Limiting Risk in Stock Selection

Overview

Last month we looked at a portfolio selection method that would predict future stock returns and pick the stocks with the highest expected return. As shown below, this method increased wealth dramatically during several years of a test simulation before plummeting back down and actually losing money in the long-run. The reason for this is that high returns are correlated with high risk, and the method of selecting the stocks with the highest expected returns will naturally also select the stocks with the highest risk.

The table below shows average returns, standard deviations, Sharpe ratios, and Sortino ratios for the methods we analyzed previously. These methods increased the standard deviation of returns by over 3 times but did not provide any incremental return versus the market. This is shown clearly on the Sharpe and Sortino ratios where both of these methods fail to outperform the market.

90-Day Returns (starting 6/2/2008)

		Log Returns			
Method	Avg.	Avg. Return	Stdev	Sharpe	Sortino
S&P 500	3.1%	2.5%	10.80%	0.2337	0.1704
Avg20(trade=20)	6.1%	-0.1%	35.13%	-0.0039	-0.0037
Avg20(trade=5)	7.5%	2.0%	32.97%	0.0604	0.0545

Obviously, selecting the stocks with the highest expected returns is not the right approach. Thus, we will modify that premise slightly. What if we select the highest return but place a limit on the amount of risk we are willing to take? This method does produce superior selection and more stable results. However, the incremental performance versus the market is typically just 1 or 2%. The methods also become somewhat sensitive to parameter selection, with some selections producing results that under-perform the market.

Limiting Risk in Stock Selection

Baseline

As we did last time, we will use the S&P 500 as a baseline for comparison. $100k invested in this index from 6/2/2008 through 5/10/2017 would have a final value of $173,005.

Approach

We will assume that we want to out-perform these returns but will impose a new restriction: the amount of risk we take on cannot exceed the S&P 500. Risk will be measured by the Sharpe ratio (using standard deviation or returns) and the Sortino ratio (using lower partial moments). We will use two methods to calculate these statistics:

Method	Description
Historical Statistics	If we are just forecasting the average historical return, we can estimate the standard deviation and lower partial moments by calculating them directly from history. This will be the same as if we had forecasted the average value for all points in history. Of course, we would not have been able to do this, since some of the values that go into that average would have been in the future, but it's a fast and simple way to get a historical view of standard deviation and use this to estimate risk.
Out-of-Sample Forecasts	The best measure of forecast error is obtained by creating forecasts and then measuring their results using out-of-sample data. Any forecast method can be used in this case (not just the average). Each forecast will be saved and measured against the actual value observed later. Whatever error is seen historically on the forecast (as measured by the standard deviation and lower partial moment) will be expected to measure our error estimates for current forecasts.

