Dynamic Weighting Rotation Strategy

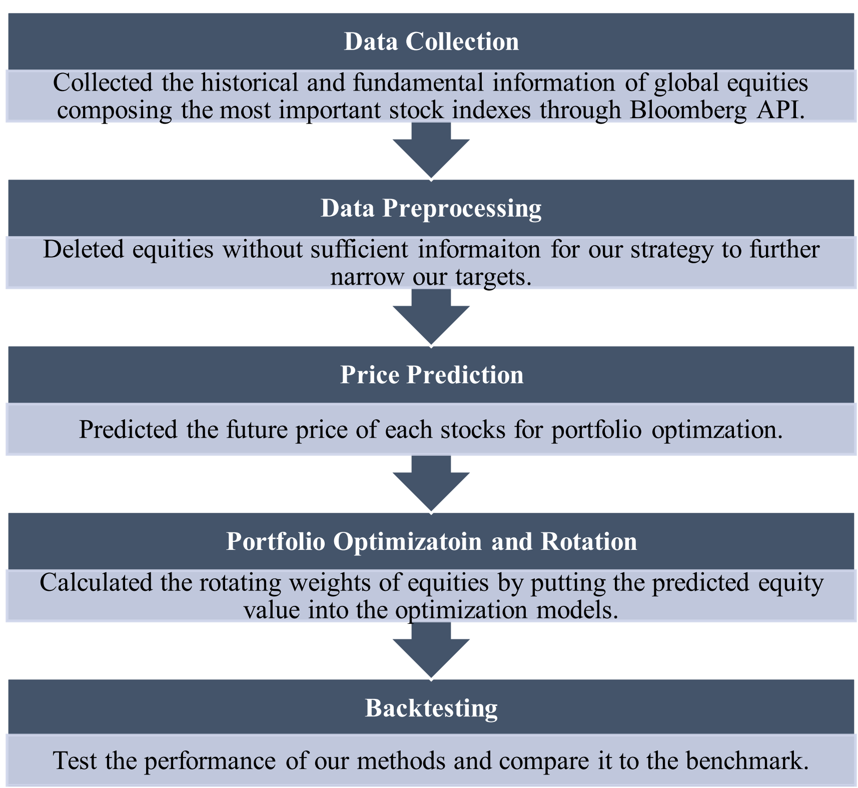
A Time Series & Multi-period Optimization Method

Sikun Xu, Sijia Sun, Yichao Lin, Yili Zhu
The Fu Foundation School of Engineering and Applied Science, Columbia University
500 W. 120th Street #510, New York, NY 10027

I. Introduction

There are more than 48,000 stocks with $70.2 trillion equity market capitalization being traded in global exchanges. How to select among them and further optimize the portfolio weights has been studied for a long time. In this paper, we proposed a dynamic weighting rotation strategy for asset diversification and risk diffusion, which takes advantage of different sectors and countries within the equity markets according to changing market conditions.

To diversify the assets among different countries, we select equities from 22 stock exchanges of both emerging and developed countries by referring to the component list of several important global indexes. Meanwhile, we control the price risks and maximize the returns by conducting a multi-period rotating optimization method, in which industry factors are also taken into consideration to prevent potential systematic losses caused by risk exposures of specific sectors.



GRAPH 1: Strategy Decision Making Process

The rest of this report is organized as follows. Section II describes the primary data we collected and preprocessed for further analysis. Section III presents how we predict the future prices of each equity by a modified ARIMA time series model along with the necessary regression test reports. Section IV introduced the optimization model for the portfolio weighting rotation and the principles of giving how much weight to a certain equity in the next trading day. We conducted the back-testing and discussed the results in Section V. Section VI concludes this report by and promoting potential improvements we could make with more market factors considered.

II. Data Preprocessing

Strong calculation power is needed to run the model for selecting equities if we take all stocks being traded into consideration, which exceeds the time requirement and is also unnecessary. Therefore, we narrowed down our target equities by referring to the components of the most significant global indexes, which are representative of the stock markets they belong to. They are also known to have relatively stable performance.

