Construction and Analysis of S&P N Indices

John E. DeForest Atanas G. Iliev Vadin Thadhani
Ameer Talha Yasser

Contents

Abstract

In this paper, we propose S&P N indices that track market cap weighted portfolios of the largest N publicly traded companies on US exchanges. First, we discuss the specifics about the construction of each portfolio. Then we calculate a historical Sharpe ratio for each index over a month for 6 months. Last, we produce a forward looking 1 month Sharpe ratio for each S&P N index and compare the results to the historical ones. We perform a Spearman correlation coefficient testing to compare the two methods of calculation.

1 Background and Motivation

The S&P 500 is considered the mainstream benchmark index for US equities. It weighs the 500 largest companies traded on US stock exchanges by their market capitalization and then tracks the performance of the resulting portfolio.

Since there is no way to *directly* invest in the S&P 500, a multitude of exchange traded funds (ETFs) have developed computer programs to continuously adjust a portfolio of stocks so that it tracks the S&P 500 as close as possible. The most famous example is the SPY, the oldest and largest US ETF.

Funds like the SPY, provide an easy way for market participants to invest in the S&P 500 and provide a way for retail investors to gain broad exposure to US public equities and share in their upside. Next, we provide a quick summary of reasons about what makes the S&P 500 a great thing to invest in.

The number one reason is returns. Historically, the average annualized return of the index has been a bit over 10% and the index has existed for over 66 years as of today. The second major reason is diversification which stems from the S&P 500 tracking - by and large - the broad public equities market instead of a specific sector. Although there have been periods where the overall US equities market has been dominated by a subset of stocks with the ‘Magnificent 7’ technology stocks being the modern day example.

1.1 A Potential Problem

ETFs such as SPY appeal to passive investors. You would usually invest in the SPY, hold it for a long period of time, and expect about 10% in terms of annual returns over the long run. While one can't predict what will happen to about 80% of the stock market over a single day, it is reasonable to expect that stocks will go up in the long run.

Passive investors are less concerned with specifics than those who want to actively manage their portfolios. The SPY reduces their exposure to risk while at the same it still allows them to partake in the stock market and promises about 10% in returns over the long run.

Let us imagine for a moment that we are a passive investor who is just a bit more interested in the market than the one we have just described. In particular we replicate the strategy of the SPY and other S&P 500 tracking ETFs. Our own investment strategy is the same, we want to weigh the largest US publicly-traded companies in a portfolio so that we are still promised results (the key being the fact that we invest in large companies) and we still aim for reducing our risk (by the same means of diversification). In fact we have but a single concern with the SPY and with the S&P 500 in general. Why 500?

We believe that the number 500 seems a bit arbitrary. Is it the S&P 500 because 10% is a really nice round number? Or is it because it contains about 80% of the total market cap of US public companies (another nice round number)? Not really.

In fact, we could not find a good explanation about why the index is tracking 500 and not 499 companies.[1] There is one logical explanation that we are going to present shortly. In a way, our paper is simply a challenge of that explanation.

1.2 Risk-Adjusted Returns and Sharpe Ratios

Let us first think about what we expect to happen once we adjust N. Perhaps the simplest case to consider is the S&P 1. As of March 9, 2024, this would be as Microsoft Corp (MSFT). One immediate prediction we can make is that the S&P 1 is going to be way more 'risky' then our usual ETFs like SPY. This is evident because all the risk-management SPY is doing comes from diversification. Given this information, it is interesting to consider what the long-term returns of MSFT look like. A quick check reveals that the trailing

[1] As it turns out, right now the S&P 500 includes 503 distinct components. That is because while it is always tracking exactly 500 *companies* sometimes one or in this case three of these companies will issue more than one class of listed *stock*. One example is Alphabet, with GOOG and GOOGL both in the index.

5 year return is about 250.5% compared to the S&P 500's 81.5%. Now, let's consider an N big enough that the resulting S&P-like index includes all listed US companies. While there isn't a straightforward way to check the returns of this index, the Russell 3000 is an index that aspires to emulate just what we want and its 5 year return is about 76%. Of course, its 'risk' is also lower than that of the standard S&P 500 because it includes all (or nearly all) of the stock market and is not completely concentrated in the technology sector.

Hence, we can reasonably conclude that as N increases both expected returns and risk seem to decrease. This leads us to the one explanation as to why the N did not matter in the first place. Simply put, one may expect that as N increases, both the risk and expected return of the index decrease in a way such that the risk-adjusted return remains the same. But how can one measure that in the first place?

One approach is to use Sharpe ratios. The Sharpe ratio S_p of a given financial portfolio p is given by:

$$S_p = \frac{R_p - R_f}{\sigma}$$

where R_p is the return of the portfolio over a given period of time, R_f is the risk-free return (for example the rate of Treasury bills/bonds) and σ is the standard deviation of the portfolio over that same period of time[4].

The Sharpe ratio is the most straightforward way to measure risk-adjusted returns. You take the return of your portfolio, subtract how much you could have made without any risk[2] and divide by the risk you are assuming. If we want to be a bit more technical, it might be helpful to express the formula as

$$S_p = \frac{R_p - R_f}{\sigma - 0}$$

to indicate that we are subtracting the risk of the risk-free portfolio[4].

Sharpe ratios, and specifically the way of calculating them, bring up an important proposition. That is, that the way to measure risk is by using standard deviation. In fact, standard deviation of measuring a portfolio's volatility which is in turn assumed to measure a portfolio's level of risk. These two assumptions that 1) standard deviation is a reliable measure of volatility[3]

[2]Of course, there are some minor implicit assumptions here such as that the US government is not going to default on its liabilities.

[3]This one seems pretty straightforward given what we usually think of when we say volatility.

and that 2) volatility is a reliable measure of risk are upheld throughout this paper.

Having discussed Shape ratios and their meaning, we can now formulate:

1.3 A Potential Explanation

Explanation 1: As N varies from 1 to 500, $S_{\text{S\&P}N}$[4]. remains constant.

There are a few problematic things with that explanation. First, this seems to be a massive assumption. It is important to note, this project is primarily motivated by the belief that this assumption is wrong. While it is a completely plausible result that as N increases both the expected return and volatility of an S&P N portfolio decrease, we want to challenge the that risk-adjusted returns stay constant. We were not able to find any literature that would support that statement in the first place, so we formulate our:

Hypothesis: As N varies from 1 to 500, S_N does not remain constant.

Before we continue to explore the implications of our hypothesis and to formulate the specific goals of this paper, we want to address one other caveat about *Explanation 1*.

1.3.1 Different Risk Appetites

That is, even if S_N remains constant as N varies from 1 to 500, one justification for the existence of different indices of the form S&P N is differential appetite for returns. While risk-adjusted returns may remain constant, we have already demonstrated that 'expected' returns[5] and risks are definitively not constant for different portfolios such as the S&P 1 (MSFT) and the S&P 500.

