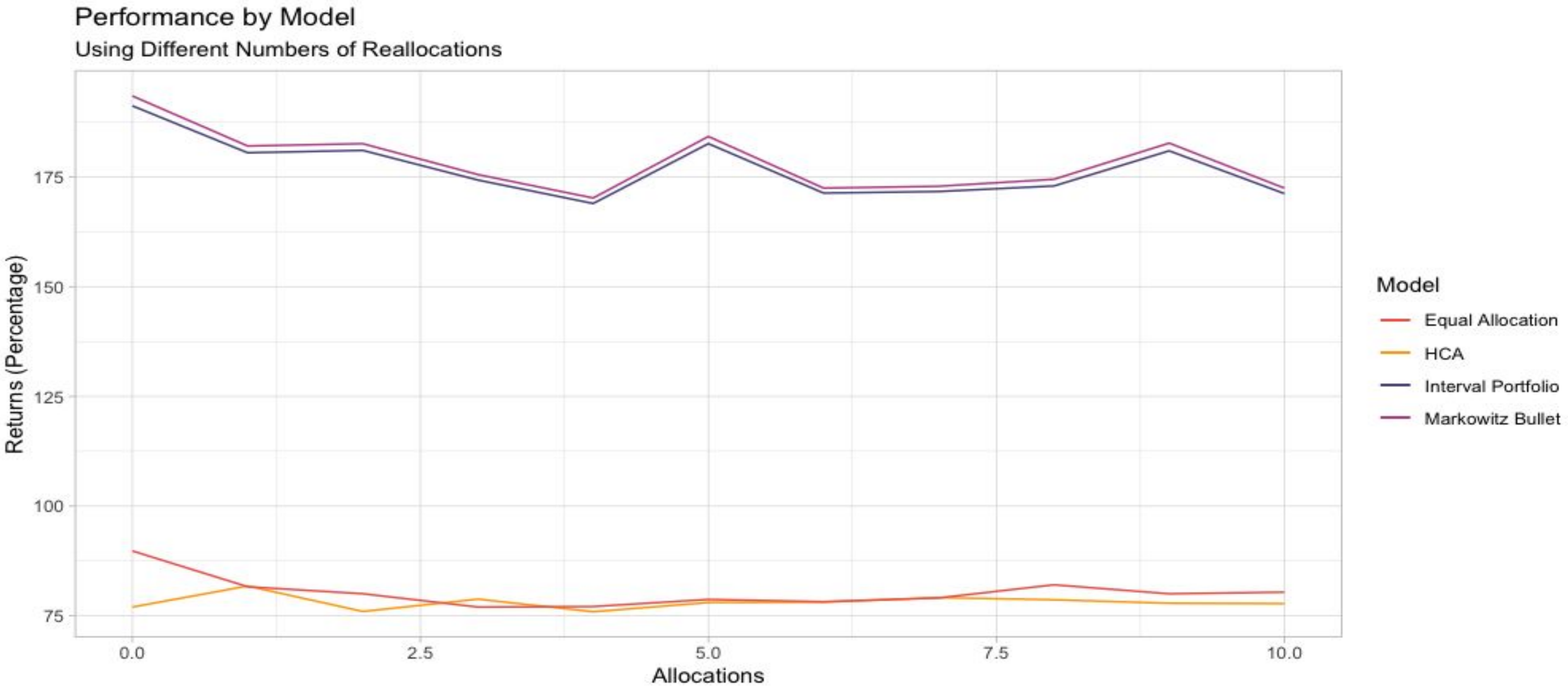
Smoothing Function Applied For Legibility, Actual Returns in Background

3-Year Rolling Window (Rebalance Halfway):

Smoothing Function Applied For Legibility, Actual Returns in Background

Static Window, Entire Dataset (Rebalancing):

