Finding coronavirus-resistant stock portfolios

Winfield Chen, Sebastian Bugal, Ishaan Bedi

April 23, 2020

1 Introduction

There are more than two million confirmed cases of COVID-19 worldwide. In addition to the global health crisis, COVID-19 is causing unprecedented economic destruction as countries take measures to prevent new cases. Over 20 million people are unemployed in the United States alone. The International Monetary Fund forecast for 2020 predicts a 3% decrease in global Gross Domestic Product (GDP), the amount of goods and services the world makes. Goldman Sachs projects a United States GDP drop of 29% in the first half of 2020. The S&P 500 Index, which tracks the 500 most valuable companies in the United States, lost 34% of its value from the all-time high of February 19 to the recent bottom of March 23. Stock markets around the world lost over 16 trillion US dollars.

2 Problem

Stock market losses are not evenly distributed. Some companies are impacted more than others.

Company	Period		Change
	Feb 19	Mar 19	
Amazon	2170.22	1880.93	-13%
Carnival	43.34	10.00	-77%

When demand for a company's products drops, so does the price of its stock. For example, COVID-19 decreased the demand for cruises more than the demand for online retail. This resulted in a -13% change in Amazon's stock price compared to -77% for Carnival Cruises. Clearly, investors should prefer companies which do well over companies which do poorly.

2.1 Portfolio Theory

The collection of stocks which an investor holds is known as a portfolio. The science of deciding which stocks an investor should purchase is known as portfolio theory. Portfolio theorists have found two opposing metrics of portfolio performance

Metric	Definition	Objective
Return	Average percent change in value over some time period	Maximize
Risk	Varies according to the approach	Minimize

known as the risk-return tradeoff. When an investor seeks a greater return, they must purchase riskier stocks. Portfolio theory seeks to optimize one metric while constraining the other to levels acceptable for the investor. We use two linear programming (LP) approaches for each choice of objective: finding minimum risk portfolios given a minimum return, and finding maximum return portfolios subject to constraints which reduce risk.

3 Minimizing Risk

One definition of risk used by portfolio theorists is mean absolute deviation (MAD) from expected return. [3] For a sequence of N times $t_1, ..., t_N$, the return at time t_n for stock i is the percent change in price from time t_{n-1} to time t_n.

$$R_i(t_n) = \frac{P_i(t_n) - P_i(t_{n-1})}{P_i(t_{n-1})}$$

The expected return is then estimated using the mean return over $t_1, ..., t_N$. The error in this estimate decreases as N grows.

$$\overline{R}_i = \frac{1}{N} \sum_{n=1}^{N} R_i(t_n)$$

The deviation from expected return at time t_n for stock i is the difference between the return and the expected return.

$$R_i(t_n) - \overline{R}_i$$

For a portfolio of a set of stocks S with each stock $i \in S$ having a weight w_i in the portfolio, the absolute deviation at time t_n is the absolute value of the weighted sum of deviations for each stock.

$$\left| \sum_{i \in S} w_i (R_i(t_n) - \overline{R}_i) \right|$$

The MAD over $t_1, ..., t_N$ is then:

$$\frac{1}{N} \sum_{n=1}^{N} \left| \sum_{i \in S} w_i (R_i(t_n) - \overline{R}_i) \right| .$$

3.1 LP Formulation

Our objective is to minimize the MAD over $t_1, ..., t_N$ by changing the weights of each stock in our portfolio.

$$\min_{w_i} \frac{1}{N} \sum_{n=1}^{N} \left| \sum_{i \in S} w_i (R_i(t_n) - \overline{R}_i) \right|$$

This is not a linear objective due to the absolute value function.

$$|x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}$$

We linearize the absolute value function by encoding it as the sum of two non-negative variables whose difference is the operand.

$$\begin{aligned} |x| &= y + z \\ y - z &= x \\ y, z &\geq 0 \end{aligned}$$

The pair of variables correspond to each case in the definition of the absolute value function when minimized. In the first case, z is minimized to 0, so $x = y$ and $|x| = x = y = y + z$. In the second case, y is minimized to 0, so $-x = z$ and $|x| = -x = z = y + z$. In both cases, $|x| = y + z$ as desired.