Hence, one reason for the existence of the different portfolios might be catering to passive investors with different (if we consider $1 \leq N < 500$, we can just say higher) risk tolerances than that of an investor in SPY. Unfortunately, for our attack on *Explanation 1*, it does have a simple response.

[4]From now on, we are going to abbreviate $S_{\text{S\&P}N}$ as S_N.

[5]Refering to our previous example, when we say expected returns, we mean this from the point of an investor on March 2019 given that the information we used technically refered to 5 year historical returns for both portfolios we discuss in this sentence.

One could just invest in a portfolio made up of MSFT and SPY in different proportions. In theory, that should be able to take the investor on any point on the line connecting the expected returns/risks of the two basic indices (S&P 1 and S&P 500). However, the MSFT is in the SPY so the covariance between the two is definitely not 0 which makes the calculations more difficult for our target audience of passive investors to do manually. And while different risk appetites is not the central point of this paper, an attack on *Explanation 1* based on them is something that may be worthwhile exploring in the future.

Now that we have addressed this little caveat, we can move to formulating the specific aims of this paper.

1.4 Two Objectives

Given our hypothesis our first objective is:

1.4.1 Actual Sharpe Ratios

Objective 1: Calculate which N such that $1 \leq N \leq 500$ has resulted in the highest Sharpe ratio S_N over a predefined period of time.

Fulfilling this objective may have intriguing theoretical implications. One (remote) possibility is that the S&P 500 has had the highest Sharpe ratio of all considered S&P N portfolios throughout a certain time period. While it is a null result, it would still imply that no passive investor in the SPY has missed on anything at all as long as they wanted to rely on the S&P N strategy in the first place. However, any other N consistently outperforming the rest might be even better news. That would mean there is a possible S&P N strategy that would result in higher risk-adjusted returns for our passive investors than the SPY.

There are two assumptions about this. First, there isn't an actual ETF or other financial instrument that tracks any of the indices we want to consider except for the S&P 1, S&P 100, and S&P 500. At this point we say little about this. Our paper is theoretical and we assume that such an ETF can be created.

The second assumption is that while the Sharpe ratio is a commonly used measure its inputs are historical[4]. Even if it turns out that some $N \neq 500$ outperforms SPY over a certain historical range of time, it would require a great deal of fundamental reasoning to assume that this outperformance persists. Hence, we might need to often adjust which N are we actually invested in at a given point in time. We are confident that an automated ETF would easily implement such adjustments just as SPY has implemented the reweighing of the actual S&P 500 when needed.

Our concern is that Sharpe ratios, in the form we have given them, cannot be used to predict (estimate) future risk-adjusted returns. This is of vital importance, because we will want our investors to know what they want to invest in before they can see using the now historical data what they should have invested in. Hence, we want to construct a modified Sharpe ratio that aspires to be at least partially forward looking for each of our indices. This leads us to the formulation of our second objective:

1.4.2 Semi-Forward-Looking Sharpe Ratios

Objective 2: Construct a Sharpe ratio S_N^{iv}[6] for each index S&P N where $1 \leq N \leq 500$ which takes forward looking volatility instead of historical standard deviation in its denominator. Then compare each S_N^{iv} to the respective actual Sharpe ratio S_N.

We call such Sharpe ratios semi-forward-looking ones[7]. In particular, we care about the risk-aspect of the ratio, or the volatility of each portfolio, to be an implied one and not a historical one. In terms of expected returns, these will remain historical in the semi-forward-looking Sharpe ratios, but we will use back data when calculating them. For example we may use data from January of a given year to calculate the implied volatility for a given S&P N in February of that year. When we compare the resulting ratio, we will use the actual Sharpe ratio for February. The returns in the semi-forward-looking one will still be historical but from January[8].

From a month-to-month timeframe, previous month returns can have some bearing on next month's returns due to the existence of momentum. Realistically, however, the momentum assumption is challenged upon the announcement of news or shocks to the market.

Having formulated our two goals for this paper, and having explained all the necessary motivation behind the project, we now proceed with our

[6]Here, iv stands for implied volatility.
[7]In the plots we present later on, this is going to be shortened as SFL Sharpe Ratios.
[8]More on that in the next section, Methodology.

methodology.

2 Methodology

In this section we discuss the methodology we used to achieve the aims of the paper. In particular, we discuss data collection, the construction of each index, the pricing of each index, the calculation of the actual Sharpe ratios, the calculation of the implied volatilities of each index, and the calculation of the semi-forward-looking Sharpe ratios for each index. Along the way, we make note of all the assumptions necessitated by the scale and scope of the project.

First we begin by discussing the data our paper required.

2.1 Required Data

The first challenged we faced was acquiring all the data we required. Most of the data was sourced from WRDS. Since we wanted to let N vary between 1 and 500, we needed data for each component of the original S&P 500. We used CRSP for our data[2].

Unfortunately, this turned out to be problematic in a multitude of ways. First, actually collecting the ticker symbols for every constituent component of the S&P 500 turned to be more difficult than expected. In fact neither CRSP, nor WRDS as a whole seem to have a nice way of quickly pulling all the tickers in the S&P 500 at an arbitrary point in time[2]. In the end, we ended up pulling all 503 stocks that are currently in the S&P (as of the end of February, 2024) from a Google search and used these for every day we needed data for. We used the same 503 ticker symbols for every day for which we pulled data from CRSP. That is we assumed that for a period of six consecutive months, no stocks entered or exited the S&P 500.

Immediately after, we put our list of tickers in CRSP we faced another problem. We couldn't use the six most recent months because OptionMetrics, what we used for pulling up the data needed to calculate the implied volatilities to be used in the semi-forward-looking Sharpe ratios, did not let us access the most recent data[9]. We ended up using data from August 2022 to February 2023. We assume that S&P 500 contained the exact same 503 ticker from the beginning of August 2022 until the end of February 2023 as it

[9]The last accessible data is for February 28, 2023[3].

did at the end of February 2024. We have decided to do all of our testing on seven[10] consecutive months instead of six arbitrary months so as to minimize the impact of our assumptions. The six month period we used was roughly flat for the S&P 500 (Cumulative return of roughly 1.6% and buy and hold return of 0.4%[2]) so our data is not biased by an overly bullish or bearish market.

The last one somewhat decreases the confidence in our backtesting, but at the same time mimics the fact that passive investors would usually invest in the S&P 500 (SPY) for a long period of time given that as discussed in the introduction the annual expected returns are reliable only in the long run. We performed all the comparisons over monthly Sharpe ratios, whereas passive investors would usually invest in our proposed products for longer periods of time.

