Using Data to Capitalize on Behavioral Finance

December 12, 2016

Wenzel Analytics Inc

For almost twenty years I have been downloading *Stock Investor Pro* (SIP) data and looking for what combination of variables, and what values for each, are most associated with consequent gains and losses. What should I buy and when should I sell? With both weekly and monthly data, I dutifully export into Excel the 80 or so SIP variables I have found to be most fruitful. I then append the data for each week or month to a large database and attach to each row the returns for four weeks, 13 weeks, 26 weeks and 52 weeks. Recently I was working with over a million rows in a file of monthly data over the last 12 years or so. The file is then imported into a powerful statistical tool called KnowledgeSEEKER.

Almost every time, the most statistically significant variable is the week or month associated with the data. Not only is aggregate price change by month, week or day highly correlated with individual stock price variations, but we can observe it every hour and minute on technical charts. Of course this should not be a surprise as most of us know that the market accounts for roughly 80% of a stock's price variance. But how do we explain this common observation, and how do we capitalize on it? Is there some kind of ether that simultaneously influences all buyers and sellers?

My own theory is that news of any kind is not pervasive enough by itself to effect this time period correlation between most stocks. I'm more inclined to look at the impact that index and benchmark buying have on the markets. John Bogle and publications such as *ETF Report* continually make the case for passive investing over active management. Research has shown that the average active investor does not beat the market. Even if the probabilities of beating the market are 50%, minus fees, most of us went through school scoring above average and presume that we can do better than average. I happen to think an individual investor can more likely beat passive indexes with low liquidity stocks than by buying large-cap stocks and trying to go up against the sophisticated algorithms of institutional and high-frequency traders.

Even if the individual investor chooses the greater reliability of buying indexed products, do the indexed and benchmarked products themselves impact and distort the market? Do we have an instance of Tragedy of the Commons where something is advantageous for each individual, but harmful in its overall effect? Is it like each of us enjoying the convenience of driving our cars while the global impact is destructive to our planet's viability? The index products that might be advantageous for each individual investor may be damaging in their overall impact on markets.

The markets are an auction, balancing sellers and buyers through supply and demand. If we look at any other auction familiar to us such as estate sales, art auctions, boat auctions, cattle auctions or even Ebay, wouldn't it "ruin" the bidding process if a third or more of the bidders are buying everything regardless of price? If everything is bought or sold regardless of price, that defeats the purpose of an auction to establish a proper valuation for each item based on the balance of supply and demand, sellers and buyers. When an index fund buys or sells all stocks in the index without weighing the fundamental valuation metrics of each stock, the buying or selling pressure is not related to the merits of each individual stock.

I'm suggesting that most stock prices are primarily driven by their correlation with all other stocks, and that extreme correlation between stocks is driven by large-scale index and benchmark buying. How else can you explain the lockstep price movements?

Lee Wenzel
(952) 944-2699
Lee@WenzelAnalytics.com
www.WenzelAnalytics.com

Wenzel Analytics, Inc.
Registered Investment Advisor
8666 Westwind Circle
Eden Prairie, MN 55344

When I ask speakers or authors advocating passive investing about the impact of index products on markets, I usually get an inference that it is an absurd question. One speaker said that the market could be 85% passive products before it would have an effect on the market.

There is so much intermingling and overlapping of investment vehicles, that I can probably get General Electric (GE) stock in any of 50 or more indexed products. Or inversely, by buying shares of General Electric I'm buying into all of these passive products. It is all one interconnected tangle.

Behavioral finance enumerates all the ways we make irrational decisions based on emotional or personality factors which do not optimize our returns. If individuals have an irrational tendency for the assurances that their returns would nearly match the market, and that distorts the market, how does that create opportunities for the knowledge-based and systematic investor? If we know that investors irrationally avoid a portfolio that has annual returns of 30%, but 30% of the positions have prices that fall showing a loss and 3% go bankrupt and are a total loss, what opportunities does that open up for the quantitative investor? How can we be like the casino owner and use the laws of large numbers to our advantage?