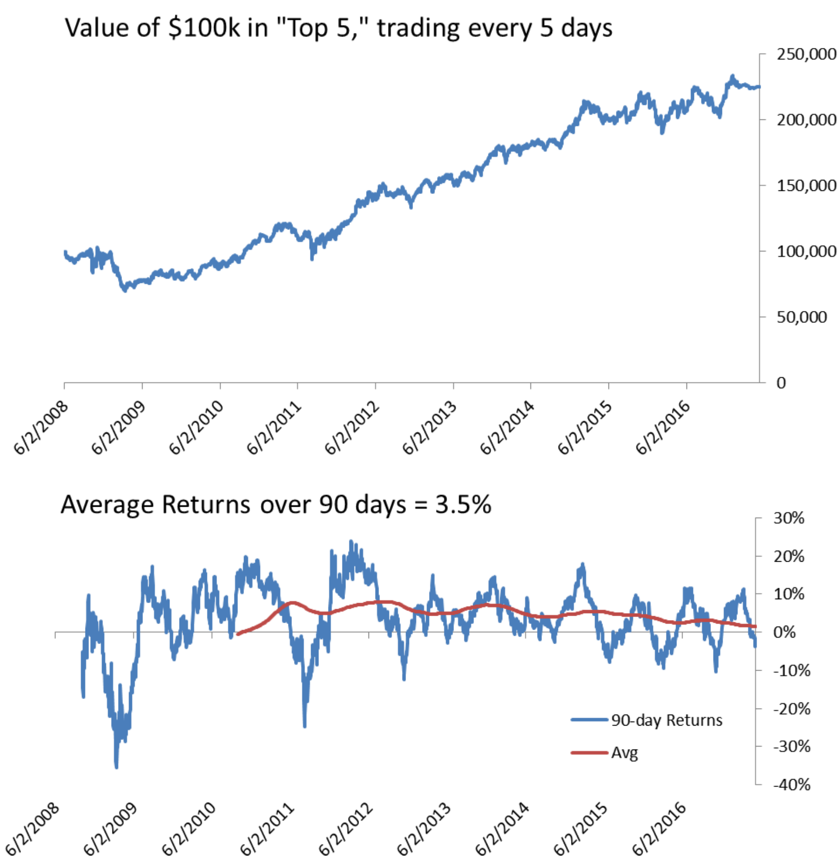
To illustrate how this works, imagine that we are looking at the last 250 days of returns and are interested in a 5-day forecast. The historical statistics method can look at 5-day returns during this period, calculate their average, standard deviation, and lower partial moment. We will assume that these describe our current expectations about the stock. The historical average will be our forecast. The historical standard deviation and lower partial moment become our expectations for the future error around this forecast. Our goal will be to find stocks that have the highest expected returns but without exceeding the risk measures for the S&P 500 as measured by the standard deviation and lower partial moment. We can re-evaluate our portfolio at any time and use this method to select new stocks.

NOTE: We could also try to simply maximize the Sharpe or Sortino ratios directly. However, this might not produce the desired results. It is possible to have a high Sharpe ratio but also low expected returns. As long as the ratio of returns to risk is high, the Sharpe ratio will still be high. If our goal is to out-perform the market, we need to ensure we are selecting stocks with the highest returns possible.

CONSIDER: We could also maximize Sharpe ratios subject to the expected returns on the portfolio being higher than the market. This would be identical to selecting stocks with the lowest volatility subject to expected returns being higher than the market. This may result in stocks with expected returns very close to that of the market though, rather than those that exceed it by a significant amount.

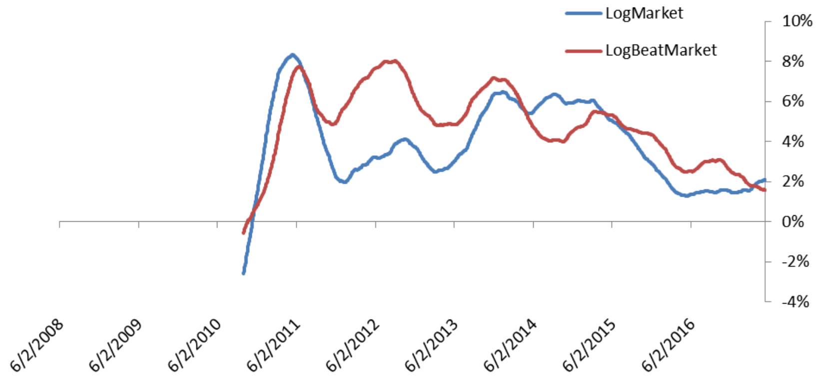
Results from Historical Statistics

There are many different calibrations for the model just described. To begin with, we'll look at a portfolio that buys 5 stocks at a time, trades every 5 days, and produces forecast using 250 days of history and examines 20-day returns. This portfolio climbs in value to $225,125 and out-performs the market:



Limiting Risk in Stock Selection

The second graph shows 90-day returns for this method with a moving average overlayed. If we compare that long-term average to the market we get the following:



This portfolio out-performs the market fairly consistently. There are just brief periods of a few months or 1 year (between 6/2/2014 and 6/2/2015) where the market outperforms.