The last point was necessitated mostly because of the time complexity of our algorithms. Simply put, the calculations would have taken too long to compile and run were we to compute the ratios over a year or even several years as we originally intended.

Now, we can list all the data we collected from CRSP. The information is presented in the following list:

CRSP Daily Data on 503 Ticker Symbols from August 1, 2022 until February 31, 2023, inclusive[2]:

- Ticker symbol
- Number of shares outstanding
- Bid price
- Asked price

We pulled all data entries into an Excel spreadsheet. Then we used OptionMetrics to pull the following data which we needed to calculate the denominators of the semi-forward-looking Sharpe ratios:

OptionMetrics Data on 503 Ticker Symbols from August 1, 2022 until February 28, 2023, inclusive[3]:

[10]The testing actually happens on six months but the first is used for the calculation of the semi-forward-looking Sharpe ratio of the second (first on which a comparison is performed) and so on.

- American call options prices expiring in $\leq$ 30 days for all strikes
- American pull options prices expiring in $\leq$ 30 days for all strikes

Finally, we needed the data for the risk-free returns. This was done for the seven months we have already discussed and we used the rate on 1-month Treasury bills as our benchmark for risk-free returns. This data was sourced from YCharts[5].

Having discussed the nature and sourcing of all the data we required, as well as its implication on our methodology, we can move to addressing the method for the construction of the portfolios.

2.2 Constructing the S&P N Portfolios

In this subsection we discuss the specific way of constructing each S&P N index[11] and the implications this results in.

As previously mentioned, we pulled the data from CRSP in an MS Excel spreadsheet. From there we needed to do two things. First, calculate a price to use for the construction and pricing of each S&P N. Second, calculate each stock's market cap to use in the weighing of the components in the portfolios. The calculated price refers to the average of the bid-ask (this is slightly different from the S&P 500 which uses closing prices when it calculates its weightings and prices.

The reason behind this is purely technical. We used CRSP for the rest of our stock data and it did not provide close price[2]. This is definitely something we would wish to correct in the future, given that averaging high and low daily prices is not good in terms of measuring spot price in the first place.

From here we calculated the average of the high and low prices for each ticker for each day. Then we multiplied that by each ticker's number of outstanding shares for a given day to find its market capitalization on that same day. Lastly we sorted the companies first by day and then by market capitalization.

Then we uploaded our spreadsheet into Python and began constructing the portfolios. At different points in our code we sometimes used data frames

[11]Since no actual ETFs or other financial instruments tracking our proposed indices exist, from here on we use 'portfolios' and 'indices' interchangeably.

and sometimes used multidimensional lists/arrays depending on the preferences of the individual programmers[12]. We used a series of nested loops to go over each day and each N. For each N first we added all the market capitalizations of the first N entries for the given day (those have been sorted by market capitalization as mentioned in the previous paragraph). Last, we divided each each ticker's market cap for each index for the first market day of every month by the respective total market capitalization for the given index on the given first market day. We copied this data for every market day in the current month. We saved all of this data in an object. This object specifies all the weights of the stocks that make a certain S&P N on a given day.

While this process may seem to be the same as the one used for the construction of the actual S&P 500, this is not the case, and brings with it, the following two closely related provisions:

First, we ignore the 'proprietary divisor'. As mentioned previously the S&P 500 always tracks 500 companies but sometimes it includes more than 500 stocks. This is usually the case when a company issues preferred shares in addition to regular ones. To correct for this Standard & Poor use what they call a proprietary divisor. At all points, we assume that this divisor is equal to 1.

What we do instead is treat each individual stock as its own company. For example, we consider GOOG and GOOGL to be different companies ignoring that they are both issued by Alphabet. This leads us to treat each different ticker as a different company to address the problem introduced by the last provision. This carries over to our calculation of N. That is, we now consider the regular S&P 500 at the end of February 2024, to be equivalent to our S&P 503.

The last one is not really something to be concerned about but more of a technicality. It helps explain the small difference in the prices of the real S&P 500 and our variant at a given end of day[13].

Finally, one might have noticed that we reweigh each of our indices monthly. As we found out after having completed that section of our algorithms, the real S&P 500 is only reweighed every 3 months. This results in

[12]Throughout the paper, we are striving to describe each algorithm without anything specific to the OOP side of the code so that this does not present a problem.

[13]Another reason for that is the usage of the average of high/low prices for individual stock instead of closing ones as already discussed.

one last portfolio-construction related difference between our calculation and the one of the actual S&P 500. In particular, we reweigh all of our portfolios monthly instead of quarterly.[14]

Before we move the pricing of our portfolios we address a small late addition to our project.

2.2.1 Equal-Weight S&P N Portfolios

As a robustness check, we have decided to also construct portfolios such that is stock has equal weight in the overall index. One reason for that is that it would be easier for one to invest in such a portfolio given the existing instruments at this time. We perform all the analysis on both value-weighted and equal-weighted versions of the portfolios.

To construct these we still needed to calculate the market capitalization so as to be able to find the largest N tickers by market capitalization for each equal-weight S&P N. From then on the process of construction is exactly the same but the weight of each stock in each portfolio at any day is just $\frac{1}{N}$.

We are now ready to move the pricing of our portfolios.

2.3 Pricing the S&P N Portfolios

The procedure for the pricing of each portfolio at any day is the same for both the value-weighted and the equal-weighted indices.

In particular, for each ticker in a given S&P N on the given day we access the average price on that day for that stock from our Excel sheet and multiply by the weight. In the value-weighted ones this is given by the respective entry in our custom object, and in the equal-weighted variant, it is simply $\frac{1}{N}$.

We do this for every stock in a given index at a given day and sum them up. We store that value in an object. This is the price for the respective S&P N at the respective day. Now, we can confidently move to the calculation of the actual (historical) Sharpe ratios for each portfolio.

[14]Some may even say this seems more precise than the actual process if a bit useless given that it is probably not needed. In fact, we are not sure as to why the S&P 500 is not reweighed monthly if only for the bit of extra precision. We also tried doing it daily but later calculations took too much time to run.

2.4 Calculating the Actual Monthly Sharpe Ratios for each S&P N

Again, the procedure is similar in nature for the value-weight and equal-weight indices. In total, we need three things to calculate the monthly Sharpe ratio S_N of each S&P N. Namely, the expected (historical) monthly return, the risk-free rate, and the monthly standard deviation of each portfolio.

The easiest one of these is the risk-free return. To do this we simply pull the data entry for the rate of 1-month Treasury bills on the first day of the month we want to calculate the ratio for. In our case, we will need six entries, one for each considered month except for August 2022.

