# Portfolio Construction

Once we can predict the EPS for the next quarter, we try to incorporate our predictions into portfolio construction techniques, and build several hypothetical investment funds, where the prediction of EPS will be our main source of asset weights. We shall see how our portfolios perform compared to the benchmark.

## Asset Universe and Benchmark

We performed EPS on 9 assets in the financial sector, and 3 assets in the technology sector. They are:

* Wells Fargo
* Goldman Sachs
* Bank of America
* Berkshire Hathaway
* Blackrock
* BNY Mellon
* Citigroup
* JP Morgan
* Morgan Stanley
* Adobe
* Apple
* NVIDIA

The reason why we would not expand the universe into more assets was because the preparation of data that is needed for the prediction involves a lot of manual steps and data cleanup. We could not find a good source / vendor of data that can give us the fundamental of many companies in a clean format.

Therefore, we chose to do our portfolio benchmarking in 2 universes: financials, and all sectors. For financials, our universe is all 9 financial assets, and the benchmark is the SPF index (the financial sector of SP500). For all sectors, our universe is all 12 assets, and the benchmark is the SP500 index.

## Time Periods

Each asset has clean data with varying starting date, but all assets have clean data starting in January 1st, 2004. Prior to that date, we observed some issues with the data that we collected from some vendors. For the sake of uniformity, that will be our starting period for all assets(*T1* = 2004/1/1).

Our fund “inception” date is *T2* = 2010/1/1, and we rebalance weekly on Fridays.

Our last rebalance date is the last Friday in November of 2019 (*T3* = 2019/11/29) for all sectors, and the last Friday in January 2018 (*T3* = 2018/1/25) for financial sector. The reason why we stopped early for financial sector is because the data for SPF that we obtained only existed until January 2018.

## Methodology

For each rebalance time *t*, where *T2* <= *t* <= *T3*, we take the data from *T1* up until *t*-1 (from beginning of training period up until the day before rebalance), and train a brand new regressor with that data. This way, we simulated the scenario where the portfolio manager would be able to see all the financial data up until the day before rebalance, in order to make his decisions about that rebalance period. That also means that the regressor always has the most up-to-date information about the company, without being able to see the future (i.e. we never train it with data AFTER the rebalance because it is not supposed to be able to see that data yet).

The EPS prediction is converted into asset weights using 3 flavors of interpretation, which we shall see below. After we obtain the asset weights, we make the trading at the end of the Friday rebalance, using the price as of the close of Friday. This is a known limitation because in reality, the effective price that we traded at would not be that price.

Another known limitation is that we also ignore transaction cost and all trading is assumed to be completed instantaneously, instead of taking several days as is common in many asset management firms. We will revisit these limitations in more detail in another section below.

For each flavor of portfolio construction, we do 2 variations (for a total of 4 combinations): long-only vs long / short, and with or without Citi.

Long-only vs long / short: we observed that a lot of our gains come from shorting stocks whose prices went down, so we tried the long-only strategy to test our hypothesis that a large portion of our gain is from shorting stocks.

With / without Citi: As we will observe later, there were three consecutive rebalance periods where we shorted Citi in large quantity (~35% of the fund’s portfolio), and as a result, profited a lot from Citi’s price falling (around the end of 2016). As per Prof. Hirsa’s suggestion, we tried to take Citi out of our portfolio and compared the result, to see whether we “got lucky” with Citi, or whether the strategy is sound even without an event like Citi.

### Flavor A

Given a predicted EPS, we can calculate the predicted P/E, by taking stock price divided by EPS. High P/E is undesirable, whereas low P/E is desirable. So our asset weights are proportional to 1 / P/E. In other words, the asset weights are proportional to EPS / price.

When we compute the long-only variation, all the negative EPS predictions are taken out of the universe, so we are only investing proportional to the positive EPS / price.

### Flavor B

Given a predicted EPS, we compare it to analyst’s prediction of EPS, which we get from FactSet. The asset weights are then proportional to (predicted EPS - analyst EPS) / predicted EPS = (1 – analyst EPS / predicted EPS).