The MSCI World Index only includes stocks of "developed" markets (US, Western Europe, Japan, etc.). Our model, nevertheless, equities of emerging markets (like China, India, etc.) are also taken into account, which makes the strategy robust and the portfolio more diversified, and thus potentially outperformed the benchmark. The following table presents the indexes we referred to.

Index	S & P	NASDAQ 100	RUSSELL	TSX	MEXBOL	IBOV
Main Countries/Regions	United States	United States	United States	Canada	Mexico	Brazil
Index	MERVAL	IPSA	BE500	SX5E	UKX	DAX
Main Countries/Regions	Argentina	Chile	Europe	Europe	United Kingdom	German
Index	SMI	CAC	IBEX	FTSEMIB	BEL20	OMX
Main Countries/Regions	Swiss	France	Spain	Italy	Belgium	Baltic Countries
Index	PSI20	NKY	HSI	SHSZ300	TWSE	KRX100
Main Countries/Regions	Portugal	Japan	Hong Kong	China	Taiwan	Korea

Table 1: Referring Indexes and Corresponding Countries/Regions

Moreover, to consider the fluctuations of foreign exchange (FX) rates, we convert the stock prices primarily demonstrated in foreign currencies into US dollars based on daily foreign exchange rates. The strategy settled in US dollars not only standardizes the portfolio net worth evaluation in backtesting, but also enables us to have less foreign exchange exposure, since it is common for some markets to bearish and for the others bullish, especially in emerging markets, whose financial system is not so advanced to provide a stable currency environment.

To test the efficiency of our strategy, over 10GB global equities data across nations and sectors are collected from Bloomberg API. Daily official closing prices ranging from January 1st, 2017 to December 31th, 2018 are used for model training. Data from January 1st, 2019, to March 4th, 2019 are applied for backtesting and performance validation evaluation. The two-stage method of portfolio selection and weight optimization were conducted after all necessary data were ready. All computations following were performed in Python 3.7 on 2.7 GHz Dual-Core Intel Core i5 processor running on MacBook Pro, with 8 GB 1867 MHz DDR3 memory, Intel Iris Graphics 6100 1536 MB Graphics and macOS Catalina (version 10.15) operating system.

III. Price Prediction Based on Historical

In this part, we proposed the ARIMA model built for each equity and forecast its return of the following few days.

It is crucial to estimate the equity expected returns for portfolio optimization. We modified the traditional ARIMA model (Auto-Regression Integrated Moving Average Model) and conducted the realization with the historical information preprocessed in the last section in Python to forecast the next five days' returns.

Since ARIMA is a quite efficient short-term prediction time series model, we choose to build our prediction algorithm based on its philosophy. The equity prices are known to be influenced by countless factors like interest rates, investors' expectations and even weather, therefore, it is rather complicated to consider all for a large pool of stocks' price projections.

While the updated historical prices are fair reflections of all factors that influence the equity prices, and in most cases, the factors that influence the stock prices will not change a lot during a short-term period, it is quite reasonable to make predictions based on equities' historical prices. Since ARIMA model only focus on the historical records of a certain target, which furthermore, has the property of inertia, applying ARIMA model to forecast short-term equity price is quite reasonable and advantageous.

In general, the equity "AAPL.0" is used in the remaining part of this section to demonstrate how we analyze the data, and the other equities' calculations follow the same process.

3.1 Stationary test

The underlying assumption of Auto-Regression and Moving Average is that the series has equal means and variances, in other words, being stationary. However, for most equities, the overall price changes have trends. The integrated part of ARIMA model uses difference to deal with the dilemma. The order (d) in ARIMA(p, d, q) refers to the times of difference. After conducting a difference calculation, we use Augmented Dickey-Fuller test (ADF test) to see whether the series has become stationary, or we will further try difference of other degrees until it satisfies the test requirement. In general, the series ought to be stationary if the p-value of the ADF test is less than 0.05. The result of ADF test for the original series of "AAPL.0", for example, is 0.277231, thus we will further try other options. While since after one-step difference, the p-value decreases to almost zero, we conclude that the order (d) of ARIMA for "AAPL.0" should be one. The table below presents the result of the whole ADF test.