We know that volatility is more persistent than price. Stocks with prices going wildly up and down will continue to have wide price swings. Even the Black Scholes options formula is built on historical volatility. A stock cannot go up without volatility, but most investors focus on volatility being a negative for stocks because it can also cause stock prices to decline. So why are we so afraid of volatility and gravitate to the low-volatility factor products such as iShares Edge MSCI Minimum Volatility ETF (USMV)?

Methodology

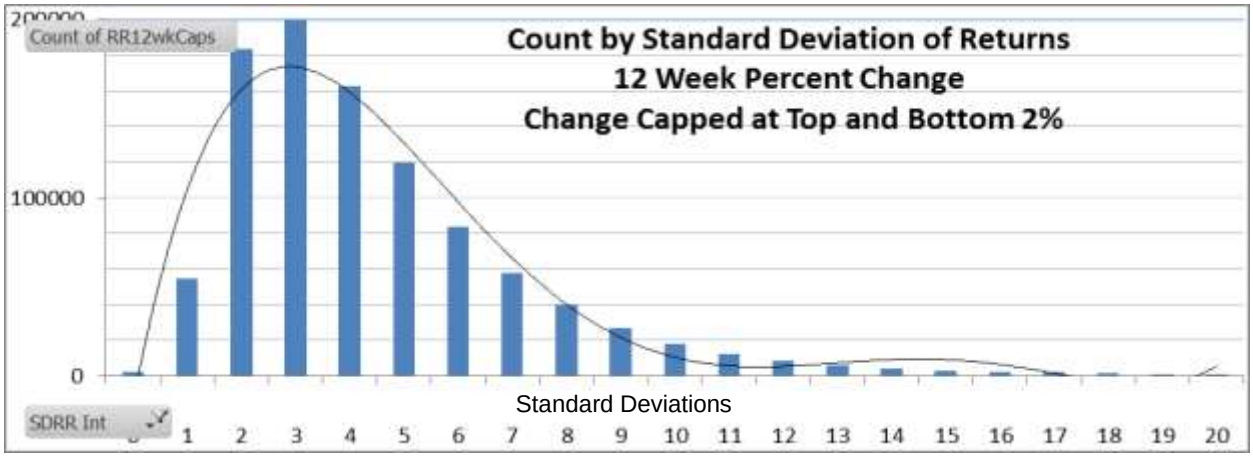
What follows are some data explorations designed to answer these questions. I will not provide a specific recipe consisting of exact criteria for you to capitalize on how the dynamics of markets are driven by indexes and the foibles of naïve investors. My experience is that any such announcements are arbitraged away before I can benefit. Besides, I tend to not trust someone else's research and recommendations as much as I trust my own crunching of numbers, analysis and conclusions. What I hope to show here are interesting explorations into relevant data that might encourage you to travel through similar terrain and find your own discoveries.

I started out looking at returns by market volatility and correlations between stocks. I will also present data on returns by price and momentum. I didn't go into the analysis in order to disprove a null hypothesis or to argue a specific conclusion. Instead, the methodology was that of an exploration to discover facts of interest. Some of the discoveries can be explained; most cannot, but even so, one can construct portfolios based on data and discoveries rather than intuition and behavioral shortcomings. The data teases us with all kinds of possible explanations and implications, most of which will be left to you the reader.

To look at volatility, I took a list of about 4,000 tickers from *Stock Investor Pro* (SIP) having a price greater than one dollar and having minimal liquidity. (Minimal liquidity was measured by average 10-day volume greater than or equal to 5 in the SIP data and dollar daily average over 3 months greater than or equal to 500 as found in SIP.) Then using Excel and the XLQ add-on, I calculated each stock's standard deviation of the price change rate over rolling 25-week periods going back six years. Of course, not all current tickers have a six-year history and I don't know what happened to the stocks that disappeared and thus did not make my current list of tickers. We have an unknown survivorship bias. Companies could have gone bankrupt, been acquired or reorganized under a different stock symbol.

