FE630 Portfolio Theory & Applications Final Project

Yuwen Jin, Minghao Kang, Shiraz Bheda

November 2020

Abstract

In this project, we aim to create a factor-based long/short global macro strategy to produce optimal returns, as well as exploring the sensitivity of strategy to variations of target Beta, look back period on covariance matrix and expected returns.

The main part of this project is composed of 3 steps: calculate the covariance matrix, derive expected return; predict expected returns every week when updating the strategy, regenerate betas of each ETF as well, and use optimizer for the updated weights.

The way we evaluate the performance of a strategy is to plot out its cumulative return curve, and display some statistics describing it with a table comparing to the benchmark, the SPY ETF.

1 Objective

Using a French Fama 3-factor model, build an investment strategy that maximizes portfolio return subject to a constraint of target beta and weekly portfolio rebalancing from March 2007 to November 2020.

In addition, evaluate portfolio sensitivity to variations in both target beta as well as length of estimators for expected returns and covariance matrix under different market scenarios.

2 Data Universe

ETFs

1. CurrencyShares Euro Trust (FXE)
2. iShares MSCI Japan Index (EWJ)
3. SPDR GOLD Trust (GLD)
4. Powershares NASDAQ-100 Trust (QQQ)
5. SPDR SP 500 (SPY)
6. iShares Lehman Short Treasury Bond (SHV)
7. PowerShares DB Agriculture Fund (DBA)
8. United States Oil Fund LP (USO)

9. SPDR SP Biotech (XBI)
10. iShares SP Latin America 40 Index (ILF)
11. iShares MSCI Pacific ex-Japan Index Fund (EPP)
12. SPDR DJ Euro Stoxx 50 (FEZ)

French-Fama History Data

Freely available for download from Ken French's website for the factors historical values:

http://mba.tuck.dartmouth.edu/pages/faculty/ken.french

3 Models

Basically, the main model we are using in this project is Capital Asset Pricing Model(CAPM) and the French Fama(FF) three factor model. Here are the notations:

3.1 CAPM

We have the formula:

$$\rho_i = r_f + \beta_i^M(\rho_M - r_f) + \alpha_i$$

where ρ_i is the expected return of stock i, r_f stands for risk-free return, corner mark M represents the market and α_i is excess return.

For each beta of individual stocks, we get $\beta_i^M = \frac{cov(r_i, r_M)}{\sigma^2(r_M)}$

3.2 Factor model

According to formula:

$$\rho_i = r_f + \beta_i^3(\rho_M - r_f) + b_i^s \rho_{SMB} + b_i^v \rho_{HML} + \alpha_i$$

where ρ_{SMB} stands for the size factor, ρ_{HML} is the value factor and β_i^3, b_i^s, b_i^v are the coefficients to be estimated.

Also we can derive the following formula:

$$\begin{aligned} cov(\rho_i, \rho_j) = {} & \beta_i^3 \beta_j^3 Var(\rho_M - r_f) + b_i^s b_j^s Var(\rho_{SMB}) + b_i^v b_j^v Var(\rho_{HML}) \\ & + (\beta_i^3 b_j^s + \beta_j^3 b_i^s) cov(\rho_M - r_f, \rho_{SMB}) + (\beta_i^3 b_j^v + \beta_j^3 b_i^v) cov(\rho_M - r_f, \rho_{HML}) \\ & + (b_i^s b_j^v + b_j^s b_i^v) cov(\rho_{SMB}, \rho_{HML}) \end{aligned}$$

4 Program structure

The whole program is contained in one python script and the workflow is composed of 3 steps: calculate the covariance matrix(Sigma), derive the array of expected returns(rho) for each asset; predict expected returns every week when updating the strategy, regenerate betas of each ETF as well, and use optimizer for the updated weights. To evaluate the performance of the strategy, we calculate and output daily return of the portfolio based on our weights and the returns of the ETFs. Finally we plot the cumulative returns of our portfolio versus the benchmark, and produce a table of performance metrics.

