Factor Models Implementation and Factor Timing

Yi Wang, Yiyang Wen, Zhiwei Zhou

March 2018

Abstract

Factor model is a financial model that employs multiple factors in its computations in order to explain market phenomena and help predict the assets returns. By introducing several factors into the optimization model, we could have a relative more structured model compared to the basic Markowitz mean-variance optimization model. To be more clearly, factor model could lower the parameters we need to estimate so that the model's stability and prediction power can be enhanced.

In this paper, we managed to generate 20 effective factors first and conduct corresponding factor analysis. Then we implemented two approaches to factor model, one is the nontraditional approach: time series factor model, the other one is traditional approach: cross sectional factor mode. Also, we explored the performance of the related strategies under different weighting methods, backtracking windows, and constraints. we used annualized return, Sharpe ratio and maximum drawdown to evaluate our strategy.

Last but not least, we introduced a new regularization term into the factor optimization model in order to time the factors. We mainly utilized information coefficient to help us construct the factor momentum and added factor momentum term in the objective function of Markowitz optimization. That method can improve the portfolio performance under the certain circumstances.

Contents

1 Data Collection

We obtained the fundamental and pricevolume data of all current S&P 500 component stocks we needed on Bloomberg. All of the data are from Jan 2010 to Feb 2018 with daily frequency, which can be summarized as follows:

Raw data	**Definition**
Best sales	Annual sales of the company
Beta raw overridable	A measure of the covariance of a Stock with the market
Capital expend	Funds used by a company to acquire physical assets
Cash to tot asset	Cash asset ratio
Dividends indicated yield	The dividend yield that a share of stock would return based on its current indicated dividend.
Dividends per share	Sum of declared dividends issued by a company for every ordinary share outstanding
EBIT	Earnings before interest and taxes
EBITDA	Earnings before interest, taxes, depreciation and amortization
FCF to total debt	Free cash flow to debt ratio
Free cash flow yield	Free cash flow yield
Geo grow net inc	5 year growth of net income
Is int expense	Interest expense
Market cap	Total dollar market value of a company's outstanding shares
Net debt	(Short-term debt + long-term debt) - cash and cash equivalents
Net income	A company's total earnings (or profit)
Price last	Stock last price
Ret fututre 1m	Forward return of next month
Return com EQY	Return on company equities
Total assets	Total assets of the company
Total assets geo growth	5 year growth of total assets
Volume	Trading Volume

2 Factor Calculation and Analysis

2.1 Factor Calculation

Based on the data we collected, we next calculated 20 factors including profitability factors, growth factors, value factors, dividends factors, momentum factors, size factors, volatility factors, technical factors and so on. The following table shows the factor names, corresponding categories as well as the formulas of these factors:

Factor name	Category	Formula
Net debt 12 month to EBITDA 12 month	Leverage	Net Debt /EBITDA
EBIT 12 month to interest expense 12 month	Leverage	EBIT / Is Int Expense
Capital expenditur to assets 12 month	Profitability	Capital Expend / Total Assets
EPS growth 12 month	Growth	(EPS - EPS 12m) / EPS 12m
Total assets 5 year growth rate	Growth	(Total Assets - Total Assets 5y) / Total Assets 5y
Free cash flow yield 12 month	Value	Free Cash Flow Yield
Dividend indicated yield	Dividends	Dividend Indicated Yield
Dividend to net income average 3 year	Dividends	Average(Dividend / Net Income)
Total return 1 month	Momentum	(Price - Price 1m) / Price 1m
Total return 6 month	Momentum	(Price - Price 6m) / Price 6m
Total return 1 year	Momentum	(Price - Price 1y) / Price 1y
Total assets	Size	Total Assets
Market cap	Size	Market Cap
Net income 12 month	Size	Net Income
Price vs 52 week high	Tech	Price / Max(Price, 52 weeks)
Price to 200 day moving average	Tech	Price / Average(Price, 200)
Volatility 1 Y	Volatility	Volume

2.2 Factor Analysis

2.2.1 Information Coefficient

We used the information coefficient (IC) to assess the predictive power of a factor. The IC of a factor is its Spearman Rank Correlation. To break it down, we calculated the IC between the factor values and the forward returns for each period. The IC assesses the monotonic relationship between factors and returns. What this means, intuitively, is that it provides a measure for whether higher factor values can be associated with higher returns. A higher IC indicates that higher factor values are more closely associated with higher return values (and lower factor values with lower return values). A very negative IC indicates that higher factor values are closely associated with lower return values. An IC of 0 indicates no relationship.

For every trading day, we compute the IC by

$$IC_t = SpearmanRankCorrelation(ForwardReturns_t,\ FactorValues_t)$$

Then we can plot the time sereis of IC for a single factor. (See Figure 2)

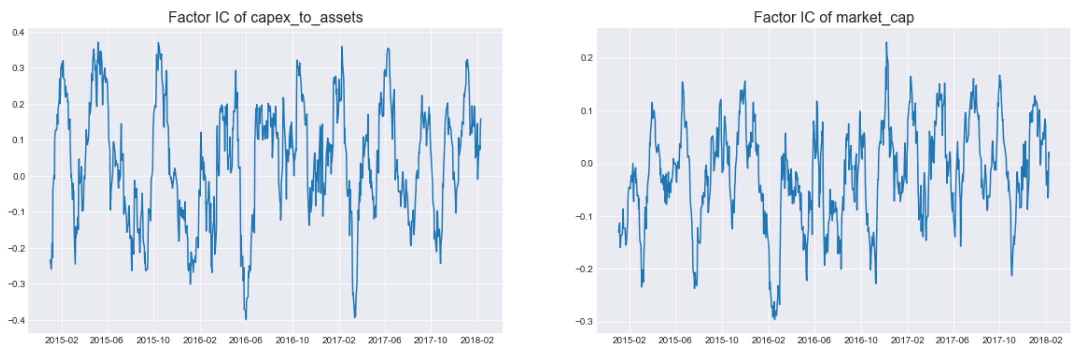

Figure 1: IC Time Series

Also, in the computation of Spearman rank correlation, hypothesis test is embedded in. If p value is less than 0.05, we concluded that the Spearman rank correlation is not equal to zero, which is defined as "*significant*". Across all the trading days, the percentage of the days when a factor is *significant* is defined as "*significant rate*". Taking both into consideration, the product of *significantrate* and *absolute value of IC* is a reasonable metric to access the overall predictive power of a factor.