$$|x| = \begin{cases} x = y - z = y & x \geq 0 \\ -x = z - y = z & x < 0 \end{cases}$$

The linearized objective substitutes the absolute value with this two-variable encoding. The substitution introduces N equality constraints which enforce the definitions of y and z for each absolute deviation at times $t_1, ..., t_N$. We add the corresponding non-negativity constraints.

$$\min_{w_i, y_n, z_n} \frac{1}{N} \sum_{n=1}^{N} (y_n + z_n)$$

$$\text{subject to}$$

$$\left[y_n - z_n = \sum_{i \in S} w_i (R_i(t_n) - \overline{R}_i) \right]_{n=1}^{N}$$

$$y, z \geq 0$$

We add an equality constraint since the total weight of the portfolio is one. We add non-negativity constraints on the weights since it is not possible to hold negative stock. We add an inequality constraint since the return of the portfolio must be at least the minimum required return specified by the investor, R_p. This total return of the portfolio is the weighted sum of the expected return for each stock. The final LP formulation is then:

$$\min_{w_i, y_n, z_n} \frac{1}{N} \sum_{n=1}^{N} (y_n + z_n)$$

$$\text{subject to}$$

$$\left[y_n - z_n = \sum_{i \in S} w_i (R_i(t_n) - \overline{R}_i) \right]_{n=1}^{N}$$

$$\sum_{i \in S} w_i = 1$$

$$\sum_{i \in S} w_i \overline{R}_i \geq R_p$$

$$w, y, z \geq 0.$$

Changing the one on the right side of the total weight equality constraint to something less than one would correspond to a portfolio with the remaining weight invested in risk-free assets like cash and cash-equivalents such as United States Treasury Bills and other short-term government bonds. Removing the non-negativity constraint on w would allow short-selling of stock. We proceed with the current LP formulation since this analysis is directed at stock-only investors who do not short-sell.

4 Maximizing Return

Our second LP formulation maximizes return subject to constraints which reduce risk. We use a different measure of risk which is more appropriate for linear constraints in this formulation. Maximum drawdown is a popular metric portfolio theorists use for the stability of a portfolio. The drawdown at time t_n is the negative percent change in value of the portfolio from time t_{n-1} to time t_n. We limit the portfolio's drawdown for each time t_n to be no greater than x. For a set of stocks S over N times $t_1, ..., t_N$ the return of stock $i \in S$ at time t_n is the percent change in price from time t_{n-1} to time t_n. [4]

$$R_i(t_n) = \frac{P_i(t_n) - P_i(t_{n-1})}{P_i(t_{n-1})}$$

This is identical to the minimum risk LP formulation. We observe that this repetition is no accident. This definition of return is a foundation of portfolio theory. Defining return as the percentage change in price of each stock normalizes price changes with respect to the magnitude of the stock's value. This normalization property allows direct comparison of the performance of stocks with different magnitudes of price, which is essential when constructing a portfolio. For example a price increase of $1 in a $100 stock is a proportionally lesser change than a $1 increase in a $50 stock. We average the return of stock i across times $t_1, ..., t_N$ as in the minimum risk formulation.

$$\overline{R}_i = \frac{1}{N} \sum_{n=1}^{N} R_i(t_n)$$

We take the weighted sum over all stocks of each average return to obtain the average return of the portfolio

$$\overline{R} = \sum_{i \in S} w_i \overline{R}_i$$

where w_i is the weight of each stock and our decision variable as previously.