Now that this is out of the way, we move to calculating the monthly returns for each portfolio. To do this we pull from our designated objects their daily price on the last and first business day of the month in question and subtract the first from the last. Then we divide by the price on the first business day and multiply by 100 to get a rate in percentage. We do that six times for each of the 500 indices (again, except for August 2022). This leads us to one small caveat: while the risk-free rate is always the 1-month rate on government securities, some monthly returns may be calculated over a slightly shorter period due to the last/first day of a month not being market days but that should have a negligible effect since the term structure should be close to flat for a few days over and under a month.

Now that we have each of the monthly returns we need, we again store them in an object.

Last, we need to calculate the monthly standard deviations for each index on each of the six months we want to do back testing on. To do that we follow the standard procedure. First, we sum all the daily prices from our object for a given portfolio for a given month. Then we divide by the number of market days for that month[15]. Next, we subtract that value from each daily price and square the result. We take the sum of all these and divide it by $m-1$ (where m equals the number of market days in the respective month). Last, we take the square root of that value to obtain the monthly standard deviation for a given portfolio. We do that for all 500 portfolios for each of the 6 months.

We now take our monthly returns, subtract from them the beginning

[15]We have checked these on Google since there are only six of them (no standard deviation calculation for August 2022).

of the month 1-month risk-free rate and divide by the monthly standard deviation to obtain about[16] 3000 actual Sharpe ratios, one for each index for each month from October 2022 to February 2023 inclusive. From here we plotted them against N and ranked the Ns by the values of their Sharpe ratios to achieve our first goal.

Next, we move to the methodology for the fulfillment of our second objective.

2.5 Calculating the Semi-Forward-Looking Monthly Sharpe Ratios for each S&P N

To calculate the semi-forward-looking monthly Sharpe ratios for each *index* for each month, we need to first calculate the monthly implied volatility of each of the 500 *stocks*. Then we need to make a variance-covariance matrix using this data in order to be able to calculate the implied volatility of each individual index. There are a few things we need to say about that before describing the actual process.

First, this was not our original plan. Initially, we planned to calculate the implied volatility (IV) of each portfolio in the same way this is done for the actual S&P 500. In fact, this is given by another index all together called the VIX. The VIX uses prices of call and put options on the SPY to measure the volatility of the S&P 500 much in the way we do in our own process[1]. However, it isn't possible to adapt this for most of the portfolios we consider. The reason is that options on most of our S&P Ns simply do not exist. We cannot simply use options on the individual constituents of every index and add their IVs up because of the covariance between any two of them. Hence, we needed to resort to the variance-covariance matrix technique[1].

Second, there is a reason we did not need to do any of that for the actual (historical) monthly Sharpe ratios. That is because we were able to calculate the monthly standard deviation of every portfolio which in turn was possible because we calculated the daily price of each index for each day in every month we considered. This is not possible when it comes to implied volatility of any sort because, again, no options on the specific indices actually exist[17].

[16] We actually constructed 503 and not 500 portfolios for reasons already mentioned in the paper.

[17] It will be an interesting continuation of this project to create options on each of these indices and calculate respective VIX-like indices for each S&P N using them.

Let us now move to the discussion on the procedure of calculating the implied volatilities of each stock. In subsection 2.1 of this paper we have already discussed the data we will be using to calculate the implied volatilities of each ticker. These will be monthly. To calculate the IVs we use options from the last market day of the previous month to the one we will be comparing the current result to the actual Sharpe ratio. For example if we want to compare a semi-forward-looking Sharpe ratio to an actual Sharpe ratio from February, the options we use to calculate the denominator of the semi-forward-looking one will be from the last market day of January of that same year (in our case, 2023). These options will be expiring within 30 days as previously discussed. Then we filter our data, which again we have pulled into an Excel spreadsheet, for options with strikes that are currently at the money. We then import that into Python and backsolve using the Black-Scholes PDE for call prices for the monthly IV (standard deviation) of the given index at the given month. We use the same equation for put options as well because while the two PDEs are inherently different, the volatility input they receive should be the same in case the prices on the options are similar. We note that we use American options exclusively but feed them to the Black-Scholes equation which is designed for European options. As a result our IVs might be overestimated, because American options are usually more expensive than their respective European alternatives given the early exercise clause they offer.

We have decided to proceed despite this given that at the end of the day we are only concerned with how close the declining orders of Ns by Sharpe ratios are and not how close the actual values are[18]. Now that we have the monthly implied volatilities for every stock for every month (no data from February 2022 is needed since the last one is from the last business day of January 2023), we can move to the construction of the variance-covariance matrix.

The matrix is of size 500 by 500 entries. First, we square the monthly IV (hence we construct six matrices in total) for each stock and feed them into the leading diagonal of the matrix. All other entries (in fact, all entries) are specified by a unique row and column. Each entry equals the covariance of the stock given by the row with the one given by the column. At this point two

[18]In reality we are only concerned about the top of the two lists because no one should invest in an index with a sub-optimal Sharpe ratio when they know the better alternative unless their risk tolerance is lower than required by the optimal N.

things seem noteworthy. Then, half of the matrix except for the diagonal is identical to the other half but inverted. Then, the stocks are listed along the rows and columns of the matrix in decreasing order by market capitalization. This allows us to quickly access the specific variance-covariance matrix for each portfolio S&P N by extracting the first N rows and columns from the big one.

We proceed by constructing a column vector from this where each entry corresponds to the weight of each stock in descending order in the respective S&P N. We then multiply the transpose of this column vector by the specific variance-covariance and then by the column vector itself to obtain the monthly variance for each individual portfolio. We do that for all 500 portfolios for all six months. Last, the square root of that value gives us the monthly implied volatility (standard deviation) of each specific portfolio at any specified month. We store these into an object.

Finally, to calculate the semi-forward-looking monthly Sharpe ratios S_N^{iv} for each S&P N for each of the six months, we look at the respective actual monthly Sharpe ratio S_N and change the data in the numerator with one from the previous one. We are able to do that because we did not calculate actual monthly Sharpe ratios for August 2022. Last, we change the denominator by the respective value from the object we have just obtained. We store all the semi-forward-looking monthly Sharpe ratios in a new object. This is a good point to add our last methodological point of interest which we already mentioned. Namely, that the 'expected' returns are still historical. We already gave the reasoning behind that in section 1 of this paper, so here we concentrate on the assumption we are making. We assume, with regards to our estimate for 'expected' returns 1-month in the future, that there is heavy month to month momentum which is not always the case given that upcoming news announcements and earnings calls can make a discontinuous change to stock prices.

At this point we have a total of 6000 monthly Sharpe ratios (3000 actual ones and 3000 semi-forward-looking ones). We then compare the similarity of the rankings produced from ordering the Ns using these Sharpe ratios in descending order using a Spearman's Rank Correlation, and Kendall's Tau. We compare rankings instead of raw ratios because we could not confidently assume that the forward looking ratio and actual ratio can be compared on an apples to apples basis. The goal is to see if the forward looking measure can somewhat reliably predict which portfolio will have the best risk adjusted returns over the next month.