	Original Series	One-step Difference Series
Test Statistics	-2.021	-24.993
p-value	0.277	0.000
Number of Lag(s) Used	1	0
Number of Observations	728	728

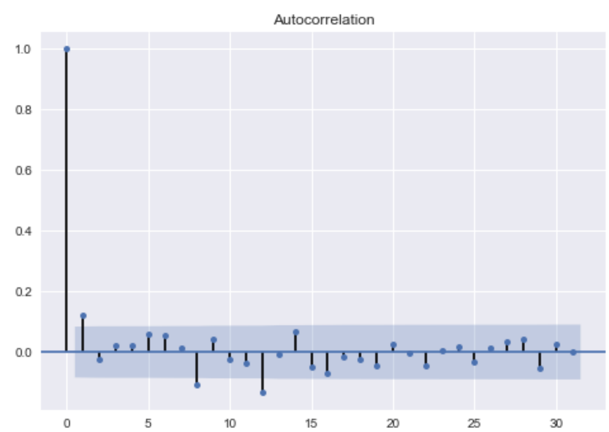
Critical Value	1%	-3.439	-3.439
	5%	-2.866	-2.866
	10%	-2.569	2.569

Table 2: Stationary Test Report

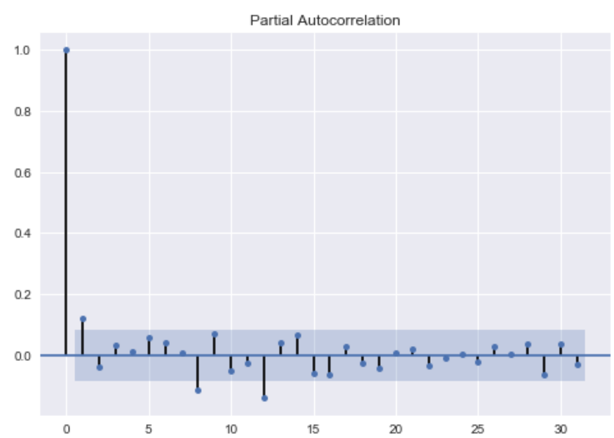
3.2 ACF and PACF

Autocorrelation and partial autocorrelation are measures of association between current and past series values and indicate which past series values are most useful in predicting future values. With this knowledge, one can determine the order of processes in an ARIMA model. Here, we can use ACF and PACF to observe the trend of the series.

In this subsection, we prove that choosing seasonal ARIMA model is a wise decision. According to the property of ARIMA, the order of an AR or MA model will be seen in the exponential decay of the lags. The ACF/PACF of the series are graphed as follows, from which we conclude that tendency to converge is obvious, and thus the time series model fits the condition.