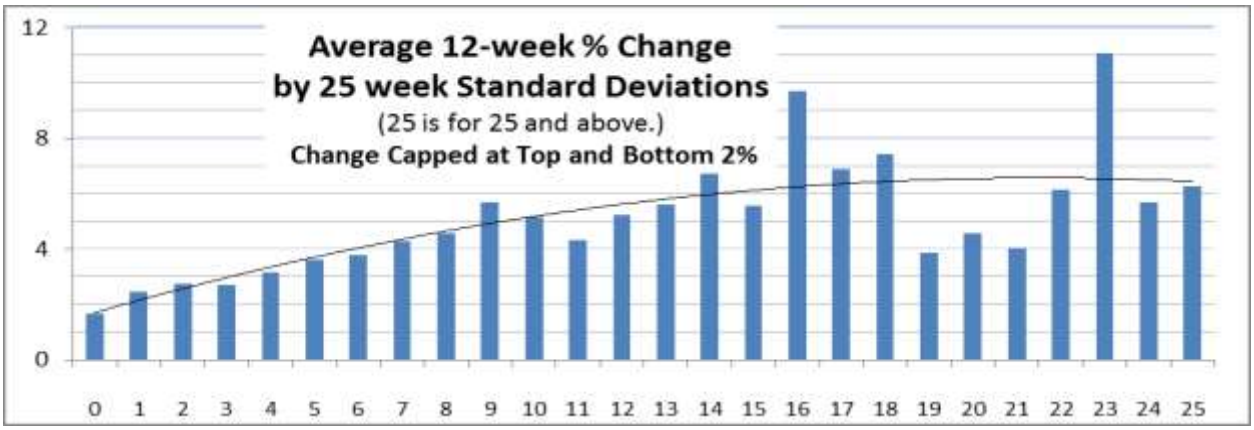
Volatility

Looking at the percent of price change over 12 weeks, the distribution of standard deviations looks like this:



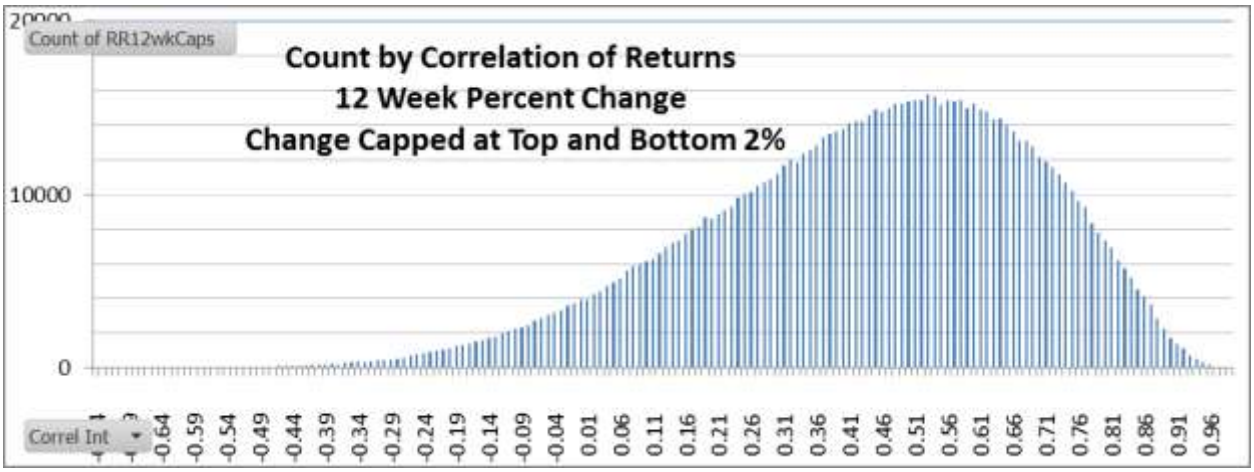
While there are extreme standard deviations above 15, they are very rare. The extremes in returns (the top and bottom 2%) are capped at that level to avoid the inordinate influence of very high and low returns on the average calculation.