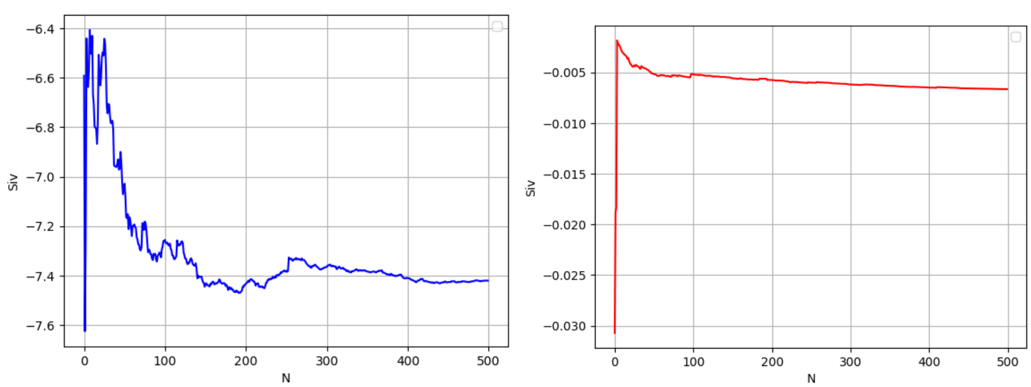
We are now ready to proceed with the revelation of our findings.

3 Results

We separate this section into two parts. First, we present several graphs to visually compare Sharpe Ratios. Then, we present two statistical results discussed at the end of the last section.

3.1 Graphical Comparisons

In particular, first we give 12 graphs in pairs for each of the six months we have considered. Again, the month range is September 2022 - February 2023. These are plots of the actual and semi-forward looking monthly Sharpe ratios for each S&P N where $1 \leq N \leq 500$. The reason we present the curves separately and not on the same graph is because of the scaling. We remind the reader that we are not concerned with the actual values of the two Sharpe ratios for each index, but only about whether the ordering by Sharpe ratios are similar enough, and in particular whether the best-performing Sharpe ratios occur at close values of N. On the graphs we want to look at where the peaks on each pair occur and to a lesser degree whether the two curves have similar shapes. Then we do the same for 12 more graphs, this time using the Sharpe ratios of the equal-weight portfolios.

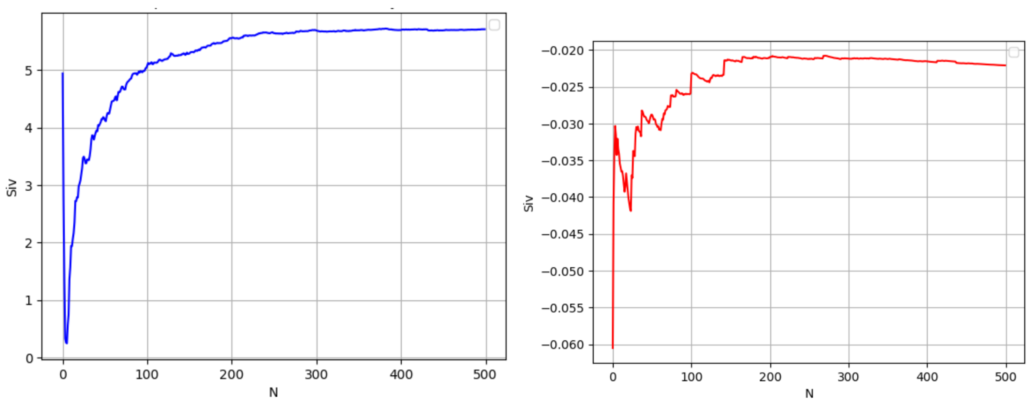


(a) Actual Sharpe Ratios Against N (b) SFL Sharpe Ratios Against N

Figure 1: September 2022

The 5[19] best performing values of N according to the actual monthly Sharpe ratios are: 8, 11, 4, 26, 27.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 4, 5, 6, 7, 8.



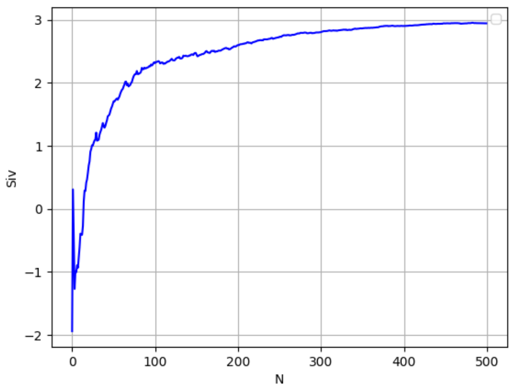
(a) Actual Sharpe Ratios Against N (b) SFL Sharpe Ratios Against N

Figure 2: October 2022

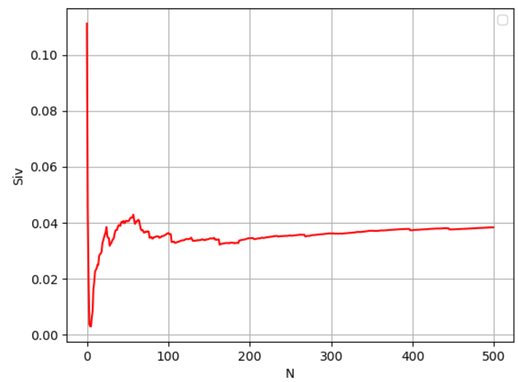
The 5 best performing values of N according to the actual monthly Sharpe ratios are: 383, 382, 384, 385, 379.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 271, 272, 269, 270, 273.

[19]Yes, the number 5 is arbitrary :D



(a) Actual Sharpe Ratios Against N

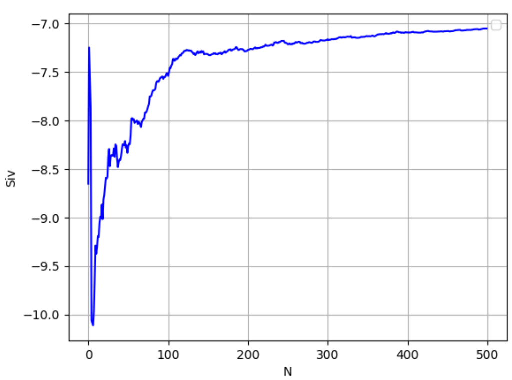


(b) SFL Sharpe Ratios Against N

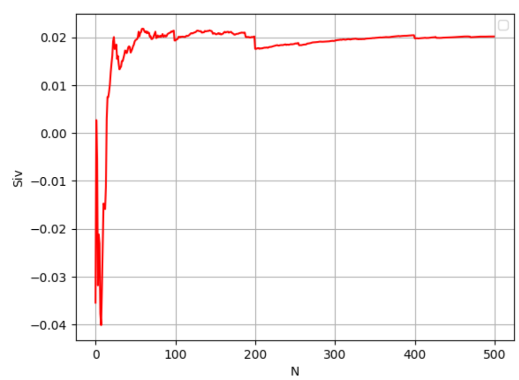
Figure 3: November 2022

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 483, 484, 482, 487, 485.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 1, 2, 58, 57, 55.



