Portfolio Optimization Through Parametric Portfolio Policies

By: Derick Hanscom, Jharold Montoya, Kaiyu Wen,
Yuqiang Zhang, Yi-Hsuan Fan, Tristan Ko

Professor Chris Kelliher & Professor Eugene Sorets
MF796 Computational Methods
May 2nd, 2023

Abstract

Parameterizing portfolios can be exploited by characteristics in the cross sectional return of assets. By considering the weights distributed to each asset as a function of its characteristics we can optimize and model the evolution of weights through time as data is accumulated. If we maximize the expected utility of our portfolio return we can generate a portfolio that can outperform the equally weighted portfolio. We present an approach to seamlessly compute and generate robust returns by investing in a broad set of indices around the world based on four main characteristics; momentum, quality, value, and trend.

1. Introduction

Financial performance and technical indicators are known to reflect the expected return and variance/covariance of equity assets. These characteristics can be used to model the weights of a portfolio by considering them to be the domain of a function that maps to the range being the portfolio weights. By exploiting this in conjecture to maximize the expected value of the utility function we can generate a portfolio that outperforms the equally weighted portfolio. The assets used consist of a broad set of indices to develop a global equity portfolio. We used Parametric Portfolio Policy as our portfolio optimization framework. This framework allows us to perform periodic weight rebalance efficiently and suffers less from insample data overfitting and less estimation error towards the parameters (Brandt). We assumed a CRRA (Constant Relative Risk Aversion) utility function to ensure our objective function is twice differentiable such that our gradient based optimization method conditions are satisfied. The CRRA model uses a relative risk aversion parameter that explains an investor's sensitivity to loss. The larger the relative risk aversion parameter, the larger the sensitivity to extreme loss. By using an expanding window we can invest and update our positions over time allowing our characteristic weights to converge to an optimal value as we accumulate data. This approach reduces time complexity seen in the Markowitz approach by avoiding the calculations of the first, second, and any other necessary moments of returns, for large dimensions this has a large impact on runtime differences. A side effect of this dimensionality reduction is that we avoid the issue of multicollinearity expressed through imprecise coefficient estimates which in turn reduces the degree that the model overfits the previous data (Brandt). For the risk matrix, we used VaR, CVaR, and Maximum Drawdown to assess the potential losses (Stan)(Admin). VaR, CVaR, and Maximum Drawdown are all useful risk measures that can provide investors with different perspectives on the potential downside risk of their investment portfolio or trading strategy (MarkD). VaR and CVaR provide estimates of the likelihood and magnitude of potential losses, while Maximum Drawdown measures the maximum loss an investor could experience from a portfolio's previous high.

2. Data

We gathered leading index prices and financial ratios from developed and emerging countries from both Bloomberg and Reuters between January 2010 and March 2023. We then used that information to calculate the factor-based strategies, those being momentum, trend, quality, and value.

While cleaning the data, we excluded indexes that had too many missing values. As a result, 28 countries were included: Austria, Belgium, Brazil, Canada, Chile, China, Czech Republic, Denmark, Finland, France, Germany, Hungary, India, Indonesia, Japan, Mexico, the Netherlands, New Zealand, Norway, the Philippines, Poland, Portugal, Sweden, Switzerland, Taiwan, Thailand, the UK, and the US.

We utilized the cross-sectional z-scores of each strategy as index characteristics for the following optimization process.

3. Method

3.1 Factor Investing

Factor investing is an investment approach that involves targeting specific drivers of return across asset classes. There are two main types of factors: macroeconomic and style. Investing in factors can help improve portfolio outcomes, reduce volatility and enhance diversification.

Building a Factor-based portfolio is a form of active investing. They purposely "tilt" portfolios towards certain characteristics, like recent momentum, higher quality, or lower stock prices to achieve specific risk and return objectives.

Factor investing may be appropriate if a long-term investor is seeking to pursue specific factors but are looking for more transparency than can be found in a traditional actively managed fund.