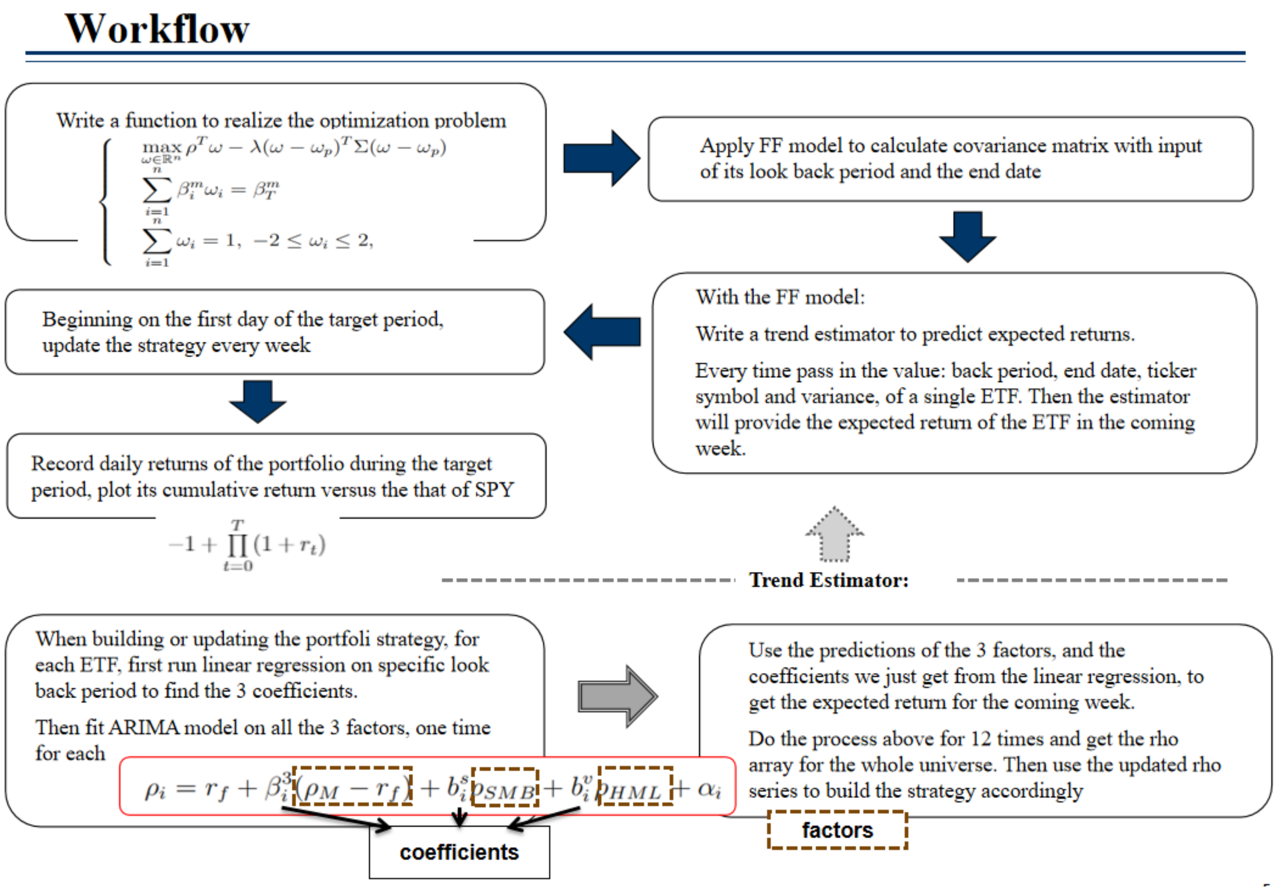


Figure 1: Workflow

One important part in our program is the trend estimator. In order not to use 'future' data to 'predict' future, we implement the following rules:

Every time when we update our strategy, we run the estimator 12 times (one time for each ETF). With the estimator, we first run linear regression on a specific look-back period to find the 3 coefficients of FF model. Then, we fit ARIMA model on the 3 factors individually. We believe this model is appropriate because first factors are time related series. In addition, the ARIMA model allows us for true factor prediction. In addition, we perform a stationary test before fitting the ARIMA model, so we are fine with the ARIMA assumption. Then, after we get predictions for each factors, we use our predictions, together with the coefficients we get, we calculate the expected return for the coming week.

```
P_value of ADF test on HML is 2.3110811537696474e-15
P_value of ADF test on SMB is 0.0
P_value of ADF test on Mkt-RF is 6.608021209349659e-27
```