(a) Actual Sharpe Ratios Against N

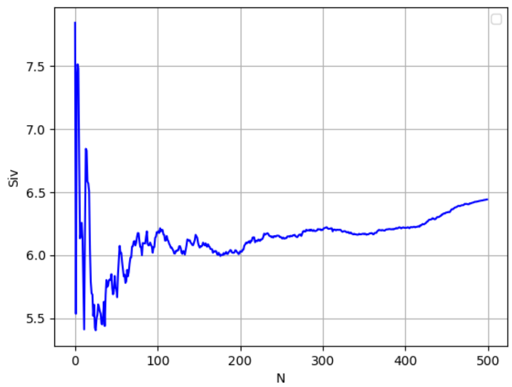


(b) SFL Sharpe Ratios Against N

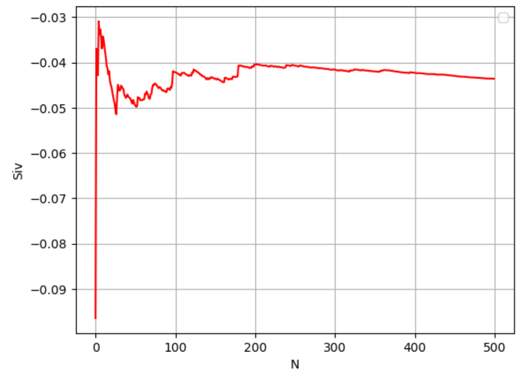
Figure 4: December 2022

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 499, 498, 500, 496, 497.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 431, 472, 436, 432, 475.



(a) Actual Sharpe Ratios Against N



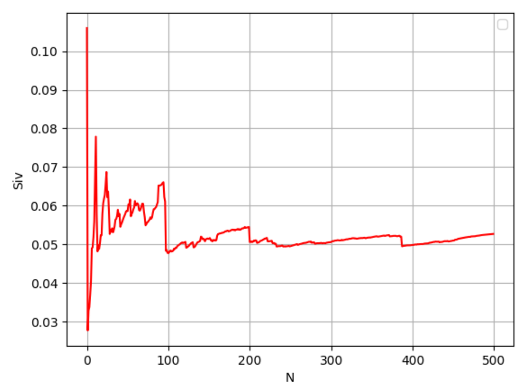
(b) SFL Sharpe Ratios Against N

Figure 5: January 2023

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 1, 4, 5, 3, 14.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 5, 7, 6, 10, 8.

(a) Actual Sharpe Ratios Against N



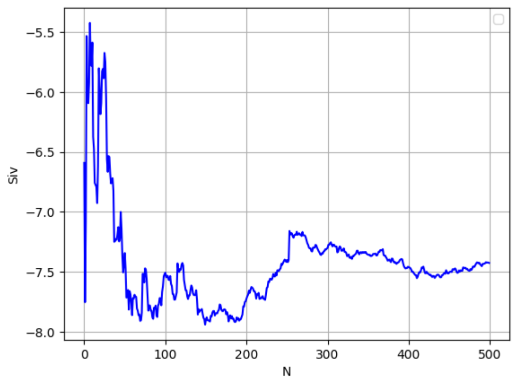
(b) SFL Sharpe Ratios Against N

Figure 6: February 2023

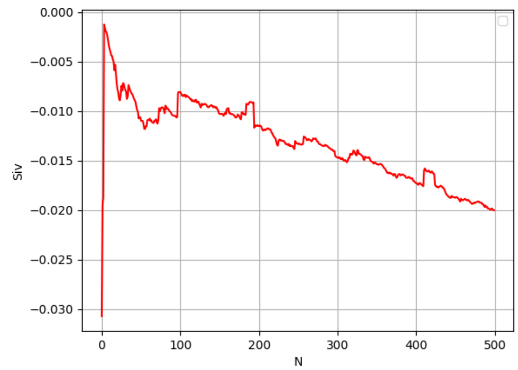
The 5 best performing values of N according to the actual monthly Sharpe ratios are:

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 1, 2, 7, 8, 9.

Next we do the same thing for the equal-weight portfolios: 1, 12, 25, 11, 95.



(a) Actual Sharpe Ratios Against N

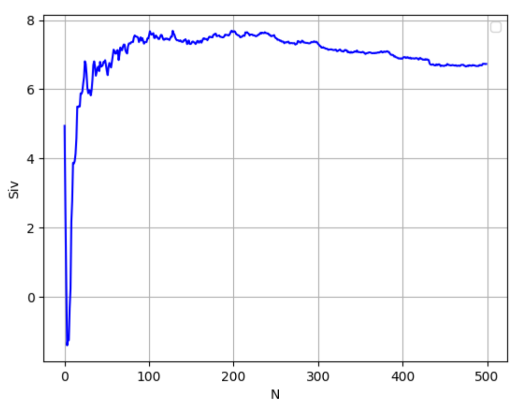


(b) SFL Sharpe Ratios Against N

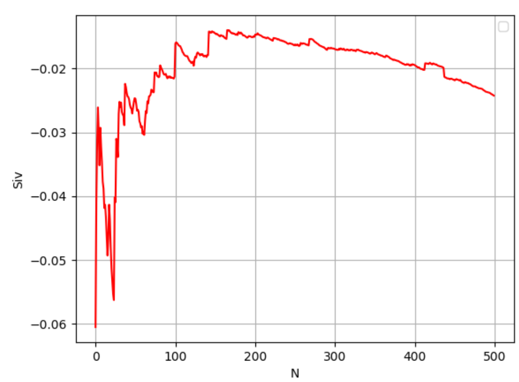
Figure 7: September 2022, Equal-Weight

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 8, 4, 11, 26, 27.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 4, 5, 6, 7, 8.