We are using four well documented Style Factors to build our portfolio:

3.1.1 Momentum

- **Description**: The momentum investment strategy is founded on the principle that stocks that have exhibited strong performance in the past will likely continue to outperform, while stocks that have performed poorly in the past will continue to underperform. Momentum investors search for stocks that have shown strong recent performance and invest in them with the expectation that they will continue to do well. Moreover, when a stock's price starts to rise, more investors become interested in it, which can lead to further price increases. Therefore, this strategy is based on the belief that positive momentum is indicative of positive investor sentiment and reflects a company's strong fundamentals.
- **Method:**

$$M_i^t = R_i^{t-1,t-12}$$

Where $R_i^{t-1,t-12}$ is the cumulative return of the index i from the month $t-1$ to $t-12$.

3.1.2 Trends

- **Description**:The trend-following investment strategy is based on the belief that assets that have exhibited strong short-term trends are likely to continue those trends in the future. This strategy involves using technical analysis tools such as exponential moving averages to identify trends in asset prices over different time periods. The exponential moving average is a weighted average that gives more weight to recent data points, making it a popular tool for identifying short-term trends. Trend-following investors seek to identify assets that have shown positive momentum over a short-term period, while also considering longer-term trends.
- **Method:**

$$T_i^t = EWMA_i^{32} - EWMA_i^{128}$$
$$EWMA^t = \alpha \cdot P_t + (1-\alpha) \cdot EWMA^{t-1}$$

Where EWMA is the exponential weighted moving average P_t is the price of the index in period t and $\alpha = 2/(t+1)$. The trend is calculated in a recursively until $EWMA^0$ is achieved.

3.1.3 Value

- **Description:** The value investment strategy is based on the principle that assets that are undervalued, or cheaper relative to some measures of fundamental value, are likely to outperform assets that are overvalued or expensive in the long run. The fundamental metrics below are used to determine whether an asset is undervalued or overvalued by the market.
- **Method:**

$$Price\ to\ earnings = \frac{Stock\ price}{earnings\ per\ share} \qquad Price\ to\ book = \frac{Stock\ price}{Book\ value\ per\ share}$$

$$Price\ to\ cash\ flow = \frac{Stock\ price}{Cash\ flow\ per\ share} \qquad Price\ to\ EBITDA = \frac{Stock\ price}{EBITDA\ per\ share}$$

3.1.4 Quality

- **Description:** The quality investment strategy is based on the belief that companies with solid financial metrics are likely to show superior future performance. This strategy involves analyzing a company's financial metrics, such as the return on equity ratio, debt to equity ratio, and trailing 12 months profit margin, to determine whether the company is financially sound and well-managed.
- **Method:**

$$Return\ on\ Equity = \frac{Net\ Income\ available\ for\ common\ shareholders}{Average\ total\ common\ equity}$$

$$Debt\ to\ Equity = \frac{Total\ debt}{Current\ Enterprise\ Value}$$

$$Trailing\ 12M\ Profit\ Margin = \frac{Trailing\ 12M\ Net\ Income}{Trailing\ 12M\ Net\ Sales}$$

3.2 Objective Function

$$\max_{\{w_{i,t}\}_{i=1}^{N_t}} E_t[u(r_{p,t+1})] = E_t[u(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1})]$$

where: $r_{p,t+1}$ is the portfolio return from time t to $t+1$

N_t is the number of assets at time t

$w_{i,t}$ is the individual asset weight being held from time t to $t+1$

$r_{i,t+1}$ is the individual asset return from time t to $t+1$

Equation above is our objective function. At the current time t period, we are trying to maximize the expected utility of the upcoming holding period portfolio return given all the information we know currently at time t, with respect to the assets' individual holding weights.

We will define the individual weight $w_{i,t}$ as follows:

$$w_{i,t} = \bar{w} + \frac{1}{N_t}\theta^T \hat{x}_{i,t}$$

The individual weight is composed by two parts: the constant equally weighted weight, $\bar{w}$, plus the zero-sum long-short weight part, $\theta^T \hat{x}_{i,t}/N_t$ which is determined by the historical factor characteristics. $1/N_t$ is the normalization constant. θ is a vector of coefficients needed to be estimated, and $\hat{x}_{i,t}$ are the factor characteristics of asset i at time t being **standardized cross-sectionally with mean zero and unit variance**. The reason we standardize the factor characteristics is because this step naturally sets up the fully invested constraint and will guarantee the weights sum up to 1.