If this number is high, it means that predicted EPS is likely to beat analyst EPS, and this stock is desirable. If this number is low (or even negative), it means that the predicted EPS is likely to miss analyst prediction, and we should short the stock. The predicted EPS itself may be “high” in an absolute sense, but it is compared to analyst prediction, and that’s our clue of whether the stock price is going to increase or decrease immediately following the earnings call.

### Flavor C

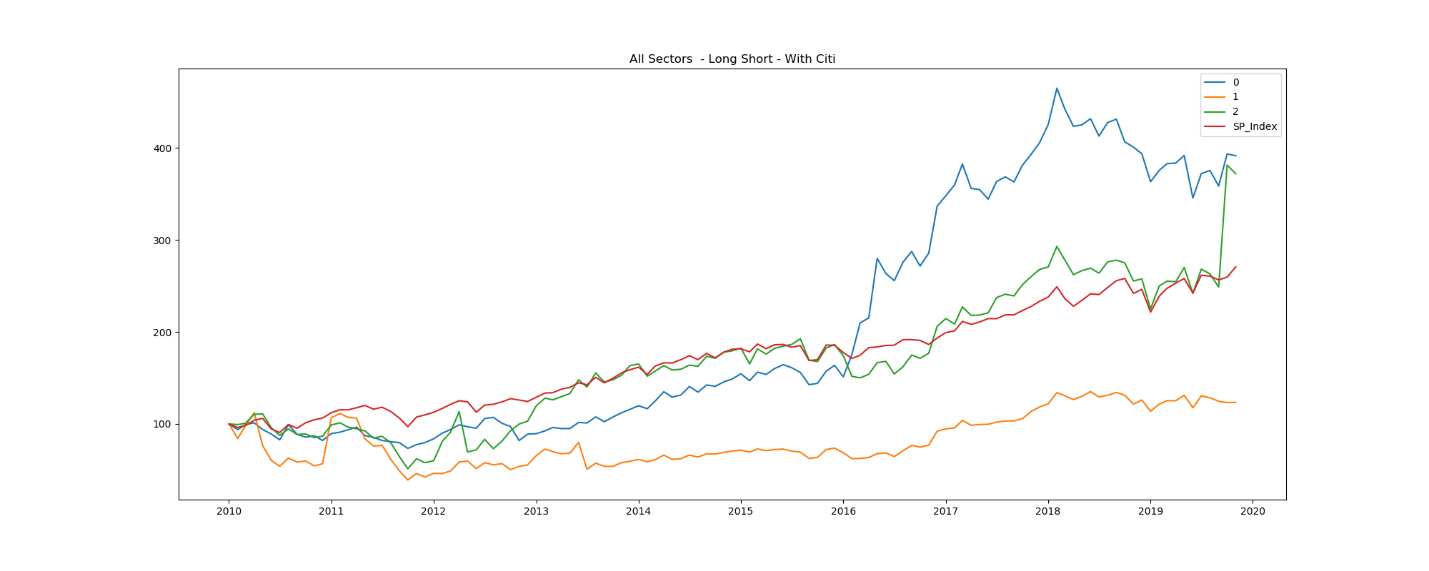
Similar to flavor B, the asset weights are proportional to (predicted EPS – average EPS) / average EPS, where average EPS is the average EPS for the last 1 year.

If this number is high, it means that the earnings are going to increase compared to how the firm has historically been performing, which would then encourage price increase. Likewise, if the number is low, the stock price is likely to decrease, and we should short the stock.

## Result

With 3 flavors, 2 universes (financials and all sectors), and 2x2 variations (long-only / long-short, and with / without Citi), we have 24 different portfolios that we simulated from 2010. The growth plot for the portfolios are shown below, with each portfolio indexed at 100 at inception.

### For All Sectors



### For Financial Sector

### Information Ratio

For each portfolio we also calculated the information ratio against the respective benchmark.