(a) Actual Sharpe Ratios Against N

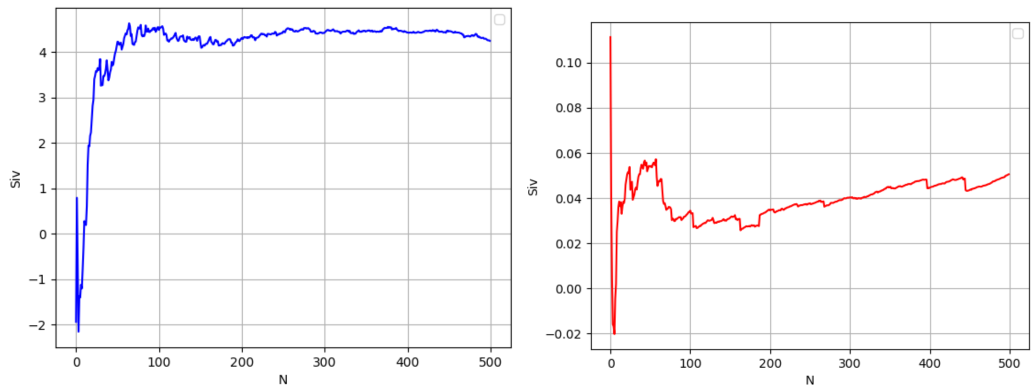


(b) SFL Sharpe Ratios Against N

Figure 8: October 2022, Equal-Weight

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 199, 129, 202, 198, 102.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 166, 168, 167, 169, 147.

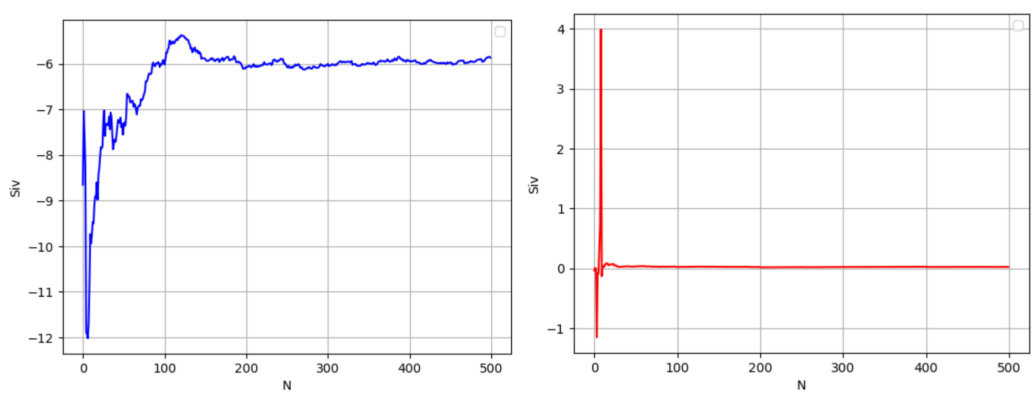


(a) Actual Sharpe Ratios Against N (b) SFL Sharpe Ratios Against N

Figure 9: November 2022, Equal-Weight

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 65, 79, 85, 105, 66.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 1, 58, 44, 43, 46.

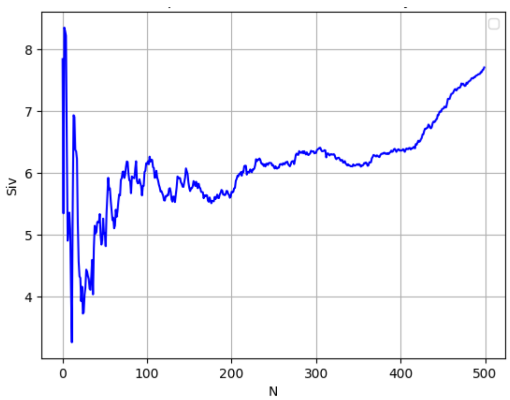


(a) Actual Sharpe Ratios Against N (b) SFL Sharpe Ratios Against N

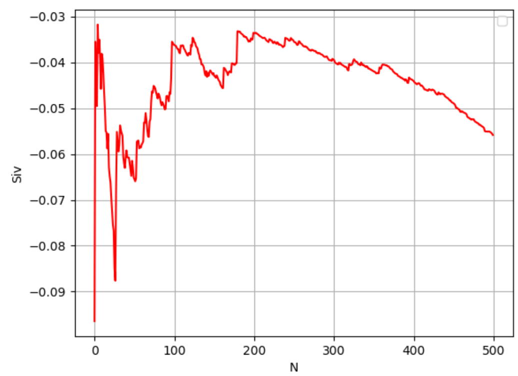
Figure 10: December 2022, Equal-Weight

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 121, 122, 123, 124, 125.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 9, 8, 7, 17, 15.



(a) Actual Sharpe Ratios Against N



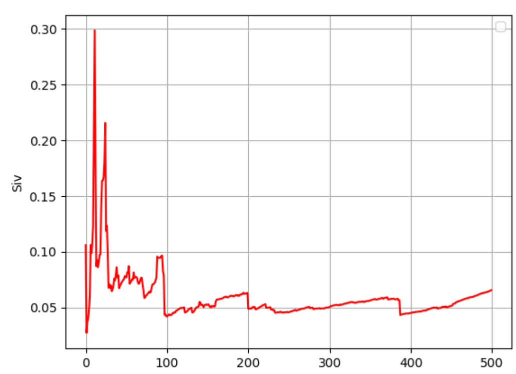
(b) SFL Sharpe Ratios Against N

Figure 11: January 2023, Equal-Weight

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 3, 4, 5, 1, 500.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 5, 181, 180, 183, 182.

(a) Actual Sharpe Ratios Against N



(b) SFL Sharpe Ratios Against N

Figure 12: February 2023, Equal-Weight

The 5 best performing values of N according to the actual monthly Sharpe ratios are: 7, 1, 8, 9, 2.

The 5 best performing values of N according to the semi-forward-looking monthly Sharpe ratios are: 12, 25, 11, 24, 13.

3.2 Statistical Results

Market-Cap	SEPiv	SEPact	OCTiv	OCTact	NOViv	NOVact	DECiv	DECact	JANiv	JANact	FEBiv	FEBact
Weighting	4	8	271	383	1	483	431	499	5	1	1	1
for S&P N	5	11	272	382	2	484	472	498	7	4	12	2
	6	4	269	384	58	482	436	500	6	5	25	7
	7	26	270	385	57	487	432	496	10	3	11	8
	8	27	273	379	55	485	475	497	8	14	95	9
	9	9	204	380	56	479	476	495	11	15	94	11
	10	24	274	381	54	492	473	494	12	6	93	12
	11	10	205	376	64	481	433	487	9	16	24	13
	12	19	275	375	59	462	484	488	2	17	92	10
	13	25	206	378	53	486	471	489	13	18	91	14
Average S&P "N"	8.5	16.3	251.9	380.3	45.9	482.1	458.3	494.3	8.3	9.9	53.8	8.7
Average Top 10 Percent Difference Normalized to 500 Rankings		1.56		25.68		87.24		7.2		0.32		-9.02
Overall Average % Difference		21.8										
Average % Difference w/o Outlier Months		4.5										