We can see that in the individual weight function, θ is the only unknown variable, and it is also invariant w.r.t time and assets. Constant θ w.r.t. asset means that the individual weight only depends on the asset's factor characteristics. And constant θ w.r.t. time means that the coefficients which maximize the conditional expectation at a given date are the same for all the historical dates. Therefore, we can simplify our maximization framework into:

$$\max_{\theta} E[u(r_{p,t+1})] = E[u(\sum_{i=1}^{N_t}(\bar{w} + \frac{1}{N_t}\theta^T \hat{x}_{i,t}) r_{i,t+1})]$$

Given the historical data of $\hat{x}_{i,t}$ and $r_{i,t+1}$, we can estimate the expectation by the sample mean:

$$\max_{\theta} \frac{1}{T}\sum_{t=0}^{T-1} u(r_{p,t+1}) = \frac{1}{T}\sum_{t=0}^{T-1} u(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1})$$

There are several advantages for this portfolio optimization framework. First is that our optimization has less estimation error and is numerically robust (Brandt). The traditional Mean-Variance approach involves estimating the expected returns and covariance matrix using the in-sample data. The estimation error for the entries of the covariance matrix and expected returns accumulates and is absorbed into the weights. In contrast, our framework suffers less estimation error and in-sample overfitting. It has less estimation error because θ has way smaller dimensions, 4 in our case, when comparing with expected returns plus covariance matrix. It is more numerically robust for out of sample since the entries of θ will only deviate from 0 if the historical factor characteristics provide an interesting combination of return and risk consistently across assets and through time.

The second advantage is that, albeit not shown in the objective function, our optimization process considers the connection between the factor characteristics and all the moments of the historical returns (Brandt). If we take a Taylor series expansion of $u(r_{p,t+1})$ around $E[r_{p,t+1}]$ and take the expectation:

$$E[u(r_{p,t+1})] \approx u(E[r_{p,t+1}]) + \frac{1}{2}u''(E[r_{p,t+1}])E[(r_{p,t+1} - E[r_{p,t+1}])^2] + \cdots$$

We can see all the order of moments of the portfolio return, and hence the distribution, are being considered. Since θ controls the assets' weights, and weights controls the portfolio return, the optimization process is definitely going to consider the joint distribution of asset historical returns when optimizing the θ (Brandt).

To have a better understanding of our portfolio under this framework, let's look at the return of the portfolio $r_{p,t+1}$:

$$r_{p,t+1} = \sum_{i=1}^{N_t} \bar{w} r_{i,t+1} + \sum_{i=1}^{N_t} (\frac{1}{N_t}\theta^T \hat{x}_{i,t}) r_{i,t+1} = r_{m,t+1} + r_{h,t+1}$$

We can think of our portfolio return as the sum of two sub-portfolio returns: an equal-weighted portfolio plus a hedge fund portfolio with weights add up to zero. Therefore, it is intuitive to compare the performance for our optimized portfolio with the equal-weighted portfolio(bench mark).

3.3 Utility Function

The utility function we are going to use is the Constant Relative Risk Aversion(CRRA) utility function.

$$u(r_{p,t+1}) = \frac{(1 + r_{p,t+1})^{1-\gamma}}{1 - \gamma}$$

$$\left\{ \begin{array}{l} u(r_{p,t+1}) = \frac{(1+r_{p,t+1})^{(1-\gamma)}}{(1-\gamma)} : \gamma > 1 \\ u(r_{p,t+1}) = \ln(1 + r_{p,t+1}) : \gamma = 1 \end{array} \right\}$$

where: γ is the relative risk aversion coefficient

The reason we chose to use this specific utility function is because it has multiple attractive properties. First is that CRRA is invariant to the wealth level, it simplifies our analysis and decision-making process for portfolio optimization. Second, it allows us to easily change the γ depending on the investors position on risk. Additionally, CRRA is twice differentiable which allows us to leverage the efficiency of gradient-based optimizer when finding the global optimal point.