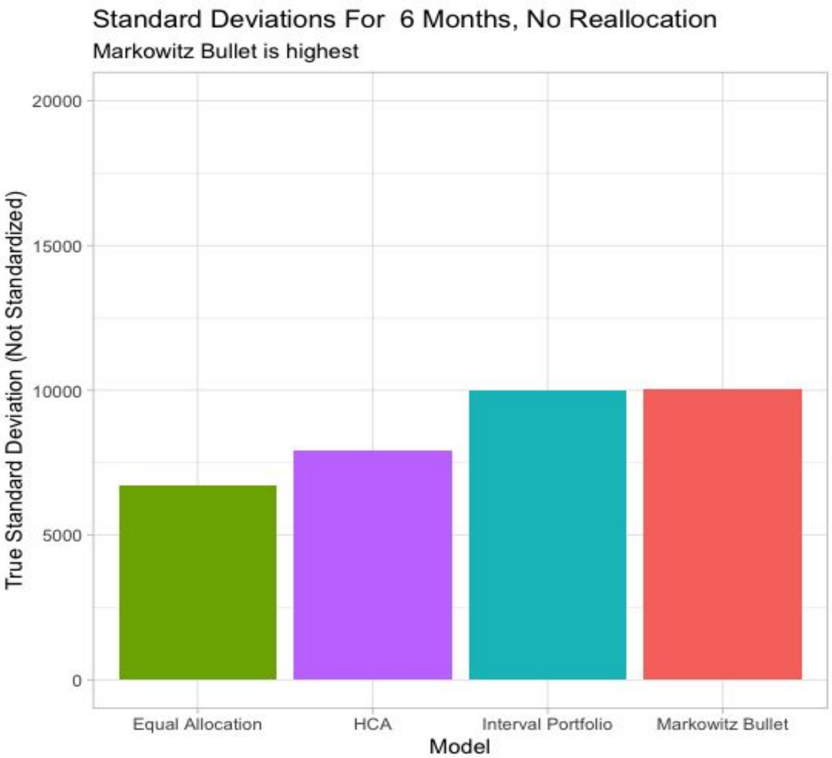
Reallocation Effect on Total Returns

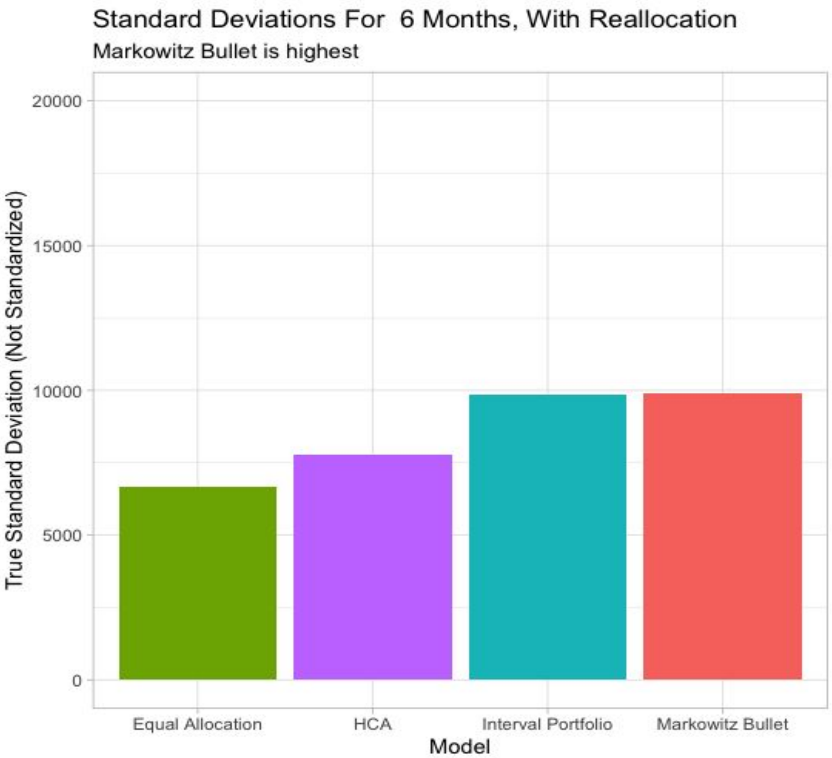


Standard Deviation of Models

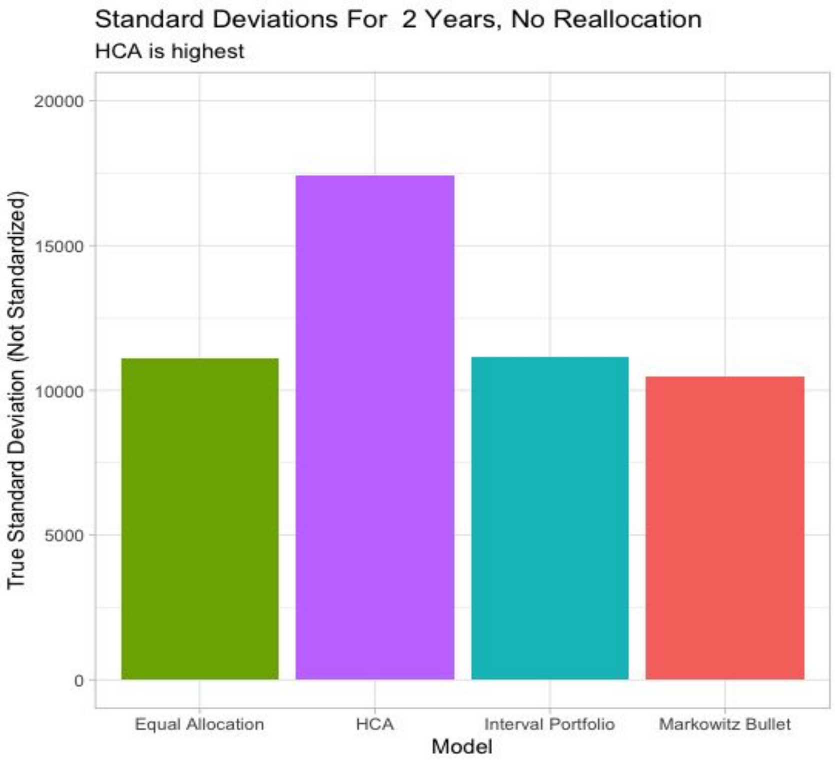
	6 Months No Rebal.	6 Months Rebal.	2 Years No Rebal.	2 Years Rebal.	3 Years No Rebal.	3 Years Rebal.