4.1 LP Formulation

We maximize the average return of the portfolio by changing the weight of each stock.

$$\max_{w_i} \left[\overline{R} = \sum_{i \in S} w_i \overline{R}_i \right]$$

We take our total weight equality constraint from the minimum risk formulation. As before, this constraint ensures the total weight of the portfolio is one. The same observations made before regarding the usefulness and meaning of upper bounds different from one apply here as well, however we proceed with one for the same reasons. We also take the non-negativity constraint on the weights in the same manner.

$$\sum_{i \in S} w_i = 1$$
$$w \geq 0$$

We now constrain the maximum drawdown. Here the ubiquity of return we emphasized before becomes evident again. Recall that drawdown is the negative *percent change in value* of the portfolio from time t_{n-1} to time t_n. This is just the negation of the *return* of the portfolio from time t_{n-1} to time t_n, which is the weighted sum of $R_i(t_n)$ for each stock i. The maximum drawdown constraint becomes a minimum return constraint. We limit the maximum drawdown to x by constraining the portfolio's return for each time in $t_1, ..., t_N$ to be at least x using this relationship.

$$\left[\sum_{i \in S} w_i R_i(t_n) \geq x \right]_{n=1}^{N}$$

This maximum drawdown constraint is meant to limit the risk of the portfolio to levels under to the investor's tolerance x, the parameter of our formulation. Tightening the constraint results in portfolios with greater diversification, which is when a portfolio includes a variety of stocks with diverse behaviour to decrease the impact of a decrease in price to any one stock. This is a natural consequence of the rule of thumb that increasing the number of stocks in a portfolio decreases the probability that all stocks will be correlated, which decreases the probability that all stocks will decrease in price at the same time. However, this also decreases the probability that all stocks will increase in price at the same time, which reduces the potential return, a corollary of the risk-return tradeoff discussed in the portfolio theory section. A value of $x = -1$ or an omission of this constraint entirely results in a single-stock portfolio holding the stock with the greatest return over the time period $t_1, ..., t_N$. The final maximum return LP formulation is then:

$$\max_{w_i} \sum_{i \in S} w_i \overline{R}_i$$

$$\text{subject to}$$

$$\sum_{i \in S} w_i = 1$$

$$\left[\sum_{i \in S} w_i R_i(t_n) \geq x \right]_{n=1}^{N}$$

$$w \geq 0.$$

There is a drawback to both LP formulations. Stocks are traded in integer quantities. It is not possible to purchase fractional shares. Investing according to these portfolio weights requires rounding to the nearest whole share. Rounding introduces differences between the portfolio's theoretical and observed performance, known in portfolio theory as "tracking error," which becomes less significant as the total monetary value of the portfolio increases. We chose to avoid (mixed) integer linear programming (ILP) formulations because we would then need to parameterize the amount of money available to invest in the portfolio, which we felt would eliminate the generalizability of the portfolios with no significant benefit (in real life, the tracking error is not consequential for the large investment funds which use these techniques). We note that in some respects, the maximum return formulation behaves in some ways like a "dual" of the minimum risk formulation. We use "dual" in a non-rigorous sense of the word (the formulations do not actually exhibit LP duality). In the minimum risk LP, we provide a return constraint and our objective optimizes risk. In the maximum return LP, we provide a risk constraint and our objective optimiizes return. This symmetry reflects a deeper connection between the two formulations (hence the "duality," in a loose sense) through the risk-reward tradeoff, which we won't elaborate on, but is interesting nonetheless.

5 Data

Wharton Research Data Services (WRDS), a division of the Wharton School of Business at the University of Pennsylvania, is highly-regarded in the academic community and has a dedicated quality control analyst at the New York Stock Exchange. [2] We use the WRDS Compustat-Capital IQ North America Index Constituents dataset to get the list of companies in the S&P 500 Index on March 20. We query the WRDS Millisecond Intraday Indicators dataset using this list to obtain the daily closing price for each S&P 500 company from February 20 to March 20 in a Pandas dataframe (Pandas is a Python data science package). This date range is chosen because it is the one-month time period which captured almost all of the peak-to-trough decrease in the S&P 500 Index due to COVID-19 (-32% out of a possible -34% from February 20 to March 23). We use the Pandas dataframe as input to both LP formulations.