However, the results can be sensitive to parameter settings. It doesn't make much sense to look at 20-day returns when we are trading every 5 days. If we instead analyze 5-day returns – which makes more sense – results get worse and we only end up with $143,772 in value. In this case we actually under-perform the market.

90-Day Returns (starting 6/2/2008)

		Log Returns			
Method	Final Value	Avg. Return	Stdev	Sharpe	Sortino
S&P 500	$173,005	2.5%	10.80%	0.2337	0.1704
Avg20(trade=5)	$225,125	3.5%	8.58%	0.4090	0.3561
Avg5(trade=5)	$143,772	1.6%	8.58%	0.1838	0.1495

Interestingly, both of the methods produce standard deviations that are lower than those of the market. This is unexpected, since the method of estimating standard error using historical standard deviation should produce results that under-estimate our out-of-sample standard error. The method may have similar effects for all stocks though, so that we are under-estimating the S&P 500 just as much as any other stock.

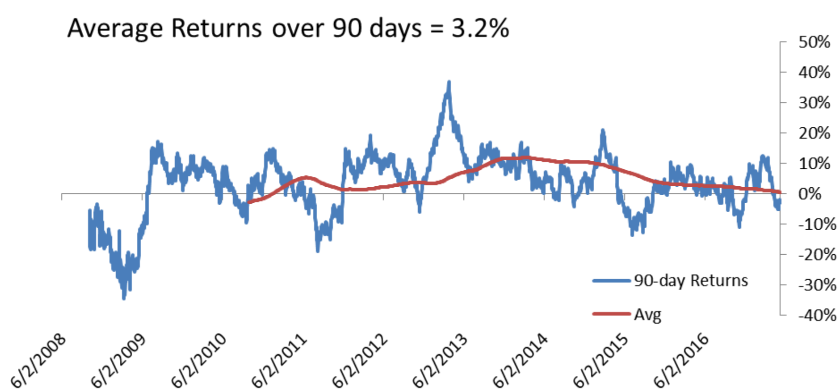
Results from Out-of-Sample Forecasts

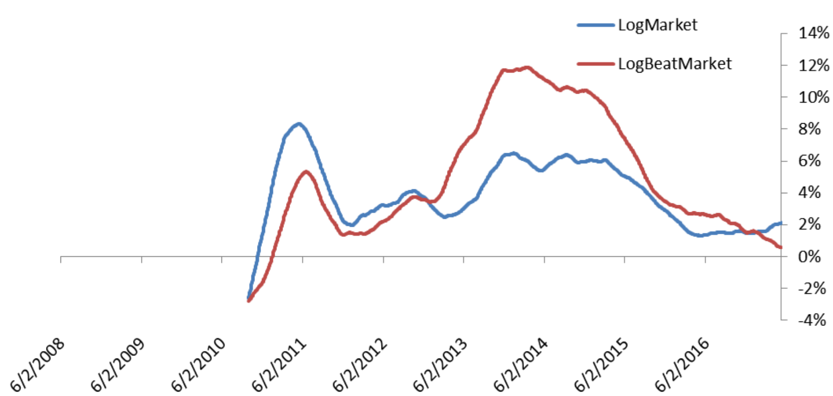
Better error and risk estimates should be obtained by measuring actual error on out-of-sample forecasts. We will still find that results are sensitive to parameter settings though. Once again, we will use 250 days of history to calculate an average return and forecast this. We will also allow the portfolio to be traded every 5 days. Instead of using historical standard deviation to estimate future error, we will instead make predictions and record the observed error on those forecasts. We now have to decide how many errors to use when estimating our errors.

The chart below shows results when we use 90 days of errors.

Limiting Risk in Stock Selection

The 90-day return profile is shown below. 90-day returns swing between positive and negative (just as before). The long-term average out-performs the market since 2013, but it under-performs longer before that and also slightly in 2017.





The table below shows statistics for different calibrations of this model. Some do better, but some do worse.

