Modern Search Techniques Applied to Equity Portfolio Construction

January 24th, 2017

Nathan Weinstein

**Abstract**

Brin and Page (1998) demonstrated that a system which ranks search results utilizing “PageRank” was superior (yielded more relevant, robust results) to prior commercial systems (which, amongst other methodologies, ranked results based on keyword frequency). PageRank determined relevancy (i.e. the order in which to present results) using 1) the number of inbound links to a website, and 2) the value of those links, based on the rank of the in-linking webpage. The equity portfolio construction method described and back-tested here considers a manager’s holding of a stock (per 13-F filings) as a “link” to that stock, and further, weights the link(s) based on the manager’s relative recent track record, and position sizing relative to assets under management (“AUM”) as a proxy for relevance (confidence). Using iterations of this simple process, a concentrated portfolio of stocks can be built from amongst a constituent set of managers, that captures the relevancy of the stocks in consideration, ranks them accordingly, and in a back-test provided respectable (top quartile) returns vs the set of constituent managers, capturing a large degree of the information value generated by the active managers without incurring their active management fees.

**1. Introduction**

Google managed to capture an unprecedented share of the consumer internet search market, and achieve an enormous market capitalization by leveraging what, in its essence, is a fairly intuitive and simple iterative process that utilizes a web crawler to aggregate all the backlinks between web pages on the internet, and, considering that network in its entirety, to rank which pages are most relevant to users on the internet based on the number of inbound links to those pages, as well as the value of those links, where the value of each specific link is determined using the rank of the in-linking website.

The more inbound links a webpage has, and the more value placed on those links, the higher the rank of that page and the more valuable its out-bound links will be. As a prerequisite to this paper, I suggest reading “The Anatomy of a Large-Scale Hypertextual Web Search Engine” written by Sergey Brin and Lawrence Page in 1998, particularly section 2.1.1 “Description of PageRank Calculation.” Brin and Page called PageRank “an objective measure of [a website’s] importance that corresponds well with people’s subjective idea of importance.”

The idea of the “modern search techniques applied to equities portfolio construction,” is to map the concepts from PageRank to the financial markets, specifically to equities. In order to accomplish this, the requisite pieces are 1) a stock [viewed metaphorically as a website], 2) an institutional asset manager [viewed as the progenitor of a backlink], and 3) the manager’s holding of a stock [viewed as the backlink]. We build up the relevant internet by pulling together a number of different managers (“the constituent managers”) which have to be A) different enough so that they’re not all exactly the same (e.g. so differences between their holdings and AUM are noticeable), and B) similar enough so that there’s overlap in style and capitalization range such that insights gleaned from each constituent can be aggregated and parsed without unmanageable dispersion.

13-F filings are a legal requirement in the United States for all SEC registered investment firms who manage $100 million or more. 13-F filings show a static snapshot of these firm’s holdings on a single day 45 days before the filing. The filings are required to be made four times a year. The filings reveal, for long-term oriented funds (e.g. those with low turnover), a reasonable proxy of their investment portfolio. There is of course a meaningful degree of error, given intermittent trading activity that often takes place before the filings are made public, but it is the best way going to see what major fund groups hold in their portfolios.

Gauging the amount of money an asset manager has in a specific equity position over the total amount of money that asset manager has under management can be considered a proxy for the fund’s confidence in that position, i.e. that position’s relevance to that fund. For example, if Warren Buffett had $1.00 under management, and put $0.80 of it in GEICO and $0.20 of it in Coca Cola, you would be able to determine which company he was most confident in owning at that point in time.