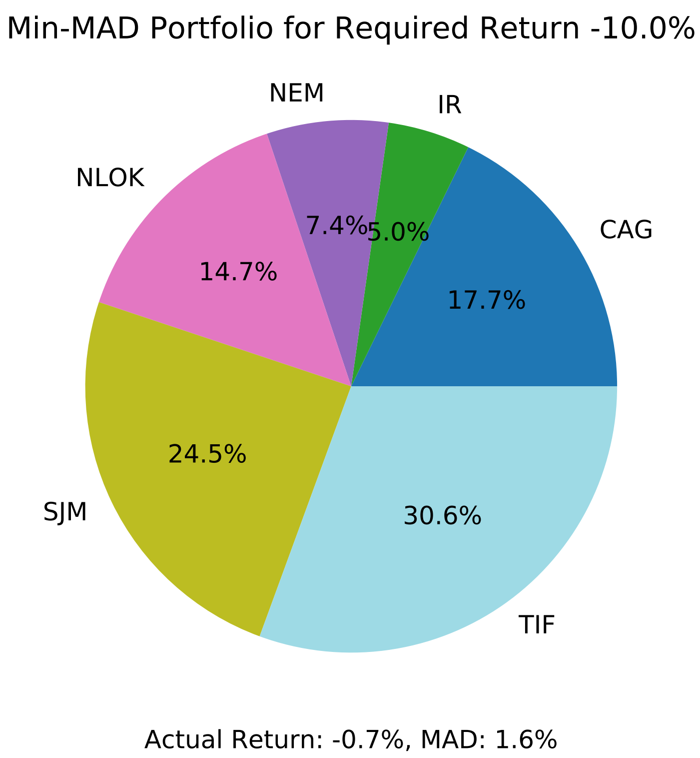
6 Results

We implement both LP formulations in Python using the Pyomo mathematical optimization package and the GNU Linear Programming Kit (GLPK) as the underlying solver. We choose a mathematical optimization package (Pyomo) in a general-purpose programming language (Python) over Solver in Excel due to the size of each LP. The minimum risk LP has 10,891 parameters, 537 variables, and 23 constraints. The maximum return LP has a similar size. Implementing these models in a spreadsheet program is clearly infeasible.

6.1 Minimizing Risk

We solve the minimum risk LP with $R_p = -0.1, 0.0, 0.005$ $(-10\%, 0\%, 0.5\%)$ to find conservative, balanced, and aggressive portfolios respectively. These values are selected for their diversely-composed portfolios from several solves in the full range of possible R_p values ($[-1, 1] = [-100\%, 100\%]$). The portfolios are named with respect to their risk level since increasing the minimum required return R_p restricts the feasible region, which results in solution weights which have greater MAD (risk).

6.1.1 Conservative Portfolio



NEM Newmont Corporation is the world's largest gold mining company. The company owns gold mines worldwide, including in the United States, Canada, Mexico, Austria, Ghana, Argentina, and Peru. Newmont also mines copper, silver, zinc, and lead.

IR Ingersoll Rand is a multinational provider of industrial equipment, technologies, and power tools. Ingersoll Rand has over thirty manufacturing facilities located in the Americas, Europe, Middle East, and Asia-Pacific.

CAG Conagra Brands is a packaged foods company which sells Act II, Hunt's Healthy Choice, Orville Redenbacher's, P.F. Chang's, and Bertolli.

TIF Tiffany & Co. (Tiffany's) is a luxury jeweller which is to be acquired by Louis Vuitton SE by the end of 2020.

SJM The J. M. Smucker Company, also known as Smucker and Smucker's, is a manufacturer of jam, peanut butter, jelly, fruit syrups, beverages, shortening, ice cream toppings, oils, and other products in North America.

NLOK Norton Life Lock is a cybersecurity software and services company.

Tiffany's The company with the greatest weight in our portfolio is Tiffany's. Tiffany's is due to merge with Louis Vuitton by the end of 2020 at a negotiated price of 135 US Dollars per share. The price is unlikely to deviate from this amount. This guarantees Tiffany's valuation to investors and results in a near-zero risk investment.