HCA	7910.5	7775.5	17407.9	16228.9	16732.8	15440.3
Markowitz Bullet	10052.3	9879.5	10483.5	9424.4	9778.9	10959.6
Interval Portfolio	10014.3	9838.1	11135.6	10031.6	9987.1	11253.4
Equal Weights	6740.3	6678.0	11116.6	10729.6	4526.5	4510.6

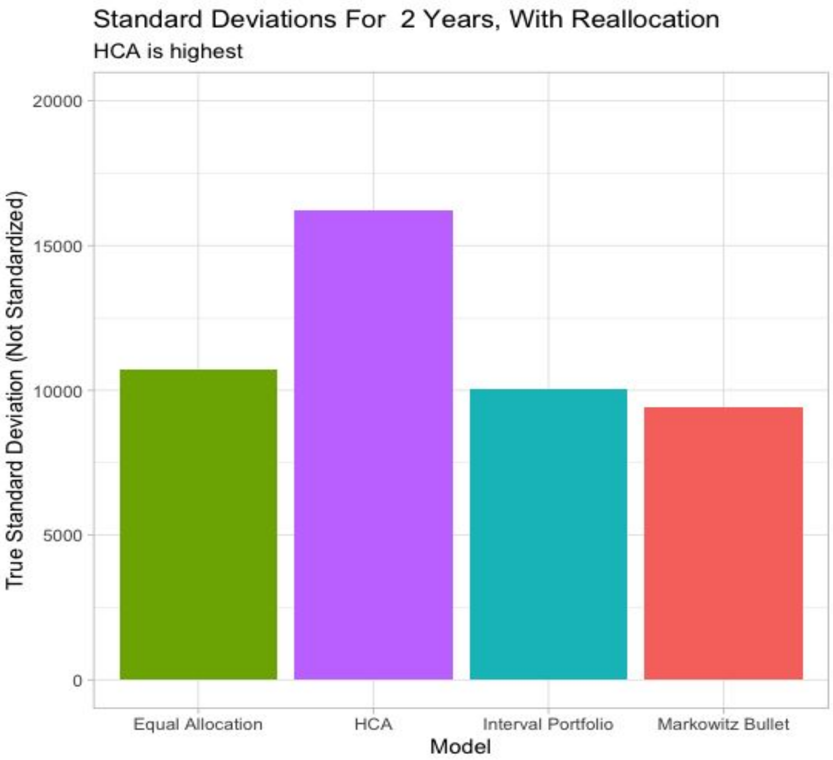
Standard Deviations Across Tests: 6 Months Test