3.4 Optimization

The quasi-newton method is a set of adaptations to Newton's method that avoids the direct and computationally expensive calculation of the inverse hessian matrix. Newton's method is used to find a stationary point in a function where the gradient is zero, either a local maximum or minimum. The inverse hessian matrix is a powerful tool commonly used in gradient approaches but can lead to some issues like computational expense and the possibility of a singular hessian matrix (non invertible). These extensions to Newton's method do not need to compute the inverse hessian directly and instead estimates and updates the inverse of the hessian using gradient vectors. The quasi-newton method becomes increasingly efficient on higher dimensional settings; the time complexity of Newton's method is O(N^3) while the time complexity of quasi-newton method BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm is O(N^2). The updating process of the inverse hessian matrix at time k+1 is expressed as:

$$\left(I - \frac{\Delta x_k y_k^{\mathrm{T}}}{y_k^{\mathrm{T}} \Delta x_k}\right) H_k \left(I - \frac{y_k \Delta x_k^{\mathrm{T}}}{y_k^{\mathrm{T}} \Delta x_k}\right) + \frac{\Delta x_k \Delta x_k^{\mathrm{T}}}{y_k^{\mathrm{T}} \Delta x_k}$$

Here we can see that the estimation of the inverse can be expressed solely with the given data without needing to take any inverses making it an ideal choice of a gradient based optimization approach. An interesting artifact of this method is that although one of the required conditions for this method to converge is that the hessian is invertible, the method can still converge if the hessian is singular because we only calculate an estimate which can sometimes possess the ability to achieve convergence.

3.5 VaR, CVaR, and Maximum Drawdown

VaR (Value at Risk), CVaR (Conditional Value at Risk), and Maximum Drawdown are all risk measures commonly used in finance to assess the potential losses that an investment portfolio or trading strategy may incur, but they measure the risk from completely different angles. The former two project the potential losses based on historical return distribution, whereas maximum drawdown computes the realized loss.

VaR is a statistical measure that quantifies the potential loss in value of a portfolio or a particular investment over a specific time horizon and at a specified confidence level. It represents the maximum amount of money that an investor could expect to lose over a given period with a certain probability.

CVaR is an extension of VaR that takes into account the severity of losses beyond the VaR level. Unlike VaR, which only measures the likelihood of a loss, CVaR provides an estimate of the expected size of losses that exceed the VaR threshold (Stan).

For VaR and CVaR, we first run tests to see if they follow Normal Distribution or not (Admin). If they are not following it, we use historical data to calculate it. The formula for historical VaR is just taking the last alpha-th percentile of the historical data, and multiplying it by the portfolio's current value. The formula for historical CVaR is the average of the data that are lower than alpha-th percentile.

If they follow **Normal distribution**:

- $VaR_p = \sqrt{\Phi^{-1}(1-\alpha)-\mu_p}$
- $CVaR_p = \sigma_p \alpha^{-1} \phi(\Phi^{-1}(1-\alpha)) - \mu_p$
- $\mu_p = w^T \mu$

Where $\Phi^{-1}(1-\alpha)$ is the quantile of the standard normal distribution, $1-\alpha$ is the confidence level, μ_p is the expected return of the portfolio, and ϕ is the standard normal density function.

Maximum Drawdown is a measure of an asset's largest price drop from a peak to a trough, before a new peak is attained. Maximum drawdown is an indicator of downside risk over a specified time period. However, it only measures the size of the largest loss, without taking into consideration the frequency of large losses (Adam Hayes).

The formula for Maximum Drawdown is : $\frac{Trough\ Value\ -Peak\ Value}{Peak\ Value}$

4. Results

We leveraged the optimization techniques previously mentioned and ran the back testing to examine the portfolio performance from **2017-07 to 2023-03**.

We chose to use an expanding window when calculating the weight because this approach will lead to a more stable θ and weights updating, which conforms with the idea of having smooth θ throughout history. Our **initial fitting data set** is from 2010-01 to 2017-06. This means that the first month's weights are computed by using the data from 2010-01 to **2017-06**, and the next month's weights are computed by using the data from 2010-01 to **2017-07**.