A few observations:

* When we force the portfolio to be long-only, our performance did suffer terribly. In some cases we were still able to beat the benchmark, by investing in stocks that were about to go up, but the weights were more evenly distributed. The huge jumps of performance that we get from long-short were because we shorted a big amount in one or two names. In other words, we were a lot more aggressive with the short strategy as opposed to with the long strategy, and that’s why we were able to make bigger gains with the short strategy.
* When we included Citi with the long-short strategy, we did the best, both in financials and all sector universe. When we included Citi but not allow short, we didn’t do as well (for the reasons we mentioned in the previous bullet point). When we remove Citi but allow short, our performance is quite good but not as spectacular as when we included Citi. This reinforces our belief that while our prediction strategy is very good with predicting the decrease in price, we also “hit the jackpot” with Citi.

## Limitations

There are a number of limitations with our simulation that make it impossible to replicate in real-life scenario. We shall address these limitations:

1. We ignored transaction cost, as well as portfolio implementation cost and timing. We assumed that each trade happened instantaneously on Fridays at the close of day, using the price as of the market close on Friday. In real life scenario, transaction cost is a non-trivial component of performance, especially for a fund that rebalances weekly like ours. It is also unlikely that the new weights can be implemented instantaneously without affecting the market or missing the target price.
   1. Our portfolio also varied wildly (with huge turnovers and sometimes changing from long a stock into short that very same stock the following week). This is because our portfolio construction methods do not take into account current holdings, therefore we could not “optimize” to try to not veer away from current holdings too much. In order to be as realistic as possible, we should either limit the turnover rate (where that limit is baked into the portfolio construction method), or incorporate transaction cost and market impact into our performance calculation.
2. We ignored borrowing cost. In real life portfolio, whenever we short a stock, there is borrowing cost associated with that short holding, and the amount is often non-trivial. One of the reasons our short portfolio performed so well was because we completely ignored borrowing cost in our performance calculator. In real life scenario, the borrowing cost is a significant “drag” on our performance. Also, shorting ~30% of a fund’s total value in a single name proves to be a good move in hindsight (during backtesting), but in real life, that move is extremely risky. We imagined that the firm’s risk manager would not easily approve of such position.

## Takeaway

1. Our EPS prediction does really well when combined with short selling strategy. Even accounting for the fact that we ignored borrowing cost and transaction cost, we were able to predict when certain stocks are going to decrease in value. It’s not necessarily that the EPS is negative, but when the EPS is going to be well below analyst expectations, we were able to predict it ahead of earnings call.
2. Our EPS prediction does moderately well when combined with long-only strategy. We can predict when the EPS is going to beat analyst expectations, but the confidence and magnitude of our prediction is less accurate than that of the short-selling scenario. Therefore, while we still profit from the increase in stock price, we did not allocate as much into that stock as we could have.
3. Our EPS prediction scheme hits the mark once in a while, but for a large amount. In other words, we “get lucky” for a name or two once in a while, but that correct prediction allows to gain a significant profit.

## Future Work

1. Future work based upon this project could include realistic portfolio implementation scenarios, such as:
   1. Adding time to implement the portfolio (For example, when we switch from holding 60% Goldman Sachs to 20%, it will have to be implemented over 5 trading days instead of instantaneously, and the price of sales will follow the closing price of those 5 days)
   2. Adding borrowing cost for short positions
   3. Adding transaction cost, especially for large turnovers, or alternatively, adding turnover restriction baked into the portfolio construction step itself
2. Adding data for more companies. If we have many (~100) companies with predictions, we could invest in, say, only 50 companies at a time, with 50 of the best prediction. In other words, our EPS prediction becomes part of our “stock picking” mechanism, and the weight construction can be based on other metrics (expected value, covariance and risk) and our EPS prediction can be included in those metrics too. We feel that this construction scheme will be a lot more robust and stable, and reduces most of the large turnovers we see in current scheme.