The IC summary for part of factors are as follows:

	Factor	IC mean	IC significance rate	$\|mean\| * sig\_rate$
0	dividend indicated yield	-0.033567	0.517879	0.017384
1	eps to price	0.001973	0.526510	0.001039
2	ebit 12m to interest expense 12m	0.057614	0.456227	0.026285
3	fcf yield 12m	0.033794	0.627620	0.021210
4	volatility 1m	0.000829	0.289766	0.000240
5	roe avg 3y	0.0086171	0.344020	0.002965

3 Models

3.1 Markowitz Portfolio Theory

In our paper, we are using the Markowitz portfolio model to find the optimal solution for our portfolio. The model formula is given below.

$$\begin{aligned} &max_x \ \mu^T x - \frac{1}{2}\gamma x^T V x \\ &s.t \quad e^T x = 1 \\ &\qquad x_i > 0 \quad for \ i = 1, 2, \cdots, n \end{aligned}$$

The constrains here mean we want long only portfolio

δ: risk aversion parameter
x: weight of the asset, our decision variable
μ: mean
V: covariances
The unknown μ and V can be calculated by the factor models listed below.

3.2 Time Series Factor Model

3.2.1 Model Overview

The goal of the time series factor model is to predict asset returns and covariances from the factors:

$$r = \alpha + Bf + \epsilon$$

f: factor returns (obtained by constructing a long-short portfolio)
B: matrix of factors loading (obtained by regression across time)
α: a constant
ϵ: residual
r: return for asset
Under the assumptions that $E(\epsilon_i) = 0$, we can get the formula for mean and covariance:

$$\mu = \alpha + BE(f)$$
$$V = BFB^T + \Delta$$

Where F is the covariance matrix of f and Δ is the diagonal matrix whose (i, i) entry is var(ϵ_i)
In our time series factor models, f is observable by our factor data and B is unknown.

- Step 1: Construct a long-short portfolio based on the ranking of the factor value. For each day and each factor, we long the top 20% of the stocks with high factor value and short bottom 20% of the stocks with low factor value, in a equal weight manner. Then, the forward return of that portfolio represents the factor return $f_j(t)$.

  Next, we did moving average on $f_j(t)$ across 5 days and adopted that as our finalized factor return.

$$f_j(t) = \frac{1}{5}\sum_{i=0}^{4} f_j(t-i)$$

- Step 2: We used Weighted Least Squares to obtain B matrix. The formula is given below:

$$min_{\beta_{i,1}\ldots\beta_{i,m}} \sum_{t=1}^{T} w(t)(r_i(t) - \sum_{j=1}^{m} \beta_{i,j} f_j(t))^2$$

  where $w(t)$ is the coefficient of exponential decay. After doing this for assets i, we can get B=$\{\beta_{ij}\}$

- Step 3: We computed α_i by using the formula

$$\alpha_i = \sum_{t=1}^{T} w(t)(r_i(t) - \sum_{j=1}^{m} \beta_{i,j} f_j(t))$$

- Step 4: $\epsilon_i(t) = r_i(t) - \alpha_i - \sum_{j=1}^{m} \beta_{i,j} f_j(t)$
- Step 5: Computed Δ: the ith diagonal entry of Δ=$\sum_{t=1}^{T} w(t)\epsilon_i(t)^2$
- Step 6: $F = \sum_{t=1}^{T} w(t)(f(t) - \sum_{s=1}^{T} w(s)f(s))(f(t) - \sum_{s=1}^{T} w(s)f(s))^T$
- Step 7: We plugged in all the data we generate above to obtain μ and V, then plugged them in Markowitz Portfolio Theory to obtain the optimal portofolio weight.

3.2.2 Backtest

We backtested the portofolio based on the optimal weight we obtained from above, but we added a diversification constraint here, now the problem becomes

$$\begin{aligned} max_x \ & \mu^T x - \frac{1}{2}\gamma x^T V x \\ s.t \quad & e^T x = 1 \\ & x_i > 0 \quad for \ i = 1, 2, \cdots, n \\ & max(\{x_i\}) \leq 0.1 \end{aligned}$$

- Backtest setting:

Universe	S&P 500
Rebalancing Frequency	20 days
Benchmark	S&P 500 index
Backtest period	12/22/2015 to 3/15/2018

- PnL Curve:

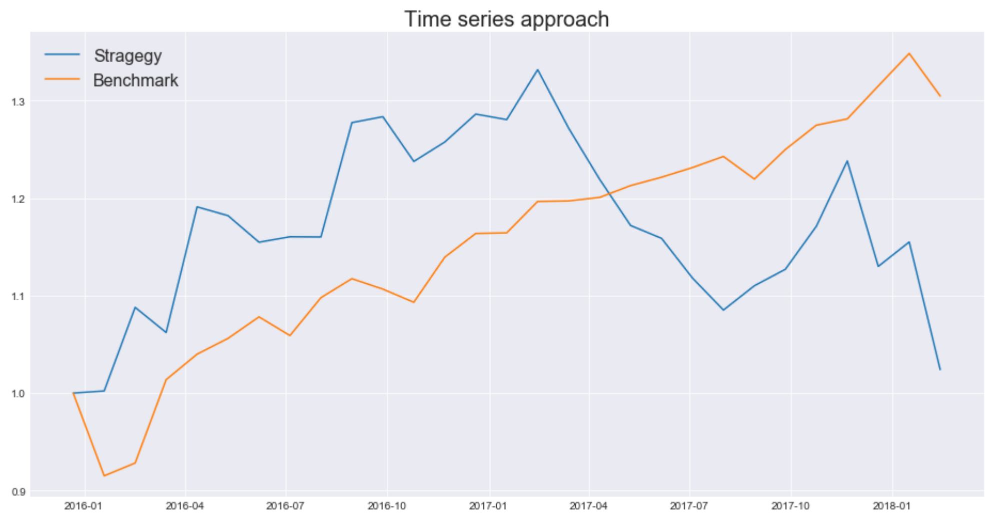

Figure 2: Gamma = 10

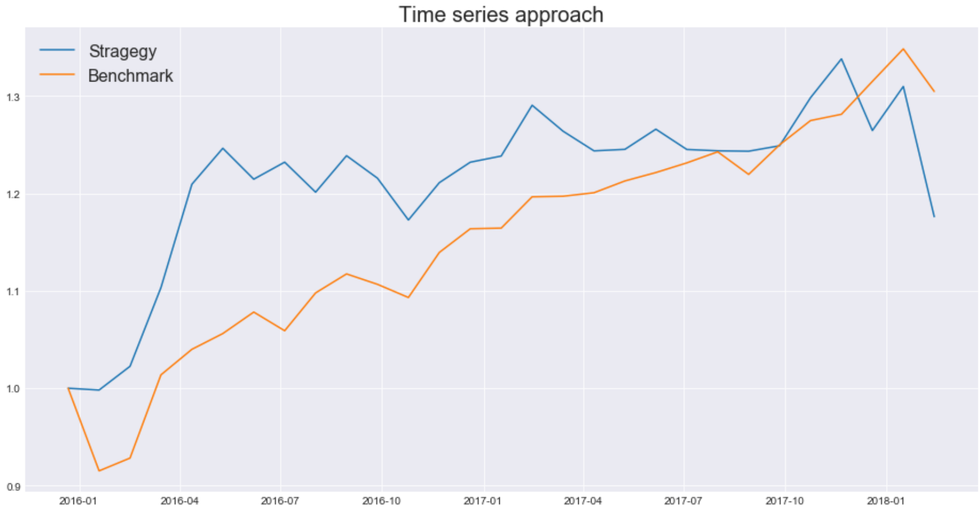

Figure 3: Gamma = 100

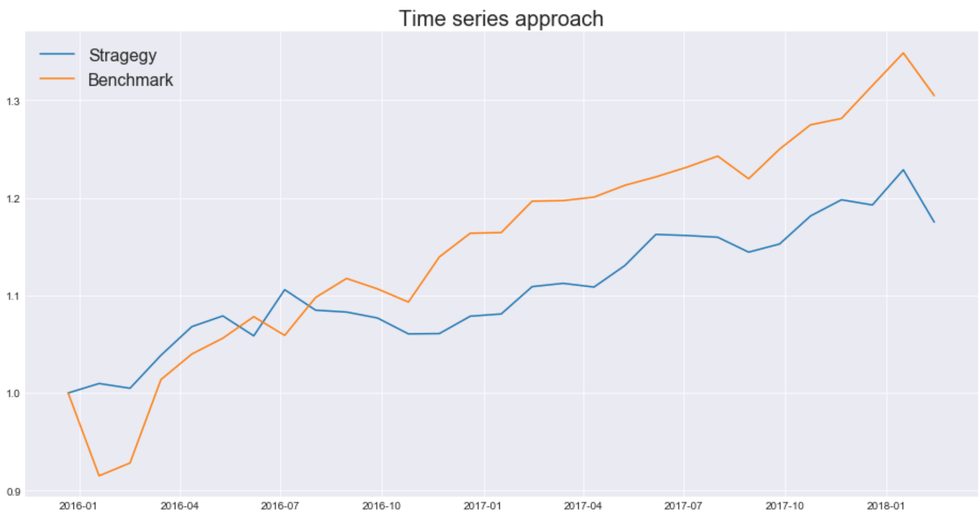

Figure 4: Gamma = 500

- Performance Measurements

	$\gamma = 10$	$\gamma = 100$	$\gamma = 500$
Annualized Return	0.010	0.075	0.075
Sharpe Ratio	0.124	1.027	1.32
Max Drawdown	0.231	0.121	0.043

From above, we perceived that as γ grew, we focused more on risk part of the objective function. That leads to lower max drawdown and higher sharpe ratio.

3.3 Cross Section Factor Model

3.3.1 Model Introduction

In this model, we are using the same factor model formula as above. However, the only difference is that in this model B is observable and f is unknown.

- Step 1: We can find f by doing the following optimization problem.

$$min_{f_1(t)...f_m(t)} \sum_{i=1}^{n} (r_i(t) - \sum_{j=1}^{m} \beta_{i,j}(t) f_j(t))^2$$

- Step 2: E(f)=$\sum_{t=1}^{T} w(t) f(t)$

- Step 3: To calculate B, we can choose B=$\sum_{t=1} Tw(t)B(t)$
- Step 4: now we can follow the steps in the time series factor model to compute α_i, $\epsilon_i(t)$, Δ and F

After achieving all the variables, we are able to calculate μ and V that are needed for the Markowitz Portfolio.

3.3.2 Model Implementation

In the initial cross-sectional model, the weighted least square rule is decided by us. In this paper, we first used the simple moving average method, which can be described as:

$$w(t) = \begin{cases} \frac{1}{n} & t \geq T - n \\ 0 & otherwise \end{cases}$$

Where n is the backtracking window we use to average variables such as factor returns, residuals, alphas, etc.

Then we run several backtests for different parameters under cross-sectional models. For all left backtests in this section, we used the same backtest setting as follows.

Universe	S&P 500
Rebalancing Frequency	20 days
Benchmark	S&P 500 index
Backtest period	10/22/2015 to 2/09/2018

3.3.2.1 A Backtracking Window of 5

When backtracking window is 5, the capital curve plots and performance measurements for different γ are:

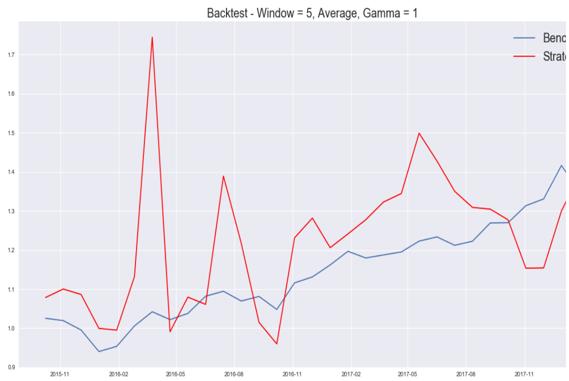

Figure 5: Window = 5, Gamma = 1

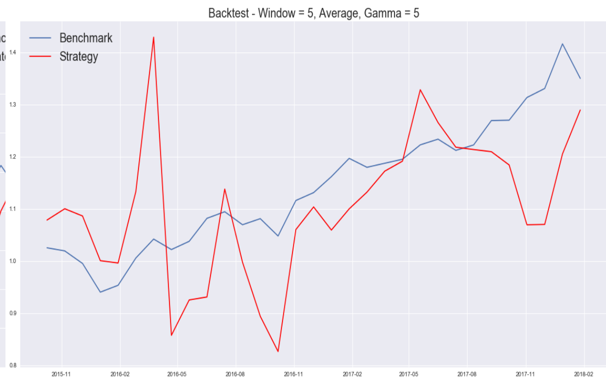

Figure 6: Window = 5, Gamma = 5

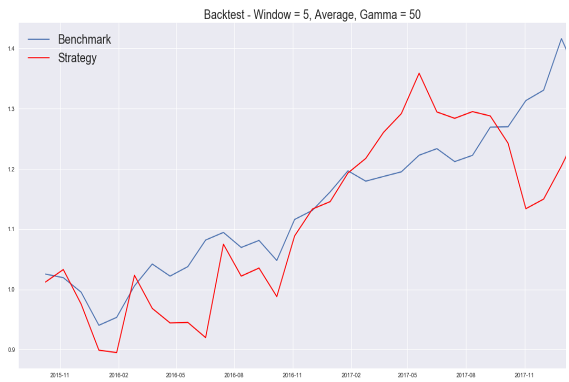

Figure 9: Window = 5, Gamma = 50

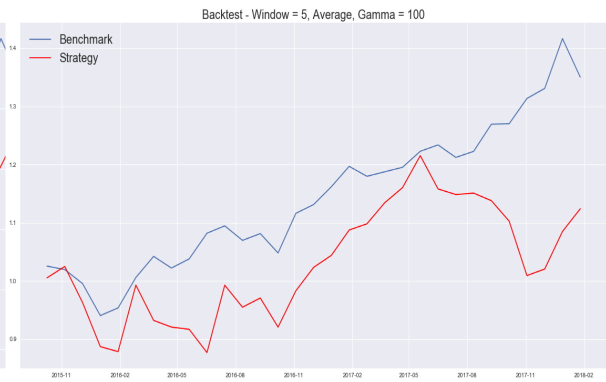

Figure 10: Window = 5, Gamma = 100

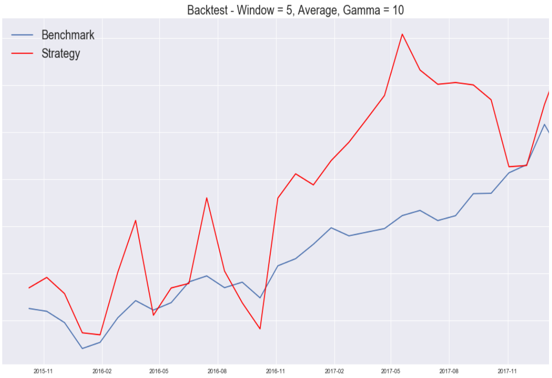

Figure 7: Window = 5, Gamma = 10

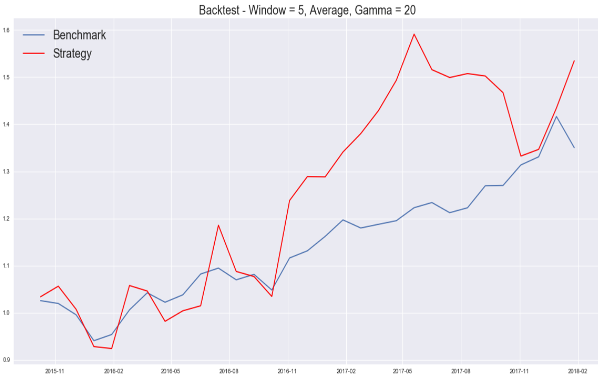

Figure 8: Window = 5, Gamma = 20

	$\gamma = 1$	$\gamma = 5$	$\gamma = 10$	$\gamma = 20$	$\gamma = 50$	$\gamma = 100$
Annualized Return	0.136	0.103	0.188	0.180	0.095	0.046
Sharpe Ratio	0.797	0.763	0.942	0.847	0.688	0.489
Max Drawdown	0.449	0.421	0.220	0.162	0.165	0.169

We found that the best gamma for backtracking window 50 is around 10 or 20.

3.3.2.2 A Backtracking Window of 50

When backtracking window is 50, the capital curve plots and performance measurements for different γ are:

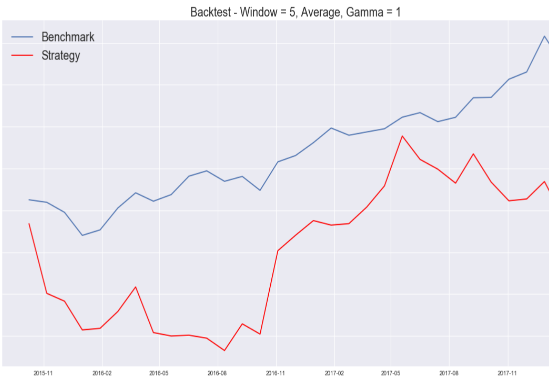

Figure 11: Window = 50, Gamma = 1

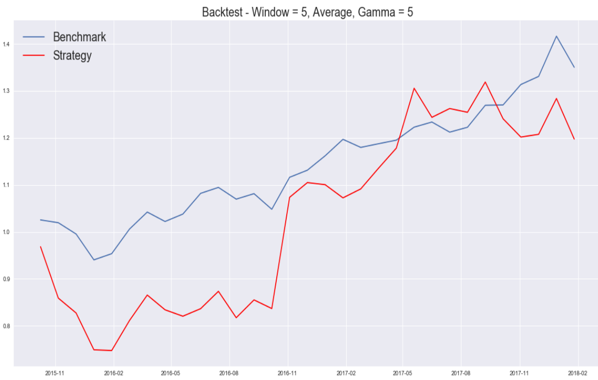

Figure 12: Window = 50, Gamma = 5

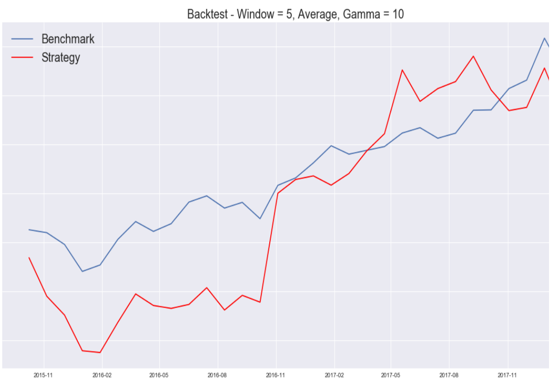

Figure 13: Window = 50, Gamma = 10

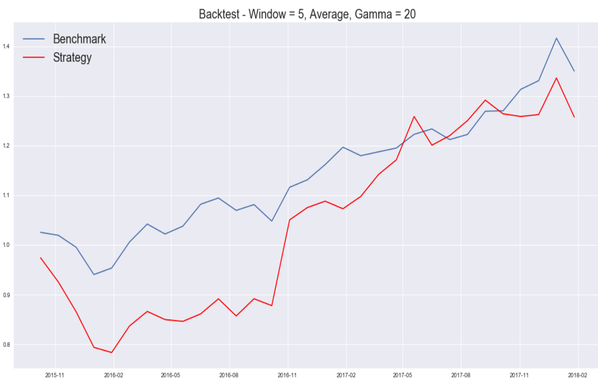

Figure 14: Window = 50, Gamma = 20

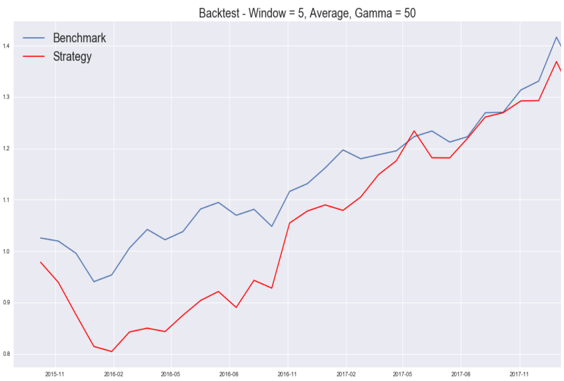

Figure 15: Window = 50, Gamma = 50

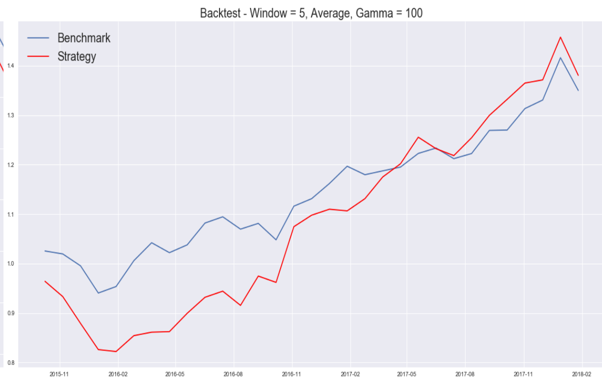

Figure 16: Window = 50, Gamma = 100

	$\gamma = 1$	$\gamma = 5$	$\gamma = 10$	$\gamma = 20$	$\gamma = 50$	$\gamma = 100$
Annualized Return	-0.004	0.072	0.096	0.093	0.106	0.133
Sharpe Ratio	-0.030	0.379	0.474	0.520	0.620	0.670
Max Drawdown	0.313	0.228	0.200	0.195	0.178	0.147

We found that the best gamma for backtracking window 5 is larger than 100.

3.3.2.3 Multiple Backtracking Windows

According to the backtest results of previous strategies, we found when window = 50, the standard deviation of strategy tends to be lower and when window = 5, the total return tends to be higher. So we guessed that maybe the estimation of covariance matrix needs larger backtracking window while the estimation of factor returns needs smaller window.

To prove our assumption, we used a backtracking window of 5 to average factor returns and a backtracking window of 50 to average covariance matrix.

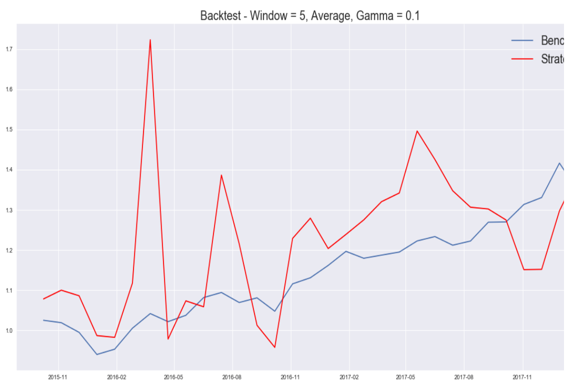

Figure 17: Window = 5&50, Gamma = 0.1

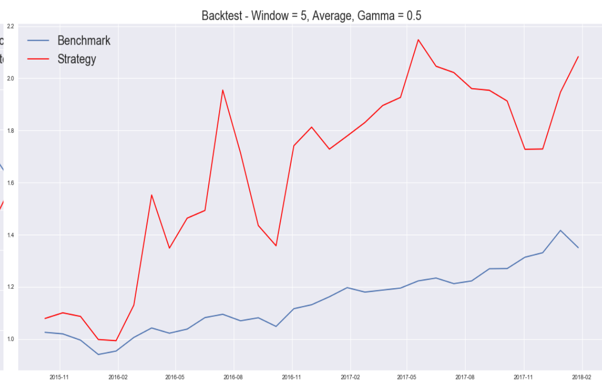

Figure 18: Window = 5&50, Gamma = 0.5

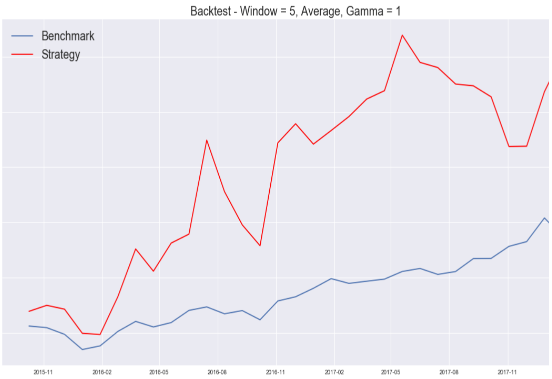

Figure 19: Window = 5&50, Gamma = 1

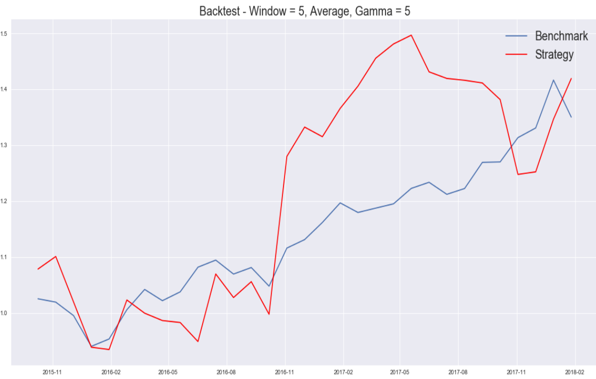

Figure 20: Window = 5&50, Gamma = 5

	$\gamma = 0.1$	$\gamma = 0.5$	$\gamma = 1$	$\gamma = 5$
Annualized Return	0.135	0.328	0.308	0.145
Sharpe Ratio	0.795	0.941	0.927	0.748
Max Drawdown	0.444	0.305	0.224	0.166

The above curve plots confirms our guesses and obviously when $\gamma = 0.5$ or $\gamma = 1$, the 2 year's return of strategy is even over 100%, which is really profitable.

3.3.2.4 Exponential Moving Average

Beyond only using simple moving average method, we also explored other moving average approaches like EMA in this project. EMA (An exponential moving average) is a type of moving average that is similar to a simple moving average, except that more weight is given to the latest data. The formula of EMA is as follows:

$$EMA_t = \frac{f_t + (1-\alpha)f_{t-1} + (1-\alpha)^2 f_{t-2} + \cdots + (1-\alpha)^3 f_{t-n+1}}{1 + (1-\alpha) + (1-\alpha)^2 + \cdots + (1-\alpha)^{n-1}}$$

Where f is the variable we hope to estimate by exponential moving averaging, n is the backtracking window and α is the decay rate. In this project we take n = 20, $\alpha = 0.5$. The performances of new strategies are as follows:

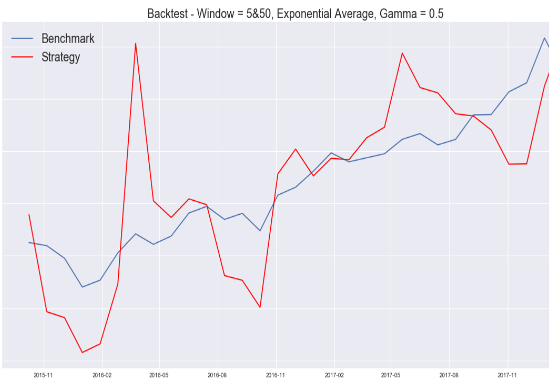

Figure 21: EMA, Gamma = 0.5

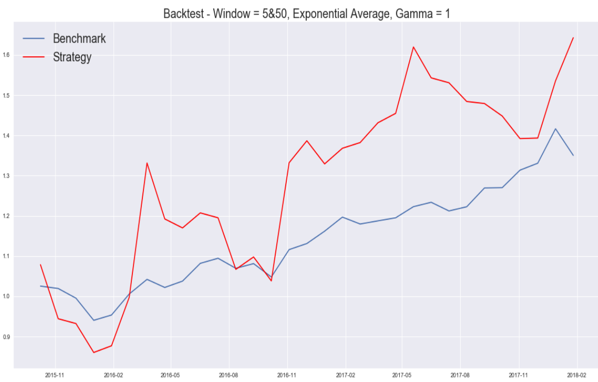

Figure 22: EMA, Gamma = 1

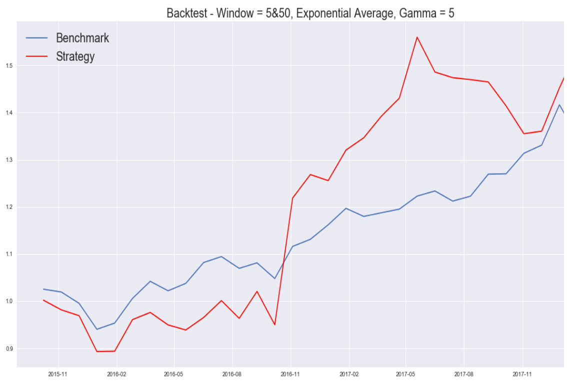

Figure 23: EMA, Gamma = 5

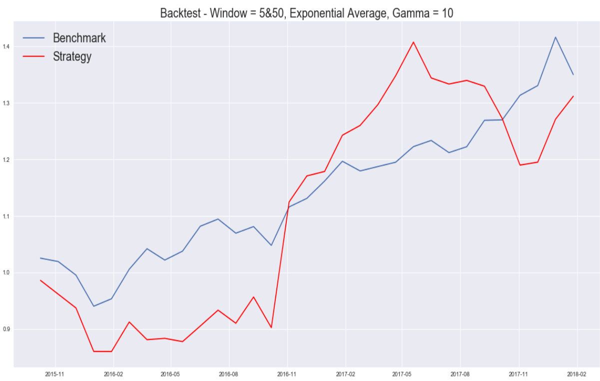

Figure 24: EMA, Gamma = 10

	$\gamma = 5$	$\gamma = 1$	$\gamma = 0.5$	$\gamma = 10$
Annualized Return	0.179	0.212	0.144	0.111
Sharpe Ratio	0.781	0.953	0.857	0.586
Max Drawdown	0.131	0.220	0.358	0.154

It seems that compared to simple moving average method, EMA doesn't enhance the performance of strategy significantly. It may because for α_i, ϵ_i or other indicators, simple moving average seems to a more stable estimation method.

3.3.2.5 With Diversification Constraint

When solving the above optimization problems, we found sometimes only 3 or 5 stocks were assigned large weights, which is not beneficial to diversification. Thus we added a new constraint into the Markowitz optimization problem, which is as follows:

$$\begin{aligned} & max_x \ \mu^T x - \frac{1}{2}\gamma x^T V x \\ s.t \quad & e^T x = 1 \\ & max(\{x_i\}) \leq 0.04 \\ & x_i > 0 \quad for \ i = 1, 2, \cdots, n \end{aligned}$$

This constraint that the maximum value of elements in the weight vector is not larger than 0.04 suggests there will be at least 25 stocks having large contribution to the final portfolio.

Then the capital curve plots of new strategy under different γ are as follows:

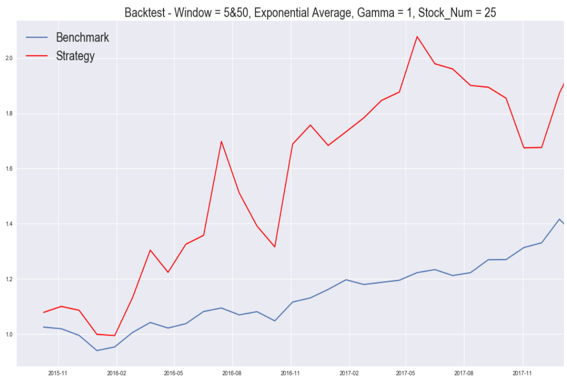

Figure 25: Constraint, Gamma = 1

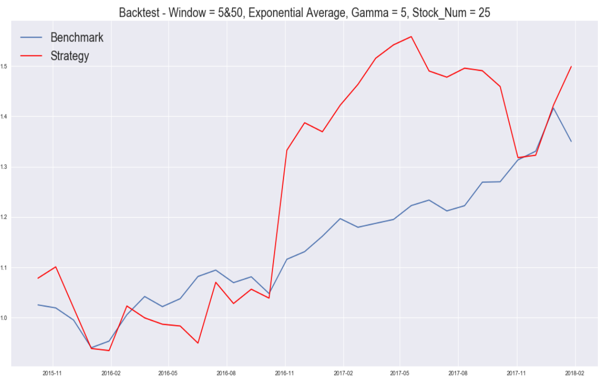

Figure 26: Constraint, Gamma = 5

	$\gamma = 1$	$\gamma = 5$
Annualized Return	0.318	0.180
Sharpe Ratio	0.997	0.789
Max Drawdown	0.204	0.134

It seems that after adding diversification constraint, the new strategy performs better than the previous strategy.

3.3.3 Conclusion

From the backtest results above all, we know that simple moving average method outperforms exponential moving average method and adding diversification constraints could enhance the performance of strategy. In addition, the optimal backtracking window for averaging factor returns tends to be lower than that for averaging other variables such as α_i, ϵ_i, etc. For different averaging methods, backtracking window or constraints, usually the optimal γ is different.

The performance measurements for all parameters and methods are as follows:

4 Factor Timing

4.1 Why Factor Timing

The notion that certain stock characteristics or stock factors drive stock returns has been proved in the above sections and we also have constructed some effective strategies based on basic factor models. However, we still might find other approaches to enhance our model. In this paper, we mainly explore the field of factor timing.

Actually, through our factor analysis in the previous sections, we noticed the highly time-varying property of factor performance. Some factors may perform well in the last 1 year but poorly in the next 3 months. For example, In the Figure-27, we could see that around Fall 2015, the "Volatility 1y" factor outperformed the other two factors, while in November 2016 the "Capex to Assets" factor was the best.