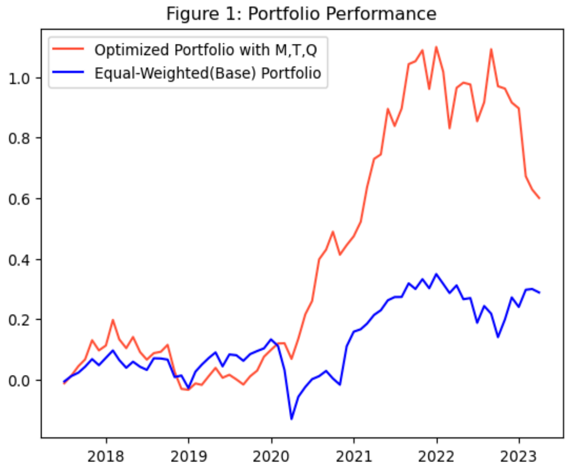
To test which combinations of factors are useful toward portfolio optimization, we have tried to optimize the portfolio with:

1. Full features: Momentum, Trend, Value and Quality
2. Momentum, Trend and Value
3. Momentum, Trend and Quality
4. Trend, Value and Quality

4.1 Return Performance

	Full Features	3 Features Only			
		M,T,V	**M,T,Q**	M,V,Q	T,V,Q
Mean	0.7050%	0.3667%	**0.7781%**	0.5760%	0.4649%
Median	1.0774%	0.9585%	**1.6286%**	1.1682%	0.6163%
Std Dev	4.6650%	3.8600%	**4.5956%**	4.6920%	4.2392%
Risk-adjusted Return	0.15112%	0.095%	**0.16931%**	0.12277%	0.10967%
Total Return(Annualized)	52.02%(7.44%)	22.73% (3.57%)	**60.03%(8.39%)**	38.54% (5.75%)	30.02% (4.60%)

From the above table, we see that the portfolio optimization with Moment, Trend and Quality performs the best. This suggests that using Value as one of the inputs worsen the performance(in back-testing period). One of the reason could be that Value does not provide any more useful information toward analyzing the joint distribution of asset historical returns when optimizing the θ.



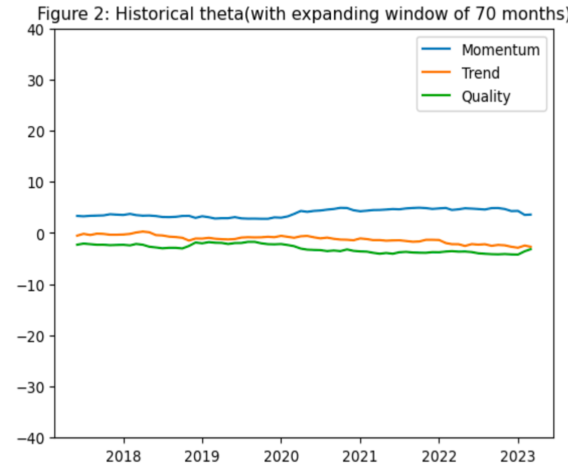
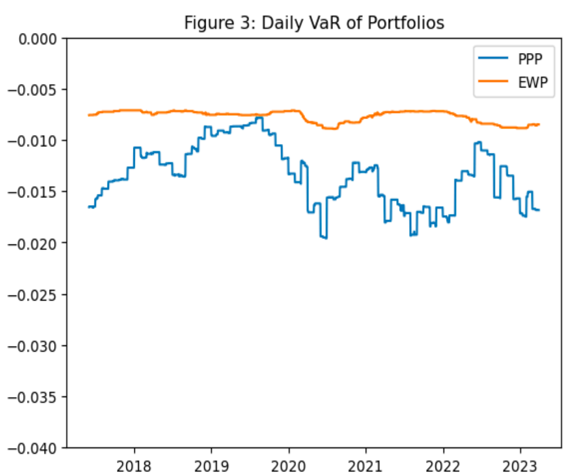