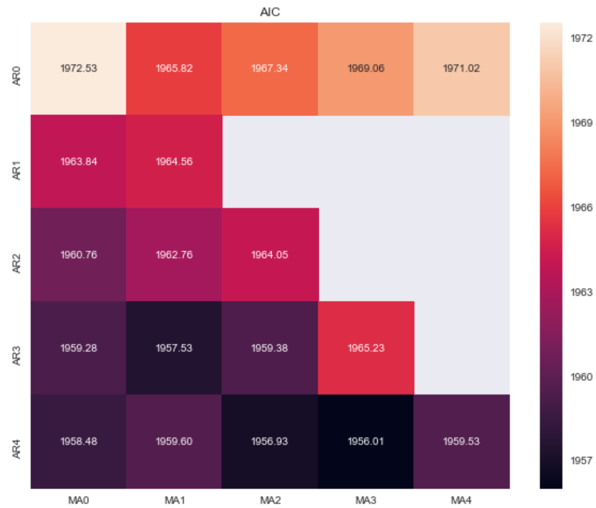
Graph 2: Auto-correlation



Graph 3: Partial Auto-correlation

3.3 Order and Parameter Estimation

There are three orders in the ARIMA model. The order d as the time of difference is determined in section 3.1. We apply Akaike Information Criterion (AIC) to calculate the value of p and d to determine the order of AR and MA. AIC is defined as 2k - 2ln(L), in which k is the number of estimated parameters in the model and L is the corresponding maximum likelihood function of the model. For all possible values of p and q, the pair which minimizes the AIC is the fittest ARIMA model. The value of p and q are chosen between 0 to 4 and the following graph presents the results.

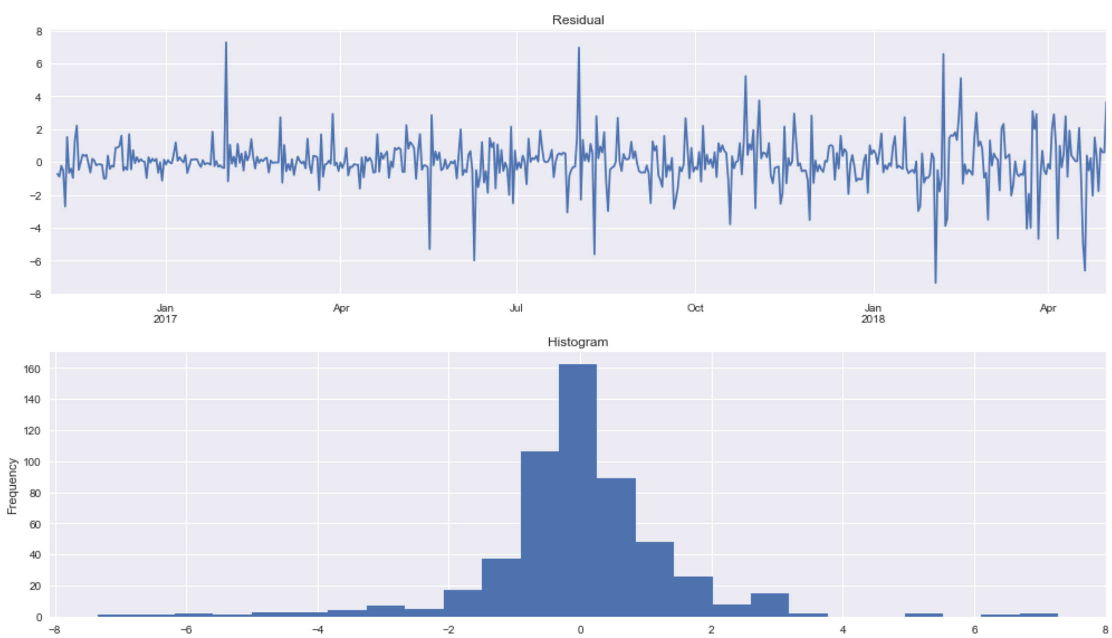


Graph 4: ARIMA Difference Calculation

As shown in the heatmap above, the part without value means the series doesn't fit the corresponding ARIMA model, which means the likelihood function doesn't converge. In addition, every square in the heatmap refers to the AIC value of the corresponding ARIMA model. We can see that ARIMA (4,1,3) fits best, which thus is used for the prediction.