The portfolio construction system described in this paper posits that active management can add value (achieve alpha) using methods which include, inter alia, better-than-consensus modeling, garnering actionable insights from management face time, site visits, scuttlebutt, channel checks, etc. Assuming those alpha-production activities are ultimately reflected in relative position sizes within a manager’s portfolio (i.e. that an active manager doesn’t size its positions randomly, but does so logarithmically according to their relative confidence in each equity), the system described here leverages PageRank-derived ranking methodology, applying it to 13-F filings (available essentially for free, or at a low cost from SEC’s EDGAR), to construct a model portfolio from a set of constituent portfolios, and this model, with its positions weighted based on the ranking system methodology’s determination of the relevant confidence(s), was found to have outperformed the majority of the constituents and the index benchmark in our back-test, allowing investors to gain access to active alpha for a fraction of the management fee. In a way, our portfolio was akin to a thriving (valued added), yet sycophantic leach (leveraging other’s work without paying a management fee to them).

**2. Method**

Step 1

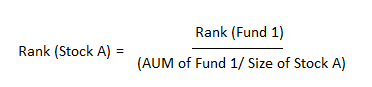
Begin with a selection of funds, preferably within the same category or ones you know will sometimes have some overlap. Choosing as ones constituents, for instance, one large cap growth fund, and one micro-cap value fund, will make it unlikely you’ll be able to aggregate position insights across the constituents – simply put, they won’t hold any of the same stocks. Because we wanted this to be a practical exercise, we chose a group of nine small-cap funds (nod on the return to size from Fama-French), which had been selected ex ante as top performers in their respective category. Note this is not a prerequisite for the method – any group of funds can be chosen, or the whole universe of funds can be chosen, but the fewer and more concentrated the group, the more likely to induce beneficial tracking error to the applicable index or benchmark.

Step 2

Collect 13-F filings for the funds. Determine the fund’s AUM by summing the disclosed positions’ values.

Step 3

The basic ranking calculation formula is:

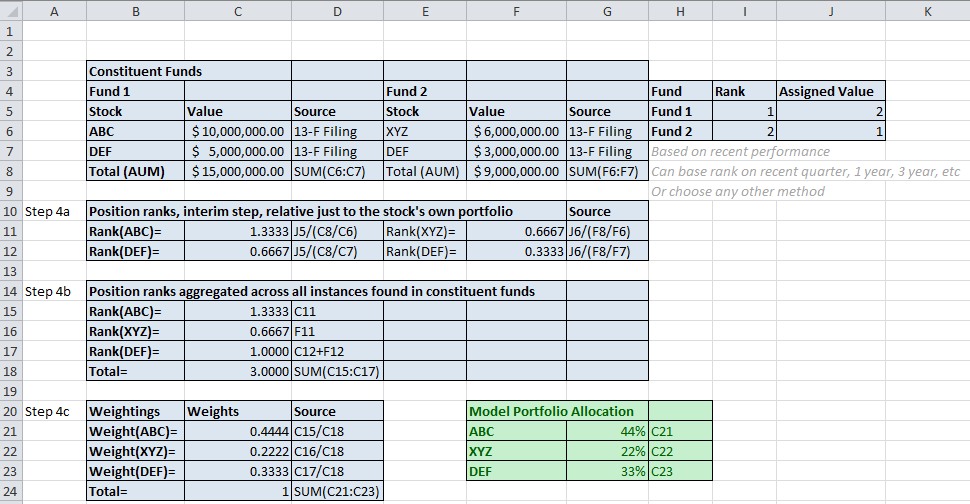


*\*Note the rank is the inverse to the total number of funds. For instance, in our selection of nine funds, the first ranked fund is assigned #9 for rank, the second #8, and so on.*

Step 4

Here we illustrate the key steps to building the model portfolio from the constituents, using a simple, hypothetical example of two funds with two positions each.

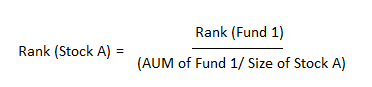
1. Determine equity position ranks using the ranking formula (from step 3)
2. Aggregate all the ranked positions and determine the denominator for the next weighting step, which is merely the summation of the aggregated ranks
3. Determine the weights by dividing all the aggregated position ranks by the denominator

**

Step 5

Each time new 13-Fs are filed, the fund automatically should rebalance to the new weights. The easy way to deal with de-listings (e.g. through bankruptcies or acquisitions) is to just exclude them going forward, and replace that stock with the next one after it in the ranking. This is true for attrition amongst the constituent funds; if one drops out just stay calm and go with the next best option in the category, if none exists, just keep doing the same process with fewer constituents.