Figure 2: Stationary test on factors

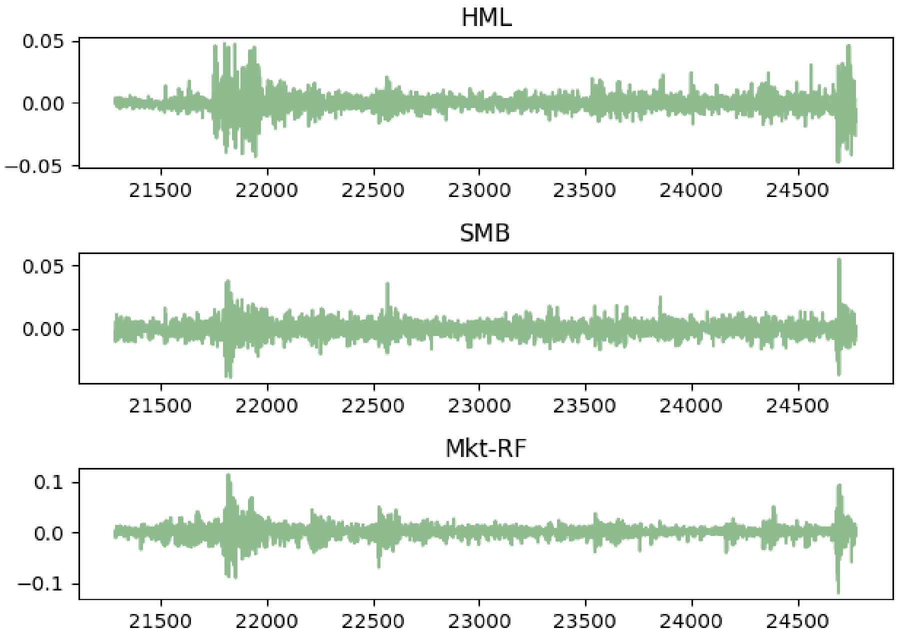


Figure 3: Plots of factors

Moreover, we have included commentary where appropriate in our code, including an introduction of how to use it, where to set up parameters, and also instructions for using defined functions.

5 Result and analysis

Basically we are taking combinations of the 3 variables to design and record the portfolio performance aside from the target period. Target beta describes how much portfolio will change relative to the benchmark; Covariance estimators are the look-back period of Sigma, and Rho estimators are that of Rho. The label for result tables are in format of $S^{rhoperiod}_{covperiod}(\beta_T)$.

By varying the 3 inputs of our portfolio strategies, we can derive the performance metrics, such as cumulative PnL or Returns, Mean returns, Maximum 10-Day Dropdown, and Standard Deviation. In order to find out how each input influences performance, we constructed 9 tables covering 3 separate periods of time: before the sub-prime crisis, during the crisis, and after the crisis. For each of these 9 tables, we changed one input in order to better understand the impact of each input. For example, in our first table, we fix the look back for both covariance matrix and expect return, only change the target beta. In this way, we can find out the portfolio sensitivity to the Target beta.

In order not to make report too long and make it convenient to read, we put most plots together will the statistics data frame versus benchmark in several word documents inside the result folder.

Here are 2 of the return series we generated compared with the SPY return series. We find that with different settings the return variate a lot. Then we going to their statistics.