The chart below collapses all the standard deviations of the rate of return above 25 into the 25 column. The standard deviations go up in a rather linear curve through a 9% price change over 12 weeks. Price changes beyond 9% over 12 weeks continue to be high as volatility increases, although the data are not very meaningful because of the low counts.



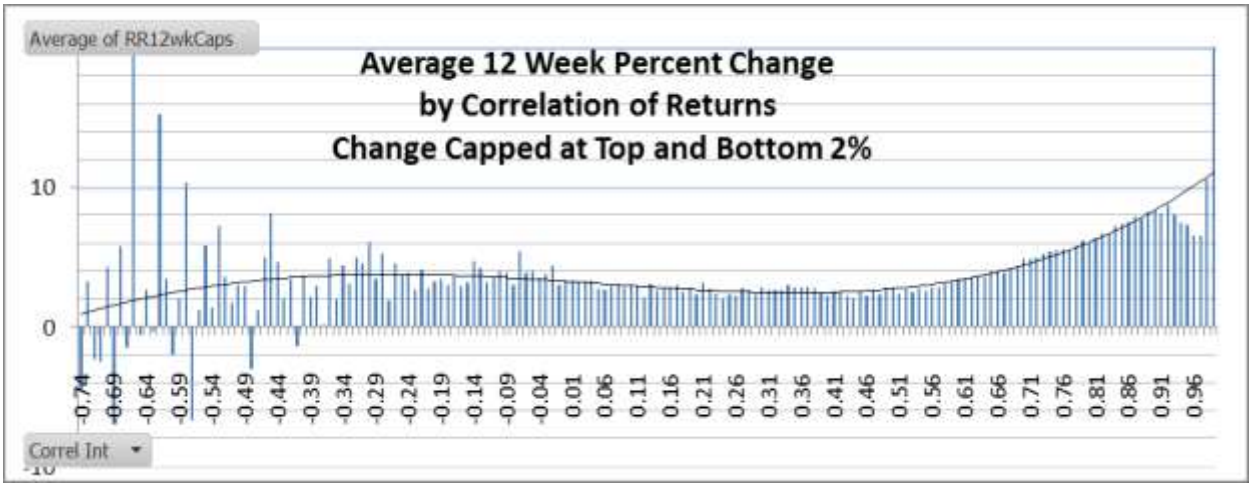
Correlation

Using the same data set of rolling weekly data, I did a correlation in Excel between each stock's percent price change and the average percent price change of all stocks. The distribution is shown in the first chart below.