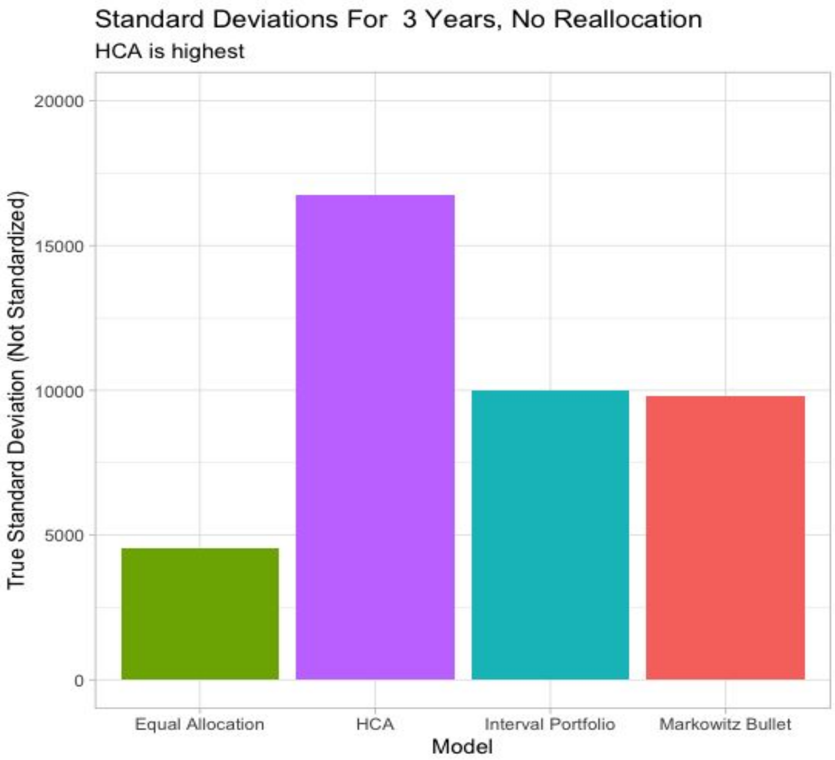


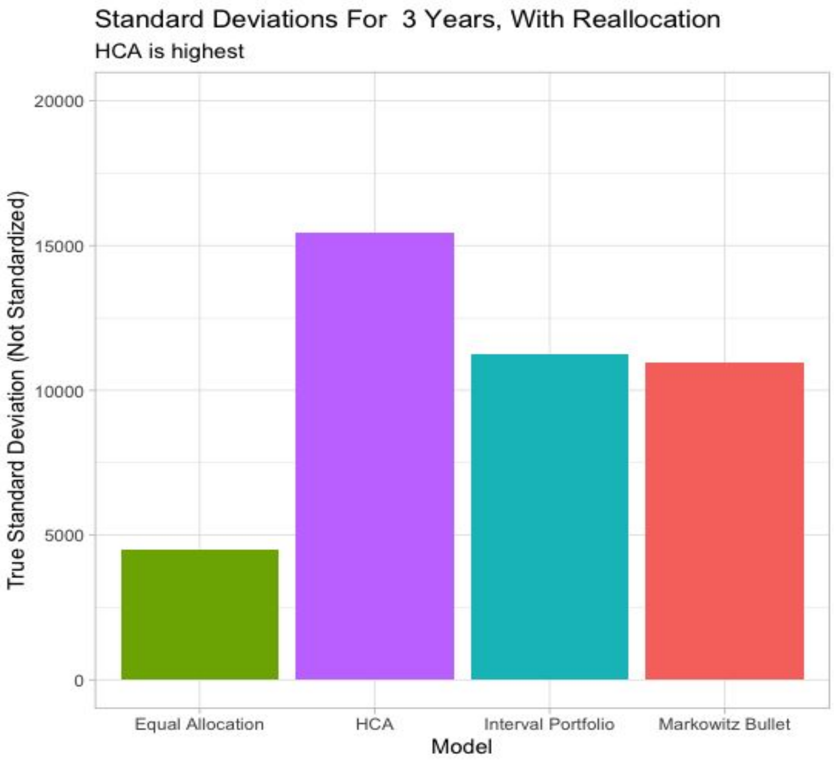
Standard Deviations Across Tests: 2 Years Test





Standard Deviations Across Tests: 3 Years Test





Significance Analysis

- Using a difference-of-means analysis via a paired t-test, we saw that most of the returns were significantly different from the baseline equal allocation portfolio. The chart below displays the t-values for the t-test.
- Due to the large data sets, we can approximate a distribution using a normal distribution, which, for an alpha of .01, leads to a two-sided critical z-score of 2.575. The significantly different portfolios are highlighted below in yellow.