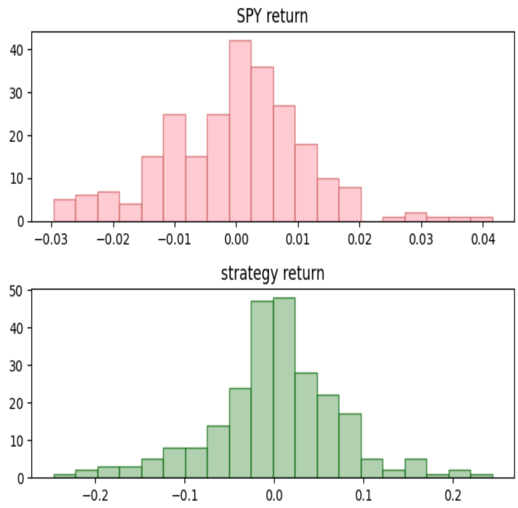


Figure 4: histogram of return series

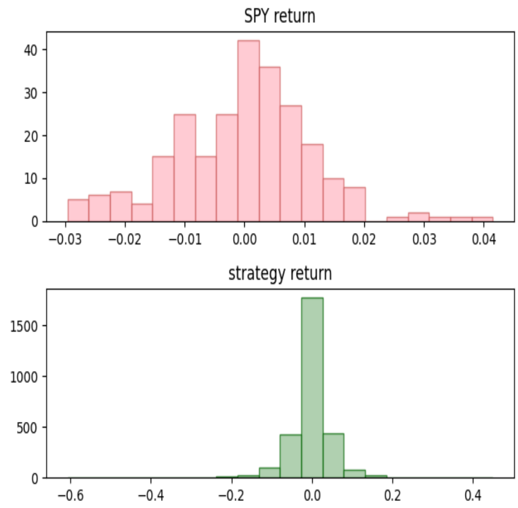


Figure 5: histogram of return series

5.1 Before the crisis: March 01, 2007 - March 31, 2008

5.1.1 Sensitivity on target beta:

	$S_{60}^{90}(-0.5)$	$S_{60}^{90}(0)$	$S_{60}^{90}(0.5)$	$S_{60}^{90}(1)$	SPY
Cum PnL	47.6626	22.455	40.8289	55.2541	-5.69
Mean Return	0.8256	0.659	0.7834	0.8808	-0.0419
Max 10 Days DD	0.6519	1.0557	0.8856	0.9859	0.0889
Standard Deviation	0.9271	0.9506	0.9383	0.9364	0.1839
Sharpe Ratio	0.8492	0.653	0.7941	0.8997	-0.4363
Skewness	0.1952	0.0695	0.4122	0.4087	0.0295
Kurtosis	2.9613	2.0736	2.2943	2.1974	0.7669
VaR	-0.0812	-0.1251	-0.0969	-0.0913	-0.0203

To analyze the result forms, for example, with assuption that look back period of expected return equals to 90 and that period of Σ equals to 60, from the table we can see, we have the lowest PnL when the Target beta is set to zero. When Target beta is set to 1, we have the highest PnL; however, we note that PnL is not the only standard by which we evaluate portfolio strategy.

By plotting performance over time, we notice that PnL depends on the setting of the end date. For example, in portfolio strategy below "6", although we lose money in the end, we see significant returns during the first quarter of 2008. And the portfolio in "7", which has close PnL as Fugure 2 but it's definitely a lot worse than "6". Thus, we realize the importance other gauging strategies by other statistics. We see here, when target beta is equal to 1, it also has the highest mean returns, which mean this strategy is better than others based on mean returns. And also, with the Target beta = 1, we have even lower risk than target beta = 0. Therefore, people can choose parameters based on their risk appetite.