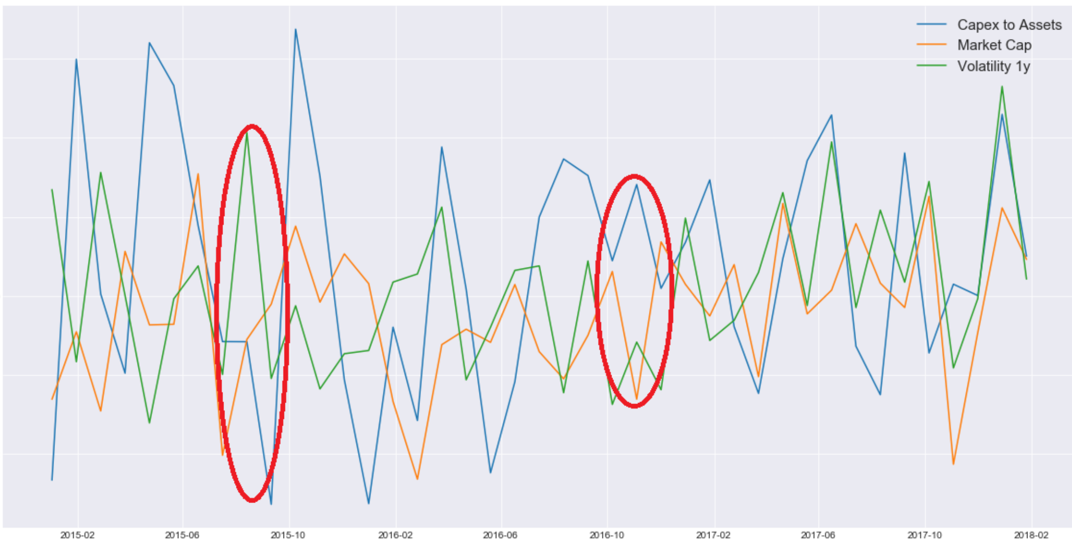

Figure 27: Factor Cyclicality

Obviously, these factors suggest strong cyclicalities in the history. As a result, if we can grasp the pattern of factor cyclicality or factor trend, we can conduct a better estimation for corresponding factor returns and build more profitable portfolio. In this project, we mainly introduces the factor timing into the cross-sectional factor model and use some indicators to describe the factor trends.

4.2 Indicator for Factor Momentum

As we have mentioned, Information Coefficient (IC) is a great indicator to reflect on the performance of factors in the past. And since the values of one factor's IC usually oscillate around its long term mean level, we could describe the factor momentum by the ratio of short-term IC over long-term IC, that is:

$$FactorMomentum = \mid \frac{ShortTermIC}{LongTermIC} \mid -1$$

In this project we chose 1 month as short-term and 6 month as long-term.

Intuitively, the higher the ratio, the better the recent performance of this factor compared to its mean performance in the history.

4.3 Factor Timing Model

The basic assumption of our factor timing model is that if this factor performed well recently it will have a large probability to continually perform well in the next term. It's because that usually it would spend more than one term (one month in thie project) for a factor to change its sign of Factor Momentum.

As a result, we can add m regularizers into the initial cross-sectional factor models to adjust the estimation of factor returns, where m is the number of factors:

$$min_{f_1(t)...f_j(t)} \sum_{t=1}^{T} w(t)\big[r_i(t) - \sum_{j=1}^{m} \beta_{i,j} f_j(t)\big]^2 - \delta \sum_{j=1}^{m} fm_j f_j(t)^2$$

In the cost function, δ is the parameter to control the whole influence of factor mining. fm_j is the factor momentum of factor j. In addition, We use $f_j^2(t)$ instead of f_j because f_j may be nagative and it remains a convex property.

The following optimization and calculation process is the same as the classic cross-sectional factor model which has been described in 4.3.

- step 2: E(f)=$\sum_{t=1}^{T} w(t)f(t)$
- step 3: To calculate B, we can choose B=$\sum_{t=1} Tw(t)B(t)$
- step 4: Now we can follow the steps in the time series factor model to compute α_i, $\epsilon_i(t)$, Δ and F

4.4 Performance of Factor Timing Model

The following two figures and table show that the performance of strategy based on the factor timing model:

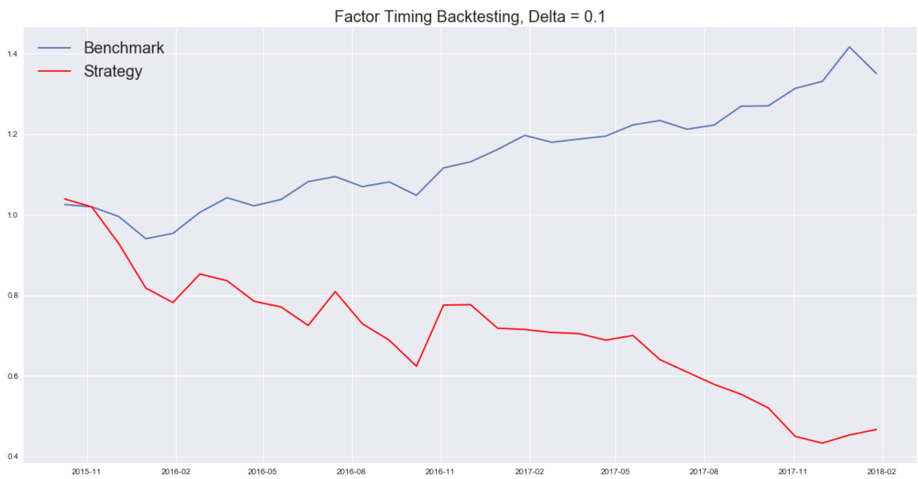

Figure 28: Performance of Factor Timing Model with Delta = 0.1

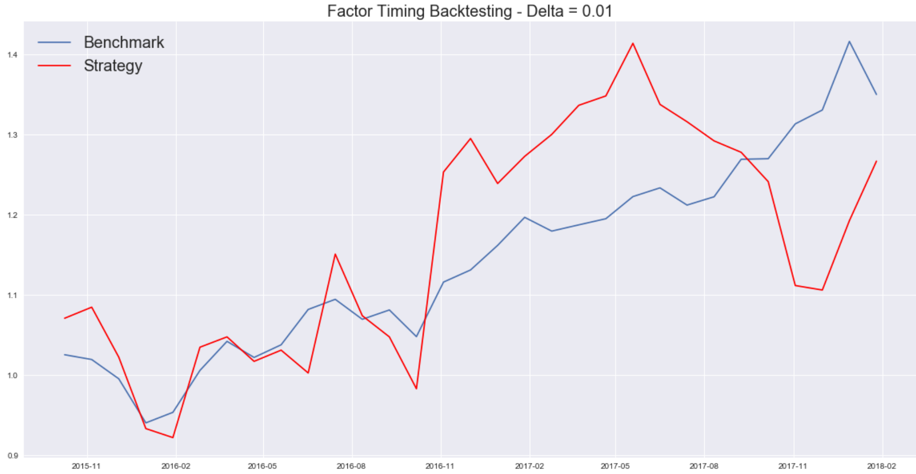

Figure 29: Performance of Factor Timing Model with Delta = 0.01

	$\delta = 0.01$	$\delta = 0.1$
Annualized Return	0.096	-0.255
Sharpe Ratio	0.696	-1.695
Max Drawdown	0.217	0.088

Unfortunately we found that the new model does not outperform the previous model without factor timing. We guess it may because that the factor momentum term change the optimization problem too much. So we tried to use the following improved factor timing model.