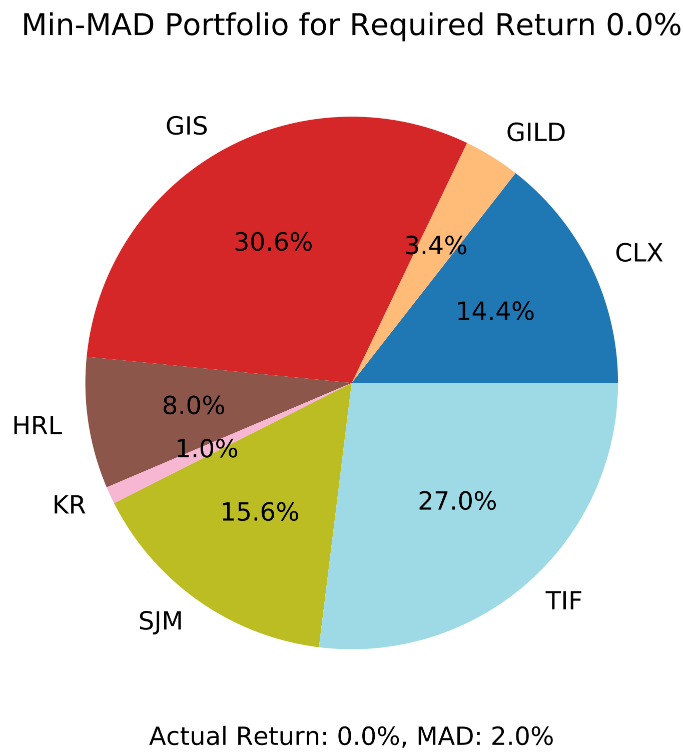
Conagra Brands and J. M. Smucker Fast-moving consumer goods (FMCG) companies, also known as consumer defensive or consumer staples companies, are low-risk due to demand for their products remaining constant despite changes in the overall economy. This property is known as "demand inelasticity" in economics and results in these stocks being "recession-proof" in portfolio theory. FMCG companies such as Conagra Brands and J. M. Smucker compose 42% of the portfolio. Since groceries are a necessity even during COVID-19 (in fact, there have been worldwide runs on grocery stores leaving store shelves bare of essential items) the demand for these companies' products remains inelastic, making them safe investments during this crisis.

Norton Life Lock As COVID-19 has driven an increase in online remote work, the demand for cybersecurity software and services offered by Norton Life Lock has increased as well, driving their company valuation up.

Newmont Newmont Corporation has benefited from an anticipated appreciation of gold during COVID-19, which normally rises in price in times of economic uncertainty as it is a safe-haven asset and long-term store of value. This appreciation is usually with respect to national currencies, which tend to be depreciated during crises as governments mint more money for economic stimulus spending.

Ingersoll Rand Investors have increased their valuations of Ingersoll Rand during COVID-19 as industrial automation in the manufacturing sector is seen as a potential consequence of COVID-19 as a means to reduce dependence on the availability of human labour. This would increase demand for their products such as precision tools and machinery.

6.1.2 Balanced Portfolio



GIS General Mills is a multinational manufacturer of branded consumer foods sold through retail stores, including Haagen-Dazs, Annie's Home Grown, Liberte, Betty Crocker, Pillsbury, Cheerios, and Old El Paso.

GILD Gilead Sciences is a biopharmaceutical company that researches, develops, and commercializes drugs. The company focuses primarily on antivirals used in the treatment of HIV, Hepatitis B, Hepatitis C, and influenza.

CLX Clorox is a manufacturer of consumer and professional chemical cleaning products.

TIF Tiffany's; see our conservative portfolio.