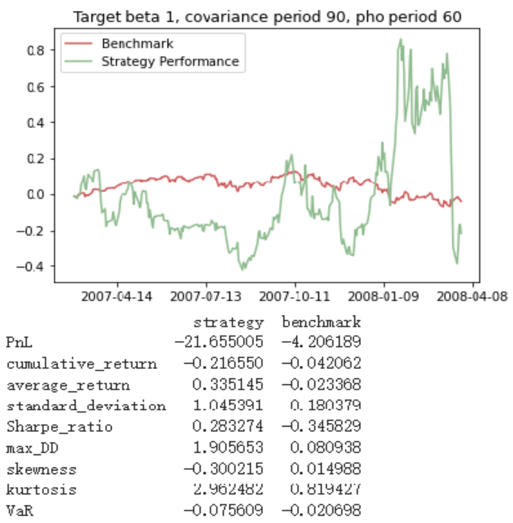


Figure 6: $S_{90}^{60}(1)$

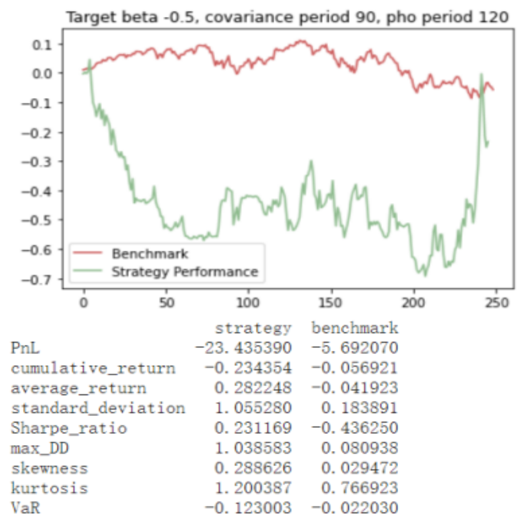


Figure 7: $S_{90}^{120}(-0.5)$

5.1.2 Sensitivity on look back period for covariance matrix :

	$S_{30}^{60}(0.5)$	$S_{60}^{60}(0.5)$	$S_{90}^{60}(0.5)$	$S_{120}^{60}(0.5)$	SPY
Cum PnL	59.5887	-61.8408	32.1388	-16.7347	-5.69
Mean Return	1.0512	-0.343	0.884	0.4339	-0.0419
Max 10 Days DD	0.6537	1.061	0.6663	1.0329	0.0889
Standard Deviation	1.0731	1.1144	1.0943	1.1067	0.1839
Sharpe Ratio	0.944	-0.3422	0.7729	0.3575	-0.4363
Skewness	0.1529	0.1269	0.1873	-0.0269	0.0295
Kurtosis	2.4027	2.4955	2.4519	2.3049	0.7669
VaR	-0.1055	-0.1152	-0.0883	-0.0933	-0.0203

With fixed $\beta = 0.5$ and loof back period for $\rho = 60$

5.1.3 Sensitivity on look back period for expected return :

	$S_{90}^{30}(1)$	$S_{90}^{60}(1)$	$S_{90}^{90}(1)$	$S_{90}^{120}(1)$	$S_{90}^{150}(1)$	SPY
Cum PnL	-1.4673	-21.655	2.5649	15.7183	-25.0802	-5.69
Mean Return	0.5186	0.3351	0.4317	0.5833	0.172	-0.0419
Max 10 Days DD	0.9322	1.9057	0.7768	0.6405	0.6883	0.0889
Standard Deviation	1.0337	1.0454	0.9044	1.9294	0.9627	0.1839
Sharpe Ratio	0.4646	0.2833	0.435	0.5864	0.1388	-0.4363
Skewness	0.2158	0.3002	0.3533	0.0463	0.1584	0.0295
Kurtosis	2.243	2.9625	2.3995	3.5201	3.3089	0.7669
VaR	-0.1039	-0.0756	-0.0858	-0.0797	-0.0729	-0.0203

With fixed $\beta = 1$ and loof back period for $\Sigma = 90$

5.2 During the crisis: April 01, 2008 - December 31, 2008

5.2.1 Sensitivity on target beta:

	$S_{60}^{90}(\beta_T^m = -0.5)$	$S_{60}^{90}(\beta_T^m = 0)$	$S_{60}^{90}(\beta_T^m = 0.5)$	$S_{60}^{90}(\beta_T^m = 1)$	SPY
Cum PnL	-72.7936	-16.0105	13.6613	16.3497	-33.634
Mean Return	-0.8655	0.6402	1.0628	1.0489	-0.4365
Max 10 Days DD	1.3751	1.1447	.7416	.8209	.3324
Standard Deviation	1.3066	1.2998	1.3165	1.28	.4538
Sharpe Ratio	-.6729	.482	.7969	.8089	-.9922
Skewness	-.0487	-.3134	-.281	-.405	.554
Kurtosis	.6593	2.5026	2.1897	1.8322	5.2836
VaR	-.1472	-.1358	-.1202	-.1157	-.0641

With fixed look back period for $\rho = 90$ and that of $\Sigma = 60$

5.2.2 Sensitivity on look back period for covariance matrix :

	$S_{30}^{90}(1)$	$S_{60}^{90}(1)$	$S_{90}^{90}(1)$	$S_{120}^{90}(1)$	$S_{150}^{90}(1)$	SPY
Cum PnL	191.6674	16.3497	-11.1577	71.9459	3.0686	-33.634
Mean Return	2.3114	1.0489	.6368	1.562	0.8082	-.4365
Max 10 Days DD	.8226	.8209	.8159	1.223	1.0964	.3324
Standard Deviation	1.3044	1.2798	1.2388	1.2263	0.6479	.4538
Sharpe Ratio	1.7615	.8089	.503	1.2215	1.2675	-.9922
Skewness	-.4331	-.405	-.4126	-.5817	-0.2055	.554
Kurtosis	1.7074	1.8322	2.3022	1.9342	1.6909	5.2836
VaR	-.0999	-.1157	-.1088	-.1101	-0.135	-.0641