Figure 13: Top 10 values for N Ranking Comparison

Equal	SEPiv	SEPact	OCTiv	OCTact	NOViv	NOVact	DECiv	DECact	JANiv	JANact	FEBiv	FEBact
Weighting	4	8	166	199	1	65	9	121	5	3	12	7
for S&P N	5	4	168	129	58	79	8	122	181	4	25	1
	6	11	167	202	44	85	7	123	180	5	11	8
	7	26	169	198	43	105	17	124	183	1	24	9
	8	27	147	102	46	66	15	125	182	500	13	2
	9	9	143	201	55	377	16	120	203	499	23	11
	10	10	144	203	56	379	24	118	200	498	22	12
	11	19	149	200	54	100	23	119	184	496	21	10
	12	24	170	217	40	104	22	127	201	497	20	13
	13	23	180	238	57	76	14	126	204	495	27	14
Average S&P "N"	8.5	16.1	160.3	188.9	45.4	143.6	15.5	122.5	172.3	299.8	19.8	8.7
Average Top 10 Percent Difference Normalized to 500 Rankings		1.52		5.72		19.64		21.4		25.5		-2.22
Overall Average % Difference		12.7										
Average % Difference w/o Outlier Months		3.2										

Figure 14: Top 10 values for N Ranking Comparison, Equal-Weight

After running statistical tests on all 6 months of data, we calculated the Spearman's rank correlation coefficient for each month's descending rankings of values of N by their associated actual and SFL monthly Sharpe ratios. We found an always positive correlation coefficient that ranges from about 0.2 in January 2023 to about 0.9 for the December 2022 predictions versus the actual Sharpe ratio rankings. The month we have the highest overall correlation (December) is the only month where the S&P 500 outperformed us (barely), which reinforces the idea that while the Spearman's rank correlation coefficient test is meaningful, we are mostly concerned with how good our SFL Sharpe ratios are at predicting the best performing values of N.

With some monthly exceptions, our top-5 and top-10 S&P N predictions do well to track the eventual realized Sharpe ratios by the constructed portfolios. In all but one of the months examined, our top pick according to the SFL monthly Sharpe ratio outperforms the S&P 500. In half of the top 10 ranking comparisons over both the market-cap weight and the equal-weight portfolios, our top prediction was in the top 7 of realized Sharpe ratios. See Figure 13. More broadly, even for an incipient model and a short range of testing data, we see strong agreement in the relative magnitude of predicted and realized Sharpe ratios - i.e. for September 2022, the average N of our top 10 ranked predictions was 8.5, and the realized average top 10 Ns was 16.3, while in December predicted was 458.3 and actual 494.3.

We compared these top 10 average rankings over a scaling of 500 total rankings for each of the month pairings. For market-cap weighted portfolios our top 10 rankings were only 21.8% different on average from the eventual rankings, and without the outliers of October and November (where the S&P 500 happened to gain 8.8% and 4.4% respectively), our top 10 ranking predictions come within 4.5% of the eventual ranks. On the scale of 500 possible rankings, this is an incredibly close agreement given the coarseness and breadth of our data.

Even if our exact top pick(s) don't perfectly line up with the eventual Sharpe ratios for whatever various reasons that might stem from our assumptions and execution of the portfolios and natural/unpredictable market events, we are able to suggest a general bounding region of N values that are very likely to be highly performant. Exceptions, of course, include some of the rankings of October and especially those of November 2022, but by and large, our market-cap weighted portfolio predictions track with the market realizations. Because our predictions rely on historical data, for months where trends find strong reversal, our predictions are weaker in accounting for future risk-adjusted return estimates. Notably, we generally tend to outperform the S&P 500 and predict successful portfolios for months where the S&P 500 has a weak gain or loss over the month (i.e. more stable), and these market-cap portfolio estimations glean us less accurate predictions and returns over months where the S&P 500 has strong growth.

Our equal-weighted portfolios performed similarly – running the same average top 10 ranking differential against the realized Sharpe ratio rankings, overall resulting in only a 12.67% difference from our predictions over the months examined (See Figure 14). These portfolios were stronger in September, October, and February (3.15% average difference from realized) and weaker in November, December, and January (22.18%). These mostly line up with the prediction strengths and weaknesses of the market-cap weighted portfolio performance of the top 10 Ns, but December and January are better predicted by market-cap weighted, while October is better predicted by equal-weight formulation.

3.3 Conclusion

Through graphical interpretations and correlation analysis of SFL versus actual Sharpe ratios, we have concluded that our predicted portfolios outperformed the S&P 500 almost consistently. Despite some particularities in our approach, our results prove that a retail investor would have been better positioned to invest in our recommended portfolios as compared to the S&P 500 for at least the six months we observe. Our data, however, was limited to six months so the magnitude of the correlation might be influenced by the specific time window we are observing.

In the future, we might try a combination of equal and market-cap weighted portfolios, since for a few of our testing months, the equally-weighted portfolios were more accurate in their top 10 predictions than our initial market-cap weighted portfolios. We might also consider cross-predictions of portfolio rankings – i.e. using equal-weight portfolio performance to infer potential market-cap portfolio performance, and of course vice versa.

For our forecasting, we used the market-cap weights from the previous month to determine the portfolio weights. In theory, an investor, at the start of a new month, could reasonably be expected to have new market caps and thus could reweigh the portfolios in accordance with the new information. We started converting our prediction dataframes, but ultimately ran out of time with long code runtimes for each month to fully explore the possibilities of updated weighting methods. Potentially a linear regression or some backward-looking weighting method for finding the best combination of historical and current weights to forecast the best future weights for the portfolios could be employed.

We are encouraged by the general agreement of both implied and actual top-5 and top-10 rankings, Spearman tests, and general shapes of the S&P N Sharpe Ratio curves – given our early testing of a modest time frame with a flat market, we find decent concordance between the two – using prior month returns and option data implied volatilities as a proxy for next-month risk-adjusted performance.

4 References

[1] CBOE. White Paper CBOE Volatility Index. https://www.sfu.ca/~poitras/419_VIX.pdf/, 2019.

[2] CRSP. Wharton Research Data Services. https://wrds-www.wharton.upenn.edu/pages/get-data/center-research-security-prices-crsp/, 2024.

[3] OptionMetrics. Wharton Research Data Services. https://wrds-www.wharton.upenn.edu/pages/get-data/optionmetrics/, 2024.

[4] William Sharpe. The Sharpe Ratio. https://web.stanford.edu/~wfsharpe/art/sr/sr.htm, 1994.

[5] YCharts. 1 month treasury rate. https://ycharts.com/indicators/1_month_treasury_rate/, 2024.