	3 Yr No Reallocation	3 Yr Reallocation	2 Yr No Reallocation	2 Yr Reallocation	6 Mo No Reallocation	6 Mo Reallocation
HCA	-11.7	-10.91	-23.22	-25.43	4.78	4.6
Markowitz	22.592	15.01	-1.52	-14.054	-2.19	-3.03
Interval	19.41	12.53	-3.58	-21.79	-2.72	-3.58

Results Summary

- There is a small but significant difference between models’ performances in 6 month rolling window.
- HCA performance drops at 2 year rolling window, but picks up greatly at 3 year rolling window.
    - Possibly HCA invests more in one industry that experiences an uptick in stock values towards the end of the 3-year window.
- Significant difference between HCA and Equal Allocation vs Interval Portfolio and Markowitz Bullet after testing for 3 years.
- Rebalancing decreases the variance of the models slightly, which is what one would expect.

Discussion

- In an investment scenario, it is helpful to rebalance your portfolio to reduce risk. However, there appears to be little benefit to rebalancing more often than once to several times a year (see graph on slide 32).
- The purpose of the Markowitz and Interval models is not necessarily to maximize returns, but to minimize risk. This was achieved in the two-years testing period, but not in the other periods, suggesting that the predictive power of historical data was not particularly high in this time period.
- From our two and three year runs, we can see that Markowitz and Interval Portfolio both produce much lower volatility overall than Hierarchical Cluster Analysis, which makes sense since HCA does not seek to minimize volatility.

Discussion Continued

- The HCA, which ascribes equal allocation to the clusters, exhibited more long-term variance, as it should since it does not, by itself, try to minimize covariance, but rather broadens exposure to different industries.
- Lastly, as a brief validation, the Markowitz model tended to have a smaller variance than the interval model, suggesting that the indeed, the relationship between returns and volatility held true in most cases.

Limitations of our Analyses

- We neglect transaction costs of purchasing assets
  - There is often a cost associated with re-allocating portfolios, so investing in a re-allocated portfolio could decrease returns.
- While the Markowitz and interval models are very strong at optimizing risk-adjusted returns under specific conditions, they assume **stock price changes are normally distributed** and that each day's returns is independent of each other day's returns.
  - In the long run, the latter assumption may be true, but the first assumption fails to take into account significant **short-term outlier events** such as Trump's election or COVID-19, which can skew the results of the 6-month tests.
  - The Markowitz model assumes that past returns are indicative of future returns (and thus volatility). If we train the model on a particularly turbulent/calm period in the markets, it misallocates the stocks in the future.
  - The QQ-plots indicate the long-run normality of the data.

Limitations Continued

- In general, the hierarchical clustering model performed significantly worse than the other models. However, we do not calculate the **risk-adjusted returns**, which according to Raffinot (2017) should illustrate that the HCA model performs better than illustrated in our analyses. It should be noted that the HCA model may sacrifice a higher volatility for higher returns, while the Markowitz and interval models seek to minimize volatility.
- To construct our portfolios, we **randomly selected 30 stocks** from our larger data set, as it is impractical to have a portfolio with 400+ stocks. However, it is quite possible our randomly selected portfolio had outlier stocks that did not perform similarly relative to the rest of the market (e.g. Netflix, Amazon, GE). Had we tested on a **different random sample** containing a different selection of stocks, our results may have been different.
    - If we had more time, it would have been interesting to **compare our results** for multiple random samples.

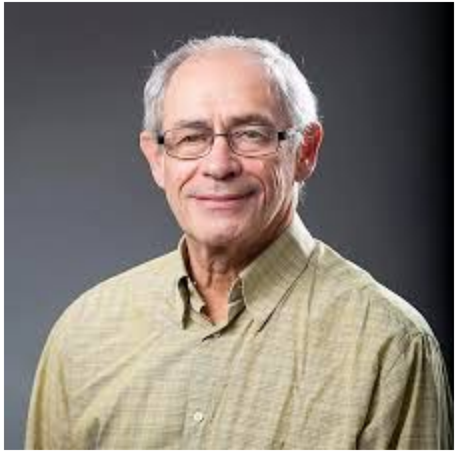
Concluding remarks

1. Rebalancing the portfolio reduces variance of all models.
2. Markowitz Bullet and Interval Portfolio perform better over long periods of time.
   - Little difference between models over a short period of time.
   - This makes sense, as the goal is to have a stable, long-term portfolio.
3. Limitations in dataset analysis and our models could impact results.

Thank you Professor Demidenko and M70!

“Mathematics is the queen and statistics is the king of all sciences”

- Professor Demidenko



Questions?