3.4 Model Diagnostic Check

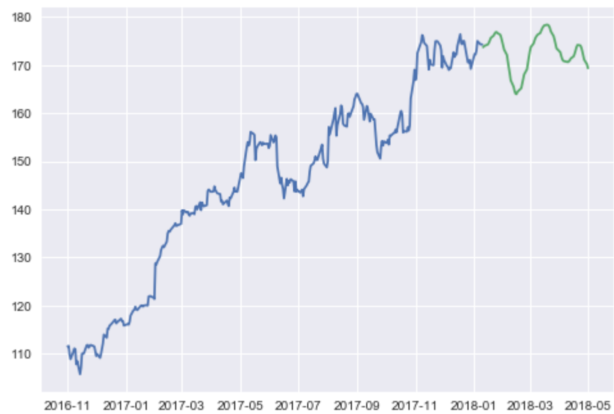
To prove the adequacy of the ARIMA models, we first checked the stationarity of our models via the residuals. The residuals of ARIMA models stand for white noises, which follows the normal distribution. Here, we conduct the Ljung-Box-Pierce Q-test and have proved that the residuals are close to a white noise since the p-test accepts the null hypothesis with a P-value less than 0.05.



Graph 5: Diagnostic Check

3.5 Forecasting with the Seasonal ARIMA Model

After identifying the order and parameters of the seasonal ARIMA model, we use it to forecast the following 5 days' prices and returns of an equity. The graph below is the price trend prediction of "AAPL.0". The rest of the equities in our asset pool are also processed following the same analysis logics.



Graph 6: "AAPL.0" Price Prediction

IV. Portfolio Optimization and Rotation

Given the expected returns, the problem has been reduced to the optimal weight calculation. The next step thus is to decide the dynamic rotating weights of each equity in our asset pool. By conducting a multi-period mean-variance optimization model, we make daily decisions based on the stocks' historical performance.

The frequency of decision making is once a day, and each day we obtain expected returns updated based on the close prices of last 30 trading days. We predicted the stock price of the decision day, regarding it as the actual prices we will observe on that day and make adjustments on the existing portfolio. Such decision is modeled as a portfolio optimization problem.

Optimization is a powerful decision-making tool, which connects the data to the final decisions. It returns the optimal solutions under various conditions based on decision makers' preferences and objectives. An optimization model consists of two parts: objective functions and constraints, which specifies goals we want to achieve with all the limits we faced.

Here are the notations of our optimization model:

Notations	Explanations
$\boldsymbol{I}$	The equity pool, consisting of all the target equities of our investment;
$\boldsymbol{M}$	Model parameter, used to linearize nonlinear constraints;
$\boldsymbol{Cov(i,j)}$	$i, j \in I$, indicating the covariance of equity i and equity j;
$\boldsymbol{\alpha_i}$	$i \in I$, indicating the expected return of equity i
$\boldsymbol{c}$	Transaction cost charged for selling a unit of equity, represented in percentage;
$\boldsymbol{w_0}$	$w_0 \in R^I$, the portfolio weight of the last decision making;
$\boldsymbol{w_{bench}}$	The equity weights in the benchmark portfolio;

LB_w, UB_w	The lower bound and upper bound of equity weights in the portfolio;
LB_N, UB_N	The lower bound and upper bound of equity weights in each market;

Table 3: Parameter Explanations

The decision variables of the optimization model are:

w_i	$i \in I$, indicating the weight of equity i in our investment portfolio;
y_i	$i \in I$, indicating the quantity of equity i we sell on the present trading day;
z_i	$i \in I$, artificial variable.

Table 4: Parameter Explanations

4.1 Objective of Portfolio Optimization Model

We strike a balance between maximizing expected return and portfolio diversity by constructing the objective functions and constraints in portfolio optimization model. Up to now, the expected return has been calculated by the time series model in the previous section, which is:

$$\max_{w} \sum_{i \in I} \alpha_i w_i$$

Because we are considering investment strategy in a long-time spam rather than a one-time decision, we need to consider the cost of portfolio change, i.e. the transaction costs:

$$\max_{w} \sum_{i \in I} \alpha_i w_i - c \sum_{i \in I} y_i$$

The influence of transaction cost is significant, because it determines the gains in each trading action. On the one hand, higher transaction costs result in a smaller portfolio adjustment each time; On the other hand, it represents the tradeoff between expected return (or drawback) of the portfolio at hand and that of the equities on the market but outside the portfolio.