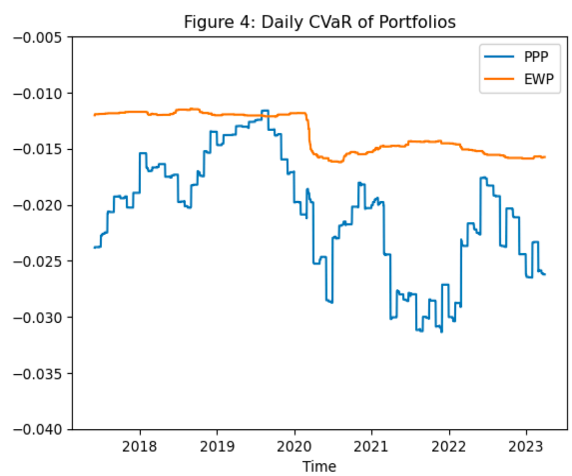
Figure 1 shows the cumulative returns of the equal-weighted portfolio(bench mark) with our optimized portfolio with Momentum, Trend and Quality as factors. We can see, because of these additional long-short positions for the factors, our optimized portfolio has a promising return.

Also Figure 2 shows the θ for those factors, and it is very stable starting from the beginning. This is because we included a long history of data to initialize the expanding window. If we shorten that initial period length, we will observe the phenomenon that the early θ has a huge variation, then slowly converges.

4.2 VaR, CVaR, and Maximum Drawdown

We used a 5-year rolling window to calculate the VaR and CVaR on a daily basis. Below is the graph for the daily VaR and CVaR for those two portfolios.

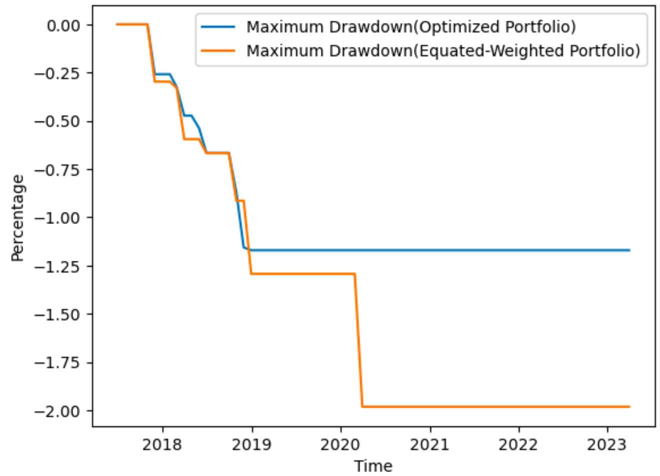




As we can see from Figure 3 and 4, the optimized portfolio has a larger VaR and CVaR, which means the optimized portfolio endures a higher potential loss risk. This is as expected since we receive the extra return generated from the long-short position for the factors, and need to bear that proportion of risk for our portfolio.

The VaR and CVaR estimates the potential losses for the portfolio, and having a larger VaR and CVaR for the optimized portfolio, compared to the equal weighted, doesn’t mean our realized loss is also larger. Let’s look at the maximum drawdown:

Figure 5: Portfolios Maximum Drawdown



As we can see in Figure 5, the optimized portfolio holding return has a significantly lower maximum drawdown, especially after 2019. This is because our optimized portfolio did not suffer the big losses due to the start of global pandemic, thanks to our long-short positions.

5. Conclusion:

Our results suggest that our optimized portfolio through parametric portfolio policies outperforms the equally weighted portfolio depending on trading patterns expressed through the assets characteristics. Newton's method with the BFGS extension is a computationally efficient gradient based optimization approach for maximizing the expected return of the CRRA utility function. The best portfolio was long momentum and short quality and value (M,T,Q), achieving an annual return of 8.39% compared to the equally weighted portfolio which achieved an annual return of 7.44%. This strategy primarily searches for reversals using a shorter term momentum characteristic and term quality and value characteristics. Basing portfolios off of characteristics allows for easy tuning of strategies to achieve profits in desired scenarios. These results show that portfolio optimization through parametric portfolio policies yield robust returns, are highly customizable, and computationally efficient. In turn this method is very practical for large datasets and strategy implementation.

6. Limitations & Future Studies

There were several limitations we faced when defining those factors. For example, the optimal lookback period for momentum strategies may change as equity and bond market behavior evolve, so there are multiple methods to determine the lookback period for momentum, each with its own strengths and weaknesses. After considering different approaches, we decided to use only the previous year's return as the method for calculating momentum throughout the entire time window. Also, when calculating the Value and Quality, we only computed the average for those financial ratios. In the future studies, we could try different approaches to calculating these factors and assess whether overall performance improves.