With fixed $\beta = 1$ and loof back period for $\rho = 90$

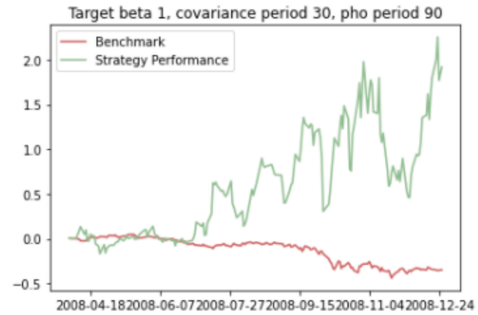
5.2.3 Sensitivity on look back period for expected return :

	$S_{60}^{30}(\beta_T^m = 1)$	$S_{60}^{60}(\beta_T^m = 1)$	$S_{60}^{90}(\beta_T^m = 1)$	$S_{60}^{120}(\beta_T^m = 1)$	SPY
Cum PnL	-96.2156	277.8341	13.6613	-2.0344	-33.634
Mean Return	-2.8985	2.8366	1.0628	.7809	-.4365
Max 10 Days DD	3.1765	.9542	.7416	.6934	.3324
Standard Deviation	1.6171	1.5099	1.3165	1.2732	.4538
Sharpe Ratio	-1.8009	1.8696	.7969	.6026	-.9922
Skewness	-.3743	1.6171	-.281	.1957	.554
Kurtosis	5.355	8.2819	2.1897	.8234	5.2836
VaR	-.1692	-.1042	-.1202	-.1142	-.0641

With fixed $\beta = 1$ and loof back period for $\Sigma = 60$

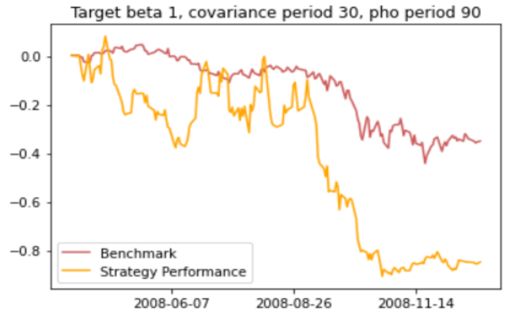
5.3 After the crisis: January 01, 2009 - June 20, 2020

Due to the time consuming problem, for time period of size large to more than 10 years, we simplified the trend estimator. And we compare the previous model's performance and the simplified one on the same setting of inputs and on the same period. Here are 2 groups for comparison, each line containing 2 plots as one group



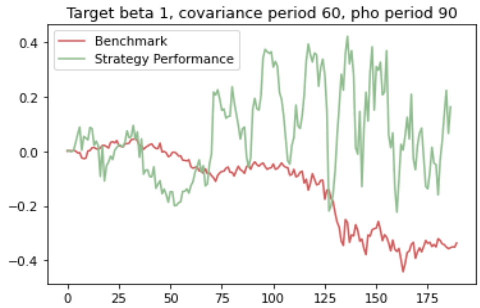
	strategy	benchmark
PnL	191.667372	-33.633977
cumulative_return	1.916674	-0.336340
average_return	2.311409	-0.436517
standard_deviation	1.304373	0.453773
Sharpe_ratio	1.761535	-0.992187
max_DD	0.822584	0.332415
skewness	-0.433124	0.554044
kurtosis	1.707430	5.283598
VaR	-0.099883	-0.064079

Figure 8: $S_{30}^{90}(1)$ Before the crisis



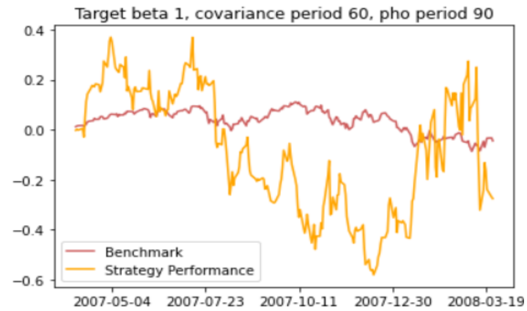
	strategy	benchmark
PnL	-84.776460	-33.633977
cumulative_return	-0.847765	-0.336340
average_return	-1.544782	-0.436517
standard_deviation	1.365872	0.453773
Sharpe_ratio	-1.141023	-0.992187
max_DD	1.537585	0.332415
skewness	-0.111491	0.554044
kurtosis	1.347718	5.283598
VaR	-0.158430	-0.054981