The second part of the objective function is about maximizing the portfolio diversity, which is measured by return covariances. The latter is a measurement of correlations between each pair of equities in our equity pool. A positive covariance means that a high return of one equity often comes with a high return of the other, and vice versa. A covariance approximately equaling to zero implies that there is little correlation between the selected pair of equities. By minimizing the covariance of each pair of equities, we achieve the objective of portfolio diversity maximization and hence reduce the portfolio risk. The covariance minimization objective is expressed as:

$$\min_{w} \sum_{i \in I} \sum_{j \in I} w_i w_j Cov(i, j)$$

Combining the objectives of maximizing the expected returns and minimizing the portfolio covariance, we build a multi-objective programming model. In addition, the return maximization objective and the covariance minimization objective are all valued between 0 and 1, we thus don't need to standardize the result further.

However, since investors all have a preference for risks and returns, we introduce a tuning parameter $\theta \in [0,1]$ representing the significance of each objective for our final multi-objective programming model. The final objective function is shown below:

$$\max_{w} \quad \theta\left(\sum_{i\in I} \alpha_i w_i - c\sum_{j\in I} y_i\right) - (1-\theta)\sum_{i\in I}\sum_{j\in I} w_i w_j Cov(i,j)$$

Minimizing the covariance represents a conservative strategy, while maximizing the expected return represents an aggressive strategy. For the model testing, we adopted a moderate one (θ=0.5).

4.2 Constraints of the Portfolio Optimization Model

The basic constraints of a portfolio optimization model are:

1. No shorting: $w_i \geq 0, \forall i \in I$
2. Weight Normalization: $\sum_{i\in I} w_i = 1$
3. Weights bounds: $LB_w \leq w_i \leq UB_w, \forall i \in I$

Since the decision variables y_i and w_i are internally connected, we introduce the relation specifically:

$$y_i = max\ \{0, w_i^0 - w_i\}$$

However, this equality is non-linear, which would make the final model harder to solve. Thus, we introduce an artificial binary variable $z_i, i \in I$ to linearize the equality constraints:

$$-y_i + (w_i^0 - w_i) \leq Mz_i, \forall \mathrm{i} \in I$$
$$y_i - (w_i^0 - w_i) \leq Mz_i, \forall i \in I$$
$$w_i^0 - w_i \leq M(1 - z_i), \forall i \in I$$

In addition, to enhance the performance of our portfolio, we introduce an extra constraint. On the one hand, to diversify our portfolio across the global markets, we limit the weight of equities of a certain region or country in our portfolio by setting a predetermined percentage allowed to be deviated from the benchmark's weight assignments. For example, we found that the benchmark puts 58% weight in the U.S. market, the total weights our portfolio put in the U.S. market can only takes on value between $58\% - LB_N$ and $58\% + UB_N$. In our model, LB_N and UB_N are all 10%. The limit indicates that we are allowed to include certain equites of the markets that are not covered by the benchmark. The constraint is expressed as:

$$LB_N \leq \sum_{i\in n} w_i - w_{bench,n} \leq UB_N, \forall n \in N$$

In conclusion, the overall portfolio optimization model is:

$$\begin{aligned}
\max_{w,y,z} \quad & \theta\left(\textstyle\sum_{i\in I} \alpha_i w_i - c\sum_{j\in I} y_i\right) - (1-\theta)\sum_{i\in I}\sum_{j\in I} w_i w_j Cov(i,j) \\
s.t. \quad & LB_N \leq \textstyle\sum_{i\in n} w_i - w_{bench,n} \leq UB_N, \forall n \in N \\
& -y_i + (w_i^0 - w_i) \leq Mz_i, \forall i \in I \\
& y_i - (w_i^0 - w_i) \leq Mz_i, \forall i \in I \\
& w_i^0 - w_i \leq M(1 - z_i), \forall i \in I \\
& LB_w \leq w_i \leq UB_w, \forall i \in I \\
& \textstyle\sum_{i\in I} w_i = 1 \\
& w_i \geq 0, \forall i \in I
\end{aligned}$$

V. Backtesting

5.1 Max Drawdown

Max drawdown rate is the maximum observed loss from a peak to a trough of a portfolio chosen, which is an indicator of downside risk and also demonstrates the robustness of a given strategy and investment. The less the max drawdown rate is, the better the investment made.

$$\mathrm{D}rawdown = \frac{Max(Day_i - Day_j)}{Day_i}, 0 \leq i < j \leq T$$

Day_i/ Day_j is the closing price of the i^{th}/j^{th} day.