In addition, we can implement different trading strategies through the chosen asset characteristics. Additionally, we can use other types of portfolios as the benchmark by simply changing the $\bar{w}$ in $w_{i,t} = \bar{w} + \theta^T \hat{x}_{i,t}/N_t$. For example, we can change $\bar{w}$ into $\bar{w}_{i,t}$being the time varying market cap weights. We can also apply this technique to other asset classes to determine how capable this method is on a broad scale. Lastly, instead of using a long expanding window when computing the weights, one can also try to set a limit on the window length and re-initialize the expanding window. In that way, the optimization can fit a more localized regime.

Appendix

Figure 6: Portfolio Performance

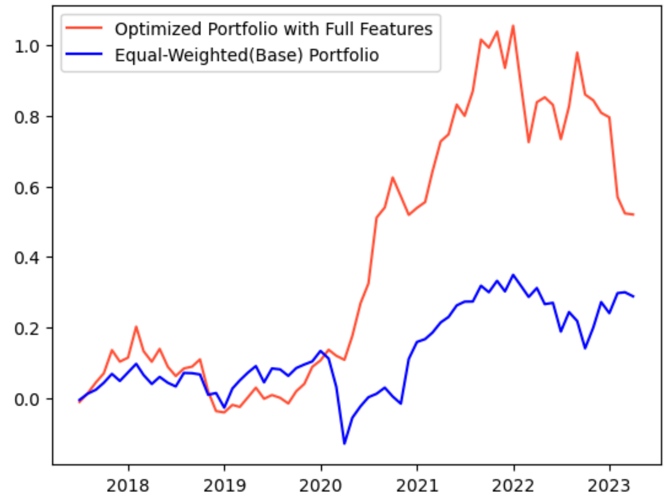


Figure 7: Portfolio Performance

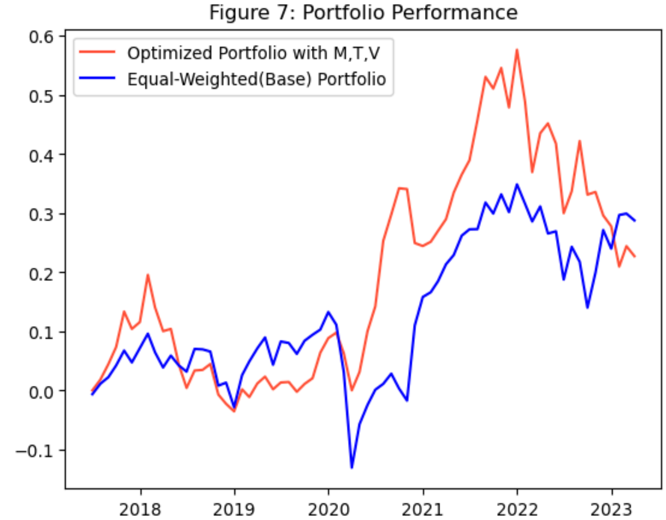


Figure 8: Portfolio Performance

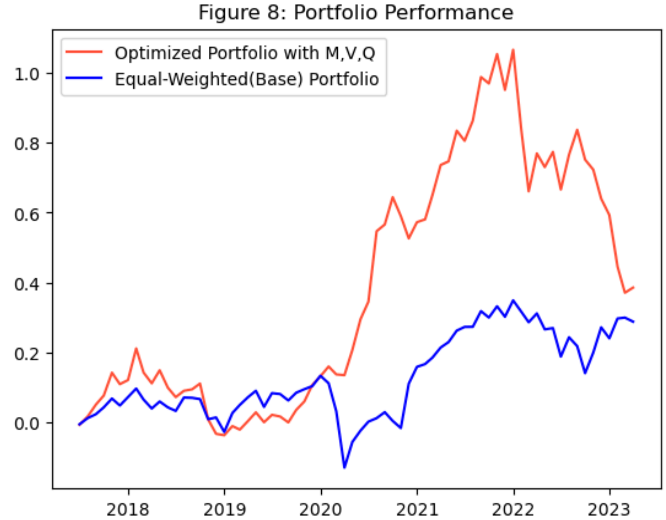


Figure 9: Portfolio Performance

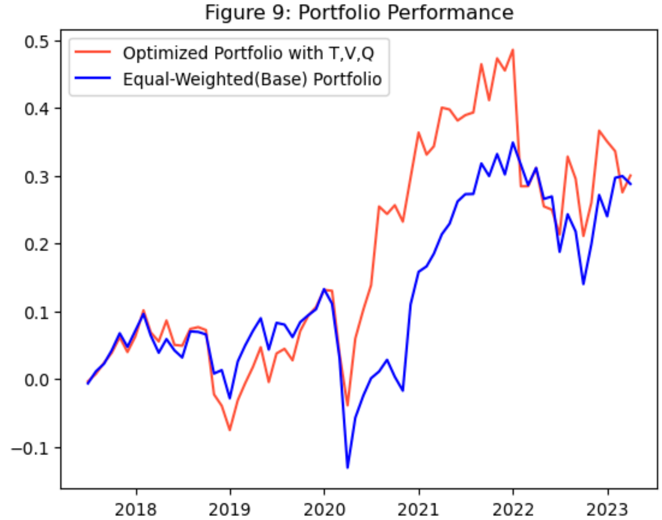


Figure 10: Historical theta(with expanding window of 150 months)

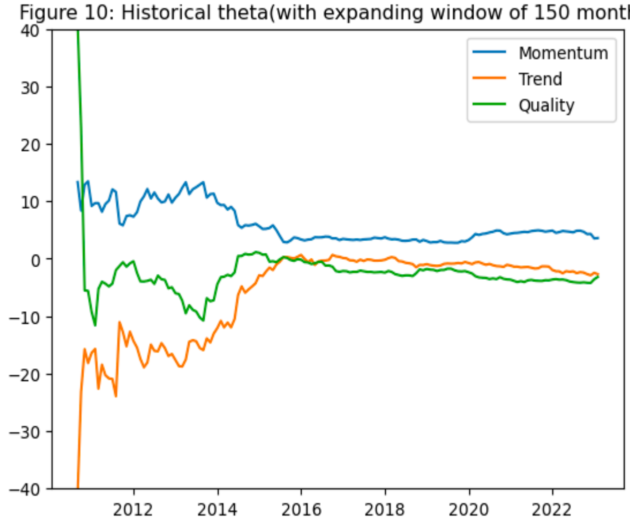


Figure 6~9 shows the backtesting results for optimized portfolios with:

1. full features