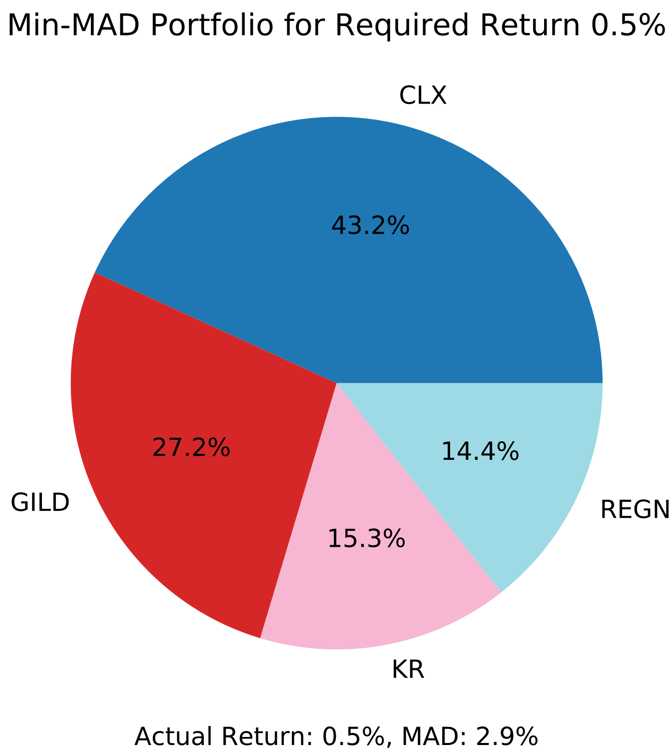
SJM J. M. Smucker; see our conservative portfolio.

KR Kroger is a chain of supermarkets.

HRL Hormel Foods is a food processor and packager which owns Spam and Stagg Chili.

FMCG companies (HRL, KR, SJM, CLX, GIS) compose 73% of the portfolio for the same reasons as in the conservative portfolio. Clorox cleaning products in particular are in high demand during COVID-19. Hormel Foods' Spam is also selling out on store shelves. The remaining 27% consists of Tiffany's and Gilead Sciences. Tiffany's is present for the same reasons as in the conservative portfolio. Gilead Sciences is currently testing remdesivir, a possible COVID-19 treatment which could save lives and propel the company to international recognition.

6.1.3 Aggressive Portfolio



GILD Gilead Sciences; see our balanced portfolio.

CLX Clorox; see our balanced portfolio.

KR Kroger; see our balanced portfolio.

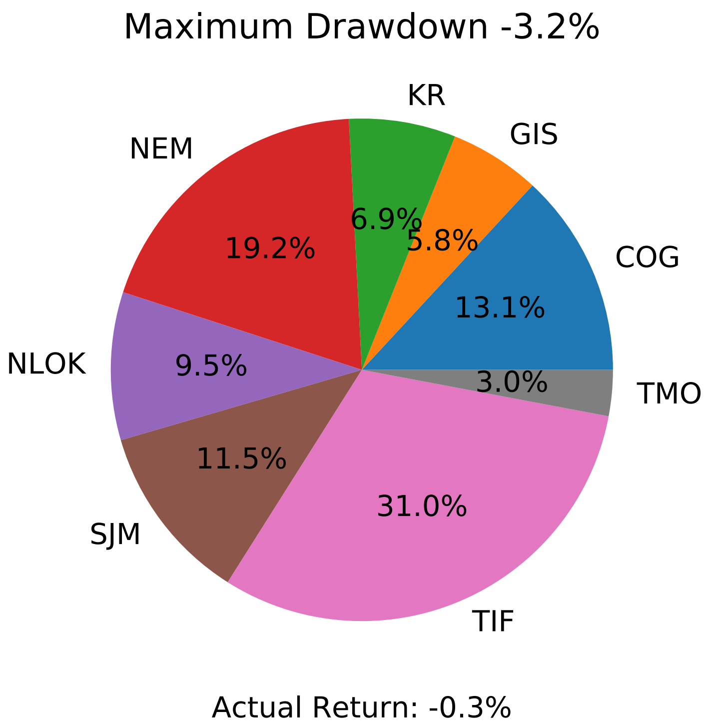
REGN Regeneron Pharmaceuticals is a biotechnology company.

The only new addition compared to the balanced portfolio is Regeneron, which isolated COVID-19 antibodies from mouse models and human survivors in mid-March that could neutralize the virus. An antibody cocktail is being prepared for clinical trials in the summer. Regeneron is also partnering with Sanofi to repurpose the rheumatoid arthritis treatment Kevzara for COVID-19, which is currently being tested in clinical trials. Like Gilead Sciences, this could be a life-saving breakthrough for the world and the company.