5.2 Sharpe Ratio

Sharpe Ratio measures the performance of a portfolio as compared to the risk-free return, after adjusting for its volatility. It represents the additional amount of return that an investor receives per unit of increase in risk.

$$SharpeRatio = \frac{E(R_p) - R_f}{\sigma_p}$$

R_p is the portfolio return, R_f is the risk-free return and we use the U.S. Treasury Yield of 1.5% in our calculation. $E(R_p) - R_f$ is the expected value of the excess of the asset return over the risk-free return, and σ_p is the standard deviation of the asset excess return.

5.3 Backtesting Configuration

Period	01/01/2019 – 03/04/2019
Initial Principal	$ 1,000,000
Transaction Cost	1‰
Benchmark	MSCI World NTR (USD)
Trading Instruments	Over 2,000 global equities
Trading Frequency	Once a day
Settlement Currency	US dollar
Data Sources	Bloomberg

Table 5: Backtesting Configuration

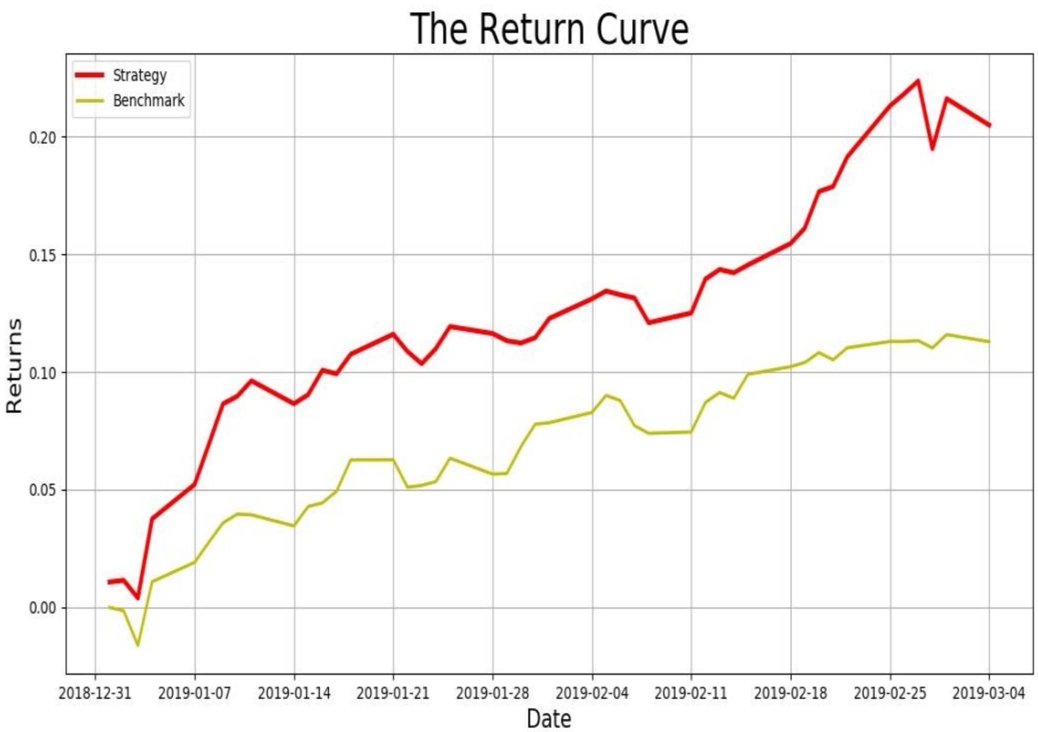
5.4 Model Performance

Every trading day, we evaluate our portfolio net worth after the trade closes and calculate the new weights of each equity in the portfolio using the close price. Afterwards, we would add the close price of the day to our dataset as updated information for the next trading day's prediction and optimization. The model is daily rerun based on the dynamic rotation methods presented in Section III and IV, after which, the optimal weights of each equity for the next trading day is decided. If the costs of weight adjustments are lower than the benefits we will get, then our model will show how to make the optimal decision.