2. Momentum and Trend and Value
3. Momentum and Value and Quality
4. Trend and Value and Quality

Figure 10 shows the best optimized portfolio's theta but with a longer expanding window. It shows that the θ converges as time passes. Whereas in the paper, the θ is stable at the beginning because of the longer initialization window.

References:

Brandt, Michael W., et al. "PARAMETRIC PORTFOLIO POLICIES: EXPLOITING CHARACTERISTICS IN THE CROSS-SECTION OF EQUITY RETURNS." *Academic.oup.com*, Dec. 2004, https://academic.oup.com/rfs/article/22/9/3411/1572695.

Stan Uryasev, "VaR vs CVaR in Risk Management and Optimization" https://www.ise.ufl.edu/uryasev/files/2011/11/VaR_vs_CVaR_CARISMA_conference_2010.pdf

Admin, "Value at Risk (VaR) and Conditional VaR (CVaR)" https://quantpy.com.au/risk-management/value-at-risk-var-and-conditional-var-cvar/

MarkD, "How can I calculate the Maximum Drawdown MDD in python" https://quant.stackexchange.com/questions/18094/how-can-i-calculate-the-maximum-drawdown-mdd-in-python

Adam Hayes, "Maximum Drawdown (MDD) Defined, With Formula for Calculation" https://www.investopedia.com/terms/m/maximum-drawdown-mdd.asp#:~:text=What%20Is%20a%20Maximum%20Drawdown,over%20a%20specified%20time%20period.