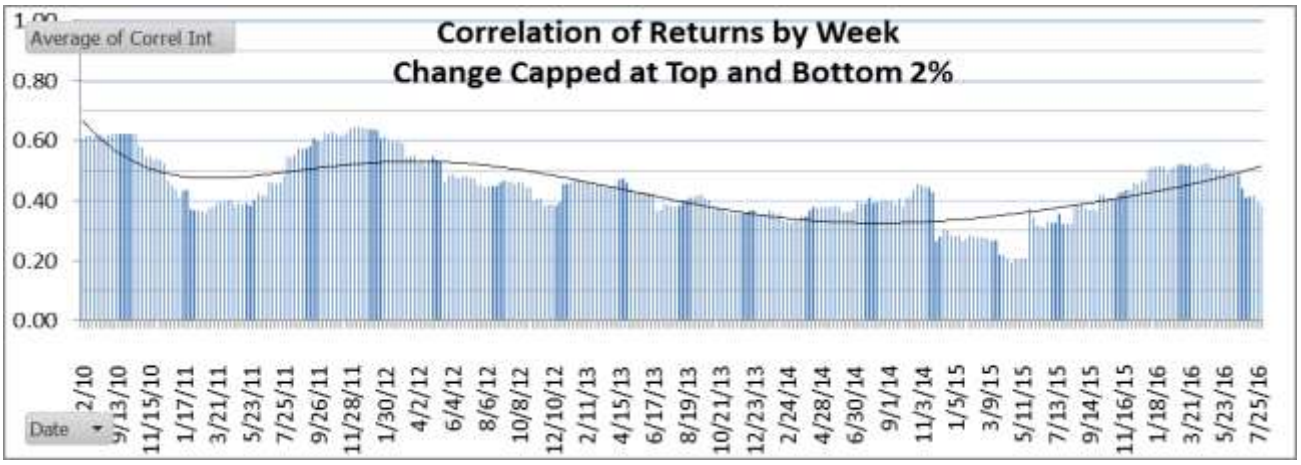


The correlations make a very smooth but skewed bell-shaped curve. The median correlation is about 0.51.

The percent change in price (Rate of Return) 12 weeks out is in the chart below.



The extremes and wide variations are where there are very few records. One can see that there are somewhat higher returns for correlations less than 0.17 and then significantly higher returns above 0.71. While there are a lot of stocks with correlations above 0.71, the outsized returns are in sync with market patterns. That is an area to explore if one is open to dependence on market swings. Using monthly SIP data supplemented with XLQ correlation data, I did not find correlation to be a useful variable in isolation. In general, with weekly data I found that correlation was not a useful tool as the primary or root variable, but helpful in its interaction effects in conjunction with other variables.

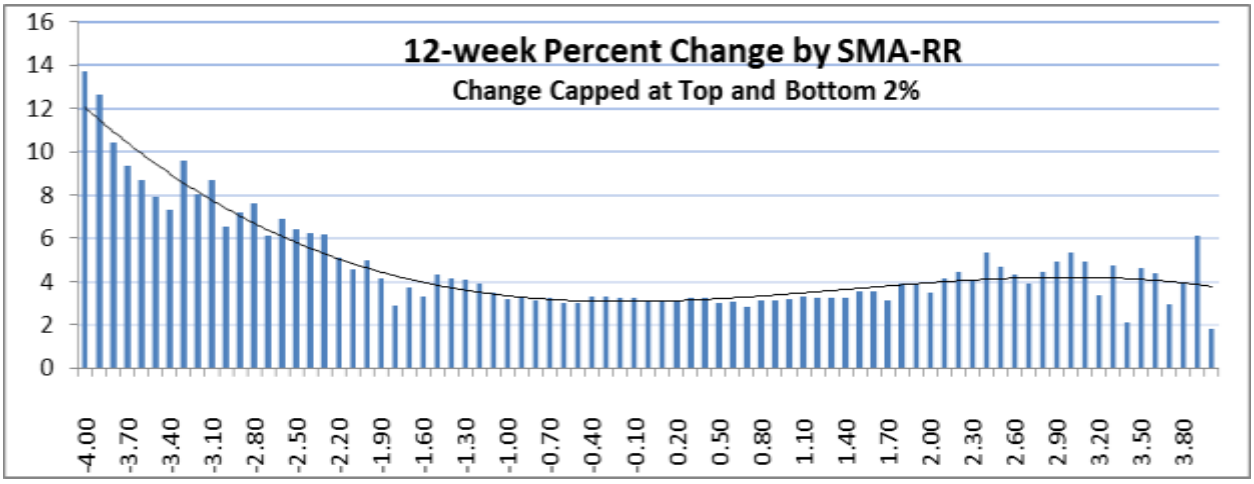


What accounts for the wide variation at different times in correlation between individual stocks and the average of all stocks?

Momentum or Reversion to the Mean?