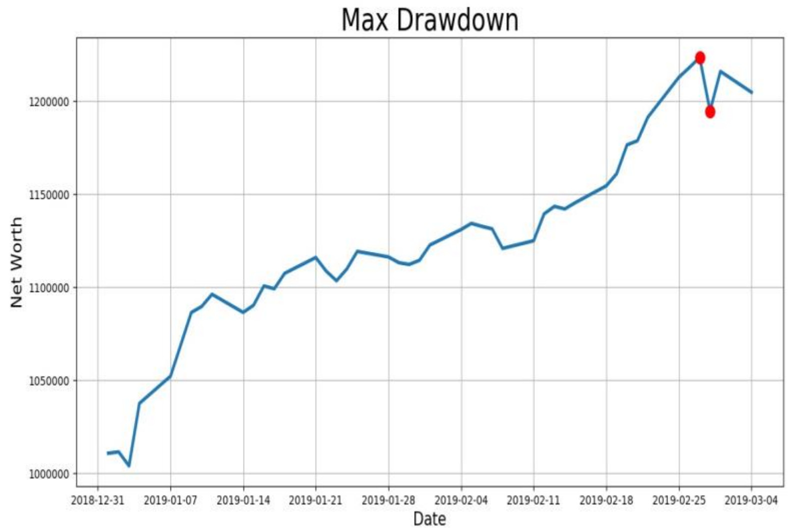
On a new trading day, we use the open price of the asset, which is the accumulated price of the bid and ask upon the opening of an exchange, to buy and sell shares of each asset in our portfolio to keep the balance.

	Total Returns	**Sharpe ratio**	**Maximum Drawdown**
Strategy	20.48%	7.07	2.35%
Benchmark	11.28%	5.43	1.62%

Table 6: Model Performance



Graph 7: The Return Curves of the Strategy and Benchmark



Graph 8: The Max Drawdown of our Strategy

VI. Conclusions and Discussions

To build the dynamic weighting rotation strategy for equity selection and portfolio optimization, we first narrowed down the targeted stocks by only considering the component stocks of several important indexes covering both emerging and developed countries as well as various industrial sectors. After filtering those stocks without qualified data for our analysis, we put the data sets of nearly 2000 stocks for analysis and allocation.

In the main part of the portfolio construction, we first predict the future stock prices by doing the ARIMA regression and then use the predicted prices for portfolio optimization. The programs we write can automatically conducted the time series regression tests required for ARIMA method and chose the prediction model with suitable difference order for each stock. Then the projected prices of the next following days are returned for the equity weight assignments. An optimization model based on equities' mean-variance are then built for the decision making.

The results are consistent with our expectations, which show that the return of our strategy outperform the benchmark (MSCI). Also, since the model is quite efficient and robust, we can lower down the decision-making frequency. As long as the portfolio current weight assignment gains money, the is not a must to change the weight every day. Otherwise, the models should be run again for new weight decision.

In the future research, we plan to include more elements that may affect the short-term performance of a stock such as investors' sentiments to predict the prices more accurately. In addition, compared to an outright purchase, we might make use of swaps, futures, forwards and Contract for Difference (CFDs), which require less cash and have less FX exposure. Furthermore, we are working on designing a more efficient pre-selection algorithm before we predict the future prices and conduct the portfolio optimizations, so that more well-performed index components can be taken into consideration. Last but not least, since the model in Section V runs a relatively long time, we will speed up the program by improving the algorithm further.