Figure 9: $S_{30}^{90}(1)$ Before the crisis



	strategy	benchmark
PnL	16.349712	-33.633977
cumulative_return	0.163497	-0.336340
average_return	1.048858	-0.436517
standard_deviation	1.279772	0.453773
Sharpe_ratio	0.808853	-0.992187
max_DD	0.820885	0.332415
skewness	-0.404956	0.554044
kurtosis	1.832238	5.283598
VaR	-0.115654	-0.050351

Figure 10: $S_{60}^{90}(1)$ During the crisis



	strategy	benchmark
PnL	-27.541347	-5.692070
cumulative_return	-0.275413	-0.056921
average_return	0.364095	-0.041923
standard_deviation	1.164652	0.183891
Sharpe_ratio	0.279736	-0.436250
max_DD	0.876126	0.080938
skewness	-0.203005	0.029472
kurtosis	1.312156	0.766923
VaR	-0.107607	-0.022176

Figure 11: $S_{30}^{90}(1)$ During the crisis

It turns out our original model which including the ARIMA model in the trend estimation works better. This is one reason that our strategy with this range of target period works much worse than before.

The other reason is that market is changing. Therefore we should actually observe the market and the portfolio performance, then modify parameters accordingly at some time points. In addition, with the simplified estimator, also this the settings we have for now, the strategy fail to modify its action in time.

5.3.1 Sensitivity on target beta:

	$S_{60}^{90}(\beta_T^m = 1)$	$S_{60}^{90}(\beta_T^m = .5)$	$S_{60}^{90}(\beta_T^m = 0)$	$S_{60}^{90}(\beta_T^m = -.5)$	SPY
Cum PnL	-94.2688	-91.2521	-76.3296	-99.2539	319.1791
Mean Return	.2114	.2276	.3072	.0152	0.1408
Max 10 Days DD	3.0209	2.9938	2.9046	2.93	0.3044
Standard Deviation	.949	.9334	.9272	.936	0.1819
Sharpe Ratio	.2176	.2386	.326	.0111	0.7475
Skewness	.1273	.2961	.2565	.3282	-0.4369
Kurtosis	8.707	7.898	6.975	6.8969	11.3844
VaR	-0.1003	-0.0895	-0.0976	-0.0975	-0.0175

5.3.2 Sensitivity on look back period for covariance matrix :

	$S_{30}^{90}(\beta_T^m = 1)$	$S_{60}^{90}(\beta_T^m = 1)$	$S_{90}^{90}(\beta_T^m = 1)$	$S_{120}^{90}(\beta_T^m = 1)$	SPY
Cum PnL	-81.8611	-83.4528	-84.7908	-93.809	319.1791
Mean Return	0.3077	0.3077	0.3097	0.2205	0.1408
Max 10 Days DD	3.4293	2.9941	2.9431	2.9741	0.3044
Standard Deviation	0.9458	0.948	0.9628	0.9548	0.1819
Sharpe Ratio	0.3201	0.3121	0.3166	0.2259	0.7475
Skewness	0.1381	0.1852	0.0944	0.2319	-0.4369
Kurtosis	8.8317	8.623	8.4089	8.0916	11.3844
VaR	-0.0911	-0.0893	-0.0969	-0.0916	-0.0175

5.3.3 Sensitivity on look back period for expected return :

	$S_{60}^{30}(\beta_T^m = 1)$	$S_{60}^{60}(\beta_T^m = 1)$	$_S60^{90}(\beta_T^m = 1)$	$S_{60}^{120}(\beta_T^m = 1)$	SPY
Cum PnL	222.5426	-97.5221	-92.1117	-96.8245	319.1791
Mean Return	.5306	.115	.2409	.1719	0.1408
Max 10 Days DD	2.5846	2.9204	3.2157	2.9712	.3044
Standard Deviation	.9296	.9285	.9531	.9612	.1819
Sharpe Ratio	.5655	.1186	.2476	.1737	0.7475
Skewness	.4214	.2796	.1672	.1247	-0.4369
Kurtosis	4.9603	8.3111	8.002	8.178	11.3844
VaR	-.0912	-.0988	-.0956	-.0959	-0.0175