6.2 Maximizing Return

We solve the maximum return LP using maximum drawdown values $x = -3.2\%, -5\%, -6.8\%$ for conservative, balanced, and aggressive portfolios respectively (as mentioned previously, the lower the maximum drawdown, the less risk is incurred by the portfolio). These values were determined from sampling from the full range of maximum drawdown values ($[-100\%, 0\%]$) and selecting diversely-composed portfolios.

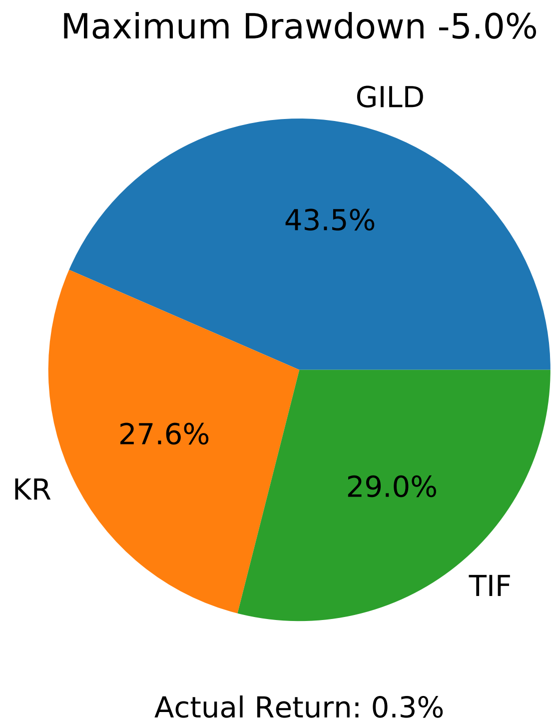
6.2.1 Conservative Portfolio



- **KR** Kroger; see the minimum risk portfolios.
- **GIS** General Mills; see the minimum risk portfolios.
- **TIF** Tiffany's; see the minimum risk portfolios.
- **SJM** J. M. Smucker; see the minimum risk portfolios.
- **NLOK** Norton Life Lock; see the minimum risk portfolios.
- **NEM** Newmont Corporation; see the minimum risk portfolios.
- **COG** Cabot Oil & Gas Corporation is engaged in natural gas exploration and drilling.
- **TMO** Thermo Fisher Scientific is a provisioner of scientific instrumentation, reagents, and consumables, software and services to healthcare, life science, and other laboratories in academia, government, and industry.

These companies are in our minimum risk portfolios (and their reasons for inclusion remain the same as they were before) except for Cabot Oil & Gas and Thermo Fisher Scientific, which together only compose 16% of the portfolio. Thermo Fisher Scientific in particular sees an increase in demand for its products as the global scientific and healthcare effort to fight COVID-19 accelerates. This portfolio is also highly diversified, with a total of eight stocks, making it the lowest risk of the maximum return portfolios (more on diversification was discussed in the LP formulation section).

6.2.2 Balanced Portfolio



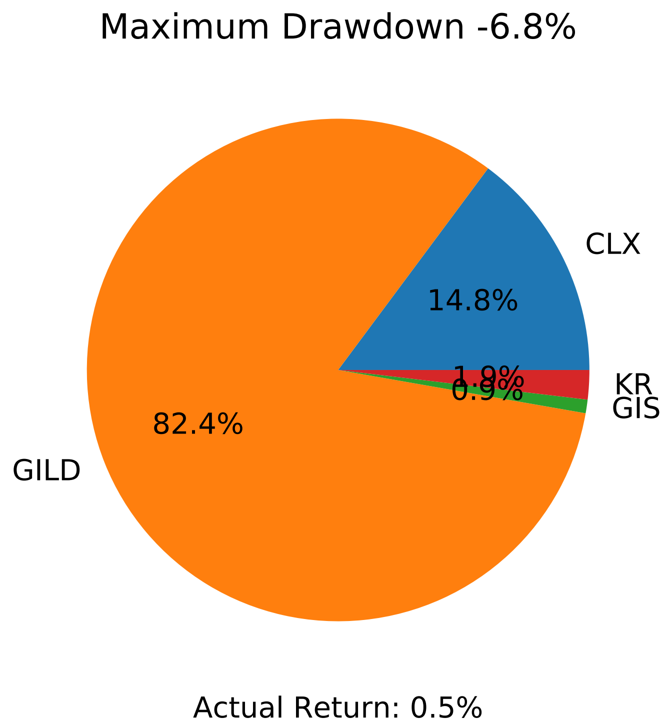
KR Kroger; see the minimum risk portfolios.