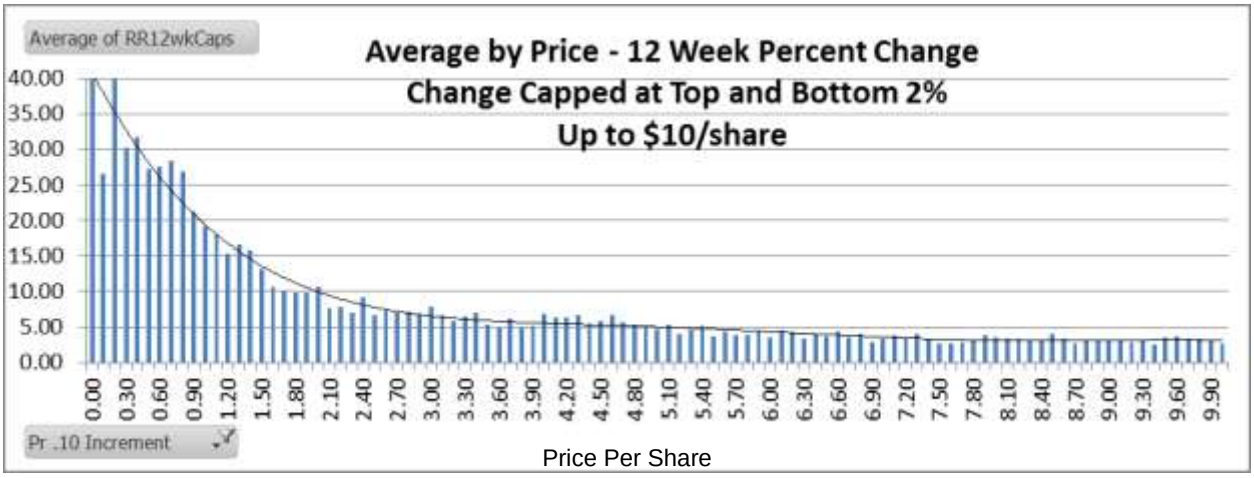
Momentum is often listed as a factor contributing to outperformance, along with quality, small-cap, low volatility and dividends. I have seen reports showing higher returns for stocks with low share turnover as a factor as well, but have not been able to replicate their results using their definitions. However, I have found very promising results from my SIP database using "Volume-Dollar Daily Avg 3 month" and some other variables, and have invested in a set of stocks using the screen.

To measure momentum, for each stock and each week I calculated the average of the percent change in price each week over the preceding 25 weeks. This gives us simple moving averages (SMAs) of the preceding rate of return. As shown in the chart below, I found that an SMA's-Return Rate (SMA-RR) above 1.5% weekly over 25 weeks have a slight increase in future 12-week returns. However the effect is minor compared to the pronounced return improvements found for reversal stocks having a SMA-RR below -1.00%.



Price

In the early 2000's I was able to make outsized returns for myself and my clients using low priced stocks. It then became a non-factor or non-productive for a number of years. Now it appears to be back.



To give perspective, the trendline crosses the 5% every 12 weeks (vertical axis) at about the $5.00 point (horizontal axis). So that is roughly 20% a year over six years (generally speaking, there are four 12-week periods in a year; 4 x 5% = 20%). Stocks priced at $1.00 averaged a 19.07% change every 12 weeks which comes to approximately 75% returns. Of course we have a survivorship bias, not knowing what happened to the stocks that disappeared. The data are also subject to distortion by reverse splits, such as when a stock splits 1:4 and the price goes from $1 to $4.

Conclusion

I have four portfolios that refine the interaction effects between price, correlation and standard deviation. They are based on quantitative data rather than on investors' intuition of quality or safety. After only two months, and markets responding to the election during those two months, it is too soon to draw conclusions. The four portfolios are ahead of the Russell 3000 benchmark.

I hope my thoughts and explorations will stimulate your thinking and research. I welcome feedback that supports, refines or negates what I have suggested and found.