1. **A Closer Look at the Ranking Formula**



The ranking formula is the key to the process, and it should appear very intuitive. The rank of a stock (in this case “stock A”) should go up if the numerator is higher, and the numerator is higher if the fund is ranked higher, in which case it – using the reciprocal method whereby rank #1 out 9 is assigned “9” and, for instance, rank #3 out of 100 would be assigned “97,” – would get the highest cardinal number of the constituents, and on down.

Furthermore, the smaller the denominator, the larger the fraction, and the higher the rank of stock A (and the more of it our end-result fund will own). The denominator will be smaller if the fraction in parentheses there is larger, which will happen if “fund 1” which holds “stock A” has more AUM relative to the size of its position in “stock A.” That’s it.

1. **Implementation**

In order to implement the model described in this paper, we selected a set of nine constituent funds. These were small-cap funds which had been selected by an internet website as having been best-in-class performers over the last three or five year period. Both of those things (who ranked them and why, and what style and cap range they traffic in) are arbitrary. Whoever implements this model should have a method to determine which group of funds they want, and how they want to rank them. In our case, we wanted small-cap exposure and those with good multi-year track records, with their rankings based on their relative performance over a meaningful number of years.

Because this was a back-test, we were able to collect the 13-F data for the entirety of the tested period, in this case we pulled seven filings, 12/31/13, 3/31/14, 6/30/14, 9/30/14, 12/30/14, 3/30/15, and 6/30/15. Nine funds and seven filings equates to 63 documents we had to download; we used FactSet to facilitate this collection, but there are other low-cost methods, such as through EDGAR.

Once we had the filings, we put them into Excel, and essentially ran through the steps demonstrated earlier in this paper. Note that the complexities of referencing nine mutual fund, each with hundreds of positions, is slightly more involved than the two fund/ two position example above, but the steps are the same.

In practice, i.e. implementing this system in real time, once the constituents are chosen and ranked, the fund will need to pull-down the new 13-Fs each time they become available, and re-rank the positions and the funds. The funds can be re-ranked based on their performance.

There are a few things you have to live with. One, you will never know what these funds own in real-time; the lower the typical turnover of the constituent group, the more often more closely the 13-F filing will be a realistic snapshot of the fund. In this regard, we do suggest a constituent group of long-only (13-F filings only show you the long side a book), long-term oriented (e.g. long holding periods with little turnover) funds. Two, you’ll probably be calculating each constituent fund’s performance based on pricing data and none of this will be exactly accurate because you’re seeing a stale snapshot of their holdings and can’t know exactly how they bought and sold. Finally, you’ll be buying up stocks after the constituent funds have already gotten in. It’s possible that in the intervening time the stock has move drastically. Knowing things about the funds helps, as does a basic rule like, “don’t add any position if it has moved +10% in the last 30 days,” for instance; but knowing that the funds in question are buying stocks for years not quarters is helpful and makes their positions relevant to those who have to wait until the new filings to see what to buy and what to sell.

In our model fund, we culled the list of hundreds of potential stocks down to 20. A concentrated portfolio of 20 stocks gives nearly the full benefit of diversification in MPT, without an undue number of satellite holdings and their associated trading costs.

1. **Performance**

We found that, for the period in question (early 2014 to late 2015), our fund returned **10.94%**. Of the nine funds in question, only one exceeded us, with an 11.29% return, the next highest was 8.19%, and the rest were 5.15% and lower. This return does include a trading cost we imposed on ourselves. No tax was considered, as it varies by individual fund holder.