Limiting Risk in Stock Selection

90-Day Returns (starting 6/2/2008)

		Log Returns			
Method	Final Value	Avg. Return	Stdev	Sharpe	Sortino
S&P 500	$173,005	2.5%	10.80%	0.2337	0.1704
Avg (h=5,trade=5,errDays=90)	$221,401	3.2%	10.00%	0.3247	0.2751
Avg(h=5,trade=5, errDays=250)	$178,566	2.4%	8.98%	0.2629	0.2056
Avg(h=5,trade=5, errDays=30)	$198,250	2.8%	9.98%	0.2800	0.2312

All of these models out-performed the market, both in returns of total value and the risk-adjusted returns. However, there were still some calibrations that under-performed the market. These were observed when the portfolio was traded every 20 days and when we forecasted 20-day returns. The final values associated with error days of 30, 90, and 250 were $142,305, $157,130, and $215,550, respectively. One calibration ended with $310,099. This was a strange one that used a minimum of 30-days to produce its first error forecasts and a maximum of 250 days. There doesn't seem to be any good reason for these calibrations to return better or worse results. Instead, it leads us to believe that the model is just sensitive to parameter settings and that we might be over-fitting the model by paying too much attention to specific settings. In general, it does seem like this method out-performs the market, but only by a percentage point or two.

Conclusion & Next Steps

This method of selecting stocks with maximum returns subject to a limit on their risk does produce results that are better than those of just selecting based on returns alone. It out-performs the market in most cases, albeit only by 1% or 2% annually. There are several different ways to extend and continue this work.

First, any method that improves our ability to forecast stock predictions should now allow us to select better portfolios. A better forecast is one that results in lower error, making more stocks viable for selection based on this method. However, if we also improve the forecast for the S&P 500, we might not recognize this benefit entirely. It may make sense to select stocks with lower risk than the S&P 500 as measured by a pre-defined forecast (such as a historical average forecast). This means we aren't always moving the bar on our baseline as we improve our forecast methods for other stocks. It also may be possible to introduce a bit of "slack" into the model. Instead, of just selecting stocks with risks lower than the market, we may decide to allow stocks if they are less than 1.1 times the market risk. This method has been tested on some calibrations and shown to produce slightly better results – although they also come with higher risk.

Second, we may improve on the method for selecting stocks. When we select stocks with standard deviations lower than the market, we are likely to form a portfolio with a much lower portfolio. This is because a selection of stocks always has lower variance than any individual stocks. This benefit of diversification is not being captured in the current model. If our goal is to form a portfolio that matches the market in terms of risk, but surpasses it in terms of returns, we may be leaving some returns "on the table" by not including riskier stocks that still might produce final portfolios with acceptable standard deviations. Forming a portfolio that takes into account the benefits of diversifications is a difficult prospect. We can take a heuristic approach of selecting one stock and then adding others one-at-a-time. This will increase the computation time for running these simulations though.

Third, we may also modify the internal methods for estimating out-of-sample forecast error. The method described above produced one forecast at a time, measured its results, and then kept a running history of past forecast errors to estimate our expectations. The method that used 250 days of history to forecast returns then had to wait for 250 forecasts to be produced to get a good understanding of the error profile. This meant that we didn’t have any useful information available until we had seen 500 observations. We have seen that just using the historical statistics produces decent results – especially if we use the same method for all stocks. We could thus experiment with using the historical standard deviation when we only have 250 samples and slowly moving to true out-of-sample results as we accumulate more history. We can also experiment with different ways to calculate this error. Instead of just using a set of historical forecasts and actual values, we can use exponential smoothing results to keep track of sum-of-squared-error and a lower partial moment variant. This would have the advantage of using lower memory, updating faster, and transitioning seamlessly from historical (in-sample) forecast error to true out-of-sample error.

Overall, the best returns are likely to come from finding ways to improve our stock return forecast. We should experiment with technical indicators such as MACD, Stochastic oscillators, and relative strength indexes. The other methods may still be interesting for trying to fine-tune these results and improving performance.