5.4 Whole period: March 01, 2007 - June 20, 2020

5.4.1 Sensitivity on target beta:

	$S_{60}^{90}(\beta_T^m = 1)$	$S_{60}^{90}(\beta_T^m = .5)$	$S_{60}^{90}(\beta_T^m = 0)$	$S_{60}^{90}(\beta_T^m = -.5)$	SPY
Cum PnL	-61.8745	53.1656	-99.9356	-97.1655	184.754
Mean Return	.2105	.3427	.0412	.0283	.1003
Max 10 Days DD	1.5771	1.4823	3.1943	1.5533	.3324
Standard Deviation	.7563	.7891	1.0833	.7949	.2079
Sharpe Ratio	.2711	.4274	.0308	.0288	.4443
Skewness	-.0844	.6122	.1257	.128	-.0302
Kurtosi s	4.3491	7.338	4.4562	6.0577	15.0612
VaR	-.0769	-.0748	-.1077	-.0749	-.0179

5.4.2 Sensitivity on look back period for covariance matrix :

	$S_{30}^{90}(\beta_T^m = 1)$	$S_{60}^{90}(\beta_T^m = 1)$	$S_{90}^{90}(\beta_T^m = 1)$	$S_{120}^{90}(\beta_T^m = 1)$	SPY
Cum PnL	-99.8726	-99.9788	-98.5749	-99.9195	184.754
Mean Return	.0828	-.0406	.2622	.0692	.1003
Max 10 Days DD	3.5665	2.5665	3.5954	2.9604	.3324
Standard Deviation	1.0759	1.0857	1.0729	1.091	.2079
Sharpe Ratio	.0696	-.0447	.237	.0562	.4443
Skewness	.1191	.1064	.0777	-.0427	-.0302
Kurtosis	3.2426	3.0078	3.6538	3.0504	15.0612
VaR	-.1081	-.106	-.1098	-.1045	-.0179

5.4.3 Sensitivity on look back period for expected return :

	$S_{60}^{30}(\beta_T^m = 1)$	$S_{60}^{60}(\beta_T^m = 1)$	$S_{60}^{90}(\beta_T^m = 1)$	$S_{60}^{120}(\beta_T^m = 1)$	SPY
Cum PnL	-99.9755	-95.5482	-99.8588	-98.5892	184.754
Mean Return	-.0297	.3459	.126	.2915	.1003
Max 10 Days DD	2.9030	2.9497	3.3108	3.2382	.3324
Standard Deviation	1.0799	1.0692	1.1065	1.0989	.2079
Sharpe Ratio	-.0348	.3161	.1067	.258	.4443
Skewness	-.1157	-.034	.1045	.0168	-.0302
Kurtosis	3.3687	3.4237	4.0425	3.477	15.0612
VaR	-.1124	-.1057	-.1167	-.1075	-.0179

6 Conclusions

1. During normal periods, target beta should be close to the benchmark for best results. During times of crisis, the impact of beta is less clear; therefore, we do not recommend using target beta as an effective metric of portfolio performance during this time.

2. Changes in the covariance look back period appears to have the greatest impact on PnL. Shorter covariance look back periods appears to be optimal for short period portfolio strategy.

3. Investors can also choose which groups of look-back periods and target beta to use, based on their risk preference and according to the statistics included in the description table.

4.The trend estimator model makes sense, and works well in short period(around one year).

5. For longer period(10 year), sue to the time consuming problem, we didn't get enough data for reliable analysis. But according to the results that a simplified estimator provides, we see first the choice of estimator matters because it has better performance at least on short period. In addition, we recommend not to use a single group of parameters(target beta and look back periods) for a very long period. Because market changes, so the 'best' portfolio setting changes as well. Even though the portfolio works well in the first 2 years, it's very likely that it lose most of the money in the latter 8 years.

7 Future work

1. Expanding the universe of asset such that the range of target betas is increased may provide greater insight into portfolio strategy.

2. Change the value of lambda in order to observe the impact on portfolio sensitivity analysis.

4. Run more simulations on more periods with different or similar feature to generalize our sensitivity analysis, and make it more reliable as well.

3. Try different models to predict factors which may lead to more efficient code.