4.5 Improved Factor Timing Model

The only difference between this improved factor timing model and initial factor timing model is that we use fm_j^* rather than fm_j in the cost function, which is as following shows:

$$min_{f_1(t)\dots f_j(t)} \sum_{t=1}^{T} w(t) \left[r_i(t) - \sum_{j=1}^{m} \beta_{i,j} f_j(t) \right]^2 - \delta \sum_{j=1}^{m} fm_j^* f_j(t)^2$$

And the formula of fm_j^* is:

$$FactorMomentum^* = \sqrt{|\frac{ShortTermIC}{LongTermIC}|} - 1$$

The goal of this new factor momentum formula is alleviate the influence of factor momentum on the optimization.

4.6 Performance of Improved Factor Timing Model

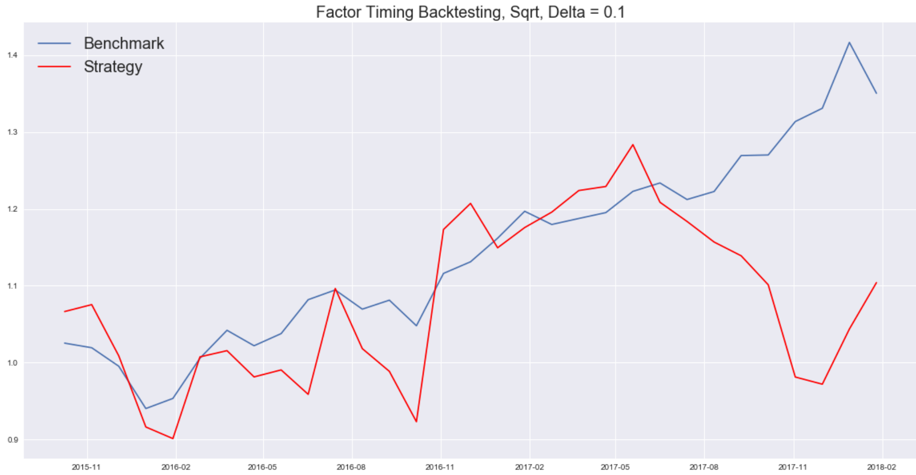

Figure 30: Performance of Improved Factor Timing Model with Delta = 0.1

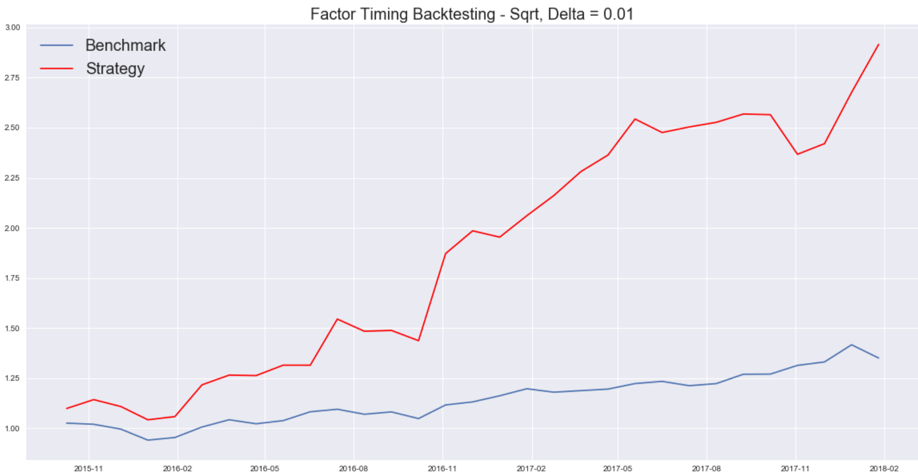

Figure 31: Performance of Improved Factor Timing Model with Delta = 0.01

	$\delta = 0.01$	$\delta = 0.1$
Annualized Return	0.052	0.039
Sharpe Ratio	1.01	0.08
Max Drawdown	0.47	0.242

We found when *delta* equals to some small value such as 0.01, the new strategy performs better than the model without factor timing.

5 Conclusion

In this paper, we implemented time series model as well as cross-sectional model and plug the mean vector μ and covariance matrix v into the Markowitz optimization framework to solve the weights vector for portfolio. After exploring different risk aversion parameter, different moving average techniques as well as different optimization constraints we found that the strategy with simple moving average and diversification constraint have the optimal performance.

In addition, we also tried to design a new regularization term in the objective function to obtain more reliable factor returns. The backtest results suggest it could enhance the performance of strategy.

6 Future Work

The followings are points of improvement for this project and areas that could be researched in more depth.

- Global strategy setting parameters: We can explore different settings such as biweekly rebalancing frequency or larger backtest period to find some useful pattern bewteen the factor model and the equity market.
- Parameter tuning approach: We can apply machine learning to learn the parameters, for example, look-back window and maximum weight.
- Factor exposure: We can set more reasonable constraints such as constraints on factor exposure (i.e. do not expose too much on size factor)
- Factor momentum modification: We can design more robust and reasonable method to define factor momentum rather than only using Information Coefficient.

References

[1] Ric Thomas and Rob Shapiro. "Dynamic Timing of SmartBeta Strategies: Is it Possible?". *IQ Insights*

[2] Cliff Asness."Factor Timing Is Hard". *https://www.aqr.com/Insights/Perspectives/Factor-Timing-is-Hard*.

[3] Hodges, Philip, Ked Hogan, Justin R. Peterson, and Andrew Ang. "Factor Timing with Cross-Sectional and Time-Series Predictors." *The Journal of Portfolio Management 44, no. 1 (2017).*

[4] Ronad Hua,Dimitri Kantsyrev and Edward Qian. "Factor-Timing Model".*http://www.master272.com/finance /memoire_2016/qian_factor_timing_model.pdf*