TIF Tiffany's; see the minimum risk portfolios.

GILD Gilead Sciences; see the minimum risk portfolios.

All of the components are in our minimum risk portfolios and for the same reasons. This portfolio is robust against COVID-19 since all 3 stocks are in different industries and 75% of weight is in pharmaceutical and FMCG companies. However, at only 3 stocks, this portfolio has a much lower level of diversification than the conservative portfolio, which is the main source of its medium risk level.

6.2.3 Aggressive Portfolio



GILD Gilead Sciences; see the minimum risk portfolios.

CLX Clorox; see the minimum risk portfolios.

KR Kroger; see the minimum risk portfolios.

GIS General Mills; see the minimum risk portfolios.

This portfolio consists of pharmaceutical and FMCG manufacturers and retailers which we have seen before in our minimum risk portfolios which have benefited from COVID-19 changes in demand (see the minimum risk portfolios for the specific reasons). However, over 80% of the portfolio is in a single company, Gilead Sciences. This single-stock emphasis reduces the diversification of the portfolio. Since the portfolio's value is greatly influenced by the value of Gilead Sciences due to this lack of diversification, it is high risk despite being the composition of four companies which are robust with respect to COVID-19.

Observations There are some interesting properties of the maximum return LP formulation. When the maximum drawdown is below -10.5% only REGN (Regeneron, a pharmaceuticals company discussed in the minimum risk portfolios) is included in the portfolio. This is because the maximum drawdown of REGN is -10.5% and REGN was the best performing stock during the time period. In general, the maximum return LP formulation quickly loses diversification and degenerates to a heavy weight in a few stocks, and the behaviour near -10.5% in this case is an example of this behaviour. There are also some drawbacks to the maximum return LP formulation. Although the portfolios possess optimal return for their given maximum drawdowns, some of the constituent companies' price increases are not from changes in the companies' underlying valuations (what portfolio theorists call the company's fundamentals) but are instead from panic buying ("fear of missing out") from retail investors and hedge funds. These emotionally-driven rallies in a company's stock are often short-lived. The maximum return portfolios do not take these effects into consideration so may not have a future performance which is similar to the theoretical return.

7 Conclusion

LP formulations are an effective method for finding minimum risk portfolios given a desired return or maximum return portfolios subject to constraints meant to limit risk to acceptable levels. The resulting portfolios are optimal for the market conditions during which they are created. Optimality means the portfolios account for whatever events are happening at the time. In the case of COVID-19, the portfolios with greatest return automatically included companies which benefited from the crisis. This automatic stock selection property is surprising considering that there is no artificial intelligence or machine learning involved. LP formulations are optimal despite knowing nothing about current events or company business models, yet simple and concise unlike complex and bulky neural networks. While the specific objectives, variables, parameters, and constraints are trade secrets, investment fund management firms such as the Vanguard Group (5.6 trillion US dollars under management) and BlackRock (7.4 trillion US dollars under management) use LP formulations to determine their asset allocations. [1] Future work includes testing our portfolios on future data for validation (an analysis of whether present performance holds in the future would be useful) and using simulated stock price data for various possible crises to construct crisis-resistant portfolios using our methods.

References

[1] Roger Aliaga-Diaz, Giulio Renzi-Ricci, Ankul Daga, and Harshdeep Ahluwalia. Vanguard asset allocation model: An investment solution for active-passive-factor portfolios — vanguard research, 2019. [Online; accessed 20-April-2020].

[2] University of Pennsylvania. Wharton research data services, 2020. [Online; accessed 20-April-2020].

[3] Wikibooks. Glpk/portfolio optimization — wikibooks, the free textbook project, 2010. [Online; accessed 19-April-2020].

[4] Martin R. Young. A minimax portfolio selection rule with linear programming solution, 1998. [Online; accessed 20-April-2020].